MORAL HAZARD, UNCERTAIN TECHNOLOGIES, AND LINEAR CONTRACTS

MARTIN DUMAV, URMEE KHAN

Abstract. We analyze a moral hazard problem where both contracting parties face non-probabilistic uncertainty about how actions translate to output and seek robust performance from a contract in relation to their respective worst-case scenarios. Linear contracts that align Principal's and Agent's pessimistic expectations are optimal.

Keywords: Moral hazard, non-probabilistic uncertainty, robustness, linear contracts

JEL Classification: D81, D82, D86

Contents

Date: November 24, 2017. The first author gratefully acknowledges the financial support from the Ministerio Economia y Competitividad (Spain) through grants ECO2014-55953-P and MDM 2014-0431.

1. Introduction

Consider a Principal who possesses only coarse information about relevant technology for a project, and wishes to hire an Agent. In particular, she knows exactly the set of actions Agent might choose from, but has less precise knowledge than does Agent about possible consequences of those actions. Out-sourcing to an 'expert' is a natural example: presumably Principal would be less knowledgeable about the technology involved, nevertheless it could be contractually possible to restrict the allowable set of actions. More generally, such issues may arise in various circumstances: dealing with brand new technology or business model usually involves a lack of past data, or information coming from conflicting sources. In such settings, it is plausible to imagine that the person to be hired, even if not himself precisely informed, might have the benefit of some prior experience to know a little bit more than his potential employer. What kind of a contract would Principal offer?

Integral to the answer of this question is the attitude of contracting parties towards such imprecise information. We analyze the case where both parties are concerned with *robustness* of their actions to the imprecision of their own perceptions, and evaluate options according to their worst-case guaranteed value. In this setting, we provide a strikingly simple answer to the question posed above: linear contracts, i.e. simple sharing rules, do as well as any arbitrarily complicated contract.

In our model, Principal and Agent both know the common set of actions available. But for any action, the consequences are *ambiguous*. Namely, consequences are perceived as *sets* of distributions. Moreover, Principal has a broader uncertainty about these *sets*, i.e., Agent's perceived set of distributions associated with each action is contained in Principal's, and that is all that she knows.

With our formulation of non-probabilistic beliefs, and robustness concerns of the parties, we show that linear contracts provide a direct and monotonic relationship between the parties' values by aligning their worst-case scenarios. The driving force behind our results is the observation that Principal faces *endogenous* uncertainty for any given contract, as she cannot uniquely infer Agent's action choice, rather has to allow for multiplicity of actions consistent with her rough information. As a result, ex-ante alignment of incentives takes center stage, as opposed to ex-post inference.

The main contribution of the paper is therefore twofold. Formally, we establish justification for linear contracts in an environment where, even though Principal can restrict the set of allowable actions, her ambiguous perception of consequences of those actions lead her to seek guarantees. In the process, the paper also provides insights about a more general question about incentive contracting, namely, what happens to 'high-powered incentives' as the information available to Principal becomes more imprecise in comparison to standard Bayesian models? Our results suggest that such incentives lose their impact.

In order to contextualize our results, at this point it is useful to take a step back and recap the standard formulation of 'moral hazard' in a principal-agent setting. These are problems involving unobserved actions taken by Agent after the contract is agreed to, actions that affect the outcome and thus the payoff to Principal. The workhorse model in solving such contract design problems comes from Grossman and Hart (1983) (we abbreviate it as G-H from now on). The 'riskiness' of the environment is captured in a common prior probability, and Principal's problem is solved in the following steps:

-calculate ex-ante expected costs of inducing Agent to undertake any available action,

-figure out which action she would like Agent to take,

-write a contract to 'implement' that action.

Let us take note of the key substantive elements of the G-H approach and the environment it is embedded in. First, Principal knows enough to decide, from an ex-ante perspective, what action(s) she would like to be implemented. Second, the information structure allows her to draw inference after the fact, so that implementation through 'reward-punishment' schemes based on some observed and verifiable signal is feasible. Third, depending on the details of a particular model, one can end up with non-linear, even non-monotone contracts (in contract theory terminology, such contracts provide 'high-powered' incentives). More often than not, the complicated form of these contracts is not robust to perturbation of the Bayesian information structure. To ensure even a basic property like monotonicity, assumptions that are only known to be sufficient, but not necessary, are imposed (for example, monotone likelihood ratio property for outcome distributions).

The above set of observations immediately give rise to a number of questions. What if the environment itself is not well-understood, at least not well enough for it to be plausibly modeled with a unique and common prior distribution? The examples we cited at the beginning are but a subset of all such possibilities. What are 'robust' solutions to moral hazard problems in an environment with such imprecise information? Furthermore, what happens to 'high-powered incentives' as the information available to Principal becomes more imprecise in comparison to standard Bayesian models? As we have already laid out, simple sharing rules turn out to be optimal when we seek robustness in the face of such imprecision. As for high-powered incentives captured in non-linear contracts, they fail to improve upon simple sharing rules. The key factor behind this result is the following: for any arbitrary non-linear contract, due to ambiguity of information, Principal and Agent differ in their perceptions of worst-case scenarios, unlike in the Bayesian model. Given that the worst-case scenarios are the criteria they use to evaluate options because of robustness concerns, aligning these worst-case scenarios improves their ex-ante welfare. Linear contracts provide the best of such improvements by fully aligning the worst-case scenarios.

