# Robust contracting under double moral hazard*

Gabriel Carroll[†] ⓡ Lukas Bolte[‡]

September 11, 2020

**Abstract**

We study contracting when both principal and agent have to exert noncontractible effort for production to take place and the principal is uncertain about what actions are available to the agent. Any contract is evaluated by the guaranteed expected payoff for the principal no matter what actions the agent may take. Both parties are risk-neutral; there is no limited liability. Linear contracts, which leave the agent with a constant share of output in exchange for a fixed fee, are optimal. This result holds both in a preliminary version of the model where the principal only chooses to supply or not supply an input, and in several suitably-formulated variants of a more general version where the principal may have multiple choices of input. The model thus generates nontrivial linear sharing rules without relying on either limited liability or risk aversion.

JEL Classification Codes: D81, D82, D86

Keywords: Uncertainty, Asymmetric Information, Principal-Agent Model, Linear Contracts, Double-Sided Moral Hazard, Robustness

---

*This work has benefited from comments from audiences at the NSF/NBER/CEME Mathematical Economics Conference, the Econometric Society European Winter Meetings, and internal workshops at Stanford. The ⓡ symbol denotes random author order (Ray ⓡ Robson, 2018). The first-listed author thanks the Research School of Economics at ANU for their hospitality during a sabbatical, and the NSF for financial support through a CAREER grant.

[†]Stanford University. E-mail: gdc@stanford.edu.

[‡]Stanford University. E-mail: lbolte@stanford.edu.

# 1 Introduction

Why do profit-sharing rules arise in agency relationships? And what determines the form that such rules take?

The bulk of the literature on principal-agent models, since Holmström (1979) and Grossman and Hart (1983), emphasizes risk aversion, and the importance of the resulting tradeoff between providing incentives and insurance. In these models, typically output results from some costly and unobserved effort provided by the agent. If the agent were risk-neutral, the optimal solution would just be "selling the firm" for a fixed fee, thereby making the agent a full residual claimant to the consequences of his effort. A separate branch of the literature focuses on limited liability constraints (Innes, 1990), which make it impossible for the principal to capture the surplus from selling the firm to the agent; then, the principal optimally gives weaker incentives in order to avoid ceding too much of the surplus.

Yet in many situations, we observe sharing rules between firms, where neither risk aversion nor limited liability seem to be key considerations. We focus here on a different issue: *double-sided moral hazard*, that is, the importance of giving incentives for both the principal and agent to provide noncontractible inputs. A leading application where this arises is in franchising (Bhattacharyya and Lafontaine, 1995): contracts typically specify that a portion of revenues should be returned to the franchisor as royalties; this sharing ensures that the franchisor has incentives to advertise and maintain the reputation of the brand, while the franchisee has incentives to exert effort in local management. Other applications where double moral hazard has been argued to be relevant in determining contract terms include warranties, where both quality provision by the producer and care by the user are subject to moral hazard (Cooper and Ross, 1985); sharecropping (Eswaran and Kotwal, 1985); effort at cost savings in supply chains (Corbett and DeCroix, 2001; Corbett et al., 2005); and collaborative business services such as consulting (Roels et al., 2010). However, our study is meant to be general and not geared toward any specific application.

In the context of one-sided moral hazard, there is by now a rich theoretical literature developing models that generate different functional forms for optimal sharing rules, such as linear contracts (Holmström and Milgrom, 1987; Diamond, 1998), debt contracts (Innes, 1990; Hébert, 2018), or threshold-based bonus contracts (Lopomo et al., 2011), and thereby understanding the advantages of each. For double moral hazard, the same questions are much less developed. The seminal model of double moral hazard is that of Bhattacharyya and Lafontaine (1995). That paper was oriented primarily toward franchising applications and noted that in practice franchising contracts are often linear. Their model indeed predicts the existence of an optimal contract that is linear. However, this prediction relies on particular structural assumptions: most importantly, that output, while random, depends only on a one-dimensional "composite effort level" that aggregates both parties' effort. The key argument is that, given any candidate contract, a linear contract with appropriately chosen slope can replicate the same first-order conditions for each party. As observed by Kim and Wang (1998), this argument is not specific to linear contracts; there are many optimal contracts, and roughly speaking, any well-behaved one-parameter family of contracts would contain some optimal contract, for the same reason. They further argue that trying to select among the optima by adding a small amount of risk aversion fails to pick out the linear contract. Anyhow, once we depart from the strong assumption of one-dimensional composite effort, linear contracts can fail to be optimal (see Example 3 in Appendix A).

In this paper, we identify a specific virtue of linear contracts, and do so with minimal structural assumptions. Our argument is based on robustness to uncertainty about details of the environment. The idea is simple: Suppose (for example) the principal leaves 1/4 of output to the agent, keeping the remaining 3/4 for herself. If the agent is known to be able to secure a payoff of, say, 1000 for himself under such a contract, then the principal is guaranteed to get at least 3000, without needing to know details about exactly what the agent can do or what his optimal action is. This intuition was previously expressed by Carroll (2015), in a one-sided moral hazard model. The principal's guarantee is formalized by a maxmin criterion, and the main result is that the highest possible such guarantee is attained by a linear contract. However, that model relied crucially on a limited liability constraint. Without such a constraint, the one-sided moral hazard model would again yield the trivial solution of selling the firm to the agent. In the present paper, we show how the same intuition can be expressed in a model with double moral hazard (and no limited liability).

Incorporating maxmin-style uncertainty into a model with noncontractible choices by both parties raises modeling questions. First: how should the possible actions of the agent be modeled, if they might interact with choices by the principal (and vice versa)? Second: the agent's participation constraint is key, since it will be binding in an optimal contract, but it is not obvious how to formulate this constraint. The agent's willingness to participate depends on his reasoning about how the principal will make her subsequent choice, but how should we model this reasoning in a framework of non-Bayesian uncertainty?

Our modeling approach cuts down the difficulties by having the parties move sequentially. In the most basic version of our model, once the contract is signed, the principal moves first and makes a binary choice: either she supplies a costly input or not. If the principal supplies the input, then the agent takes his action; if not, no output can be produced and the relationship ends. This structure allows us to model an action by the agent simply via its effort cost and the resulting probability distribution over output, as in Carroll (2015). A backward-induction argument then leads us to focus on contracts where, at the signing stage, the agent can be confident that the principal will have enough incentive to supply the input. With this in mind, we show that linear contracts can provide the optimal guarantee.

This simple model also delivers some intuitive predictions. The optimal contract is one whose slope (the share paid to the agent) is as high as possible, subject to leaving enough to the principal to incentivize her to provide the input. This maximizes the total surplus, which the principal can then fully extract by setting the fixed fee appropriately.

The all-or-nothing input assumption is a strong one, so in Section 3 we introduce the more general version of the model, where the principal has multiple choices of input. If we straightforwardly duplicate the model and repeat the same backward-induction analysis, linear contracts may no longer be optimal: Essentially, the principal may prefer to choose nonlinear contracts that are vulnerable to gaming by the agent after some low choices of input; this can provide a way for her to credibly commit to choose high input, and thereby improve the agent's willingness to accept the contract. However, as our analysis shows, this interpretation of backward induction relies on much stronger assumptions about the principal's behavior at the input-choice stage (and about the agent's knowledge thereof) than it did in the single-input model.

It may not be obvious what alternative approach should be used instead of this backward-induction analysis. Rather than commit to a single alternative, we consider a number of variations, three of which are more agnostic (to varying degrees) about how the agent expects the principal to behave, and another one in which the agent may worry that the *principal* has new, unforeseen choices of input. Each approach, by adding more uncertainty into the participation decision, restores linear contracts as the optimal way of aligning interests.

The goal of our exercise is twofold: to offer a tractable general-purpose model of double moral hazard; and to specifically express the robustness intuition underlying linear contracts, with as little reliance on functional form assumptions as possible. The sequential-move structure is a significant difference from most existing models of double moral hazard, but it has been used before, e.g. Demski and Sappington (1991). Arguably, moving sequentially is no more or less of a gross oversimplification of agency relationships than the one-shot simultaneous structure usually assumed. And this timing has the advantage of allowing for a simple (albeit customized) approach to modeling the principal's behavior under uncertainty that has no obvious counterpart in a simultaneous-move model.

A paper closely related to ours is that of Dai and Toikka (2018), who consider robust incentives for teams of agents who must simultaneously choose costly actions. Indeed, a portion of their paper considers surplus-maximizing contracts. One could view double moral hazard through the lens of their model, with a team consisting of two parties (our principal and agent). They adopt a simultaneous-move setup, and their model delivers stark conclusions: In particular, for *any* nonlinear sharing contract, there is the potential for a "race to the bottom" that leads to no output being produced at all. In our model, nonlinear contracts can still deliver some guarantees (even in our variant with uncertainty on both sides), but linear contracts turn out to be optimal.

Aside from this, our work fits into the broader literature on robustness foundations for linear incentive contracts. This includes mostly Bayesian models (Holmström and Milgrom, 1987; Diamond, 1998; Barron et al., 2020). Chassang (2013) gives a related maxmin-optimality result.

The remainder of the paper proceeds as follows. Section 2 introduces the basic version of our model, with the binary choice by the principal (supply the input or not). We show that linear contracts provide the best guarantee to the principal. Section 3, building on the machinery developed in Section 2, introduces the more general version of the model, where the principal has multiple choices of input, and shows that linear contracts remain optimal under several variant formulations. In Section 4, we analyse how the environment determines the parameters of the optimal contract.

# 2 Single-input model

## 2.1 Setup

We begin by describing the simple single-input version of the model.

First, some notational conventions: Let $\Delta(\mathcal{X})$ denote the space of Borel distributions on $\mathcal{X} \subseteq \mathbb{R}^k$, let $\delta_x$ be the degenerate distribution with weight 1 on $x$ for $x \in \mathcal{X}$, let $\text{conv}(\mathcal{D})$ denote the convex hull of set $\mathcal{D}$, and let $\mathbb{R}^+$ denote the nonnegative reals.

A principal and an agent, who are both risk-neutral, may jointly participate in a production process. The principal may supply some input to the agent at a cost $c_P \in \mathbb{R}^+$. If she does not, then no production takes place, and output is zero (at no cost to either party). If the principal does supply the input, then the agent can take an action that (stochastically) produces output. Note that, although both parties make costly contributions to production, we use the asymmetric language ("input"/"action") to reflect their asymmetric roles in the model. There is some set $\mathcal{Y}$ of possible output realizations, which we assume is a compact subset of $\mathbb{R}^+$, with $0 \in \mathcal{Y}$ as the lowest possible output.

An action of the agent is modelled as a pair $(F, c)$ where $F \in \Delta(\mathcal{Y})$ is the resulting distribution over the output space $\mathcal{Y}$, and $c \in \mathbb{R}^+$ is the cost of taking this action incurred by the agent. We use the term *technology* to denote a nonempty, compact set of possible actions. We assume that there is some given technology $\widehat{\mathcal{A}}$, representing the actions that the principal knows the agent can take. The agent's true technology is a superset, $\mathcal{A} \supseteq \widehat{\mathcal{A}}$. The agent knows $\mathcal{A}$, but the principal knows only that it contains $\widehat{\mathcal{A}}$. Both $\widehat{\mathcal{A}}$ and $c_P$ are common knowledge at the time of contracting.

Incentives are provided by a contract that specifies how the output is divided between the principal and agent. Neither the principal's input nor the agent's action are contractible; only the output is. Thus, we define a *contract* $w$ as a continuous function from the output space $\mathcal{Y}$ to the reals.[1] By convention, $w(y)$ is the share received by the agent. A contract $w$ is *linear* if it is of the form $w(y) = \alpha y + \beta$; such a contract will be denoted by $w[\alpha, \beta]$. We may also sometimes use the terms *share parameter* to refer to $\alpha$ and *fee* to refer to $-\beta$ (with the interpretation that the principal sells the agent a share $\alpha$ of output in exchange for fee $-\beta$). A couple special cases are worth noting: The zero contract, $w[0, 0]$, pays a wage of 0 for any output level; a contract of the form $w[1, -p]$ entails selling the firm to the agent at price $p$.

The timing is summarized below:

1. The principal offers a contract $w$;
2. the agent, knowing his technology $\mathcal{A}$, accepts or rejects $w$. If the contract is rejected, both parties receive a payoff of 0.[2] If the contract is accepted; then
3. the principal chooses whether or not to supply the input. If she does not supply the input, output is 0, so her payoff is $-w(0)$ and the agent's payoff is $w(0)$. If she does supply the input, then
4. the agent chooses an action $(F, c) \in \mathcal{A}$;
5. output $y \sim F$ is realized;
6. payoffs are received: $y - w(y) - c_P$ to the principal and $w(y) - c$ to the agent.

Our focus is on the choice of contract $w$ at step 1. We wish to identify contracts that guarantee the principal as high a payoff as possible in the ensuing subgame. This of course requires a solution concept to specify what may happen after any given choice of $w$. Our

[1] Continuity serves only to ensure existence of best replies and is not a substantive restriction; see Carroll (2015, footnote 1) for more discussion.

[2] One may consider more general outside options $(\underline{u}_P, \underline{u}_A)$. If $\underline{u}_P + \underline{u}_A < 0$ then there may be trivial cases where it is optimal to sign a contract but then not supply the input. Otherwise, allowing more general outside options does not significantly change our results.

analysis is essentially based on backward induction, but we will have to be precise about what this means, in view of the uncertainty about $\mathcal{A}$.

We will find it useful to define a class of "eligible" contracts, those that guarantee the principal some positive payoff (the formal definition will appear shortly). Contract $w$ will be eligible if, in the subgame after $w$ is offered, the following strategies are consistent with backward induction:

- At step 2, the contract is accepted by all types of agent (i.e. for all possible $\mathcal{A}$);
- at step 3, the principal supplies the input;
- at step 4, the agent chooses his action optimally given $\mathcal{A}$,

and these strategies give the principal a positive payoff for all $\mathcal{A}$. Note that, given that the agent must get a payoff at least 0, the above requirements at steps 2 and 3 are indeed necessary for $w$ to guarantee the principal some positive payoff level.

To formalize this, we develop some notation. Consider the agent at step 4, who has accepted a contract $w$ and received the input. Denote the actions the agent might optimally choose, and his expected payoff associated with taking them, as

$$\mathcal{A}^*(w|\mathcal{A}) = \arg\max_{(F,c)\in\mathcal{A}} \{E_F[w(y)] - c\} \qquad \text{and} \qquad V_A(w|\mathcal{A}) = \max_{(F,c)\in\mathcal{A}} \{E_F[w(y)] - c\},$$

respectively. If the agent is indifferent between two actions, we will assume that he takes the action that maximizes the principal's payoff.

Since the principal does not know the agent's technology at step 1, we evaluate the *guarantee* of a contract $w$ (for the principal) by the worst case over all possible technologies. If $w$ is accepted and the principal supplies the input, then this worst-case expected payoff is

$$V_P(w|\widehat{\mathcal{A}}, c_P) = \inf_{\mathcal{A}\supseteq\widehat{\mathcal{A}}} \left( \max_{(F,c)\in\mathcal{A}^*(w|\mathcal{A})} \{E_F[y - w(y)]\} - c_P \right).$$

We may simply write $V_P(w)$ rather than $V_P(w|\widehat{\mathcal{A}}, c_P)$ when there is no ambiguity. Notice that this definition implicitly reflects the requirement that all types of agent be willing to accept the contract; if only some agent types were to accept, then the inf should be taken over this restricted space of types.

We may also refer to $V_A(w|\widehat{\mathcal{A}})$ as the *agent's guarantee*, since it evidently is the agent's payoff under the worst technology for him.

Now, we can give our formal definition of eligibility:

Definition 1. *A contract $w$ is* eligible *if*

*(E1)* $V_P(w) > 0$*;*

*(E2)* $V_P(w) \geq -w(0)$*; and*

*(E3)* $V_A(w|\widehat{\mathcal{A}}) \geq 0$*,*

*that is, (E1) the principal's guarantee is positive, (E2) she prefers to supply the input to not supplying it; and (E3) the agent's guarantee is nonnegative.*

Let us argue more systematically that this formal criterion corresponds to our backward-induction description.

If the principal offers an eligible contract at step 1, then backward induction implies that she is willing to supply the input at step 3, therefore the agent accepts the contract at step 2, and so the principal is indeed guaranteed at least $V_P(w)$. Conversely, with an ineligible contract, she cannot be guaranteed a positive payoff: If (E2) fails, the principal will not supply the input; if (E2) holds but (E3) fails then the agent may not accept the contract; and if (E2) and (E3) hold but (E1) fails then the principal will get payoff $V_P(w) \leq 0$.

Notice also that, although we are adopting a maxmin criterion to evaluate contracts ex ante, we do *not* need to commit to modeling the principal as a maxmin decision maker, in particular in step 3. Indeed, once the principal has offered an eligible contract, (E2) ensures that she will be willing to supply the input for any possible $\mathcal{A} \supseteq \widehat{\mathcal{A}}$, so will remain willing to do so, for example, if she is actually an expected-utility maximizer with some prior over $\mathcal{A}$. As long as the agent can see this, he is willing to accept at step 2 without any assumptions about the principal's attitude toward uncertainty over $\mathcal{A}$.

With this background in mind, we focus on the case where a positive guarantee is possible, and thus study how to maximize the guarantee over the space of eligible contracts.

## 2.2 Analysis

### 2.2.1 Existence of an eligible contract

It is not obvious when an eligible contract exists. In the one-sided moral hazard setting, Carroll (2015) makes the following assumption: There exists $(F, c) \in \widehat{\mathcal{A}}$ such that

$$E_F[y] - c > 0.$$

This assumption is enough to guarantee a positive total surplus and the existence of eligible contracts in that setting. In our setting, accounting for the cost to the principal of supplying the input, a positive total surplus is feasible if there exists $(F, c) \in \widehat{\mathcal{A}}$ such that

$$E_F[y] - c - c_P > 0. \tag{1}$$

Existence of such an action is certainly a necessary condition for existence of an eligible contract (this can be formally seen by adding (E1) and (E3)). Our first question is whether this condition is also sufficient. The answer is *no* as Example 1 shows.