Both in our formal approach and our focus on robustness, Carroll (2015) is our main inspiration. It poses a related, but different question. Principal is unsure about the possible actions that Agent might take, although she knows, and agrees with Agent on the consequences of the actions as quantified in terms of an unique probability distribution on outcomes. A rationalization for linear contracts is provided, as Principal seeks guarantees robust to Agent's action set. Although we adapt a similar mathematical approach based on the affine geometry of the space of outcomes and wages, the substantive differences in our primitive environment necessitate a different set of arguments leading to our results. While some of the details are analogous to Carroll (2015), the key differences arise from allowing non-probabilistic beliefs for both parties. This leads to, in an *endogenous* manner, a rich set of actions Agent can choose for a given contract. The intuitive justification for linearity of contracts in our setting is therefore different: it reduces the impact of this endogenous uncertainty about Agent's optimal action on Principal's ex-ante guaranteed payoff.

Our paper fits more broadly in the growing literature on design of contracts in ambiguous environments. A desire for robust mechanisms often motivates the use of simple contracts in this literature, starting with Hurwicz and Shapiro (1978), Mukerji (1998) and more recently, for instance, Chassang (2013), Antic (2014), Garrett (2014), in addition to Carroll (2015)

Given the ubiquity of linear sharing rules, they have received considerable attention from contract theorists from the very early days. Holmstrom and Milgrom (1987) and Diamond (1998) offer foundations for linear contracts in the standard Bayesian setting. The intuition provided in Diamond (1998) is close to the spirit of our paper: it aligns Principal's and

Agent's objectives. In our setting with imprecise non-probabilistic beliefs, linear contracts are optimal in moral hazard relations because they align the pessimistic expectations of the contracting parties.

Our paper also relates to the literature on robust mechanism design with ambiguity, including Bergemann and Morris (2005), Bergemann and Schlag (2011), Garrett (2014), Frankel (2014). Robustness of high powered incentives have also been analyzed in specific applications. Mukerji (1998) shows that in a vertical relationship optimal contracts feature incompleteness and low powered incentives. Ghirardato (1994) shared a similar motivation. His model of moral hazard problem with ambiguity modeled with Choquet capacities illustrates the difficulties involved in analyzing the moral hazard problem with ambiguity by standard method of considering expected benefit of an action and the agency cost of implementing it, as in G-H framework. In our setting, the main challenge is that contracting parties can disagree about the worst-case scenarios, which themselves are endogenous objects. In confronting this challenge we directly cast the contracting problem in output and payoff space and analyze the complications with ambiguity in a tractable manner.

Lastly, we want to make a remark about our use of the MaxMin criteria as a formalization of attitude towards imprecise information or non-probabilistic beliefs. One class of axiomatizations of MaxMin preferences come from an approach that takes the state space as a primitive. We do our analysis on the space of outcomes and payments. In general, the state-space approach towards ambiguity is not conceptually amenable to agency problems, as the probabilities of outcomes in the latter are conditional on actions (see Karni (2006) and Karni (2009) for discussions on this issue). We therefore do not invoke such axiomatizations as a foundation for our formulation, instead we see this simply as a plausible way to capture the idea of robustness concerns when designing contracts. Our take is that such concerns are more appropriately modeled as evaluations of outcome prospects.

2. Model

2.1. **Outcomes and Actions.** Let $\mathcal{Y} \subseteq \mathbb{R}_+$ denote a compact convex set of outcomes, $\Delta(\mathcal{Y})$ the set of Borel distributions on $\mathcal{Y}$ with the weak* topology and $\mathbb{K}_{\Delta(\mathcal{Y})}$ denote the class of non-empty, compact, convex subsets of $\Delta(\mathcal{Y})$.

Agent chooses an unobservable action a from a compact set $\mathcal{A}$[1]

[1] This plays no more of a role than that of an index set. A closed, bounded set in $\mathbb{R}$, including a finite set, works.

Let $g : \mathcal{A} \mapsto \mathbb{R}_+$ be a continuous, bounded function that describes the cost of effort to Agent.

Assumption 1. *There exists* $a \in \mathcal{A}$ *such that* $g(a) = 0$.[2]

2.2. **Information.** Principal's information about technology is characterized as an upper-hemicontinuous set-valued mapping from set of actions $\mathcal{A}$ to $\mathbb{K}_{\Delta(\mathcal{Y})}$ given by $Q^P(\cdot) : \mathcal{A} \mapsto \mathbb{K}_{\Delta(\mathcal{Y})}$. We interpret this set-valued mapping as follows: for each action $a \in \mathcal{A}$, Principal believes that this action induces a convex and compact set of distributions over outcomes denoted by $Q^P(a) \in \mathbb{K}_{\Delta(\mathcal{Y})}$. Similarly, Agent's information is characterized by another upper-hemicontinuous mapping, $Q^A(\cdot) : \mathcal{A} \mapsto \mathbb{K}_{\Delta(\mathcal{Y})}$. The key assumption we make is that Agent has (weakly) more precise knowledge about technology than Principal.

Assumption 2. $Q^A(a) \subset Q^P(a)$ *for each* $a \in \mathcal{A}$.

Let $\mathcal{Q}^A(a) = \{M \in \mathbb{K}_{\Delta(\mathcal{Y})} : M \subset Q^P(a)\}$ denote the collection of all possible sets of distributions induced by action a that is consistent with Principal's information. The family of all such collections of sets, for all possible actions, denoted by $\mathcal{Q}^A = \{\mathcal{Q}^A(a) : a \in \mathcal{A}\}$ describes Principal's perception of Agent's information about technological possibilities. $\mathcal{Q}^A$ thus contains the range of any possible mapping $Q^A(\cdot)$ that Principal might imagine for Agent to possess.

2.3. **Payoffs and Timing.** We assume both Principal and Agent evaluate contracts by their worst-case expected payoff. In our primary analysis, we maintain the assumption that both are risk-neutral over monetary payoffs. In an extension, we show that introducing risk aversion for Agent does not qualitatively change our results.

A contract is a continuous, bounded function $w : \mathcal{Y} \to \mathbb{R}_+$ that specifies output contingent payments and protects Agent with limited liability (i.e. $w(0) \geq 0$). In particular, we allow the zero contract: $w(y) = 0$, $\forall y \in \mathcal{Y}$.

The timing of the contracting game is as follows:

(i) Principal offers a contract w;
(ii) Agent, knowing Q^A, chooses action $a \in \mathcal{A}$;
(iii) output y is realized;
(iv) payoffs are received: $y - w(y)$ to Principal and $w(y) - g(a)$ to Agent.

[2] This zero cost action does not play any major role in our analysis, but it does help us formalize the case where it is optimal for Principal to offer a zero contract. We interpret it as the option of doing nothing.

3. Symmetric Ambiguity

We begin with the benchmark case in which Principal and Agent both have the same, ambiguous information. Formally, $Q^A(a) = Q^P(a), \quad \forall a \in \mathcal{A}$. In this symmetric case, we assume that none of the sets $Q^i(\cdot)$ are singletons, so as not to let the problem collapse into the unique and common prior case. This is an important assumption for our technical approach to work, so we note it here.

Assumption 3. $Q^A(a) = Q^P(a) = Q(a)$, $\forall a \in \mathcal{A}$, *with each* $Q(a)$ *non-singleton.*

Consider $\mathcal{C}(w) = \mathsf{co}\{(y, w(y)) : y \in \mathcal{Y}\}$, where co denotes convex hull. Given that $w(y)$ is a continuous, bounded function on convex, compact $\mathcal{Y}$, the graph of $w(y)$, denoted by $gr(w)$, is closed and compact. Hence $\mathcal{C}(w)$ is closed and compact as the convex hull of $gr(w)$.

Consider also any action $a \in \mathcal{A}$ that Agent might choose while working under $w(y)$. Let

$$\mathcal{S}_a(w\big|Q) = \big\{(\mathbb{E}_q y, \mathbb{E}_q w(y))\big|q \in Q(a),\, y \in \mathcal{Y}\big\}$$

Clearly, $\mathcal{S}_a(w\big|Q)$ is fully contained in $\mathcal{C}(w)$. We note also that $\mathcal{S}_a(w\big|Q)$ is a convex, compact set in $\mathbb{R}^2$.

We are going to use the two sets defined above extensively. Our first lemma is about the dimensionality of $\mathcal{S}_a(w\big|Q)$.

Lemma 1. *For an arbitrary non-linear contract* $w(y)$, $\mathcal{S}_a(w\big|Q)$ *can be characterized in one of the following two ways:*

(i) it is a convex compact subset of $\mathbb{R}^2$ *with non-empty relative interior in* $\mathbb{R}^2$*; hence* $\mathbb{R}^2$ *is its affine hull and it has dimension* 2*.*
(ii) it is a line segment, with dimension 1*.*

Proof. (i) follows from the fact that $\mathcal{S}_a(w\big|Q)$ is a convex compact set, hence by Minkowski's theorem is the convex hull of its extreme points.[3] For any $Q(a)$ that has more than two points in its support, a 2-dimensional $\mathcal{S}_a(w\big|Q)$ would arise.

[3]For most of the definitions and results on convex analysis and affine geometry used in this paper, we refer to Hiriart-Urruty and Lemaréchal (2012). See Ch 4 for this theorem and discussion.

As for (ii), relative interior in $\mathbb{R}^2$of $\mathcal{S}_a(w\big|Q) = \emptyset$ implies that $\mathcal{S}_a(w\big|Q)$ is a linear segment, which is only possible for a non-linear $w(y)$, and convex, compact $Q(a)$ only if $Q(a)$ contains distributions with only a two-point support.[4] □

We have an immediate corollary of case (i) in Lemma 1. Define Π as the projection from $\mathbb{R}_2$ to $\mathbb{R}_1$, with $\Pi(u, v) = u$. Define the *slice* of $\mathcal{S}_a(w\big|Q)$ along u as

$$\sigma(u) = \left\{v \in \mathbb{R}_+ \big| (u, v) \in \mathcal{S}_a(w\big|Q)\right\}$$

Lemma 2. *For any non-linear $w(y)$, and any $Q(a)$ that has more than two points in the support, the slices of $\mathcal{S}_a(w\big|Q)$ along y are not singletons, unless at an extreme point of $\mathcal{S}_a(w\big|Q)$.*

Proof. The implication follows directly from the observation that slices along y being singletons except at extreme points would imply $\mathcal{S}_a(w\big|Q)$ is a line segment. □