Example **1.** *Let us consider a simple environment with $\mathcal{Y} = \{0, \bar{y}\}$, and only one known action, $\widehat{\mathcal{A}} = \{(\delta_{\bar{y}}, c)\}$, where $\bar{y} > c > 0$. Suppose $w$ is an eligible contract in this environment. Denote $w(\bar{y}) = \bar{w}$ and $w(0) = \underline{w}$. Suppose $\bar{w} < c$. Then, (E3) is violated. Suppose $\bar{w} > c$. Then, the agent always has a strictly positive payoff. The principal can reduce both $\bar{w}$ and $\underline{w}$ by some positive amount $\epsilon$, strictly increasing her guarantee while preserving eligibility. Thus, the principal optimally sets $\bar{w} = c$. If $\underline{w}$ is nonnegative and the agent's technology is given by $\mathcal{A} = \{(\delta_{\bar{y}}, c), ((1-\epsilon)\delta_0 + \epsilon\delta_{\bar{y}}, 0)\}$, for some small $\epsilon > 0$, then the agent would choose action $((1-\epsilon)\delta_0 + \epsilon\delta_{\bar{y}}, 0)$. But then $V_P(w) \leq 0$, as $V_P(w) \leq \epsilon\bar{y} - \underline{w}$ for all $\epsilon > 0$ and $\underline{w}$ is nonnegative by assumption, violating (E1). Thus, $\underline{w} < 0$, and for every such $\underline{w}$, we will now determine the principal's guarantee. For any technology of the agent, he will not choose an action for which his expected payoff is less than 0, since he can achieve this payoff with the known action. For any action, the expected output and expected wage lie on the dashed line connecting $(0, \underline{w})$ and $(\bar{y}, \bar{w})$ which is depicted in Figure 1 for different values of $\underline{w}$: $\underline{w}^1, \underline{w}^2$ and $\underline{w}^3$. As the agent can guarantee himself an expected payoff of 0, the expected wage paid to the agent has to be nonnegative. Furthermore, we may assume that the slope is at most 1 as*

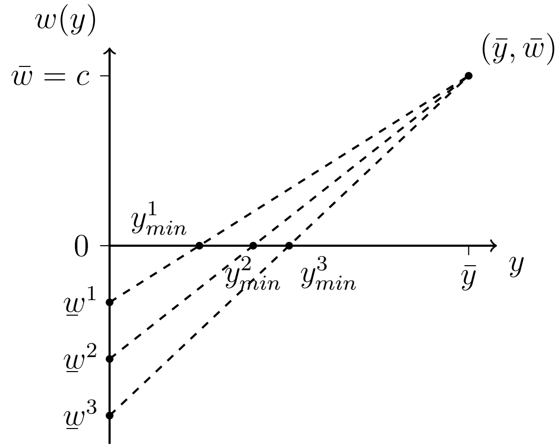



Figure 1: Positive surplus is not enough to guarantee the existence of an eligible contract

*otherwise the principal would prefer not to supply the input. Then, the worst-case expected output is given by the intersection of the dashed line and the horizontal line indicating the agent's minimum expected payoff; we call this worst-case output* $y_{min}$*, and it is depicted again for different values of* $\underline{w}$ *as* $y^1_{min}, y^2_{min}$ *and* $y^3_{min}$ *in Figure 1. Algebraically,* $y_{min}$ *is given by*

$$y_{min} = -\frac{\underline{w}\bar{y}}{c - \underline{w}}.$$

*For* $w$ *to be eligible, we require that (E1) and (E2) hold. At the worst-case expected output, the expected wage of the agent is* 0 *so that these conditions become*

$$-\frac{\underline{w}\bar{y}}{c - \underline{w}} - c_P > 0 \qquad \textit{and} \qquad -\frac{\underline{w}\bar{y}}{c - \underline{w}} - c_P \geq -\underline{w}.$$

*The conditions above are satisfied for some weakly negative* $\underline{w}$ *if and only if*

$$\bar{y} - c - c_P \geq 2\sqrt{cc_P} \qquad \textit{and} \qquad \bar{y} - c - c_P > 0. \tag{2}$$

*For* $c > 0$ *and* $c_P > 0$*,* (2) *is a stronger condition than the existence of positive total surplus as in* (1)*.*

In Example 1, (2) is a necessary and sufficient condition in a stylized environment. However, we can generalize the result to arbitrary environments.

Proposition 1. *An eligible contract exists if and only if there exists* $(F, c) \in \widehat{\mathcal{A}}$ *such that*

$$E_F[y] - c - c_P \geq 2\sqrt{cc_P} \qquad \textit{and} \qquad E_F[y] - c - c_P > 0.$$

The proof of Proposition 1 is postponed to Appendix B, after the proofs of the remaining results in this section (on which it relies). All other proofs provided appear chronologically in Appendix B.

For some intuition behind why (1) is not sufficient, note that in general, some amount of the output needs to be given to the principal to incentivize her to supply the input. This means that the agent will have to be made a less-than-full residual claimant, and so the surplus available must be large enough that even without receiving all of it, the agent is still motivated to exert effort.

### 2.2.2 Optimality of linear contracts

The next question we ask is how optimal contracts look like, provided they exist.

THEOREM **1.** *If an eligible contract exists, then among all eligible contracts there exists a linear contract that maximizes the principal's guarantee.*

There may also exist nonlinear contracts that attain the optimum: In particular, we can start from the linear contract and then change its shape at points outside the support of (known) actions. By adding an assumption to rule out this trivial multiplicity, we can ensure that only linear contracts can be optimal. Specifically, we say that $\widehat{\mathcal{A}}$ satisfies the full-support condition if for every action $(F, c) \neq (\delta_0, 0)$ in $\widehat{\mathcal{A}}$, $F$ has full support on $\mathcal{Y}$.

COROLLARY **1.** *If $\widehat{\mathcal{A}}$ satisfies the full-support condition, then every eligible contract that maximizes $V_P(w)$ is linear.*

The proof of Theorem 1 builds on the main proof from Carroll (2015), although we organize the ingredients of the proof a bit differently. The arrangement here will allow us to quickly leverage the same tools for the multiple-input versions of the model in Section 3.

For any contract $w$, we first characterize the *fundamental relationship* it induces between the principal's and the agent's guarantee, as the known technology varies. To do so, we need to introduce some further notation. For a fixed $w$, write

$$\mathcal{S} = \operatorname{conv}\left(\{(w(y)-c, y-w(y)) : y \in \mathcal{Y}, c \in \mathbb{R}^+\}\right); \tag{3}$$

$$\underline{\mathcal{Y}} = \arg\min_{y \in \mathcal{Y}} \{y - w(y)\};$$

$$\bar{\mathcal{Y}} = \arg\max_{y \in \mathcal{Y}} w(y).$$

Also let

$$y_0 = \arg\max_{y \in \underline{\mathcal{Y}}} w(y); \qquad y_1 = \arg\min_{y \in \bar{\mathcal{Y}}} \{y - w(y)\}; \qquad y_2 = \arg\max_{y \in \bar{\mathcal{Y}}} \{y - w(y)\}$$

(note that the maximizers and minimizers exist by continuity of $w$, and they are indeed unique since the values of $y - w(y)$ and $w(y)$ pin down $y$ uniquely).

Finally, put

$$\mathcal{F} = \{(u,v) \in \mathcal{S} : \nexists (u',v') \in \mathcal{S} \text{ such that } u' > u, v' < v\}. \tag{4}$$

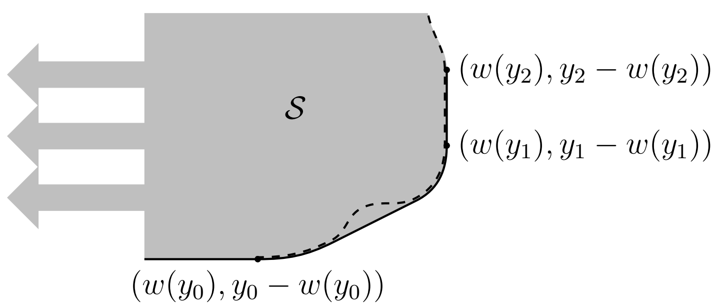



Figure 2: The dashed black line consists of points $(w(y), y - w(y))$ for $y \in \mathcal{Y}$. The solid black line describes the fundamental relationship between the principal's and the agent's guarantee. Set $\mathcal{S}$ is represented by the gray area and extends infinitely far to the left.

$\mathcal{F}$ is depicted in Figure 2 and describes the fundamental relationship between the principal's and the agent's guarantee for a given contract $w$ as follows.

Let $\mathcal{T}$ be the set of all technologies. Let $\mathcal{R}$ denote the collection of pairs of the agent's guarantee and the principal's guarantee (ignoring the $c_P$ term) as the known technology varies, i.e.

$$\mathcal{R} = \{(V_A(w|\widehat{\mathcal{A}}'), V_P(w|\widehat{\mathcal{A}}', 0)) : \widehat{\mathcal{A}}' \in \mathcal{T}\}.$$

Lemma 1. *For any contract $w$,*

$$\mathcal{R} = \mathcal{F}.$$

For an intuition behind the lemma, note that the pair of (agent's, principal's) expected payoffs for any possible action must lie in $\mathcal{S}$. Any known technology $\widehat{\mathcal{A}}'$ imposes a lower bound on the payoff that the agent can get. This corresponds to an assurance that the payoff pair lies to the right of some vertical line in the figure. Given this, the worst possible payoff for the principal is determined by the point where this vertical line intersects the lower boundary of $\mathcal{S}$, which is exactly a point on the frontier $\mathcal{F}$.

The proof of Theorem 1 then quickly follows: The lemma shows that the worst case for the principal under $w$ (and known technology $\widehat{\mathcal{A}}$) must involve some action for which the resulting expected (agent, principal)-payoff pair lies on the boundary of the convex hull of $w$. Hence, either this action is degenerate, or more generally all points in its support lie along some line that is tangent to the convex hull. Replace $w$ with this tangent line, which itself can be viewed as a linear contract $w'$. We show that $w'$ guarantees at least the same expected payoff for the principal as $w$. This implies that (E1) for the linear contract $w'$ is satisfied. We also have the comparison $w'(y) \geq w(y)$ for all $y$ which implies conditions (E2) and (E3). Hence, $w'$ is eligible and guarantees at least the same expected payoff as $w$. The full proofs of the lemma, the theorem, and Corollary 1) are in Appendix B.

One missing detail above is verifying that an optimum among linear contracts actually exists. This is done in the analysis below, which not only shows that the optimum exists but characterizes it.

### 2.2.3 Optimal linear contracts

The lemma below allows us to consider linear contracts $w[\alpha, \beta]$ with $\alpha \in [0, 1]$ only and identifies the principal's guarantee for the two boundary cases.

Lemma 2. *Consider any linear contract $w[\alpha, \beta]$.*

*A) $w[\alpha, \beta]$ can only be eligible if $0 \leq \alpha \leq 1$.*

*B) If $\alpha = 0$ and $w[\alpha, \beta]$ is eligible, then the principal's guarantee is given by* $\max_{(F,0)\in\widehat{\mathcal{A}}} E_F[y]-$

*$\beta - c_P$. (If no action of the form $(F, 0)$ exists in $\widehat{\mathcal{A}}$, then no contract with $\alpha = 0$ is eligible.)*

*C) If $\alpha = 1$ and $w[\alpha, \beta]$ is eligible, then the principal's guarantee is given by $-\beta$. (This case corresponds to the principal "selling the firm" which is only eligible if $c_P = 0$.)*

It remains to evaluate linear contracts $w[\alpha, \beta]$ with $\alpha \in (0, 1)$. Consider any such contract. For any action $(F, c)$ the agent can take, the principal's expected payoff is given by

$$E_F[(1-\alpha)y] - \beta - c_P,$$

which is increasing in $E_F[y]$. The agent, in turn, takes action $(F, c)$ only if

$$E_F[\alpha y] + \beta - c \geq \max_{(F',c') \in \widehat{\mathcal{A}}} \{E_{F'}[\alpha y] + \beta - c'\}$$

implying a lower bound of the expected level of output given by

$$\max \left\{ \frac{1}{\alpha} \max_{(F',c') \in \widehat{\mathcal{A}}} \{E_{F'}[\alpha y] - c'\}, 0 \right\}, \tag{5}$$

which is achieved for actions of the form $(F, 0)$. Because we know that an eligible contract must lead to the agent always producing a positive total surplus, expected output in particular is positive, so for such contracts, we can drop the outer maximization in (5).

The principal's guarantee for eligible linear contracts $w[\alpha, \beta]$ with $\alpha \in (0, 1)$ is therefore given by

$$V_P(w[\alpha, \beta]) = \frac{1-\alpha}{\alpha} \max_{(F,c) \in \widehat{\mathcal{A}}} \{E_F[\alpha y] - c\} - \beta - c_P. \tag{6}$$

(6) also holds for contracts of the form $w[1, \beta]$ and $w[0, \beta]$ if we define $\frac{c}{\alpha} = 0$ for $\alpha = c = 0$ (and interpret $\frac{1-\alpha}{\alpha} E_F[\alpha y]$ as $E_F[y]$ when $\alpha = 0$).

It follows that a linear contract $w[\alpha, \beta]$ is eligible if and only if

$$\frac{1-\alpha}{\alpha} \max_{(F,c) \in \widehat{\mathcal{A}}} \{E_F[\alpha y] - c\} - \beta - c_P > 0; \tag{7}$$

$$\frac{1-\alpha}{\alpha} \max_{(F,c) \in \widehat{\mathcal{A}}} \{E_F[\alpha y] - c\} - \beta - c_P \geq -\beta; \tag{8}$$

$$\max_{(F,c) \in \widehat{\mathcal{A}}} \{E_F[\alpha y] + \beta - c\} \geq 0. \tag{9}$$

Notice that for any given $\alpha$, we can decrease $\beta$ until (9) binds; doing so will increase $V_P$ and will not break (7–8). Hence, we define

$$\beta(\alpha) = -\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\}. \tag{10}$$

and focus on eligible contracts of the form $w[\alpha, \beta(\alpha)]$. Note also that for such contracts, the principal's guarantee is given by

$$V_P(w[\alpha,\beta(\alpha)]) = \frac{1}{\alpha}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\} - c_P = \max_{(F,c)\in\widehat{\mathcal{A}}}\left\{E_F[y] - \frac{c}{\alpha}\right\} - c_P. \tag{11}$$

Evidently, this expression is weakly increasing in $\alpha$, and strictly increasing wherever the relevant maximizer satisfies $c > 0$.

It remains to choose $\alpha$ to maximize this expression, subject to eligibility of the contract $w[\alpha, \beta(\alpha)]$. It then suffices to check (8), since (7) automatically holds at the maximum as long as some eligible contract exists. Because (8) carves out a closed set of possible values of $\alpha$, and $V_P$ is weakly increasing in $\alpha$, it is now immediate that the maximum does indeed exist, and we have the explicit characterization, stated in the following result.[3]

Lemma **3.** *If an eligible linear contract exists, then either the zero contract is an optimal eligible linear contract or the unique optimum in the class of eligible linear contracts is given by $w[\alpha^*, \beta(\alpha^*)]$, where*

$$\alpha^* = \max\left\{\alpha \in [0,1] : \frac{1-\alpha}{\alpha}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\} - c_P \geq 0\right\}. \tag{12}$$

*Furthermore,*

$$\frac{1-\alpha^*}{\alpha^*}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha^* y] - c\} - c_P = 0 \qquad \text{and} \qquad V_P(w[\alpha^*,\beta(\alpha^*)]) = -\beta(\alpha^*). \tag{13}$$

We note in passing an implication of the results so far: With a small change in the specification of the environment, the outcome can change discontinuously between having a contract with a large positive guarantee and not having an eligible contract exist at all. Indeed, (13) implies that the principal's guarantee equals the amount that the agent gains

[3] Although the result is stated as optimizing over linear contracts, recall from Theorem 1 that the resulting contract is then optimal among *all* eligible contracts.

by taking his best known action instead of producing zero output. When the condition for existence of an eligible contract (see Proposition 1) is just barely met, this gain must be bounded away from zero: otherwise, if the agent had an action available where he could produce very low output at zero cost, he would deviate to do so; but then the resulting total surplus (accounting for the input cost $c_P$) would be negative, which is impossible.

In Appendix C.1, we give a more explicit characterization of the optimal linear contract $w[\alpha^*, \beta(\alpha^*)]$, motivated by (13), which may be useful in computing examples.

# 3 Formulations for multiple inputs

The single-input assumption is a strong one, and it delivers a correspondingly extreme conclusion: The optimal contract is such that the principal's incentive to supply the input is binding. We now extend the model to allow the principal more choices, which may be interpreted as different types of input, or different quantities or qualities of input. For each choice that the principal makes, there is some cost to herself, and some resulting set of (known) actions $\widehat{\mathcal{A}}$ the agent can take in response. To the extent that $\widehat{\mathcal{A}}$ varies across inputs, this can be interpreted as variation in the set of physical actions that the agent can take, or as variation in the consequences (and perhaps the costs) of a given action by the agent; the difference in interpretation is immaterial. Thus, we will model an input choice directly as an ordered pair $(\widehat{\mathcal{A}}, c_P)$, describing the resulting known actions available to the agent, and the principal's cost of supplying the input. We will use the phrase *input space* to denote a finite set of such pairs, interpreted as the set of inputs from which the principal can choose.

Let $\mathcal{W}$ be an input space. For each $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$, the agent's true technology is given by some $\mathcal{A}$ that is a superset of $\widehat{\mathcal{A}}$.

The timing is summarized below:

1. The principal offers a contract $w$;

2. the agent, knowing his technology $\mathcal{A}$ corresponding to each input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$, accepts or rejects $w$. If the contract is rejected, both parties receive a payoff of 0. If the contract is accepted, then

3. the principal chooses whether to supply an input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ or not. If she does not supply any input, her payoff is $-w(0)$ and the agent's payoff is $w(0)$. If she does supply input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$, then

4. the agent chooses an action $(F, c) \in \mathcal{A}$, where $\mathcal{A} \supseteq \widehat{\mathcal{A}}$ is the agent's corresponding technology;

5. output $y \sim F$ is realized;

6. payoffs are received: $y - w(y) - c_P$ to the principal and $w(y) - c$ to the agent.

We will use some of our analysis from the single-input environment (Section 2).

Definition **2.** *A contract $w$ is* locally eligible via $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$, *if $w$ is eligible in the single-input environment where the known technology is $\widehat{\mathcal{A}}$ and the cost of supplying the input is $c_P$.*

## 3.1 Weakly eligible contracts

We first consider the most straightforward extension of the solution concept from our single-input model, essentially extending the backward induction approach to the multiple-input case. However, as we will soon see, this approach relies on strong assumptions that may be unpalatable in the multiple-input environment; the shortcomings we will observe here will motivate several alternative notions of eligibility ahead in Section 3.2.

Here, we consider contracts $w$ such that, in the subgame after $w$ is offered, the following behavior is consistent with backward induction and guarantees the principal a positive payoff:

- At step 2, the contract is accepted by all types of agent;

- at step 3, the principal chooses some particular input $(\widehat{\mathcal{A}}, c_P)$;

- at step 4, the agent chooses his action optimally given the corresponding $\mathcal{A}$.

Notice that the only change relative to the single-input model is that at step 3, the principal should prefer the particular input $(\widehat{\mathcal{A}}, c_P)$ over any other input. A natural way to formalize this idea is that this choice of input should maximize the principal's payoff, as given by the worst-case objective, $V_P(w|\cdot,\cdot)$. This reasoning leads us to the following approach to defining eligibility and specifying the guarantee from a contract.