3.1. **Agent's Optimization Problem.** Agent solves the following problem:

$$\max_{a \in \mathcal{A}} \left(\min_{q \in Q(a)} \mathrm{E}_q \left[w(y)\right] - g(a) \right)$$

Compactness of $Q(a)$ in weak* topology ensures the *min* is well defined, and our upper-hemicontinuity assumption on the correspondence Q ensures the same for the *max*.[5]

Given symmetric ambiguity Principal can fully infer Agent's decision rule[6]:

$$a^*(w|Q) = \operatorname*{argmax}_{a \in \mathcal{A}} \left(\min_{q \in Q(a)} \mathrm{E}_q \left[w(y)\right] - g(a) \right) \tag{1}$$

The best of the worst-case scenarios perceived by Agent is thereby $q_A^*(w|Q) \in Q(a^*(w|Q))$.[7]

The next lemma characterizes the max-min value for Agent in terms of a supporting hyperplane of $\mathcal{S}_{a^*}(w\big|Q)$.

Lemma 3. *Let $V_A(w\big|Q)$ be the value of the problem defined in* (3.1). *This value defines a supporting hyperplane for $\mathcal{S}_{a^*}(w\big|Q)$*

[4]or if $w(y)$ has a linear segment, which is an uninteresting case in our context. In general, when we refer to a contract as being *non-linear*, we mean it to be strictly so.

[5]See Appendix for a detailed argument.

[6]For simplicity, in the main body of the paper we take this argmax to be a singleton. There are several ways of dealing with case of multiplicity, as we discuss in the Appendix, but our qualitative results do not change.

[7]If there are multiple such distributions, Agent would be payoff-indifferent among them.

Proof. Let $H_{A,w}$ be the hyperplane defined as

$$H_{A,w} = \{(u, v) \in \mathbb{R}^2 \big| v = V_A(w\big|Q)\} \tag{2}$$

It exposes the face of $\mathcal{S}_{a^*}(w\big|Q)$ with $\mathbb{E}_{q^*}(w(y)) = V_A(w\big|Q) + g(a^*)$ and is thereby a supporting hyperplane for $\mathcal{S}_{a^*}(w\big|Q)$. □

3.2. **Principal's Optimization Problem.** For an arbitrary contract w, Principal can infer agent's decision rule given the symmetry of ambiguous perceptions. Hence Principal's ex-ante guaranteed value

$$V_P(w\big|Q) = \min_{q \in Q(a^*(w|Q))} \mathrm{E}_q[y - w(y)] \tag{3}$$

His worst-case scenario corresponds to those distributions that attain the minimum, denoted by $q^*_P(w|Q) \in Q(a^*(w|Q))$.[8]

Observation: Principal and Agent do not have to agree on the worst-case scenario, i.e., $q^*_P(w|Q)$ and $q^*_A(w|Q)$ can, and typically will, differ. This point turns out to be central to much of our subsequent analysis.

Principal offers a contract w that solves

$$\sup_w V_P(w\big|Q) \quad \text{subject to} \quad V_A(w|Q) \geq 0$$

For the subsequent analysis we assume that contracting relation is viable: for some contract w, we have that $V_p(w|Q) \geq 0$ and $V_A(w|Q) \geq 0$. Since we allow zero contracts, and zero cost actions, the set of contracts satisfying the two constraints is non-empty.

The next set of lemmas characterize another supporting hyperplane for $\mathcal{S}_{a^*}(w\big|Q)$ representing Principal's guaranteed value associated with $Q(a)$.

Define $H_{P,w}(\lambda)$ as the hyperplane

$$H_{P,w}(\lambda) = \{(u, v) \in \mathbb{R}^2 \big| u - v = \lambda\} \tag{4}$$

[8]Once again, we can ignore the issue of multiplicity without loss of generality, because of payoff-indifference. In any case, it would be a compact set of distributions.

A family of these hyperplanes can be used to represent Principal's iso-profit curves corresponding to expected profit level λ, i.e.

$$H_{P,w}(\lambda) = \left\{(\mathbb{E}y, \mathbb{E}w(y)) \in \mathbb{R}^2 \middle| \mathbb{E}y - \mathbb{E}w(y) = \lambda\right\} \tag{5}$$

Geometrically in $\mathbb{R}^2$, these are rays through the origin defined by the equation $y - w(y) = \lambda$. In particular, Principal's guaranteed expected payoff corresponding to any a^* chosen by Agent under a contract $w(y)$ is represented by $H_{P,w}(V_P(w|Q))$. Recall that Principal's guarunteed value corresponds to the worst-case scenario related to a^*, hence it is represented by an extreme point or face of $\mathcal{S}_{a^*}(w|Q)$. In other words, we have the following.

$$H_{P,w}(\pi) = \left\{(u, v) \in \mathbb{R} \middle| v - u = \pi\right\} \tag{6}$$

Lemma 4. *$H_{P,w}(V_P(w|Q))$ is a supporting hyperplane of $\mathcal{S}_{a^*}(w|Q)$.*

Proof. Consider the family of all half-spaces $\left\{(u, v) \in \mathbb{R}^2 \middle| u - v \geq \lambda\right\}$ that contains $\mathcal{S}_{a^*}(w|Q)$. The half-space generated by the intersection of this family of half-spaces corresponds to the smallest λ for which $\mathcal{S}_a(w|Q)$ is fully contained in it. The defining hyperplane has $\lambda = V_P(w|Q)$, and is by construction a supporting hyperplane for $\mathcal{S}_{a^*}(w|Q)$ □

We now have an affine set containing all $(\mathbb{E}y, \mathbb{E}w(y))$ pairs that would be incentive compatible for implementing a^*, while guaranteeing Principal at least $V_P(w|Q)$, and also a relationship between $V_A(w|Q)$ and $V_P(w|Q)$ under any contract $w(y)$ that induces a^* as the optimal action.

Lemma 5. *$\mathcal{S}_{a^*}(w|Q)$ is contained in the 2-dimensional affine convex cone generated by $H_{A,w}$ and $H_{P,w}$ hyperplanes.*

Proof. $H_{A,w}$ and $H_{P,w}$ are supporting hyperplanes of $\mathcal{S}_a(w|Q)$; the intersection of the two associated half-spaces that both contain $\mathcal{S}_a(w|Q)$ is an affine convex cone. □

As an immediate implication we have the following:

Lemma 6. *For any contract $w(y)$ that induces a^* as the optimal action,*

$$V_P(w|Q) = \min_{q \in Q(a^*)} \mathbb{E}_q[y - w(y)] \quad \textit{subject to} \quad \min_{q \in Q(a^*)} [\mathbb{E}_q w(y) - g(a^*)] \geq V_A(w|Q)$$

Proof. By construction, for any contract that implements a^* as the optimal action, $\mathcal{S}_{a^*}(w\big|Q)$ contains all feasible $\{\mathbb{E}y,\ \mathbb{E}w(y)\}$ pairs. By definition, $V_P(w\big|Q)$ and $V_A(w\big|Q)$ are the worst-cases respectively for Principal and Agent over this feasible set with expectations taken over their respective worst-case distributions. $\square$

Note that the solution to the problem specified in Lemma 6 exists, given a compact $\mathcal{Y}$ and bounded $w(y)$. It is essentially a linear programming problem in the $(y, w(y))$ space. The program feasible and bounded by construction.

Next we establish that a linear contract can be found that will implement a^*, while making Principal and Agent both as at least as well-off as under $w(y)$. In fact, within the family of linear contracts, there is an optimal one.

3.3. **Optimality of Linear Contracts.**

Note that whatever optimal action a^* Agent chooses, with the associated 'best worst-case' for Agent $q_A^* \in Q(a^*)$, Agent's expected payments satisfies

$$\mathbb{E}_{q_A^*} w(y) \geq \mathbb{E}_{q_A^*} w(y) - g(a^*) \equiv V_A(w\big|Q) \tag{7}$$

A linear contract ℓ is of the form $\ell(y) = \alpha y$ with $\alpha \in (0, 1]$. For this contract Principal's ex-post payoffs are

$$y - \ell(y) = \frac{1-\alpha}{\alpha}\ell(y). \tag{8}$$

The next Proposition lays out a key inequality characterizing the relationship between the guaranteed values of Principal and Agent for a linear contract.

Proposition 1. *For a linear contract $\ell(y) = \alpha y$, under symmetric ambiguity, the maxmin guaranteed values for Principal and Agent are related as follows:*

$$\mathbb{E}_{q_P^*}[y - \ell(y)] = \frac{1-\alpha}{\alpha}\mathbb{E}_{q_P^*}\ell(y) \geq \frac{1-\alpha}{\alpha}\mathbb{E}_{q_A^*}\ell(y) \geq \frac{1-\alpha}{\alpha}V_A(\ell\big|Q) \tag{9}$$

Proof. The first equality follows by taking expectation of both sides in equation (8).[9]

[9] For a linear contract $\ell(y)$, for any action $a \in \mathcal{A}$, $\mathcal{S}_a(\ell\big|Q)$ is a linear segment [case (ii) in Lemma 1]

The second inequality relation follows because the worst-case of expected payments *made* according to Principal's pessimism is at least as severe than Agent's conservative perception of worst-case expected payments *received*.

Formally, for an arbitrary contract $w(y)$, any a and $Q(a)$, Agent's worst-case scenario minimizes expected payments received: $q_A = \arg\min_{q \in Q(a)} \mathbb{E}_q[w(y)]$. Equivalently, $\mathbb{E}_{q_A}[w(y)] \leq \mathbb{E}_q[w(y)]$ for all $q \in Q(a)$.

In particular, this holds for $q_P \in Q(a)$ – Principal's worst-case distribution minimizing expected profits $\mathbb{E}_q[y - w(y)]$ over $Q(a)$. Setting $\ell(y) = w(y)$, we get the inequality.

The last inequality is derived by applying equation (7) to the linear case:

$$\frac{1-\alpha}{\alpha} \mathbb{E}_{q_A^*} \ell(y) \geq \mathbb{E}_{q_A^*} \ell(y) - g(a^*) \equiv V_A(w \big| Q) \tag{10}$$

□

The second inequality in (9) plays an important role in our analysis.

Let V_p be the value of Principal's problem as defined in (3.2). We then have the following theorem establishing optimality of a linear contract.

Theorem 1. *Suppose $Q^A(a) = Q^P(a)$ for all $a \in \mathcal{A}$. There exists a linear contract that maximizes V_p.*

Proof. We make the argument in several steps.

(a) Let a^* be the optimal action chosen by Agent under $w(y)$, that provides a guarantee $V_P(w \big| Q) = \lambda$ to Principal. Let $\gamma(a^*)$ be the associated $\mathbb{E}_{q_A^*} w(y)$. First we establish that there is an affine contract that implements a^* and guarantees Principal at least λ.