Definition **3.** *Let $w$ be a contract. Define*

$$V_P(w|\mathcal{W}) = \max_{(\widehat{\mathcal{A}}, c_P)\in\mathcal{W}} V_P(w|\widehat{\mathcal{A}}, c_P).$$

*We say that an input $(\widehat{\mathcal{A}}^*, c_P^*)$ is* an optimal input (given $w$) *if*

$$V_P(w|\widehat{\mathcal{A}}^*, c_P^*) = V_P(w|\mathcal{W}).$$

Definition **4.** *A contract $w$ is* weakly eligible*, if it is locally eligible via some input $(\widehat{\mathcal{A}}^*, c_P^*)$ that is optimal given $w$.*

*We define the* guarantee *of such a contract $w$ as the corresponding value of $V_P(w|\mathcal{W})$.*

The fact that $(\widehat{\mathcal{A}}^*, c_P^*)$ is optimal given $w$ implies that the guarantee-maximizing principal is willing to supply $(\widehat{\mathcal{A}}^*, c_P^*)$ at step 3. Since the agent at step 2 anticipates this input choice by the principal, local eligibility is indeed enough to make him willing to accept the contract.

How does an optimal contract look like in this environment? As Example 2 below shows, even if weakly eligible linear contracts exist, nonlinear contracts may be preferable. The intuition is that nonlinear contracts may give the principal a form of commitment power over the choice of input that linear contracts cannot. In particular, the principal can specify a low flat wage over part of the output space, giving the agent insufficient incentives to exert effort following some inputs. This in turn will make those input choices unappealing to the principal, providing a way to commit herself to choosing higher inputs instead.

Example **2.** *Suppose that there is a "high" input that is costly to supply and a "low" input that is cheap to supply: $\mathcal{W} = \{(\widehat{\mathcal{A}}^h, c_P^h), (\widehat{\mathcal{A}}^l, c_P^l)\}$. Let $\mathcal{Y} = [0, 30]$. To be concrete, let*

$$\widehat{\mathcal{A}}^h = \{(\delta_{24}, 8)\} \text{ and } c_P^h = 4 \qquad \text{and} \qquad \widehat{\mathcal{A}}^l = \{(\delta_{12}, 3)\} \text{ and } c_P^l = 2.$$

*We want to show that, among weakly eligible contracts, linear ones do not attain the optimum. To this end, we find an upper bound on the principal's guarantee for such contracts, show that no linear contract attains this upper bound, and finally construct a nonlinear contract that does and is thus optimal.*

*The maximal guarantee of the principal over weakly eligible contracts cannot be more than the corresponding maximal guarantee over all contracts that are locally eligible for some input (not necessarily an optimal input). Thus, let us consider contracts that are locally eligible for input $\widehat{\mathcal{A}}^h$ at cost $c_P^h$. By Lemma 3 and (10), an optimal contract, and the unique optimum that is linear, is given by $w[\alpha, \beta]$ with $(\alpha, \beta) = (2/3, -8)$. The principal's guarantee from supplying $(\widehat{\mathcal{A}}^h, c_P^h)$ given contract $w[\alpha, \beta]$ is*

$$V_P(w[\alpha, \beta]|\widehat{\mathcal{A}}^h, c_P^h) = -\beta = 8.$$

*This is an upper bound for the guarantee provided by any weakly eligible contract. (Note that a contract that is locally eligible for $(\widehat{\mathcal{A}}^l, c_P^l)$ cannot do better, since the total known surplus there is only $7 < 8$.) Can this upper bound be achieved by a linear contract? The only candidate contract is $w[\alpha, \beta]$. By construction, $w[\alpha, \beta]$ is locally eligible via $(\widehat{\mathcal{A}}^h, c_P^h)$. However, the principal's guarantee from supplying $(\widehat{\mathcal{A}}^l, c_P^l)$ given contract $w[\alpha, \beta]$ is*

$$\begin{aligned} V_P(w[\alpha, \beta]|\widehat{\mathcal{A}}^l, c_P^l) &= \max\left\{\frac{1-\alpha}{\alpha} \max_{(F,c)\in\widehat{\mathcal{A}}^l} \{E_F[\alpha y] - c\}, 0\right\} - \beta - c_P^l \\ &= \frac{1-2/3}{2/3}\{2/3 \cdot 12 - 3\} + 8 - 2 \\ &= \frac{1}{2}(8-3) + 8 - 2 > V_P(w[\alpha, \beta]|\widehat{\mathcal{A}}^h, c_P^h). \end{aligned}$$

*Hence, $w[\alpha, \beta]$ is not weakly eligible, as $(\widehat{\mathcal{A}}^h, c_P^h)$ is not an optimal input given $w$, and if the agent anticipates the principal's choice of $(\widehat{\mathcal{A}}^l, c_P^l)$ then he would reject the contract.*

*Finally, we design a nonlinear contract that is weakly eligible and achieves the upper bound, and that therefore is optimal. We will take the contract $w[\alpha, \beta]$ above and modify it. Specifically, let contract $w$ be given by*

$$w(y) = \begin{cases} \beta & \text{for } y \leq 12 \\ \beta + \frac{8-\beta}{24-12}(y - 12) & \text{for } 12 \leq y \leq 24 \\ \alpha y + \beta & \text{for } 24 \leq y. \end{cases}$$

*$w$ decreases the wage of the agent for low levels of output. As a result, when the principal supplies input $\widehat{\mathcal{A}}^l$ instead of $\widehat{\mathcal{A}}^h$, the principal is not guaranteed any positive expected output and $(\widehat{\mathcal{A}}^h, c_P^h)$ is indeed optimal to choose.*

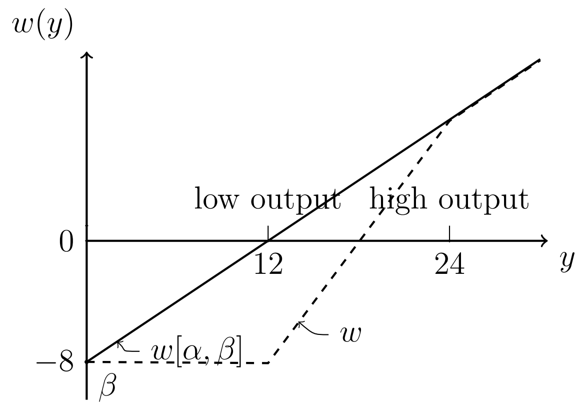



Figure 3: Contract $w[\alpha, \beta]$ is not weakly eligible because the principal has an incentive to supply the low input. By decreasing the wage for low levels of output, the resulting contract, $w$, does not guarantee any positive level of expected output for the low input. As a result, the principal optimally supplies the high input and $w$ is weakly eligible.

*Contracts $w$ and $w[\alpha, \beta]$ are depicted in Figure 3.*

*Formally, let $\mathcal{A}^l = \widehat{\mathcal{A}}^l \cup \{(\delta_0, 0)\}$ be the technology of the agent when the principal supplies input $\widehat{\mathcal{A}}^l$. Given contract $w$, the agent chooses action $(\delta_0, 0)$ and the principal's payoff is given by $-\beta - c_P^l = 8 - 2 = 6$ providing an upper bound for the principal's guarantee when supplying the low input when the contract is $w$. However, the principal's guarantee when supplying the high input when the contract is $w$ remains unchanged and $(\widehat{\mathcal{A}}^h, c_P^h)$ is indeed optimal to choose. Thus, $w$ is weakly eligible.*

## 3.2 Four proposed notions of eligibility

The backward induction underlying weak eligibility implicitly relies on strong assumptions about the agent's prediction of the principal's behavior. In particular, it requires confidence that the agent at step 2 expects the principal to choose her subsequent input to maximize $V_P(w|\cdot, \cdot)$. This makes the solution concept much more demanding than in the single-input model: Recall from the discussion at the end of Section 2.1 that the eligibility criterion there did not embody any such assumption on the principal's input choice.

In the following, we relax the assumption that the agent expects the principal to choose the input that maximizes her guarantee. Instead, we postulate four different kinds of assumptions that the agent could make about the principal's behavior which ultimately lead to four different notions of eligibility. In particular, the four different versions allow that the agent: 1) might make any assumption about the principal's input choice; 2) assumes the principal maximizes her guarantee but might expect the principal to have additional knowledge of his available actions beyond the original $\widehat{\mathcal{A}}$'s; 3) still assumes the principal maximizes her guarantee but might find new inputs to be available after the contract is proposed; or 4) assumes the principal has full knowledge of the agent's technology associated with each input and maximizes her expected utility accordingly. Each of these approaches has different consequences for what inputs the principal might be predicted to choose (and when) under any given contract, and therefore, for the agent's willingness to accept the contract in the first place. The resulting participation constraints then give rise to four different notions of eligibility, and we refer to contracts satisfying each as *maximally eligible*, *eligible with further actions*, *eligible with further inputs*, *eligible with full knowledge*, respectively.

To see more concretely how each notion of the agent's reasoning matters, return to Example 2, supposing that the principal offers the nonlinear contract $w$ that we proposed there. We discuss the four proposed assumptions in turn:

1) *Maximal eligibility:* Clearly, if the agent could expect the principal to supply any input, the proposed contract $w$ fails to assure the principal a positive guarantee, because the agent simply would not accept it if he predicts the principal will choose the low input. Indeed, under the low input, the agent can only be assured an expected net payoff of $-8 < 0$.

   (Note that in fact, the linear contract $w[\alpha, \beta]$ would not be acceptable to the agent either. However, our emphasis for the moment is on the nonlinear contract $w$. The point is that the finding from Example 2—that the nonlinear contract resolves the principal's commitment problem—is overturned here.)

2) *Eligibility with further actions:* Next, suppose that the agent's technology $\mathcal{A}^l$ when the principal chooses the low input also includes an action $(\delta_{24}, 9)$. Suppose, moreover, that the agent at step 2 thinks that the principal *knows* that $(\delta_{24}, 9) \in \mathcal{A}^l$. In this

case, a lower bound on the principal's guarantee from choosing the low input can be calculated via an analogue of the argument for (6): The agent is able to obtain a payoff of $8-9=-1$; since he always receives *at most* $\alpha y+\beta$, any action he could take that gives him at least this high a payoff would have to generate expected output at least $(-1-\beta)/\alpha = 10.5$, and so would give the principal at least $(1-\alpha)\cdot 10.5-\beta-c_P^l = 9.5$. This is higher than the guarantee of 8 from choosing the high input. Thus the principal would rather choose the low input at step 3. But this would leave the agent with a payoff of $-1$, worse than rejecting the contract. So the agent would not accept.

3) *Eligibility with further inputs:* Now suppose that, after the contract $w$ is proposed, the principal learns about a new input that includes an action $(\delta_{24}, 9)$ and is costless for the principal to supply. Similar to the calculations above, the principal's guarantee associated with providing this input is now given by 11.5 whereas the agent may still only receive a payoff of $-1$. Thus, if the principal allows that the agent might expect her to discover new inputs, then again she cannot be assured that contract $w$ would be accepted.

4) *Eligibility with full knowledge:* Lastly, suppose that the agent thinks the principal has full knowledge about the agent's technology associated with each input choice. For example, the principal knows that agent's technology is given by $\mathcal{A}^l = \widehat{\mathcal{A}}^l \cup \{(\delta_{24}, 11)\}$ and $\mathcal{A}^h = \widehat{\mathcal{A}}^h$ associated with the low input and high input respectively.

   With the low input provided, the agent would prefer his higher action over his lower one. Thus, as the output is 24 for both inputs, the principal would again choose the low (and cheaper) input, leaving the agent with a negative payoff; hence, he would not want to accept the contract in the first place.

The four extensions thus show various ways in which the backward induction analysis underlying weak eligibility is fragile. They give rise to four different notions of eligibility, each capturing a version of the idea that the principal should be confident that the agent finds the contract acceptable, not only irrespective of his true technology (for each input) but also irrespective of his reasoning about the principal's subsequent input choice.

It is difficult to justify any one of the above variations as resulting in *the* natural notion of eligibility. Instead, we will consider each in turn. As we shall see, we can give results on how the optimality of linear contracts is restored in each case.

The different proofs demonstrating optimality of linear contracts in each context will build heavily on the machinery developed to prove Theorem 1. To avoid repetition, we will provide some proofs in greater detail than others.

Note that in some variations, the principal's choice of contract could potentially convey information to the agent that is relevant to her choice of input at step 3. For example, in variation 2, the principal might choose different contracts depending on what additional knowledge she has about the agent's actions. This would turn the interaction between principal and agent into a signaling game, which would be considerably more complex. Our treatment here ignores this issue; one interpretation is that the principal's additional information is realized only after she has proposed the contract (note that our formulation of the objective remains applicable with this timing).

### 3.2.1 Maximal eligibility

We first consider the most demanding notion of eligibility: The principal makes no assumptions about which input the agent expects her to supply.

Definition 5. *A contract $w$ is* maximally eligible*, if it is locally eligible via some input $(\widehat{\mathcal{A}}, c_P)$ and, for all inputs $(\widehat{\mathcal{A}}', c_P') \in \mathcal{W}$,*

$$V_A(w|\widehat{\mathcal{A}}') \geq 0.$$

Note that the above definition entails the minimal restriction that the agent definitely expects the principal to supply some input rather than none. However, the definition above can incorporate dropping this minimal restriction on the agent's behavior if the trivial input, $(\mathcal{A}_{triv}, 0)$ with $\mathcal{A}_{triv} = \{(\delta_0, 0)\}$, is added to the input space.

Proposition 2. *If a maximally eligible contract exists, then among all maximally eligible contracts there exists a linear contract that maximizes the principal's guarantee.*

A characterization of when a maximally eligible contract exists and how the optimal one can be found is provided in Appendix C.2. The reasoning is straightforward: for any candidate slope $\alpha$, we use the definition of maximal eligibility to determine the lowest possible $\beta$; we then choose $\alpha$ to maximize the resulting value of $V_P$ (given by (6)).

The proof of Proposition 2 is a replication of the proof of Theorem 1 with minor modifications and thus only briefly sketched out. In essence, we begin with a (nonlinear) maximally eligible contract $w$ and construct a linear one, $w'$, as in Theorem 1 in the single-input environment using an optimal input. The new linear contract will remain locally eligible via the previously optimal input and also provide the principal with a greater guarantee. Furthermore, as $w' \geq w$ pointwise, the agent's payoff is still nonnegative for every input so that $w'$ is still maximally eligible, completing the proof.

#### 3.2.2 Eligibility with further actions

Here, we allow the agent to believe that the principal has additional knowledge about the inputs while maintaining the assumption that she maximizes her guarantee. In particular, the principal is believed to know of additional actions the agent has.

To develop a notion of eligibility in this context, we define the notion of *worrisome inputs*. These are the input choices such that, for some realization of additional knowledge the principal might have about the agent's possible actions, the principal would choose this input, *and* the anticipation of this choice would make the agent not want to accept the contract.

Definition **6.** *Given contract $w$, input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ is* FA-worrisome *if there exists $\widehat{\mathcal{A}}' \supseteq \widehat{\mathcal{A}}$ such that*

$$V_P(w|\widehat{\mathcal{A}}', c_P) > V_P(w|\mathcal{W}) \qquad \textit{and} \qquad V_A(w|\widehat{\mathcal{A}}') < 0.$$

This allows us to define a second stronger version of eligibility.

Definition **7.** *A contract $w$ is* eligible with further actions (EFA)*, if it is locally eligible via some optimal input and no $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ is FA-worrisome.*

*We define the guarantee of such a contract $w$ as the corresponding value of $V_P(w|\mathcal{W})$.*

Notice that this definition of the guarantee is the same as was used under weak eligibility. Indeed, we wish to say that an EFA contract guarantees the principal a certain level of payoff if the principal can choose her subsequent input in a way that ensures her that payoff no matter what the agent's technology is; and this definition of the guarantee captures that condition. Again, this should not be interpreted to mean that the principal actually will choose the optimal input or that the agent expects her to do so.

We now have the following proposition.

Proposition **3.** *If an EFA contract exists, then among all EFA contracts there exists a linear contract that maximizes the principal's guarantee.*

We again postpone the characterization of when such a contract exists and the determination of the optimal one to the appendix. The discussion is mainly in Section C.3, though it draws on some calculations contained within the proof of Proposition 3.

The proof of Proposition 3 builds on the fundamental relationship and other machinery developed earlier to prove Theorem 1. A key step is to characterize EFA contracts in terms of this machinery, as we now describe. Given a contract $w$, we can define a frontier for each choice of input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$:

$$\mathcal{F}_{(\widehat{\mathcal{A}}, c_P)} = \{(u, v) : u \geq V_A(w|\widehat{\mathcal{A}}), v \geq V_P(w|\widehat{\mathcal{A}}, c_P), (u, v + c_P) \in \mathcal{F}\}$$

where $\mathcal{F}$ is as was defined in (4).

Note that for any $(u, v) \in \mathcal{F}_{(\widehat{\mathcal{A}}, c_P)}$, there exists $\widehat{\mathcal{A}}' \supseteq \widehat{\mathcal{A}}$ such that

$$(V_A(w|\widehat{\mathcal{A}}'), V_P(w|\widehat{\mathcal{A}}', c_P)) = (u, v).$$

Indeed, the existence of some technology $\widehat{\mathcal{A}}'$ with these guarantees is given by Lemma 1, and then we can ensure $\widehat{\mathcal{A}}' \supseteq \widehat{\mathcal{A}}$ by simply replacing $\widehat{\mathcal{A}}'$ by $\widehat{\mathcal{A}}' \cup \widehat{\mathcal{A}}$ if necessary; the fact that this does not change $V_A$ or $V_P$ follows from the bounds on $u$ and $v$.

Define the set of feasible outcomes $\mathcal{U}_{\mathcal{W}}$ as

$$\mathcal{U}_{\mathcal{W}} = \cup_{(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}} \mathcal{F}_{(\widehat{\mathcal{A}}, c_P)}.$$

Define the *critical region* $\mathcal{C}$ as

$$\mathcal{C} = \{(u, v) : u < 0, v > V_P(w|\mathcal{W})\}.$$

This region consists of all the payoff pairs such that the principal would be willing to choose them if she knew she could, but such that the agent would then prefer to reject the contract. More simply put, payoff pairs in this region are the ones that could make an input worrisome.

We display two examples in Figures 4a and 4b. Here, $\mathcal{W} = \{(\widehat{\mathcal{A}}^1, c_P^1), (\widehat{\mathcal{A}}^2, c_P^2), (\widehat{\mathcal{A}}^3, c_P^3)\}$; $c_P^1 < c_P^2 < c_P^3$. The principal's optimal input is $(\widehat{\mathcal{A}}^2, c_P^2)$ which delineates the lower boundary to the critical region.