Construct the proposed affine contract as follows:

i. If $V_P(w \big| Q) < 0$, choose $w(y) = 0$

ii. If $V_P(w \big| Q) \geq 0$, let $w'(y) = \beta + \alpha y$, with $\alpha \in (0, 1)$, and $\beta = (1-\alpha)\gamma - \alpha\lambda \geq 0$

It is straightforward to check that the line $\beta + \alpha y$ passes through the vertex of the affine convex cone containing $\mathcal{S}_{a^*}(w \big| Q)$, and thereby has non-empty intersection with $\mathcal{S}_{a^*}(w \big| Q)$. In fact, the line segment defined by $gr(w'(y)) \cap \mathcal{S}_{a^*}(w \big| Q)$ represents the set $\{(\mathbb{E}_q y, \mathbb{E}_q w'(y)) \big| q \in Q(a^*)\}$. Recall also that any of these pairs $(\mathbb{E}_q y, \mathbb{E}_q w'(y))$ fall within the region representing both $y - w(y) \geq \lambda$ and $w(y) \geq \gamma$ [see Lemma 6]. In other words,

a^* is still a possible optimal choice for Agent, while Principal is at least as well-off as before.

To finish this step of the argument, we need to check for any possible deviation from a^* by Agent under the new contract $w'(y)$. We first note that Agent would not deviate to any action a' with $\min_{q\in Q(a')} \mathbb{E}_q w'(y) < \min_{q\in Q(a^*)} \mathbb{E}_q w(y)$. To see this, recall that Agent's indifference sets in the $(y, w(y))$ space are of the form $\{(u,v) \big| v \,\text{constant}\}$ [geometrically, these are horizontal lines, representing a particular value of $\mathbb{E}w(y)$]. By revealed preference, a^* had $\min_{q\in Q(a^*)} \mathbb{E}_q w(y) > \min_{q\in Q(a')} \mathbb{E}_q w(y)$ for all a' with $g(a') < g(a^*)$, and possibly even some actions with $g(a') > g(a^*)$. With $\min_{q\in Q(a^*)} \mathbb{E}_q w'(y) \geq \min_{q\in Q(a^*)} \mathbb{E}_q w(y)$, we must have, if Agent were to deviate to some a' under w',

$$\min_{q\in Q(a')} \mathbb{E}_q w'(y) \geq \min_{q\in Q(a^*)} \mathbb{E}_q w'(y) \geq \min_{q\in Q(a^*)} \mathbb{E}_q w(y) \tag{11}$$

Hence we conclude that with this new affine contract w', Agent either still chooses a^*, or deviates to an a' with $\min_{q\in Q(a')} \mathbb{E}_q w'(y) \geq \min_{q\in Q(a^*)} \mathbb{E}_q w'(y)$. Given the affine nature of w', any such deviation would also improve Principal's worst-case. We have,

$$y - w'(y) = \frac{1-\alpha}{\alpha} w'(y) - \frac{1+\alpha}{\alpha}\beta \tag{12}$$

Taking expectations on both sides, and then taking minima, establishes the last claim. Since this is true for any a^*, we have found an affine contract that dominates $w(y)$.

(b) We next show that, for any affine contract $w'(y)$, there is a linear contract $\ell(y)$ that does at least as well, or strictly better, for Principal, and implements the same optimal action as $w'(y)$.

If $\beta = 0$, then $w'(y)$ is already linear and we are done. If $\beta > 0$, then note that $w'(0) = \beta$. Define $\ell(y) = \alpha y = w'(y) - \beta$. This improves Principal's payoff, with Agent's incentives unchanged.

(c) The previous steps, taken together, imply that any action that is implemented by an arbitrary contract, $w(y)$, can be implemented by a linear contract, that makes Principal and Agent both at least as well off as under $w(y)$. In other words, it is enough to only optimize within the class of linear contracts.

From (a), we can strengthen (9) in Proposition 1 to equality:

$$\mathbb{E}_{q_P^*}[y - \ell(y)] = \frac{1-\alpha}{\alpha}\mathbb{E}_{q_P^*}\ell(y) = \frac{1-\alpha}{\alpha}\mathbb{E}_{q_A^*}\ell(y) \tag{13}$$

This follows from noting that with a linear contract, for any action a, Principal's and Agent's worst-cases are perfectly aligned, i.e. $q_P^* = q_A^*$.

In other words, with a linear contract, the best worst-case guarantees for Principal and Agent are related via a linear relationship. Principal's value V_p is continuous in $\alpha \in [0, 1]$, and an $\alpha^*(a^*)$ exists that maximizes it.[10]

$\square$

4. Asymmetric Ambiguity

In this section we will take up the case where Principal has strictly less precise information than Agent, i.e. $Q^A(a) \subset Q_P(a)$ for all $a \in \mathcal{A}$, and that is all that Principal knows. In particular, Principal does not know exactly what Agent knows, i.e. the exact subsets $Q^A(a) \subset Q_P(a)$ are not known to Principal. With a slight abuse of notation, we are going to use Q^i to denote the range of the mapping $Q^i(\cdot)$ in the following analysis. As we shall see, with a few tweaks, the results derived for the symmetric case go through for the most part.