Our characterization is then:

LEMMA **4.** *The contract $w$ is EFA if and only if (i) it is locally eligible via some optimal input and (ii) $\mathcal{U}_{\mathcal{W}} \cap \mathcal{C} = \emptyset$.*

*Moreover, if $w$ is linear, then (i) can be weakened to require only that $w$ be locally eligible via* some *input.*

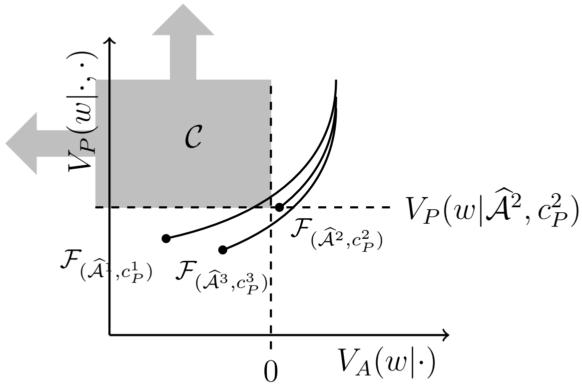
For the two cases depicted in Figure 4a and Figure 4b, Lemma 4 then implies that the contract shown in Figure 4a is not EFA, as input $(\widehat{\mathcal{A}}^1, c_P^1)$ is worrisome, whereas the contract shown in Figure 4b is EFA provided it is locally eligible via input $(\widehat{\mathcal{A}}^2, c_P^2)$.

Once this is established, to show that linear contracts are optimal, we can now begin with an arbitrary EFA contract, and replace it by a linear one as in the proof of Theorem 1. The new contract dominates the old one pointwise and also gives the principal a better guarantee if the choice of input is held fixed. From these properties, it quickly follows that the critical region $\mathcal{C}'$ for the new contract is smaller than the old one, while the input-specific frontiers $\mathcal{F}'_{(\widehat{\mathcal{A}}, c_P)}$ can only have moved downward and rightward; consequently, if the feasible outcomes and the critical region did not intersect previously, they still do not intersect, and the contract remains EFA by Lemma 4. This is shown in Figure 4c.

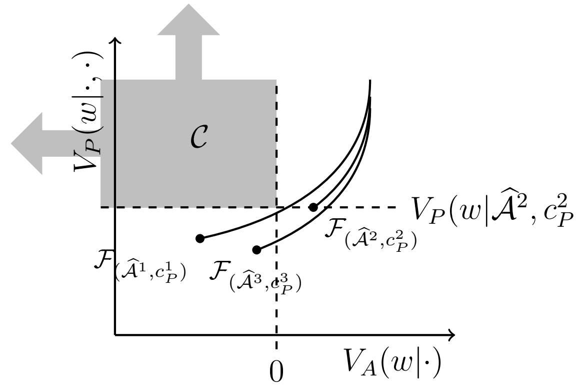
The full proof of Proposition 3 is in Appendix B.

### 3.2.3 Eligibility with further inputs

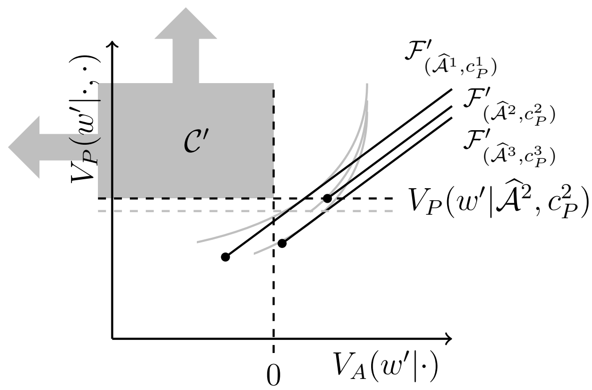
In our third variation, we return to the backward induction approach and its implicit assumptions of no new actions and maxmin decision-making at step 3, but we now introduce uncertainty about the *principal's* available choices. When contracting, the principal and the




(a) Contract $w$ is not EFA.




(b) Contract $w$ is EFA.




(c) Contract $w'$ is EFA if contract $w$ is.

Figure 4: In all three figures, jointly the black lines represent the sets of feasible outcomes; and the grey regions represent the critical regions. The contract in Figure 4a is not EFA as input $(\widehat{\mathcal{A}}^1, c_P^1)$ is worrisome because $\mathcal{F}_{(\mathcal{A}^1,c_P^1)}$ intersects the critical region $\mathcal{C}$. Contract $w$ in Figure 4b is EFA as long as $w$ is locally eligible via $(\widehat{\mathcal{A}}^2, c_P^2)$. Figure 4c depicts contract $w$ from Figure 4b replaced by a linear one as in the proof of Theorem 1. Observe that the new critical region, $\mathcal{C}'$, is smaller than the old one; and that all input-specific frontiers, $\mathcal{F}'_{(\mathcal{A}^1,c_P^1)}, \mathcal{F}'_{(\mathcal{A}^2,c_P^2)}$ and $\mathcal{F}'_{(\mathcal{A}^3,c_P^3)}$, moved downward and rightward.

agent know that the principal can choose inputs from $\mathcal{W}$. However, after the contract is proposed, a set $\widetilde{\mathcal{W}} \supseteq \mathcal{W}$ which is previously unknown to both the principal and the agent is revealed. The principal can supply any input $(\widehat{\mathcal{A}}, c_P) \in \widetilde{\mathcal{W}}$. The principal only supplies inputs that maximize her guarantee, as in Section 3.1, and the agent accepts only if he still foresees a nonnegative payoff.

The timing is summarized below:

1. The principal, knowing $\mathcal{W}$, offers a contract $w$;
2. the input space $\widetilde{\mathcal{W}}$ is revealed, where $\widetilde{\mathcal{W}} \supseteq \mathcal{W}$;
3. the agent, knowing his technology $\mathcal{A}$ corresponding to each input $\widehat{\mathcal{A}} \in \widetilde{\mathcal{W}}$, accepts or rejects $w$. If the contract is rejected, both parties receive a payoff of 0. If the contract is accepted, then
4. the principal chooses whether to supply an input $(\widehat{\mathcal{A}}, c_P) \in \widetilde{\mathcal{W}}$ or not. If she does not supply any input, her payoff is $-w(0)$ and the agent's payoff is $w(0)$. If she does supply input $(\widehat{\mathcal{A}}, c_P) \in \widetilde{\mathcal{W}}$, then
5. the agent chooses an action $(F, c) \in \mathcal{A}$, where $\mathcal{A} \supseteq \widehat{\mathcal{A}}$ is the agent's technology;
6. output $y \sim F$ is realized;
7. payoffs are received: $y - w(y) - c_P$ to the principal and $w(y) - c$ to the agent.

Applying backward induction leads to the following definition.

Definition **8.** *A contract $w$ is* eligible with further inputs (EFI), *if it is locally eligible via some optimal input and for all $\widetilde{\mathcal{W}} \supseteq \mathcal{W}$ and $(\widehat{\mathcal{A}}, c_P) \in \widetilde{\mathcal{W}}$,*

$$\text{if} \qquad V_P(w|\widehat{\mathcal{A}}, c_P) > V_P(w|\mathcal{W}), \qquad \text{then} \qquad V_A(w|\widehat{\mathcal{A}}) \geq 0.$$

*We define the guarantee of such a contract $w$ as the corresponding value of $V_P(w|\mathcal{W})$.*

It turns out that eligibility with further inputs is formally equivalent to a special case of eligibility with further actions. (Recall that $\mathcal{A}_{triv} = \{(\delta_0, 0)\}$.)

Proposition **4.** *A contract $w$ that is locally eligible via some optimal input is EFI if and only if it is EFA under input space $\mathcal{W}' = \mathcal{W} \cup \{(\mathcal{A}_{triv}, 0)\}$.*

*Furthermore, the principal's guarantee is equal in these two environments.*

Thus, Proposition 3 implies that it is without loss to optimize over the space of EFI contracts that are linear and further characterizes an optimal contract.

Corollary **2.** *If an EFI contract exists, then among all EFI contracts there exists a linear contract that maximizes the principal's guarantee.*

*Furthermore, the optimum among EFI linear contracts is given by the optimum among EFA linear contracts under input space* $\mathcal{W}' = \mathcal{W} \cup \{(\mathcal{A}_{triv}, 0)\}$ *instead of* $\mathcal{W}$*.*

Corollary 2 follows directly from Proposition 4 and its proof is omitted.

### 3.2.4 Eligibility with full knowledge

In our last variation, we consider the case when the principal fully knows the agent's technology associated to each input, and she chooses the input based on this knowledge.

Similar to EFA, to develop a notion of eligibility, we define a notion of inputs that are chosen under some full knowledge resulting in a negative payoff for the agent.

Definition **9.** *Given contract* $w$*, input* $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ *is* FK-worrisome *if there exists a technology* $\mathcal{A} \supseteq \widehat{\mathcal{A}}$ *such that*

$$\max_{(F,c)\in\mathcal{A}^*(w|\mathcal{A})} \{E_F[y - w(y)]\} - c_P > V_P(w|\mathcal{W}) \qquad \text{and} \qquad V_A(w|\mathcal{A}) < 0.$$

We then define our final stronger version of eligibility as follows.

Definition **10.** *A contract* $w$ *is* eligible with with full knowledge (EFK)*, if it is locally eligible via some optimal input and no* $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ *is FK-worrisome.*

*We define the guarantee of such a contract* $w$ *as the corresponding value of* $V_P(w|\mathcal{W})$*.*

How does an optimal contract look like in this environment? When contracts are restricted to be *monotone* (i.e. $w$ weakly increasing in $y$), then linear contracts are optimal.

Proposition **5.** *If a monotone EFK contract exists, then among all monotone EFK contracts there exists a linear contract that maximizes the principal's guarantee.*

When a monotone EFK contract exists and how the optimal monotone EFK contract can be found is described in Appendix C.4.

Why do we need the restriction to monotone contracts? Nonlinear contracts, when they are nonmonotone, can help the principal to credibly commit to not deviating to inputs that are bad for the agent, by making sure that any additional actions that could potentially make those inputs tempting for the principal will be so low-paying that the agent would never choose them. Example 4 in Appendix A illustrates this more concretely.

The proof of Proposition 5 proceeds similarly to that of Proposition 3, building on tools developed to prove Theorem 1.

In Section 3.2.2 the set of feasible outcomes consisted of the union of frontiers for each choice of input. With full knowledge of the technology, the outcome need not lie on the frontier; thus, the analogous set of feasible outcomes builds on $\mathcal{S}$, as defined in (3), as opposed to its frontier only.

Given a contract $w$, define for each choice of input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$:

$$\mathcal{S}_{(\widehat{\mathcal{A}},c_P)} = \{(u, v) : u \geq V_A(w|\widehat{\mathcal{A}}), v \geq V_P(w|\widehat{\mathcal{A}}, c_P), (u, v + c_P) \in \mathcal{S}\}.$$

Similar to before, for any $(u, v) \in \mathcal{S}_{(\widehat{\mathcal{A}},c_P)}$, there exists $\mathcal{A} \supseteq \widehat{\mathcal{A}}$ such that

$$\left(V_A(w|\mathcal{A}), \max_{(F,c)\in\mathcal{A}^*(w|\mathcal{A})} \{E_F[y - w(y)]\} - c_P\right) = (u, v).$$

The set of feasible outcomes is now the union of such sets: $\tilde{\mathcal{U}}_{\mathcal{W}} = \cup_{(\widehat{\mathcal{A}},c_P)\in\mathcal{W}} \mathcal{S}_{(\widehat{\mathcal{A}},c_P)}$; the critical region $\mathcal{C}$ is defined as before and an analogue of Lemma 4 is given by the following.

Lemma 5. *The contract $w$ is EFK if and only if it is locally eligible via some input and*

$$\tilde{\mathcal{U}}_{\mathcal{W}} \cap \mathcal{C} = \emptyset.$$

We omit the detailed proofs of Lemma 5 and Proposition 5; the arguments proceed almost identically to the proof of Proposition 3 which is given in detail in Appendix B. The restriction to monotonicity is needed to ensure that the sets $\mathcal{S}_{(\widehat{\mathcal{A}},c_P)}$ move rightward and downward when the initial contract $w$ is replaced with a linear contract $w'$: without this restriction, the sets may become taller, so that the change of contracts creates an intersection with the critical region where none existed previously.

# 4 Comparative statics

What determines the parameters of the optimal contract? In previous sections, we formally showed that linear contracts, those consisting of a fixed share in exchange for a fee, are robust to uncertainty (as in Carroll, 2015). However, how do the fee and the share parameter vary with the environment? In this section we will carry out some comparative statics exercises in some simple parameterized versions of the model, both single-input and multiple-input specifications.

These exercises serve three purposes. First, they help to build intuition and illustrate the mechanics of the models. Second, the comparative statics of our models can be compared to existing empirical evidence on contract terms in situations of double moral hazard. The literature seems to be deepest on franchising applications, so we focus our comparisons on this area, and particularly Lafontaine (1992) which is the leading empirical analysis of contract terms in franchising. And, third and relatedly, our comparative statics can be compared to previous theoretical models of double moral hazard. In particular, Lafontaine summarizes previous models (Rubin, 1978; Eswaran and Kotwal, 1985; Lal, 1990; Bhattacharyya and Lafontaine, 1995) as finding that the share of profit paid to the agent is increasing in the size of the agent's incentive problem and decreasing in the size of the principal's incentive problem. (Cooper and Ross (1985) find similar results in a simple parameterization of their warranty model.) The empirical evidence that Lafontaine considers also provides some support for these predictions.

In order to compare the predictions of our model against this general description, one has to make some choices about what is meant by the "size" of the principal's (or agent's) incentive problem. Such choices are necessarily subjective.

We will give some formal results for both the share parameter $\alpha$ and the fee $-\beta$ in the optimal contract, but for brevity, will give interpretive discussion only for the $\alpha$ results. Rubin (1978) also looks at the fraction of principal's revenue from a franchising agreement arising from the fee (as opposed to the share) as a relevant outcome for comparative statics, but studying this fraction is not substantially different from studying $\alpha$ itself (see Propositions 9 and 10 in Appendix A) so we will not consider it separately.

Throughout this section, we assume that, for all environments under consideration, an eligible contract exists.

## 4.1 Single-input model

We begin by analyzing the simplest case: There is one input (thus we study the model from Section 2), and the agent has one (known) action $(F, c)$. Let $c_P$ denote the cost to the principal of providing the input, and $\mu := E_F[y]$. The environment is summarized by the tuple $(c_P, \mu, c)$.

Let $w[\alpha, \beta]$ denote an optimal contract in this environment as defined in Lemma 3. We may write its parameters as $\alpha(c_P, \mu, c)$ and $\beta(c_P, \mu, c)$ to make the dependencies more explicit.

Proposition **6.** *The share parameter, $\alpha(c_P, \mu, c)$, and the fee, $-\beta(c_P, \mu, c)$, are decreasing in $c_P$ and $c$ and increasing in $\mu$.*

We might interpret $c_P$ as a measure of the size of the principal's incentive problem, and likewise $c$ for the agent. Then, the fact that the agent's share is decreasing in $c_P$ is natural, but the fact that it is *also* decreasing in $c$ may seem surprising. However, Lemma 3 tells us that in this model, the slope of the optimal contract is not determined by a tradeoff between the two parties' incentives. Instead, the principal wants to maximize the incentives set for the agent to exert effort (and then fully extracts the surplus via appropriate choice of $\beta$), subject only to her own incentive constraint to provide the input; thus, the latter constraint is binding at the optimal contract. An increase in the agent's cost $c$ makes it easier for the agent to be tempted by less-productive actions (if they are available), thus reducing the principal's guaranteed gain from supplying the input. Thus, an increase in $c$ actually *also increases the principal's moral hazard problem*, requiring the principal to keep a larger share of output for herself. (A decrease in $\mu$ also results in a decrease in the agent's share for a very similar reason.)

This particular result stems from the binary choice the principal takes to provide the input or not and the fact that the principal is indifferent between her choices in an optimal contract. To the extent that this fails to accord with the evidence, it makes sense to proceed to the multiple-input model to see if it makes a different prediction.

## 4.2 Multiple-input model

Recall that we offered several versions of the multiple-input model and solution concepts. Rather than exhaustively discuss comparative statics for all of them, we will focus on one to study here. We choose EFA as being most representative of the issues of interest to us: in contrast, weak eligibility was a dispreferred solution concept; maximal eligibility captures less of the double moral hazard problem (since the contract does not reflect the principal's incentive to deviate, only the possible consequences for the agent if she does so); EFI was a special case of EFA, and EFK is conceptually similar to EFA but a bit more complicated. Thus, in the below, we discuss comparative statics of the optimal EFA contract.

### 4.2.1 Fixed choice of input

We continue to keep the parameterization as simple as possible: Let the input space $\mathcal{W}$ consist of two inputs, $(\widehat{\mathcal{A}}, c_P)$, with $\widehat{\mathcal{A}}$ in turn consisting of a single action $(F, c)$, and $(\mathcal{A}_{triv}, \underline{c}_P)$, where as before $\mathcal{A}_{triv} = \{(\delta_0, 0)\}$. We again let $\mu := E_F[y]$. A multiple-input environment is then summarized by the tuple $(c_P, \underline{c}_P, \mu, c)$. This specification looks like the single-input model, but there is a crucial difference: there is no longer mutual certainty that the trivial input produces zero output, so the agent may worry that the principal will deviate to this input, and has to be offered enough that he will still be willing to accept the contract.

Let $w[\alpha, \beta]$ with $\alpha(c_P, \underline{c}_P, \mu, c)$ and $\beta(c_P, \underline{c}_P, \mu, c)$ denote the optimal linear contract in this environment as defined in Proposition 3. As described in Appendix C.3, there are two possible cases for this contract, depending on the environment. If the principal's incentive to provide the (nontrivial) input is binding, then the analysis is much as in the single-input model. We therefore focus on the complementary case, which we call "eligibility nonbinding" for short.[4]

[4]In Appendix C.3, we show that for a fixed input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ and action $(F, c) \in \widehat{\mathcal{A}}$, eligibility does not bind if $E_F[y] - c - c_P - 2\sqrt{c}\sqrt{c_P - \underline{c}_P} - \frac{\sqrt{c}}{\sqrt{c_P - \underline{c}_P}}\underline{c}_P \geq 0$ and that in this case the principal's guarantee is given by $E_F[y] - c - c_P - 2\sqrt{c}\sqrt{c_P - \underline{c}_P}$ (equations (35) and (33), respectively). Hence, an example of a numerical specification for which eligibility does not bind is given by $c_P > \underline{c}_P = 0$; and $\mu$ sufficiently large

In this case, as we show in the appendix, the optimal contract is determined by the maximizing the following objective:

$$(E_F[y] - c - c_P) - \frac{1-\alpha}{\alpha}c - \frac{\alpha}{1-\alpha}(c_P - \underline{c}_P). \tag{14}$$

For simplicity we will assume throughout that the solution to this problem is unique.

**Proposition 7.** *The share parameter, $\alpha(c_P, \underline{c}_P, \mu, c)$, is decreasing in $c_P$, increasing in $\underline{c}_P$ and $c$ and independent of $\mu$; the fee, $-\beta(c_P, \underline{c}_P, \mu, c)$, is decreasing in $c_P$, increasing in $\underline{c}_P$ and $\mu$ and ambiguous in $c$.*

Propositions 6 and 7 are summarized in Table 1.

| | single input | | multiple inputs | |
|---|---|---|---|---|
| | $\alpha$ | $-\beta$ | $\alpha$ | $-\beta$ |
| $c$ | $-$ | $-$ | $+$ | ambig. |
| $\mu$ | $+$ | $+$ | indep. | $+$ |
| $c_P$ | $-$ | $-$ | $-$ | $-$ |
| $\underline{c}_P$ | N.A. | N.A. | $+$ | $+$ |

Table 1: Comparative statics results in a stylized environment with a single nontrivial input and a single known action.

For intuition, note that the objective to be maximized in (14) consists of a term for total surplus, minus two terms representing possible losses: one representing the principal's losses due to the agent deviating to a less-productive, lower-cost action, and a second term for rent that must be left to the agent to guard against the *principal's* possible deviation to the trivial input. Each party has less incentive to deviate when she receives a higher share of output, therefore the first of these two losses is decreasing in the agent's share $\alpha$, while the second is increasing in $\alpha$. When eligibility is nonbinding, the optimal $\alpha$ is determined by the tradeoff between these two forces. Indeed, solving the first-order condition of (14) for $\alpha$ shows this clearly:

$$\alpha((F, c), c_P) = \frac{\sqrt{c}}{\sqrt{c} + \sqrt{c_P - \underline{c}_P}}. \tag{15}$$

relative to $c$.

This now implies that the share paid to the agent is increasing in $c$ and decreasing in $c_P$. If we think of $c$ as a measure of the size of the agent's incentive problem and $c_P$ as the size of the principal's incentive problem, then these effects go in the expected directions. Essentially, when $c$ is increased, the losses due to the agent's incentive problem become more sensitive to $\alpha$, making the corresponding term predominate in determining the optimal $\alpha$; likewise for $c_P$ and the rents due to the principal's incentive problem.[5]

#### 4.2.2 Varying optimal input

The previous example illustrates the key tradeoffs in determining $\alpha$, but is still oversimplified because there is a single nontrivial input for the principal and then a single known action for the agent. More generally, as the parameters of the environment vary, the optimal input $(\mathcal{A}^*, c_P^*)$ for the principal could change, as well as the optimal action within $\mathcal{A}^*$ to target.

The consequences for the optimal contract could be complicated. Here, we briefly explore a particular specification for which the analysis remains relatively simple.

We consider a parameterization in which only the principal has multiple choices; the agent still has only one (known) action. The principal's choice of input determines the probability $p$ that the agent's action successfully generates output. The environment depends on a parameter $\gamma$ that controls the cost of the input. Specifically, let $\widehat{\mathcal{A}}(p)$ be given by

$$\widehat{\mathcal{A}}(p) = \{(pF + (1-p)\delta_0, c)\}$$

for some $F$, and let $\mathcal{W}$ be given by

$$\mathcal{W} = \{(\widehat{\mathcal{A}}(p), g(\gamma p)) | p \in \mathcal{P}\}$$

where $\mathcal{P}$ is some finite subset of the unit interval, $g(0) = \underline{c}_P = 0, g' > 0$ and $g'' > 0$.[6]

Let $\mu \coloneqq E_F[y]$. The environment is summarized by the tuple $(\gamma, \mu, c)$.

[5] A similar analysis for maximally eligible contracts, and to some extent for EFK contracts, would instead find that the optimal contract is locally insensitive to $c_P$ and $\underline{c}_P$ as they do not affect the agent's inference about which input will be supplied. We note also that some other comparative statics depend on the solution concept used; for example, comparative statics with respect to $\mu$ and $c$ are also different if we use maximal eligibility rather than EFA.

[6] We restrict $\mathcal{P}$ to be finite so that $\mathcal{W}$ is finite and thus in line with our previous definition of input space.

Let $w[\alpha, \beta]$ denote the optimal linear contract in this environment as defined in Proposition 3, and write $\alpha(\gamma, \mu, c), \beta(\gamma, \mu, c)$ to make the dependencies explicit. Furthermore, let $p(\gamma, \mu, c)$ denote the optimal choice of input the principal supplies. We again assume eligibility is nonbinding.

PROPOSITION **8.** *The share parameter, $\alpha(\gamma, \mu, c)$, is decreasing in $\mu$ and increasing in $\gamma$ and $c$; and the optimal choice of input, $p(\gamma, \mu, c)$, is decreasing in $\gamma$ and $c$ and increasing in $\mu$.*

Table 2 summarizes our findings.

| | $\alpha$ | $p$ |
|---|---|---|
| $\mu$ | $-$ | $+$ |
| $c$ | $+$ | $-$ |
| $\gamma$ | $+$ | $-$ |

Table 2: Comparative statics results in a stylized environment with multiple inputs.

In this model, the effect of the principal's input choice is reflected in the comparative statics with respect to $\mu$: As $\mu$ goes up, the principal prefers to choose higher inputs $p$. But doing so is more costly to her, thus her incentive problem becomes more severe, which is reflected in a lower share $\alpha$ going to the agent in the optimal contract. As for comparative statics with respect to $c$, $\alpha$ is increasing, but now for two reasons: the direct increase in the size of the agent's incentive problem (as before), but also a decrease in the principal's incentive problem because the principal now prefers lower inputs.

Finally, since $\gamma$ might be thought of as a measure of the size of the principal's incentive problem, why is $\alpha$ increasing in $\gamma$, not decreasing? The answer is that the relevant measure of the size of the incentive problem depends not only on $\gamma$ but also on the endogenous choice of input. Intuitively, holding fixed the *total* cost $\gamma p$ spent on the input, a rise in $\gamma$ increases the marginal cost of input, and thereby makes the principal prefer to spend less. So when $\gamma$ goes up, the principal in equilibrium chooses to spend less on input, and thus her incentive to deviate actually becomes smaller, not larger.

# 5 Conclusion

We have studied a contracting problem with moral hazard on both sides: The principal and the agent both need to exert effort for production to take place. Our interest is in developing insights into what forms of contract can perform well, in parallel with the literature on this question in one-sided moral hazard models; and more specifically, in seeing whether the idea that linear contracts are robust to uncertainty about the agent's possible actions can be expressed in such a setting. We have captured this focus on robustness by evaluating any proposed contract by its worst-case payoff guarantee for the principal, and we have presented several versions of a model in which the maximum such guarantee is indeed attained by a linear contract. We performed comparative statics exercises that explore the intuitive workings of the model and related them to the available evidence.

Defining the guarantee of a contract poses modeling challenges: How should the unknown actions of the agent be modeled, and what solution concept is appropriate to describe behavior in the game between the principal and agent after the contract has been accepted? Our approach has been to model the game as sequential, with the principal moving first. This allows us to model actions of the agent as in the one-sided moral hazard model of Carroll (2015), to give a simple intuition parallel to that model about how linear contracts can provide guarantees for the principal, and to give a simple definition for the guarantee of any contract, provided that the contract is accepted. The new subtlety in the model here is how to delineate the set of contracts that the agent is guaranteed to accept; this depends on what assumptions we make about the agent's expectation of the principal's later behavior. We offered a simple version (eligibility), based on backward induction, in our preliminary model with only a binary choice of input. For our more general model, we saw that a direct generalization of the backward induction logic led to a notion (weak eligibility) under which linear contracts were not optimal. However, we proposed four less stringent approaches that lead to four different notions of eligibility—varying the objective and knowledge the principal is assumed to have at the input choice stage—restoring optimality of linear contracts in each.

Although the various eligibility concepts developed here are tailored specifically to this model, the modeling approach we have taken may provide future inspiration for other models of robust contracting that require interaction among multiple agents.

# References


Barron, Daniel, George Georgiadis, and Jeroen Swinkels (2020) "Optimal contracts with a risk-taking agent," *Theoretical Economics*, **15** (2), 715–761.

Bhattacharyya, Sugato and Francine Lafontaine (1995) "Double-Sided Moral Hazard and the Nature of Share Contracts," *The RAND Journal of Economics*, **26** (4), 761–781.

Carroll, Gabriel (2015) "Robustness and Linear Contracts," *American Economic Review*, **105** (2), 536–63.

Chassang, Sylvain (2013) "Calibrated Incentive Contracts," *Econometrica*, **81** (5), 1935–1971.

Cooper, Russell and Thomas W Ross (1985) "Product warranties and double moral hazard," *The RAND Journal of Economics*, 103–113.

Corbett, Charles J and Gregory A DeCroix (2001) "Shared-savings contracts for indirect materials in supply chains: Channel profits and environmental impacts," *Management science*, **47** (7), 881–893.

Corbett, Charles J, Gregory A DeCroix, and Albert Y Ha (2005) "Optimal shared-savings contracts in supply chains: Linear contracts and double moral hazard," *European journal of operational research*, **163** (3), 653–667.

Dai, Tianjiao and Juuso Toikka (2018) "Robust Incentives for Teams," Unpublished manuscript, Mass. Inst. of Technology, Cambridge, MA.

Demski, Joel S and David EM Sappington (1991) "Resolving double moral hazard problems with buyout agreements," *The RAND Journal of Economics*, 232–240.

Diamond, Peter (1998) "Managerial incentives: on the near linearity of optimal compensation," *Journal of Political Economy*, **106** (5), 931–957.

Eswaran, Mukesh and Ashok Kotwal (1985) "A Theory of Contractual Structure in Agriculture," *American Economic Review*, **75** (3), 352–367.

Grossman, Sanford J and Oliver D Hart (1983) "An analysis of the principal-agent problem," *Econometrica*, 7–45.

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.

Holmström, Bengt and Paul Milgrom (1987) "Aggregation and linearity in the provision of intertemporal incentives," *Econometrica*, 303–328.

Innes, Robert D (1990) "Limited liability and incentive contracting with ex-ante action choices," *Journal of Economic Theory*, **52** (1), 45–67.

Kim, Son Ku and Susheng Wang (1998) "Linear Contracts and the Double Moral-Hazard,," *Journal of Economic Theory*, **82**, 342–378.

Lafontaine, Francine (1992) "Agency Theory and Franchising: Some Empirical Results," *The RAND Journal of Economics*, **23** (2), 263–283.

Lal, Rajiv (1990) "Improving Channel Coordination Through Franchising," *Marketing Science*, **9** (4), 299–318.

Lopomo, Giuseppe, Luca Rigotti, and Chris Shannon (2011) "Knightian uncertainty and moral hazard," *Journal of Economic Theory*, **146** (3), 1148–1172.

Ray, Debraj ⓡ Arthur Robson (2018) "Certified Random: A New Order for Coauthorship," *American Economic Review*, **108** (2), 489–520.

Roels, Guillaume, Uday S Karmarkar, and Scott Carr (2010) "Contracting for collaborative services," *Management Science*, **56** (5), 849–863.

Rubin, Paul H. (1978) "The Theory of the Firm and the Structure of the Franchise Contract," *The Journal of Law & Economics*, **21** (1), 223–233.

# A Additional Results

Here we give the example referred to in the introduction, showing that in a model without uncertainty (so that the production technology is known), linear contracts are typically not optimal without imposing specific functional form assumptions. We retain here the timing of the single-input model in Section 2, but one could easily give a similar example in a simultaneous-move setup as in Bhattacharyya and Lafontaine (1995).

EXAMPLE **3.** *Suppose that output can range between* 0 *and* 2*, and the agent's choice of effort $e$ (if the principal has supplied the input) is a number between* 1 *and* 2*. For each such $e$, let $F(y|e)$ be a distribution on $[1,2]$ with mean equal to $e$. Assume the agent's cost of effort is given by an increasing, differentiable convex function $c(e)$ with $c(1)=0$ and $c'(1)<1/2<c'(2)$. Assume the cost of providing the input, $c_P$, is small but positive.*

*If the principal supplies the input, then the agent chooses $e$, and output $y$ is determined as follows: With probability $1/2$, output is drawn from a uniform distribution on $[0,1]$; with complementary probability $1/2$, output is drawn from $F(y|e)$ on $[1,2]$. (If the principal does not supply the input, then output $y$ is zero.)*

*Evidently, the first-best outcome is generated when the agent chooses $e^{FB}$ given by $c'(e^{FB})=1/2$. If the principal chooses a piecewise-linear contract of the form $w(y)=\beta$ for $y\leq 1$ and $w(y)=(y-1)+\beta$ for $y>1$, then she makes the agent a full residual claimant for his effort, so induces the first-best effort $e^{FB}$. By choosing $\beta$ appropriately, the principal can appropriate the full surplus for herself, leaving the agent with a payoff of zero. Moreover, as long as $c_P$ is small, the principal will indeed be willing to supply the input, since she is the residual claimant for output until it surpasses* 1*.*

*In contrast, a linear contract $w(y)=\alpha y+\beta$ cannot induce the first-best: by the first-order condition, the agent cannot be made to choose $e^{FB}$ unless $\alpha=1$, but a contract with a slope of* 1 *cannot motivate the principal to supply the input.*

Next, here is the example showing that nonmonotone contracts can outperform linear contracts in the multi-input model under eligibility with full knowledge.

EXAMPLE **4.** *Let $\mathcal{W}=\{(\widehat{\mathcal{A}}^1,c_P^1),(\widehat{\mathcal{A}}^2,c_P^2)\}$, with $\widehat{\mathcal{A}}^1=\{(\delta_{24},8)\},c_P^1=4,\widehat{\mathcal{A}}^2=\{(\delta_0,2)\}$ and $c_P^2=48$, and $\mathcal{Y}=[0,150]$. As in Example 2, the optimal contract subject only to local* *eligibility via some input is given by* $w[\alpha, \beta]$ *with* $(\alpha, \beta) = (2/3, -8)$. *However,* $w[\alpha, \beta]$ *is not EFK: if the agent's technology for input* $(\widehat{\mathcal{A}}^2, c_P^2)$ *is given by* $\mathcal{A}^2 = \widehat{\mathcal{A}}^2 \cup \{(\delta_{150}, 100)\}$, *the agent chooses the high action giving a greater payoff to the principal than her guarantee from the first input while leaving himself still with a negative payoff. Consider now a nonlinear, and in particular nonmonotone, variation of contract* $w[\alpha, \beta]$. *Let contract* $w$ *be given by*

$$w(y) = \begin{cases} \alpha y + \beta & \text{for } y \leq 24 \\ \alpha \cdot 24 + \beta - (y - 24) & \text{for } y > 24. \end{cases}$$

*With the second input provided, the agent only chooses actions for which his expected payoff is at least as high as the expected payoff he is guaranteed through action* $(\delta_0, 2)$, $2/3 \cdot 0 - 2 - 8 = -10$. *Given the nonmonotonicity of* $w$, *the agent would then only take actions with expected output in* $[0, 42]$. *As the cost of providing this input is* $c_P^2 = 48$, *it cannot be that the agent receives an expected payoff of at least* $-10$ *and the principal of at least* $8$, *the guarantee from the optimal input. Hence, contract* $w$ *is EFK.*

*Note that the linear contract* $w[\alpha, \beta]$ *is EFA: For example, if the principal is aware of action* $(\delta_{150}, 100)$ *available to the agent, she would still not provide the second input as the agent may also have action* $(\delta_1, 0)$ *available which he prefers.*

And here are the results concerning the share of guaranteed revenue that comes from the fee, mentioned in Section 4.

In the single-input environment, let the principal's guaranteed revenue for an eligible linear contract $w[\alpha, \beta]$ be defined as the sum of the fee $-\beta$ and her share of the guaranteed output $\frac{1-\alpha}{\alpha}\{E_F[\alpha y] - c\}$.

Proposition **9.** *For an optimally chosen fee, the share of the principal's guaranteed revenue coming from the fee is* $\alpha$ *and through her share of the guaranteed output is* $1 - \alpha$.

*Proof.* The share of the guaranteed revenue coming from the fee is given by

$$\frac{-\beta(\alpha)}{\frac{1-\alpha}{\alpha}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\} - \beta(\alpha)}$$
$$= \frac{\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\}}{\frac{1-\alpha}{\alpha}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\} + \max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\}} = \alpha.$$

□

Proposition **10.** *If in the multiple-input environment, there exists some input that is costless to provide ($\underline{c}_P = 0$), then for any EFA linear contract with optimally chosen fee for some target input $w[\alpha, \beta_{EFA}(\alpha, (\widehat{\mathcal{A}}, c_P))]$, the share of the principal's guarantee coming from the fee is $\alpha$.*

*Proof.* We draw on the characterization in Appendix C.3. The optimal fee and the principal's guarantee are given by (30) and (28) respectively so that, with $\underline{c}_P = 0$, the share of the guarantee coming from the fee is given by

$$\frac{-\beta_{EFA}(\alpha, (\widehat{\mathcal{A}}, c_P))}{\frac{1-\alpha}{\alpha} \max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\} - c_P - \beta_{EFA}(\alpha, (\widehat{\mathcal{A}}, c_P))}$$
$$= \frac{-c_P \frac{\alpha}{1-\alpha} + \max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\}}{\frac{1-\alpha}{\alpha} \max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\} - c_P - c_P \frac{\alpha}{1-\alpha} + \max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y] - c\}} = \alpha.$$

□

# B Proofs

*Proof of Lemma 1.* Take any $\widehat{\mathcal{A}} \in \mathcal{T}$. (For brevity, we will write $\widehat{\mathcal{A}}$ throughout this proof rather than $\widehat{\mathcal{A}}'$ as in the definition of $\mathcal{R}$; note the $\widehat{\mathcal{A}}$ from the main model is never needed for this lemma.)

We make the following three claims:

(i) $V_P(w|\widehat{\mathcal{A}}, 0)$ is bounded below as

$$V_P(w|\widehat{\mathcal{A}}, 0) \geq \min E_F[y - w(y)] \text{ over } F \in \Delta(\mathcal{Y}) \text{ such that } E_F[w(y)] \geq V_A(w|\widehat{\mathcal{A}}); \quad (16)$$

(ii) if $V_A(w|\widehat{\mathcal{A}}) < w(y_1)$, then (16) holds with equality; and

(iii) if $V_P(w|\widehat{\mathcal{A}}, 0) > y_0 - w(y_0)$, then whenever $F$ attains the minimum in (16),

$$E_F[w(y)] = V_A(w|\widehat{\mathcal{A}}).$$

To show (i), note that for any $F \in \Delta(\mathcal{Y})$, action $(F, c)$ is only chosen by the agent if his expected payoff is at least the expected payoff from choosing an optimal action in $\widehat{\mathcal{A}}$, i.e. only if

$$E_F[w(y)] - c \geq V_A(w|\widehat{\mathcal{A}}).$$

As $c \geq 0$, it follows that $V_P(w|\widehat{\mathcal{A}}, 0)$ cannot be smaller than the minimum of $E_F[y - w(y)]$ over $F \in \Delta(\mathcal{Y})$ such that $E_F[w(y)] \geq V_A(w|\widehat{\mathcal{A}})$.