Let us recap here the essential definitions and notations related to Principal's and Agent's optimization problems and suitably modify them for the asymmetric case. Given an output contingent contract $w(y)$, a risk neutral, ambiguity averse Agent with technology $Q^A \subset \mathcal{Q}^A$ chooses an action that solves his optimization problem:

$$a^*(w|Q^A) = \operatorname*{argmax}_a \left[\min_{q \in Q^A(a)} \mathbb{E}_q w(y) - g(a) \right] \tag{14}$$

Let $q_A^*(w|Q^A) \in Q^A(a^*(w|Q^A))$ be Agent's 'best worst-case' distribution facing w and with information Q^A. The associated guaranteed value to Agent:

$$V_A(w|Q^A) = \mathbb{E}_{q_A^*(w|Q^A)} w(y) - g(a^*(w|Q^A)) \tag{15}$$

Relative to the information Q^P, Principal's value :

$$V_p(w|Q^P) = \min_{p \in Q^P(a^*(w|Q^P))} \mathbb{E}_p[y - w(y)] \tag{16}$$

[10]There are some subtleties regarding $\alpha = 0$, that we take up in Section 4.2

Notice here that we now have a possible discrepancy between $a^*(w|Q^A)$ and $a^*(w|Q^P)$, as Principal can only use information available to him. The key point here is that any action chosen by Agent would guarantee at least as much expected payoff to him as $a^*(w|Q^P)$ does.

Now, given that $Q^A(a) \subset Q^P(a)$ for all $a \in \mathcal{A}$, Principal's worst-case expected payoff over all information mappings Agent might possess,

$$V_P(w) := \inf_{Q \subset \mathcal{Q}^P} V_P(w|Q) \leq V_P(w|Q^P) \leq V_P(w|Q^A)$$

The last two terms refer to the cases when Principal has information $Q = Q^P$ and $Q = Q^A$ respectively, and the second inequality follows as $Q^A(a) \subset Q^P(a)$ for all a, since more precise information would improve the guaranteed value.[11]

The first inequality follows from the fact that for any given contract Principal cannot infer precisely Agent's decision rule and has *endogenous* ambiguity about the induced set of probabilities over outputs. This observation plays a key role in the determination of Principal's guarantee below.

Note also that whatever optimal action a^* Agent chooses, with the associated 'best worst-case' for Agent $q_A^* \in Q^A(a^*)$, Agent's expected payments recieved satisfies

$$E_{q_A^*}[w(y)] \geq E_{q_A^*}[w(y)] - g(a^*) = V_A(w|Q^A) \geq V_A(w|Q^P). \tag{17}$$

The last term denotes Agent's guaranteed value if he were to possess the coarser information that Principal possesses; Principal in fact would compute Agent's guaranteed value this way. The second inequality holds because having more precise information can only make a pessimistic Agent (weakly)better off.[12]

The above analysis shows that in case of asymmetric ambiguity, Principal faces a substantial inference problem in anticipating ex-ante agent's optimal action choice, and also has a more pessimistic estimation of his maxmin guarantee than he would have if had the same information that Agent does. Reduction of this disadvantage is where linear contracts pay their role.

[11]Formally, $V_P(w|Q^A)$ is the minimal value over smaller sets of distributions than $V_P(w|Q^P)$.
[12]Again, the comparison is between minima over smaller as opposed to bigger sets.

4.1. **Linear Contracts.** Suppose Principal offers a linear contract: $\ell(y) = \alpha y$ with $\alpha \in (0, 1]$.

For the linear contract Principal's ex-post payoffs are:

$$y - w(y) = \frac{1-\alpha}{\alpha} w(y). \tag{18}$$

Combining this with (17) gives a lower bound on Principal's expected payoff

$$\mathbb{E}_{q_P^*}[y - \ell(y)] = \frac{1-\alpha}{\alpha} \mathbb{E}_{q_P^*}[\ell(y)] \geq \frac{1-\alpha}{\alpha} E_{q_A^*}[\ell(y)] \geq \frac{1-\alpha}{\alpha} V_A(w|Q^P), \tag{19}$$

where q_P^* is the worst-case scenario for Principal that minimizes expected profit $\mathbb{E}_q[y - \ell(y)]$ over the set of probability distributions in $Q^A(a^*)$.

Using this observation for a linear contract, we then derive a relationship between the guarantees to Principal and to Agent. Since (19) holds regardless of the technology Q^A available to Agent, taking infimum over all such technologies, Principal's worst-case expected value satisfies:

$$V_p(\ell) \geq \frac{1-\alpha}{\alpha} V_A(w|Q^P).$$

This shows how to obtain a payoff guarantee from a linear contract. Turns out, the optimal guarantee to Principal comes from a linear contract:

Theorem 2. *Suppose $Q^A(a) \subset Q^P(a)$ for all $a \in \mathcal{A}$. There exists a linear contract that maximizes V_p.*

Proof. The analysis of the contract design is driven by Principal's coarse information. For a given contract, unlike in the symmetric case Principal is not able to uniquely infer Agent's decision rule. Instead, Principal can infer the set of "rationalizable" actions that the Agent might choose consistent with Principal's coarse information. The proof of this Theorem is analogous to that of Theorem 1, with some subtle adjustments. We begin with a series of Lemmas.