Now suppose that $V_A(w|\widehat{\mathcal{A}}) < w(y_1)$. Suppose that $F$ achieves the minimum in (16). We need to show that $V_P(w|\widehat{\mathcal{A}}, 0)$ cannot be strictly greater than $E_F[y-w(y)]$. If $\text{supp}(F) \nsubseteq \bar{\mathcal{Y}}$, then let $\mathcal{A}$ be given by $\mathcal{A} = \widehat{\mathcal{A}} \cup \{(F', 0)\}$ where $F'$ is the mixture of $F$ with weight $1 - \epsilon$ and $\delta_{y_1}$ with weight $\epsilon$. For any $\epsilon > 0$, the agent chooses action $(F', 0)$ and as $\epsilon \to 0$, $E_{F'}[y - w(y)] \to E_F[y - w(y)]$ implying that $V_P(w|\widehat{\mathcal{A}}, 0) \leq E_F[y - w(y)]$. Suppose now that $\text{supp}(F) \subseteq \bar{\mathcal{Y}}$. Then by assumption, $V_A(w|\widehat{\mathcal{A}}) < w(y_1) = E_F[w(y)]$. Thus, for $\mathcal{A}$ given by $\mathcal{A} = \widehat{\mathcal{A}} \cup \{(F, 0)\}$, the agent chooses $(F, 0)$ and the principal's expected payoff is $E_F[y-w(y)]$, again, bounding $V_P(w|\widehat{\mathcal{A}}, 0)$ by $E_F[y - w(y)]$ from above.

Now suppose that $V_P(w|\widehat{\mathcal{A}}, 0) > y_0 - w(y_0)$. Let $F$ attain the minimum in (16). Suppose that $E_F[w(y)] > V_A(w|\widehat{\mathcal{A}})$. If $\text{supp}(F) \nsubseteq \underline{\mathcal{Y}}$, then let $\mathcal{A}$ be given by $\mathcal{A} = \widehat{\mathcal{A}} \cup \{(F', 0)\}$ where $F'$ is the mixture of $F$ with weight $1 - \epsilon$ and $\delta_{y_0}$ with weight $\epsilon$. For $\epsilon$ small enough, the agent chooses $(F', 0)$ contradicting minimality. Suppose now that $\text{supp}(F) \subseteq \underline{\mathcal{Y}}$. Then, the agent chooses action $(F, 0)$ if $\mathcal{A}$ is given by $\mathcal{A} = \widehat{\mathcal{A}} \cup \{(F, 0)\}$ bounding $V_P(w|\widehat{\mathcal{A}}, 0)$ above by $y_0 - w(y_0)$, a contradiction. Thus (i) – (iii) are shown.

If the conclusions in (ii) and (iii) hold, then for any $F$ attaining the minimum in (16) $V_A(w|\widehat{\mathcal{A}}) = E_F[w(y)]$ and $V_P(w|\widehat{\mathcal{A}}, 0) = E_F[y - w(y)]$. Hence, $(V_A(w|\widehat{\mathcal{A}}), V_P(w|\widehat{\mathcal{A}}, 0)) \in \mathcal{F}$. If the conclusion in (ii) holds but $V_P(w|\widehat{\mathcal{A}}, 0) = y_0 - w(y_0)$, then $V_A(w|\widehat{\mathcal{A}}) \leq w(y_0)$ and again $(V_A(w|\widehat{\mathcal{A}}), V_P(w|\widehat{\mathcal{A}}, 0)) \in \mathcal{F}$. If the conclusion in (iii) holds but $V_A(w|\widehat{\mathcal{A}}) = w(y_1)$, then $V_P(w|\widehat{\mathcal{A}}, 0)$ is bounded below by the minimum in (16) and above by $y_2 - w(y_2)$, as the agent does not take any action $(F, c)$ where $\text{supp}(F) \nsubseteq \bar{\mathcal{Y}}$, and again $(V_A(w|\widehat{\mathcal{A}}), V_P(w|\widehat{\mathcal{A}}, 0)) \in \mathcal{F}$. Lastly, if $V_A(w|\widehat{\mathcal{A}}) = w(y_1)$ and $V_P(w|\widehat{\mathcal{A}}, 0) = y_0 - w(y_0)$, then there exists action $(F, 0) \in \widehat{\mathcal{A}}$ such that $\text{supp}(F) \subseteq \bar{\mathcal{Y}}$, but for all such actions $\text{supp}(F) \subseteq \underline{\mathcal{Y}}$ as otherwise the agent would choose the action preferred by the principal. For such $F$, $V_A(w|\widehat{\mathcal{A}}) = E_F[w(y)]$ and $V_P(w|\widehat{\mathcal{A}}, 0) = E_F[y - w(y)]$. Hence, $(V_A(w|\widehat{\mathcal{A}}), V_P(w|\widehat{\mathcal{A}}, 0)) \in \mathcal{F}$.

Thus

$$\mathcal{R} \subseteq \mathcal{F}.$$

Now, take any $(u, v) \in \mathcal{F}$.

If $v = y_0 - w(y_0)$, then $u \leq w(y_0)$ so that $c := w(y_0) - u \geq 0$. Let $\widehat{\mathcal{A}} = \{(\delta_{y_0}, c)\}$. Clearly,

$V_A(w|\widehat{\mathcal{A}}) = w(y_0) - c = u$. Furthermore, $V_P(w|\widehat{\mathcal{A}}, 0)$ is bounded above by $y_0 - w(y_0)$, e.g. for $\mathcal{A} = \widehat{\mathcal{A}}$, and bounded below by $y_0 - w(y_0)$ by definition of $y_0$. Hence, $V_P(w|\widehat{\mathcal{A}}, 0) = y_0 - w(y_0) = v$ and $(u, v) \in \mathcal{R}$.

If $u = w(y_1)$, then $y_2 - w(y_2) \geq v \geq y_1 - w(y_1)$ so that there exists $x \in [0, 1]$ for which $F := x\delta_{y_2} + (1 - x)\delta_{y_1}$ satisfies $E_F[y - w(y)] = v$. Let $\widehat{\mathcal{A}} = \{(F, 0)\}$. Clearly, $V_A(w|\widehat{\mathcal{A}}) = w(y_1) = u$. Furthermore, $V_P(w|\widehat{\mathcal{A}}, 0)$ is bounded above by $E_F[y - w(y)]$, e.g. for $\mathcal{A} = \widehat{\mathcal{A}}$, and bounded below by $E_F[y - w(y)]$ as the agent will always choose the action preferred by the principal if he is indifferent. Hence, $V_P(w|\widehat{\mathcal{A}}, 0) = E_F[y - w(y)] = v$ and $(u, v) \in \mathcal{R}$.

Lastly, suppose that $v > y_0 - w(y_0)$ and $u < w(y_1)$. Pick

$$F^* \in \arg\min E_F[y - w(y)] \text{ over } F \in \Delta(\mathcal{Y}) \text{ such that } E[w(y)] = u. \tag{17}$$

(Note that this can only be done if we know that $u \geq \min_y w(y)$, but this is indeed the case: in fact $u \geq w(y_0)$ since otherwise the existence of $(u', v') = (w(y_0), y_0 - w(y_0))$ would contradict $(u, v) \in \mathcal{F}$.)

Let $\widehat{\mathcal{A}} = \{(F^*, 0)\}$. Clearly, $V_A(w|\widehat{\mathcal{A}}) = E_{F^*}[w(y)] = u$. Let us characterize $V_P(w|\widehat{\mathcal{A}}, 0)$.

Note that $F^*$ still attains the minimum in (17) when the constraint $E[w(y)] = u$ is replaced by $E[w(y)] \geq u$: if it did not, then $(u, v)$ would not belong to $\mathcal{F}$. Thus, $F^*$ attains the minimum in (16).

Lastly, if (16) does not hold with equality, then, by (ii), it must be that $u = V_A(w|\widehat{\mathcal{A}}) \geq w(y_1)$ contrary to our initial assumption. Thus, (16) holds with equality implying that $V_P(w|\widehat{\mathcal{A}}, 0) = E_{F^*}[y - w(y)] = v$ and $(u, v) \in \mathcal{R}$. □

*Proof of Theorem 1.* Consider any contract $w$. Note that $V_P(w|\widehat{\mathcal{A}}, c_P) = V_P(w|\widehat{\mathcal{A}}, 0) - c_P$. Let $\mathcal{T}$ consist of points $(u, v)$ such that $u > V_A(w|\widehat{\mathcal{A}})$ and $v < V_P(w|\widehat{\mathcal{A}}, 0)$. Lemma 1 tells us that $\mathcal{S}$ and $\mathcal{T}$ are disjoint.

By the separating hyperplane theorem, there exist constants $\kappa, \lambda$ and $\mu$ with $(\lambda, \mu) \neq (0, 0)$ such that

$$\kappa + \lambda u - \mu v \leq 0 \text{ for all } (u, v) \in \mathcal{S}, \tag{18}$$

$$\kappa + \lambda u - \mu v \geq 0 \text{ for all } (u, v) \in \mathcal{T}. \tag{19}$$

(19) implies that $\lambda, \mu$ are nonnegative as otherwise the inequality is not satisfied for large $u$ or small $v$. Rearranging (18) implies that

$$\mu(y - w(y)) \geq \kappa + \lambda w(y) \qquad \text{for all } y \in \mathcal{Y}. \tag{20}$$

$(V_A(w|\widehat{\mathcal{A}}), V_P(w|\widehat{\mathcal{A}}, 0))$ is in the closures of both $\mathcal{S}$ and $\mathcal{T}$ implying that

$$\mu V_P(w|\widehat{\mathcal{A}}, 0) = \kappa + \lambda V_A(w|\widehat{\mathcal{A}}). \tag{21}$$

Define a linear contract $w'$ by

$$w'(y) = \frac{\mu}{\mu + \lambda} y - \frac{\kappa}{\mu + \lambda}. \tag{22}$$

$w'$ satisfies (20) as an equality. For any technology $\mathcal{A} \supseteq \widehat{\mathcal{A}}$, let $(F, c)$ be an action that the agent takes under contract $w'$. Taking expectation over $y$ distributed according to $F$, (20) for $w'$ implies

$$\mu E_F[y - w'(y)] \geq \kappa + \lambda E_F[w'(y)].$$

$w' \geq w$ pointwise which implies that the agent's expected payoff if his technology is just $\widehat{\mathcal{A}}$ is at least as large under $w'$ as under $w$. As $c \geq 0$, we have

$$\begin{aligned} \mu E_F[y - w'(y)] &\geq \kappa + \lambda E_F[w'(y)] \geq \kappa + \lambda \left( E_F[w'(y)] - c \right) \\ &\geq \kappa + \lambda V_A(w'|\widehat{\mathcal{A}}) \geq \kappa + \lambda V_A(w|\widehat{\mathcal{A}}). \end{aligned}$$

Combining the above with (21) gives

$$\mu E_F[y - w'(y)] \geq \mu V_P(w|\widehat{\mathcal{A}}, 0)$$

for any $(F, c)$ the agent might choose given $w'$.

If $\mu > 0$, this implies

$$V_P(w'|\widehat{\mathcal{A}}, 0) \geq V_P(w|\widehat{\mathcal{A}}, 0).$$

If $\mu = 0$, then it must have been the case that $V_A(w|\widehat{\mathcal{A}}) = w(y_1)$. In this case, $w'$ is constant with value $-\frac{\kappa}{\lambda} = w(y_1)$. Given $\widehat{\mathcal{A}}$, the agent chooses an action $(F, 0)$ with $F$ having support on $\bar{\mathcal{Y}}$. Thus, as he chooses the action of the actions above that guarantees the principal the highest expected payoff, increasing all wages to $w(y_1)$, as $w'$ does, only increases the principal's guarantee. Thus, in either case

$$V_P(w'|\widehat{\mathcal{A}}, 0) \geq V_P(w|\widehat{\mathcal{A}}, 0). \tag{23}$$

Furthermore, $w'$ is eligible because

$$\begin{array}{ccccc} V_P(w|\widehat{\mathcal{A}}, c_P) > 0 & \Longrightarrow & V_P(w'|\widehat{\mathcal{A}}, c_P) > 0 & \text{by (23);} \\ V_P(w|\widehat{\mathcal{A}}, c_P) \geq -w(0) & \Longrightarrow & V_P(w'|\widehat{\mathcal{A}}, c_P) \geq -w'(0) & \text{by (23) and as } w'(0) \geq w(0); \\ V_A(w|\widehat{\mathcal{A}}) \geq 0 & \Longrightarrow & V_A(w'|\widehat{\mathcal{A}}) \geq 0 & \text{because } w' \geq w \text{ pointwise.} \end{array}$$

Thus, we have an eligible linear contract $w'$ delivering at least as high a guarantee as $w$.

It remains only to show that an optimum among eligible linear contracts exists; this is done in Section 2.2.3. □

*Proof of Corollary 1.* Assume that $w$ is a nonlinear eligible optimal contract and define $w'$ as in (22). Suppose first that $\mu > 0$ and $\lambda > 0$. Define $w'$ as in Theorem 1 by (22). $w'$ satisfies (20) as an equality. Thus,

$$\mu V_P(w'|\widehat{\mathcal{A}}, c_P) \geq \kappa + \lambda V_A(w'|\widehat{\mathcal{A}}) = \mu V_P(w|\widehat{\mathcal{A}}, c_P) + \lambda(V_A(w'|\widehat{\mathcal{A}}) - V_A(w|\widehat{\mathcal{A}})). \tag{24}$$

$w' \geq w$ pointwise, with strict inequality for some output levels. Thus

$$V_A(w'|\widehat{\mathcal{A}}) > V_A(w|\widehat{\mathcal{A}})$$

because $\widehat{\mathcal{A}}$ satisfies the full-support condition. As $\lambda$ in (24) is strictly positive, it follows that

$$V_P(w'|\widehat{\mathcal{A}}, c_P) > V_P(w|\widehat{\mathcal{A}}, c_P)$$

contradicting the assumption that $w$ is optimal.

If $\lambda = 0$, then $V_P(w|\widehat{\mathcal{A}}, 0) = y_0 - w(y_0)$. This case can only arise if $c_P = 0$, since otherwise the principal would not have the incentive to supply the input. In this case, as argued in Section 2.2.3, the principal can extract the full (known) surplus with a linear contract of slope 1, so $V_P(w|\widehat{\mathcal{A}}, 0)$ must equal this full surplus in order for $w$ to be optimal. But then, since $V_A(w'|\widehat{\mathcal{A}}) > V_A(w|\widehat{\mathcal{A}}) \geq 0$ and $V_P(w'|\widehat{\mathcal{A}}, 0) \geq V_P(w|\widehat{\mathcal{A}}, 0)$, it follows that the sum of the two parties' guarantees under $w'$, $V_A(w'|\widehat{\mathcal{A}}) + V_P(w'|\widehat{\mathcal{A}}, 0)$, exceeds the full surplus, which is impossible.

Finally, if $\mu = 0$ and $\lambda > 0$, then $V_A(w|\widehat{\mathcal{A}}) = w(y_1)$. By the full support assumption, it must be that $\bar{\mathcal{Y}} = \mathcal{Y}$ implying that contract $w(y)$ is constant, again a contradiction. $\square$

*Proof of Lemma 2.* A) $\alpha$ cannot be strictly greater than 1 because (E2) would be violated; take $\mathcal{A} = \widehat{\mathcal{A}} \cup \{(\delta_{\bar{y}}, 0)\}$ for $\bar{y} = \max(\mathcal{Y})$. $\alpha$ cannot be strictly less than 0 because the agent would take action $(\delta_0, 0)$ if available so that (E2) would imply that $c_P = 0$, $V_P(w[\alpha, \beta]) = -\beta$ and $V_A(w[\alpha, \beta]|\widehat{\mathcal{A}}) \leq \beta$ which implies that not both (E1) and (E3) can be satisfied.

B) If $\alpha = 0$, there are two cases to consider. If there are no actions of the form $(F, 0)$ in $\widehat{\mathcal{A}}$, then for $\mathcal{A} = \widehat{\mathcal{A}} \cup \{(\delta_0, 0)\}$ the agent chooses action $(\delta_0, 0)$, thus failing to generate a positive total surplus; as noted earlier, this contradicts eligibility. If there are actions of the form $(F, 0)$ in $\widehat{\mathcal{A}}$, the principal's guarantee is given by $V_P(w[0, \beta]) = \max_{(F,0) \in \widehat{\mathcal{A}}} E_F[y] - \beta - c_P$.

C) If $\alpha = 1$, (E2) implies that $c_P = 0$ for such contract to be eligible; if it is costly to supply the input and the principal does not receive any share of the output, she will abstain from supplying the input. The principal's guarantee is thus given by $V_P(w[1, \beta]) = -\beta$. $\square$

*Proof of Lemma 3.* The arguments preceding the lemma statement show that, if any eligible linear contract exists, then $w[\alpha^*, \beta(\alpha^*)]$ is an optimal eligible linear contract. If for some $(F^*, c^*) \in \arg\max_{(F,c) \in \widehat{\mathcal{A}}} \left\{E_F[y] - \frac{c}{\alpha^*}\right\}$, we have $c^* = 0$, then the zero contract is an optimal contract and, in fact, $V_P(w[\alpha, \beta(\alpha)]) = V_P(w[\alpha^*, \beta(\alpha^*)])$ for all $\alpha \in [0, \alpha^*]$; and further $w[\alpha, \beta(\alpha)]$ is eligible. If for all such $(F^*, c^*)$, we have $c^* > 0$, then for all $\alpha < \alpha^*$, $V_P(w[\alpha, \beta(\alpha)]) < V_P(w[\alpha^*, \beta(\alpha^*)])$ and uniqueness follows.

The second claim holds as for $\alpha = 1$

$$\frac{1 - \alpha}{\alpha} \max_{(F,c) \in \widehat{\mathcal{A}}} \{E_F[\alpha y] - c\} - c_P = -c_P \leq 0$$

implying (13) and thus

$$V_P(w[\alpha^*, \beta(\alpha^*)]) = \frac{1 - \alpha^*}{\alpha^*} \max_{(F,c) \in \widehat{\mathcal{A}}} \{E_F[\alpha^* y] - c\} - \beta(\alpha^*) - c_P = -\beta(\alpha^*).$$

$\square$

*Proof of Proposition 1.* Suppose there exists $(F, c) \in \widehat{\mathcal{A}}$ such that

$$E_F[y] - c - c_P \geq 2\sqrt{c c_P} \qquad \text{and} \qquad E_F[y] - c - c_P > 0.$$

We draw on the alternative characterization of the optimal linear contract in Appendix C.1. Let $r$ be given by (27) and let

$$\alpha = r(F, c) \text{ and } \beta = -\{E_F[\alpha y] - c\}.$$

Note that $\alpha \in (0, 1]$. $w[\alpha, \beta]$ is eligible if and only if (7) – (9) hold. (8) and (9) are satisfied because

$$\frac{1-\alpha}{\alpha} \max_{(F',c') \in \widehat{\mathcal{A}}} \{E_{F'}[\alpha y] - c'\} - c_P \geq \frac{1-\alpha}{\alpha}\{E_F[\alpha y] - c\} - c_P = 0 \tag{25}$$

and

$$\max_{(F',c') \in \widehat{\mathcal{A}}} \{E_{F'}[\alpha y] + \beta - c'\} \geq E_F[\alpha y] + \beta - c = 0.$$

(7) is satisfied if $\beta < 0$. Suppose $\beta \geq 0$. Then $c_P = 0$ by (25). $c_P = 0$ implies that selling the firm guarantees an expected payoff of $\max_{(F',c') \in \widehat{\mathcal{A}}}\{E_{F'}[y] - c'\} - c_P \geq E_F[y] - c - c_P > 0$ to the principal and is thus eligible.

Suppose now that an eligible contract exists. By Theorem 1, there exists a contract $w[\alpha, \beta]$ that maximizes the principal's guarantee. By Lemma 3, one such optimal eligible linear contract is given by $w[\alpha^*, \beta(\alpha^*)]$ defined by (12).

Let

$$(F, c) \in \arg\max_{(F',c') \in \widehat{\mathcal{A}}} \{E_{F'}[\alpha^* y] - c'\}.$$

It is immediate that $E_F[y] - c - c_P > 0$ as $E_F[y] - c - c_P \geq E_F[\alpha^* y] - c - c_P$ and $E_F[\alpha^* y] - c - c_P$ is an upper bound on the sum of the principal's and the agent's guarantee which is strictly positive for an eligible contract.

By definition of $\alpha^*$ and $(F, c)$, and using (13),

$$\frac{1-\alpha^*}{\alpha^*}\{E_F[\alpha^* y] - c\} - c_P = 0$$

implying further that $E_F[y] - c - c_P \geq 2\sqrt{c c_P}$ for $\alpha^*$ to be real. □

*Proof of Lemma 4.* For the first statement: Suppose that $w$ satisfies the stated conditions but is not EFA. Then there exists some FA-worrisome input $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ and $\mathcal{A} \supseteq \widehat{\mathcal{A}}$ such that

$$V_P(w|\mathcal{A}, c_P) > V_P(w|\mathcal{W}) \qquad \text{and} \qquad V_A(w|\mathcal{A}) < 0.$$

As $\mathcal{A} \supseteq \widehat{\mathcal{A}}$, $(V_A(w|\mathcal{A}), V_P(w|\mathcal{A}, c_P)) \in \mathcal{F}_{(\widehat{\mathcal{A}}, c_P)}$. Furthermore, as $V_A(w|\mathcal{A}) < 0$ and $V_P(w|\mathcal{A}, c_P) > V_P(w|\mathcal{W})$, $(V_A(w|\mathcal{A}), V_P(w|\mathcal{A}, c_P)) \in \mathcal{C}$, contradicting $\mathcal{U}_{\mathcal{W}} \cap \mathcal{C} = \emptyset$.

Suppose now that $\mathcal{U}_{\mathcal{W}} \cap \mathcal{C} \neq \emptyset$ and consider $(u, v) \in \mathcal{U}_{\mathcal{W}} \cap \mathcal{C}$. Then $(u, v) \in \mathcal{F}_{(\widehat{\mathcal{A}}, c_P)}$ for some $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$. Using Lemma 1, there exists $\mathcal{A}'$ such that $(V_A(w|\mathcal{A}'), V_P(w|\mathcal{A}', c_P)) = (u, v)$. Let $\mathcal{A} = \widehat{\mathcal{A}} \cup \mathcal{A}'$ and note that $V_A(w|\mathcal{A}) = \max\{V_A(w|\widehat{\mathcal{A}}), V_A(w|\mathcal{A}')\} = u < 0$, while $V_P(w|\mathcal{A}, c_P) \geq V_P(w|\mathcal{A}', c_P) = v > V_P(w|\mathcal{W})$. Thus, $(\widehat{\mathcal{A}}, c_P)$ is FA-worrisome given contract $w$.

Finally, for the last statement, suppose that $w = w[\alpha, \beta]$ is linear, is locally eligible via some non-optimal input $(\widehat{\mathcal{A}}, c_P)$, and $\mathcal{U}_{\mathcal{W}} \cap \mathcal{C} = \emptyset$. We need to show that $w$ is locally eligible via an optimal input as well. Clearly (E1) and (E2) still hold, so we only need to show that (E3) holds for an optimal input. Given Lemma 2 part A), we know $\alpha \in [0, 1]$, so consider two cases:

- If $\alpha \in [0, 1)$, let $(\widehat{\mathcal{A}}^*, c_P^*)$ be an optimal input, and assume for contradiction that $V_A(w|\widehat{\mathcal{A}}^*, c_P^*) < 0$. First suppose $V_A(w|\widehat{\mathcal{A}}^*, c_P^*) \geq w(0)$. Let $F$ be a worst-case distribution for the principal under this input, so that $V_A(w|\widehat{\mathcal{A}}^*, c_P^*) = E_F[w(y)]$ and $V_P(w|\widehat{\mathcal{A}}^*, c_P^*) = E_F[y - w(y)] - c_P^*$. Take $\epsilon$ small, let $F' = (1 - \epsilon)F + \epsilon\delta_{\bar{y}}$ where $\bar{y} = \max(\mathcal{Y})$, and let $\mathcal{A} = \widehat{\mathcal{A}}^* \cup \{(F', 0)\}$. Note that $F \neq \delta_{\bar{y}}$ since otherwise the agent could not get a positive payoff under $w$ at all. Then, under $(\mathcal{A}, c_P^*)$, the principal's guarantee is strictly higher than $V_P(w|\widehat{\mathcal{A}}^*, c_P^*)$ (since the principal receives a share $1 - \alpha > 0$ of improved output relative to $F$) and the agent's payoff is still below 0; this gives a point lying in $\mathcal{F}_{(\widehat{\mathcal{A}}^*, c_P^*)} \cap \mathcal{C}$, contradicting the assumption that this intersection was empty.

  This leaves the possibility $V_A(w|\widehat{\mathcal{A}}^*, c_P^*) < w(0)$. In this case, the agent would produce $\delta_0$ if it comes at cost 0, so the principal's guarantee is $V_P(w|\widehat{\mathcal{A}}^*, c_P^*) = -\beta - c_P^* \leq -\beta$. But we know that the contract is locally eligible via $(\widehat{\mathcal{A}}, c_P)$, so the guarantee from this input is $\geq -\beta$, so this input is also optimal and we are done.

- If $\alpha = 1$, then as we saw in Section 2.2.3, $w$ can only be locally eligible via $(\widehat{\mathcal{A}}, c_P)$ if $c_P = 0$, and then any input with $c_P = 0$ is optimal, since the principal's payoff is always $-\beta$.

$\square$

*Proof of Proposition 3.* Take any contract $w$ that is locally eligible and let $(\widehat{\mathcal{A}}^*, c_P^*)$ be an optimal input given $w$. Define $\kappa, \lambda, \mu, y_0, y_1, y_2$, and $w'$ as in the proof of Theorem 1. As in that earlier proof, $w'$ is locally eligible (via the same input). Furthermore, the critical region decreases in the sense of set inclusion. Let $\bar{y} = \max_{y \in \mathcal{Y}} y$ and, similar to before,

$$\mathcal{S}' = \operatorname{conv}\left(\{(w'(y) - c, y - w'(y)) : y \in \mathcal{Y}, c \in \mathbb{R}^+\}\right).$$

The fundamental relationship between the principal's and the agent's guarantee given $w'$ is now given by

$$\mathcal{F}' = \{(u, v) \in \mathcal{S}' : \nexists (u', v') \in \mathcal{S}', u' > u, v' < v\}.$$

Let $\mathcal{F}'_{(\widehat{\mathcal{A}}, c_P)}$ be defined as

$$\mathcal{F}'_{(\widehat{\mathcal{A}}, c_P)} = \{(u, v) : u \geq V_A(w'|\widehat{\mathcal{A}}), v \geq V_P(w'|\widehat{\mathcal{A}}, c_P), (u, v + c_P) \in \mathcal{F}'\}.$$

Take any $(u', v') \in \mathcal{F}'$ with $u' < 0$. The frontier $\mathcal{F}$ contains some point $(u'', v'')$ with $u'' \leq u'$ (for example, it contains all such points with $v'' = y_0 - w(y_0)$ and $u''$ sufficiently low). But as $w$ is locally eligible, $\max_y w(y) \geq 0$, so $\mathcal{F}$ also contains some point whose first coordinate is positive. Hence, there exists some intermediate $(u, v) \in \mathcal{F}$ with $u = u'$.

Let $F, F' \in \Delta(\mathcal{Y})$ satisfy

$$(u, v) = (E_F[w(y)], E_F[y - w(y)]) \text{ and } (u', v') = (E_{F'}[w'(y)], E_{F'}[y - w'(y)]).$$

If $E_F[y] < E_{F'}[y]$, then

$$E_{F'}[w'(y)] \geq E_F[w'(y)] \geq E_F[w(y)]$$

by linearity of $w'$ and as $w' \geq w$ pointwise. The first inequality is strict unless $\mu = 0$ in which case $w'$ is constant at $-\frac{\kappa}{\lambda} = w(y_1)$. But then, $w'$ is constant and nonnegative, contradicting our earlier statements $E_{F'}[w'(y)] = u' < 0$.

Thus, $\mu > 0$, and it must be that $E_F[y] \geq E_{F'}[y]$ which implies that $v \geq v'$. As $V_A(w'|\widehat{\mathcal{A}}) \geq V_A(w|\widehat{\mathcal{A}})$ as $w' \geq w$ pointwise, it follows that if $(u', v') \in \mathcal{F}'_{(\widehat{\mathcal{A}}, c_P)}$, there exists some $(u, v) \in \mathcal{F}_{(\widehat{\mathcal{A}}, c_P)}$ with $u = u'$ and $v \geq v'$.

It follows that the new critical region and the feasible region given $w'$ still do not intersect, so that Lemma 4 assures EFA for $w'$. (Note that because $w'$ is linear, we need not know whether $(\widehat{\mathcal{A}}^*, c_P^*)$ remains optimal under $w'$.)

Thus, whenever a nonlinear EFA contract exists, so does a linear EFA contract that furthermore guarantees the principal a weakly greater payoff.

Thus, the principal's maximum guarantee (assuming it exists) among EFA contracts is attained by a linear contract, and to complete the proof of the proposition, it suffices to show that the maximum over linear contracts is attained. In fact, we will identify it.

Let $(\underline{\widehat{\mathcal{A}}}, \underline{c}_P) \in \arg\min_{(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}} c_P$. Note that a contract is EFA if and only if it is locally eligible via some optimal input and $(\underline{\widehat{\mathcal{A}}}, \underline{c}_P)$ is not a FA-worrisome input; thus we do not need to concern ourselves with other inputs being FA-worrisome. To see this suppose that $(\underline{\widehat{\mathcal{A}}}, \underline{c}_P)$ is not a FA-worrisome input but $(\widehat{\mathcal{A}}, c_P)$ is. Then, there exists $(u, v) \in \mathcal{F}_{(\widehat{\mathcal{A}}, c_P)}$ such that $u < 0$ and $v > V_P(w|\mathcal{W})$. As before, using Lemma 1, there exists $\mathcal{A} \supseteq \widehat{\mathcal{A}}$ such that $V_A(w|\mathcal{A}) = u$ and $V_P(w|\mathcal{A}, c_P) = v$. But then let $\mathcal{A}' = \mathcal{A} \cup \underline{\widehat{\mathcal{A}}}$ and note that $V_P(w|\mathcal{A}', \underline{c}_P) = V_P(w|\mathcal{A}, \underline{c}_P) \geq V_P(w|\mathcal{A}, c_P) = v$, where the first equality follows as otherwise $V_P(w|\underline{\widehat{\mathcal{A}}}, \underline{c}_P) > v$, a contradiction. Thus the two points on the frontier $\mathcal{F}_{(\underline{\widehat{\mathcal{A}}}, \underline{c}_P)}$ corresponding to technologies $\mathcal{A}'$ and $\mathcal{A}$ have the same value of $V_P$. If they have the same value of $V_A$ as well, we have $V_A(w|\mathcal{A}') = V_A(w|\mathcal{A}) < 0$ and $(\underline{\widehat{\mathcal{A}}}, \underline{c}_P)$ is a FA-worrisome input, a contradiction. Otherwise, both points must lie on a flat segment of $\mathcal{F}_{(\underline{\widehat{\mathcal{A}}}, \underline{c}_P)}$, and then the minimum of $V_P$ along the frontier is also attained on this segment, implying $V_P(w|\underline{\widehat{\mathcal{A}}}, \underline{c}_P) = V_P(w|\mathcal{A}', \underline{c}_P) \geq v > V_P(w|\mathcal{W})$, contradicting the definition of $V_P(w|\mathcal{W})$.

With this in mind, we consider each $\alpha$ and identify the optimal $\beta$ for which $w[\alpha, \beta]$ is EFA, then argue that the resulting optimization over $\alpha$ has a solution.

Consider a linear contract $w[\alpha, \beta]$. For any $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$, as long as the contract is locally eligible, the principal's and the agent's guarantee are related as

$$V_P(w[\alpha, \beta]|\widehat{\mathcal{A}}, c_P) = \frac{1-\alpha}{\alpha}(V_A(w[\alpha, \beta]|\widehat{\mathcal{A}}) - \beta) - \beta - c_P.$$

As previously argued, it suffices to optimize over linear contracts such that $(\underline{\widehat{\mathcal{A}}}, \underline{c}_P)$ is not a FA-worrisome input. This is guaranteed if for any $\mathcal{A} \supseteq \underline{\widehat{\mathcal{A}}}$, the relation

$$V_P(w[\alpha, \beta]|\mathcal{A}, \underline{c}_P) = \frac{1-\alpha}{\alpha}(V_A(w[\alpha, \beta]|\mathcal{A}) - \beta) - \beta - \underline{c}_P > V_P(w[\alpha, \beta]|\mathcal{W})$$

implies

$$V_A(w[\alpha,\beta]|\mathcal{A}) \geq 0.$$

A decrease in $\beta$ has no effect on the former of these conditions and tightens the latter, so given $\alpha$, it is optimal to decrease $\beta$ until the former condition just becomes satisfiable with $V_A(w[\alpha,\beta]|\mathcal{A}) = 0$. Thus, the optimal $\beta$, call it $\beta_{EFA}(\alpha)$, satisfies

$$\frac{1-\alpha}{\alpha} \cdot 0 - \frac{\beta_{EFA}(\alpha)}{\alpha} - \underline{c}_P = V_P(w[\alpha, \beta_{EFA}(\alpha)]|\mathcal{W}).$$

Let $(\widehat{\mathcal{A}}^*, c_P^*)$ be an optimal input. As

$$V_P(w[\alpha,\beta]|\mathcal{W}) = V_P(w[\alpha,\beta]|\widehat{\mathcal{A}}^*, c_P^*) = \frac{1-\alpha}{\alpha} \max_{(F,c)\in\widehat{\mathcal{A}}^*} \{E_F[\alpha y] - c\} - c_P^* - \beta,$$

$\beta$ is given by

$$\beta_{EFA}(\alpha) = (c_P^* - \underline{c}_P)\frac{\alpha}{1-\alpha} - \max_{(F,c)\in\widehat{\mathcal{A}}^*} \{E_F[\alpha y] - c\}.$$

With this choice of $\beta$ for each $\alpha$, then, the guarantee is

$$\frac{1}{\alpha} \max_{(F,c)\in\widehat{\mathcal{A}}^*} \{E_F[\alpha y] - c\} - \frac{\alpha}{1-\alpha}(c_P^* - \underline{c}_P) - c_P^*.$$

We maximize this over all choices of $\alpha$ that arise in some EFA contract, taking $(\widehat{\mathcal{A}}^*, c_P^*)$ to be the optimal input for given $\alpha$. Notice, moreover, that for any fixed $\alpha$, the choice of input $(\widehat{\mathcal{A}}^*, c_P^*)$ that is optimal is in fact the same one that maximizes the expression above. So the problem is equivalent to maximizing

$$\frac{1}{\alpha} \max_{(F,c)\in\widehat{\mathcal{A}}} \{E_F[\alpha y] - c\} - \frac{\alpha}{1-\alpha}(c_P - \underline{c}_P) - c_P$$

over $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ and $\alpha \in [0,1]$ such that

$$\frac{1-\alpha}{\alpha} \max_{(F,c)\in\widehat{\mathcal{A}}} \{E_F[\alpha y] - c\} - c_P \geq 0$$

to satisfy local eligibility (recall (8)).

(In the special case $\alpha = 1$, local eligibility is possible only if $c_P = \underline{c}_P = 0$, and then the contract is automatically EFA; the formula remains valid in this case with the $\frac{\alpha}{1-\alpha}(c_P - \underline{c}_P)$ term interpreted as zero. In the case $\alpha = 0$, the formula is also correct as long as the contract is locally eligible, as discussed in Section 2.2.3.)

The objective function is continuous in $\alpha$, so it obtains its maximum. $\square$

*Proof of Proposition 4.* Take any contract $w$ that is locally eligible via some optimal input. Suppose $w$ is not EFI. Then there exists $\widetilde{\mathcal{W}} \supseteq \mathcal{W}$ and $(\widehat{\mathcal{A}}, c_P) \in \widetilde{\mathcal{W}}$ such that

$$V_P(w|\widehat{\mathcal{A}}, c_P) > V_P(w|\mathcal{W}) \qquad \text{and} \qquad V_A(w|\widehat{\mathcal{A}}) < 0. \tag{26}$$

Recall that $\mathcal{A}_{triv} = \{(\delta_0, 0)\}$. Let $\mathcal{A} = \widehat{\mathcal{A}} \cup \mathcal{A}_{triv}$. Consider $(\mathcal{A}_{triv}, 0) \in \mathcal{W}'$. Clearly, $\mathcal{A} \supseteq \mathcal{A}_{triv}$. If

$$V_A(w|\mathcal{A}) \geq 0,$$

then the agent must be choosing action $(\delta_0, 0)$ implying that the principal's guarantee from $(\widehat{\mathcal{A}}, c_P)$ is at most 0, contradicting (26) as $V_P(w|\mathcal{W}) > 0$.

Thus,

$$V_A(w|\mathcal{A}) < 0 \qquad \text{and} \qquad V_P(w|\mathcal{A}, 0) \geq V_P(w|\widehat{\mathcal{A}}, c_P) > V_P(w|\mathcal{W}),$$

i.e. $(\mathcal{A}_{triv}, 0) \in \mathcal{W}'$ is a FA-worrisome input given contract $w$.

Suppose now that $w$ is not EFA under input space $\mathcal{W}'$. Then there exists some $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}'$ that is FA-worrisome given $w$, i.e. there exists $\mathcal{A} \supseteq \widehat{\mathcal{A}}$ such that

$$V_P(w|\mathcal{A}, c_P) > V_P(w|\mathcal{W}) \qquad \text{and} \qquad V_A(w|\mathcal{A}) < 0.$$

Let $\widetilde{\mathcal{W}} = \mathcal{W} \cup \{(\mathcal{A}, c_P)\}$. Clearly, $\widetilde{\mathcal{W}} \supseteq \mathcal{W}$ and $(\mathcal{A}, c_P) \in \widetilde{\mathcal{W}}$. As

$$V_P(w|\mathcal{A}, c_P) > V_P(w|\mathcal{W}) \qquad \text{and} \qquad V_A(w|\mathcal{A}) < 0,$$

$w$ is not EFI. $\square$

*Proof of Proposition 6.* Let $c'_P \geq c_P, c' \geq c$ and $\mu' \leq \mu$. Then $\alpha(c'_P, \mu', c') \leq \alpha(c_P, \mu, c)$ follows from

$$\frac{1-\alpha}{\alpha}\{\alpha\mu - c\} - c_P \geq \frac{1-\alpha}{\alpha}\{\alpha\mu' - c'\} - c'_P \qquad \text{for all } \alpha \in [0, 1]$$

and Lemma 3. $-\beta(c'_P, \mu', c') \leq -\beta(c_P, \mu, c)$ then follows as

$$-\beta(c'_P, \mu', c') = \alpha'\mu' - c' \leq \alpha\mu - c = -\beta(c_P, \mu, c),$$

with $\alpha = \alpha(c_P, \mu, c)$ and $\alpha' = \alpha(c'_P, \mu', c')$, and by (10). $\square$

*Proof of Proposition 7.* Using the characterization of the optimal linear contract detailed in Appendix C.3, we have

$$\alpha(c_P, \underline{c}_P, \mu, c) = \frac{\sqrt{c}}{\sqrt{c} + \sqrt{c_P - \underline{c}_P}}$$

and

$$-\beta(c_P, \underline{c}_P, \mu, c) = -\sqrt{c_P - \underline{c}_P}\sqrt{c} + \left\{ \frac{\sqrt{c}}{\sqrt{c} + \sqrt{c_P - \underline{c}_P}} \mu - c \right\}.$$

Note that $-\beta(c_P, \underline{c}_P, \mu, c)$ increases in $c$ for large $\mu$ as long as $c_P > \underline{c}_P$, and such a numerical specification is consistent with eligiblity nonbinding.[7] To show that $-\beta(c_P, \underline{c}_P, \mu, c)$ can also decrease in $c$, consider the following numerical example. Let $c_P = 8, \underline{c}_P = 4, \mu = 40$ and consider two values for $c$, 4 and 9. First,

$$-\beta(c_P, \underline{c}_P, \mu, c = 4) = 12 > 9 = -\beta(c_P, \underline{c}_P, \mu, c = 9),$$

that is, the fee is decreasing in $c$. Furthermore, one can easily check that the left-hand side of (35) equals 16 and 5 for $c$ equal to 4 and 9, respectively, so that eligibility is indeed nonbinding.

The rest of the claims follow immediately. □

*Proof of Proposition 8.* We again use the characterization of the optimal contract detailed in Appendix C.3. Following (33), the objective function is given by

$$p\mu - c - g(\gamma p) - 2\sqrt{cg(\gamma p)}$$

and is supermodular in $(p, \mu)$. Thus, as $\mu$ increases, the principal chooses higher-quality inputs, i.e. increases $p$. Following (15) the optimal share is given by

$$\alpha(\gamma, \mu, c) = \frac{\sqrt{c}}{\sqrt{c} + \sqrt{g(\gamma p)}}.$$

Thus, the share parameter decreases.

The objective function is also supermodular in $(p, -c)$. Thus, as $c$ increases, the principal chooses lower-quality inputs, i.e. decreases $p$. Furthermore, as the share parameter is increasing in $c$ and decreasing in $p$, it increases.

[7] For example, let $c_P > \underline{c}_P = 0$ and see footnote 4.

Lastly, to look at the effects of changing $\gamma$, think of the principal's choice variable as being $\zeta = \gamma p$ rather than just $p$. Thus, the principal is choosing $\zeta$ to maximize $\frac{\zeta\mu}{\gamma} - c - g(\zeta) - 2\sqrt{cg(\zeta)}$. This objective is supermodular in $\zeta$ and $-\gamma$. Therefore, when $\gamma$ increases, the optimal choice of $\zeta$ decreases. Since $\alpha$ is given by $\sqrt{c}/(\sqrt{c}+\sqrt{g(\zeta)})$ (and this expression does not depend directly on $\gamma$), it follows that $\alpha$ is increasing in $\gamma$. $\square$

# C Characterizations of optimal contracts

## C.1 Eligibility in single-input model

We give an alternative characterization of the optimal linear contract in the single-input model. Let the function $r : \widehat{\mathcal{A}} \to \{-1\} \cup [0, 1]$ be defined as

$$r(F,c) = \begin{cases} \frac{E_F[y]+c-c_P}{2E_F[y]} + \sqrt{\left(\frac{E_F[y]+c-c_P}{2E_F[y]}\right)^2 - \frac{c}{E_F[y]}} & \text{if } E_F[y] - c - c_P \geq 2\sqrt{cc_P} \\ & \quad \text{and } E_F[y] - c - c_P > 0 \\ -1 & \text{otherwise.} \end{cases} \tag{27}$$

Fixing an action $(F, c)$ in $\widehat{\mathcal{A}}$, function $r$ returns the larger root of the equation

$$\frac{1-\alpha}{\alpha}\left\{E_F[\alpha y] - c\right\} - c_P = 0.$$

(The second branch of (27) corresponds to the case where there are no real roots.) Combining this with (12) leads us to the following result.

Lemma **6.** *If an eligible linear contract exists, then*

$$\max_{(F,c)\in\widehat{\mathcal{A}}} r(F, c) = \alpha^*.$$

*Proof.* First suppose $c_P = 0$. Then, an eligible contract exists as long as there is some action with strictly positive surplus, and for any such action, the formula (27) for $r$ simplifies to 1, which indeed is the value of $\alpha^*$.

Now suppose $c_P > 0$. Let

$$\alpha = \max_{(F',c')\in\widehat{\mathcal{A}}} r(F', c') \qquad \text{and} \qquad (F, c) \in \arg\max_{(F',c')\in\widehat{\mathcal{A}}}\{E_{F'}[\alpha y] - c'\}.$$

(Note that the max in defining $\alpha$ exists: it could only fail to exist if the sup were approached by a sequence of actions whose limit fails to satisfy the strict inequality constraint in (27), that is, $E_F[y]-c-c_P=0$, but that still satisfy the weak inequality constraint. This requires the limit to satisfy $c=0$. In this case, the limiting value of the formula (27) is zero, which cannot be the supremum.)

To show that $\alpha \geq \alpha^*$, note that

$$\alpha = \max_{(F',c')\in\widehat{\mathcal{A}}} r(F',c') \geq r(F^*,c^*) = \alpha^*$$

where $(F^*,c^*) \in \arg\max_{(F',c')\in\widehat{\mathcal{A}}}\{E_{F'}[\alpha^* y]-c'\}-c_P$.

To show that $\alpha^* \geq \alpha$, note that

$$0 = \frac{1-\alpha}{\alpha}\{E_F[\alpha y]-c\} - c_P \leq \frac{1-\alpha}{\alpha}\max_{(F',c')\in\widehat{\mathcal{A}}}\{E_{F'}[\alpha y]-c'\}-c_P.$$

□

## C.2 Maximal eligibility

Here and for the rest of the appendix, we turn to the multiple-input model.

The optimal maximally eligible contract can be identified as follows. First, for a given $\alpha$, we derive the optimal $\beta$ for which the agent would accept the contract regardless of the input he expects the principal to supply. Second, we maximize the principal's guarantee over $\alpha \in [0,1]$ and $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ while ensuring local eligibility.

Thus, for any $\alpha \in [0,1]$, define

$$\beta_{ME}(\alpha) = -\min_{(\widehat{\mathcal{A}},c_P)\in\mathcal{W}}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y]-c\}.$$

Let $\alpha^*$ and $(\widehat{\mathcal{A}}^*, c_P^*)$ jointly maximize

$$\frac{1-\alpha}{\alpha}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y]-c\} - \beta_{ME}(\alpha) - c_P$$

over $\alpha \in [0,1]$ and $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ such that

$$\frac{1-\alpha}{\alpha}\max_{(F,c)\in\widehat{\mathcal{A}}}\{E_F[\alpha y]-c\} - c_P \geq 0.$$

Then an optimal maximally eligible contract (if any such contract exists) is given by $w[\alpha^*, \beta_{ME}(\alpha^*)]$, and is locally eligible for $(\widehat{\mathcal{A}}^*, c_P^*)$.

A maximally eligible contract exists if and only if

$$V_P(w[\alpha^*, \beta_{ME}(\alpha^*)] | \widehat{\mathcal{A}}^*, c_P^*) > 0.$$

## C.3 Eligibility with further actions

The optimal EFA linear contract can be identified as follows. As in Section C.2, we choose $\beta$ optimally for each $\alpha$ and then optimize over $\alpha$. This leads to the following formulas (the details of the derivation are given in the proof of Proposition 3): Maximize

$$\frac{1}{\alpha}(E_F[\alpha y] - c) - \frac{\alpha}{1-\alpha}(c_P - \underline{c}_P) - c_P \tag{28}$$

over $\alpha \in [0, 1]$, $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$, and $(F, c) \in \widehat{\mathcal{A}}$ satisfying

$$\frac{1-\alpha}{\alpha}(E_F[\alpha y] - c) - c_P \geq 0. \tag{29}$$

Here, $\underline{c}_P = \min_{(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}} c_P$. Refer to the optimal values in the above problem as $\alpha^*$ and $(\widehat{\mathcal{A}}^*, c_P^*)$. Then, set

$$\beta^* = (c_P^* - \underline{c}_P)\frac{\alpha^*}{1-\alpha^*} - (E_F[\alpha^* y] - c). \tag{30}$$

(We take the first term in (30) to be zero if $\alpha^* = 1$ and $c_P^* = \underline{c}_P$.)

Then an optimal EFA contract (if any such contract exists) is given by $w[\alpha^*, \beta^*]$, and it is locally eligible for the optimal input $(\widehat{\mathcal{A}}^*, c_P^*)$.

An EFA contract exists if and only if the value of (28) is positive, and if so, that expression gives the corresponding optimal value of $V_P$.

While the argument in the proof of Proposition 3 proves that an optimal contract exists, we now characterize the optimal contract further; this characterization will be useful for the comparative statics exercises in Section 4.

We proceed as follows. For any fixed choice of input and action in (28), we optimize over $\alpha$. These artificial optimization problems are easy to solve as the objectives will be concave and subject to an interval constraint. Thus, their solution falls into one of two cases: 1) The constraint holds with equality; or 2) the solution is characterized by a first-order condition.

If the solution to the overall optimization falls under case (1), then we can equivalently find it by pretending the constraint holds with equality for *all* input-action pairs (because this will only decrease the objective for other such pairs while leaving it unchanged for the optimal pair). Otherwise, if it falls under case (2), then we can plug in the globally optimal value of $\alpha$ for each input-action pair, and characterize the solution by maximizing (28) over the set of pairs for which the corresponding $\alpha$ satisfies the constraint.

To begin carrying out the above argument, fix $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ and $(F, c) \in \widehat{\mathcal{A}}$. We can rewrite the objective function (28) to get (14) as given in Section 4.2.1. The constraint to ensure local eligibility is still (29). For a fixed input and action, the objective is concave and $\alpha$ is constrained to lie in some interval. Thus, we can divide into two cases as above.

**Case 1:** First, suppose that the constraint binds in which case

$$\frac{1-\alpha}{\alpha}\{E_F[\alpha y] - c\} - c_P = 0.$$

The objective function evaluated at this contract is then given by

$$\frac{\alpha}{1-\alpha}c_P,$$

which is increasing in $\alpha$. Thus, we can identify the optimal contract and input, if the constraint binds, by a variation of function $r$ defined in (27) and Lemma 6 in Appendix C.1.

Define $r^m$, for multiple-input environment, as

$$r^m((F,c),c_P) = \begin{cases} \frac{E_F[y]+c-c_P}{2E_F[y]} + \sqrt{\left(\frac{E_F[y]+c-c_P}{2E_F[y]}\right)^2 - \frac{c}{E_F[y]}} & \text{if } E_F[y]-c-c_P \geq 2\sqrt{cc_P} \\ & \text{and } E_F[y]-c-c_P > 0 \\ -1 & \text{otherwise.} \end{cases} \tag{31}$$

$r^m$ is identical to $r$, except that $c_P$ is now an argument rather than a constant. Let

$$(\widehat{\mathcal{A}}^b, c_P^b) \in \underset{(\widehat{\mathcal{A}},c_P)\in\mathcal{W}}{\arg\max} \max_{(F,c)\in\widehat{\mathcal{A}}} r^m((F,c),c_P)$$

and

$$\alpha^b = \max_{(F,c)\in\widehat{\mathcal{A}}^b} r^m((F,c),c_P^b)$$

where superscript $b$ stands for binding.

The optimal contract is given by $w[\alpha^b, \beta_{EFA}(\alpha^b)]$, the principal provides input $(\widehat{\mathcal{A}}^b, c_P^b)$ and her guarantee is given by

$$\frac{\alpha^b}{1-\alpha^b}\underline{c}_P. \tag{32}$$

**Case 2:** Now, suppose that at the optimal contract, the constraint does not bind and can be ignored in the maximization over $\alpha$. Note that the objective function given in (14) is concave in $\alpha$. The first-order condition yields the optimal $\alpha$ given by (15). Evaluating (14) and $\beta^*$ as defined in (30) at this $\alpha$ gives

$$E_F[y] - c - c_P - 2\sqrt{c}\sqrt{c_P - \underline{c}_P} \tag{33}$$

and

$$\sqrt{c}\sqrt{c_P - \underline{c}_P} - \left\{ E_F\left[\frac{\sqrt{c}}{\sqrt{c} + \sqrt{c_P - \underline{c}_P}}y\right] - c\right\} \tag{34}$$

as the principal's guarantee and negative of the fee respectively.

To check whether the constraint holds, rewrite the condition as

$$\frac{1}{\alpha}\{E_F[\alpha y] - c\} - \frac{1}{1-\alpha}c_P \geq 0$$

and note that

$$\begin{aligned}\frac{1}{\alpha}\{E_F[\alpha y] - c\} - \frac{1}{1-\alpha}c_P =& \frac{1}{\alpha}\{E_F[\alpha y] - c\} - \frac{\alpha}{1-\alpha}(c_P - \underline{c}_P) - c_P - \frac{\alpha}{1-\alpha}\underline{c}_P \\ =& E_F[y] - c - c_P - 2\sqrt{c}\sqrt{c_P - \underline{c}_P} - \frac{\sqrt{c}}{\sqrt{c_P - \underline{c}_P}}\underline{c}_P.\end{aligned}$$

Thus, if

$$E_F[y] - c - c_P - 2\sqrt{c}\sqrt{c_P - \underline{c}_P} - \frac{\sqrt{c}}{\sqrt{c_P - \underline{c}_P}}\underline{c}_P \geq 0, \tag{35}$$

then local eligibility is satisfied.

Thus, the optimal contract can be found by maximizing (33) over $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ and $(F, c) \in \widehat{\mathcal{A}}$ such that (35) is satisfied; and then comparing the guarantee it provides to (32).

## C.4 Eligibility with full knowledge

The optimal EFK linear contract can be identified as follows. First, for a given $\alpha$, we derive the optimal $\beta$ for which the agent would accept the contract regardless of the technology he has and the input choice by the principal that it induces. Second, we maximize the objective function over $\alpha \in [0,1]$ and $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ while ensuring local eligibility. Thus, for any $\alpha \in [0,1]$ and $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$, define $\mathcal{W}^w(\alpha, (\widehat{\mathcal{A}}, c_P)) \subseteq \mathcal{W}$ as $\mathcal{W}^w(\alpha, (\widehat{\mathcal{A}}, c_P)) \coloneqq \{(\widehat{\mathcal{A}}', c'_P) \in \mathcal{W} : (1-\alpha)\bar{y} - c'_P > V_P(w[\alpha, 0] | \widehat{\mathcal{A}}, c_P)\} \cup \{(\widehat{\mathcal{A}}, c_P)\}$, the set of inputs the principal may supply if the comparison point to define worrisome inputs is $(\widehat{\mathcal{A}}, c_P)$ and the slope of the contract is $\alpha$.

Define

$$\beta_{EFK}(\alpha, (\widehat{\mathcal{A}}, c_P)) = - \min_{(\widehat{\mathcal{A}}', c'_P) \in \mathcal{W}^w(\alpha, (\widehat{\mathcal{A}}, c_P))} \max_{(F,c) \in \widehat{\mathcal{A}}'} \{E_F[\alpha y] - c\}.$$

Let $\alpha^*$ and $(\widehat{\mathcal{A}}^*, c_P^*)$ jointly maximize

$$V_P(w[\alpha, \beta_{EFK}(\alpha, (\widehat{\mathcal{A}}, c_P))] | (\widehat{\mathcal{A}}, c_P)) \tag{36}$$

over $\alpha \in [0,1]$ and $(\widehat{\mathcal{A}}, c_P) \in \mathcal{W}$ such that

$$\frac{1-\alpha}{\alpha} \max_{(F,c) \in \widehat{\mathcal{A}}} \{E_F[\alpha y] - c\} - c_P \geq 0.$$

Then an optimal monotone EFK contract (if any such contract exists) is given by $w[\alpha^*, \beta_{EFK}(\alpha^*, (\widehat{\mathcal{A}}^*, c_P^*))]$, and it is locally eligible for $(\widehat{\mathcal{A}}^*, c_P^*)$. (Note that, for a given $\alpha$, $\beta_{EFK}$ is indeed minimized—and so (36) maximized—by taking $(\widehat{\mathcal{A}}, c_P)$ to be an optimal input.)

A monotone EFK contract exists if and only if

$$V_P(w[\alpha^*, \beta_{EFK}(\alpha^*, (\widehat{\mathcal{A}}^*, c_P^*))] | (\widehat{\mathcal{A}}^*, c_P^*)) > 0.$$