Let $q_A^*(w|Q^P) \in Q^P(a^*(w|Q^P))$ be Agent's 'best worst-case' distribution facing w from Principal's perspective, that is consistent with her coarse understanding according to Q^P. The associated guaranteed value to Agent:

$$V_A(w|Q^P) = \mathbb{E}_{q_A^*(w|Q^P)} w(y) - g(a^*(w|Q^P)) \tag{20}$$

Step 1. Construction of Agent's Supporting Hyperplane

Using Lemma 3 with Q^P in the role of Q the guarantee value of $V_A(w \big| Q^P)$ Agent's problem defined in (20) yields a supporting hyperplane for $\mathcal{S}_{a^*}(w \big| Q^P)$. Analogous to (2) in Lemma 3 let $\widehat{H_{A,w}}$ be the hyperplane defined as

$$\widehat{H_{A,w}} = \left\{(u,v) \in \mathbb{R}^2 \big| v = V_A(w \big| Q^P)\right\} \tag{21}$$

It exposes the face of $\mathcal{S}_{a^*}(w \big| Q^P)$ with $\mathbb{E}_{q_A^*(w|Q^P)} w(y) = V_A(w \big| Q^P) + g(a^*)$ and is thereby a supporting hyperplane for $\mathcal{S}_{a^*}(w \big| Q^P)$.

Step 2. Construction of Principal's Supporting Hyperplane

From Principal's perspective the set of rationalizable actions $\hat{\mathcal{A}}$ are those that she can rationalize as optimal choice by Agent on the basis of his relatively more precise information, which can take arbitrary forms consistent with Principal's information. Formally

$$\widehat{\mathcal{A}} = \{a \in \mathcal{A} : \mathcal{S}_a(w \big| Q^P) \bigcap V_A(w|Q^P)^+ \neq \varnothing\} \tag{22}$$

where $V_A(w|Q^P)^+ = \{(y,v) \in \mathcal{C}(w) : v \geq V_A(w|Q^P)\}$.

Relative to the information Q^P, Principal's guarantee value is the worst-case value from the set of rationalizable actions $\widehat{\mathcal{A}}$:

$$V_p(w|Q^P) = \min_{a \in \widehat{\mathcal{A}}} \min_{p \in Q^P(a)} \mathbb{E}_p[y - w(y)] \tag{23}$$

Let $\widehat{a}$ be the action in $\widehat{\mathcal{A}}$ that attains the minimum. Recalling the definition of the hyperplanes for Principal's isoprofit curves in (6), Principal's guarantee in the case of asymmetric perceptions $V_p(w|Q^P)$ is alternatively characterized as:

Lemma 7. $\widehat{H_{P,w}}(V_p(w|Q^P))$ *is a supporting hyperplane of* $\mathcal{S}_{\widehat{a}}(w \big| Q^P)$.

Proof follows from Lemma 4 taking Q^P and $\widehat{a}$ in the roles of Q and a^*, respectively.

Proof. Consider the family of all half-spaces $\left\{(u,v) \in \mathbb{R}^2 \big| u - v \geq \lambda\right\}$ that contains $\mathcal{S}_{\widehat{a}}(w \big| Q^P)$. The half-space generated by the intersection of this family of half-spaces corresponds to the smallest λ for which $\mathcal{S}_{\widehat{a}}(w \big| Q^P)$ is fully contained in it. The defining hyperplane has $\lambda = V_p(w|Q^P)$, and is by construction a supporting hyperplane for $\mathcal{S}_{\widehat{a}}(w \big| Q^P)$. ∎

Step 3. 2-dimensional affine cone generated by $\widehat{H_{A,w}}$ and $\widehat{H_{P,w}}$ hyperplanes for $Q = Q^P$ and the relationship between the values.

Observe that from the definition of Principal's value in (23) as the minimum value over the actions in $\widehat{\mathcal{A}}$ the half-space defined by the hyperplane $\widehat{H_{P,w}}(V_p(w|Q^P))$ supports $\mathcal{S}_{a^*}(w\big|Q^P)$ and contains all of the $\mathcal{S}_a(w\big|Q^P)$ for all $a \in \widehat{\mathcal{A}}$.

Let $\widehat{a}^*$ be the optimal action that Agent with more precise information chooses. Since his optimal choice gives Agent at least $V_A(w\big|Q^P)$ consistent with coarse information, the optimal choice is amongst the rationalizable actions, that is $\widehat{a}^* \in \widehat{\mathcal{A}}$.

We now have an affine set containing all $(\mathbb{E}y, \mathbb{E}w(y))$ pairs that would be give Agent the value at least $V_A(w\big|Q^P)$ from Step 1 while guaranteeing Principal at least $V_P(w\big|Q)$ from Step 2, and also a relationship between $V_A(w\big|Q)$ and $V_P(w\big|Q)$ for any $Q \subset Q^P$ under any contract $w(y)$ that induces a rationalizable action $\widehat{a}^* \in \widehat{\mathcal{A}}$ as the optimal action.

Lemma 8. *$\mathcal{S}_{\widehat{a}^*}(w\big|Q)$ is contained in the 2-dimensional affine convex cone generated by $\widehat{H_{A,w}}$ and $\widehat{H_{P,w}}$ hyperplanes.*

Proof. $\widehat{H_{A,w}}$ and $\widehat{H_{P,w}}$ are supporting hyperplanes of $\mathcal{S}_{a^*}(w\big|Q^P)$ and $\mathcal{S}_{\widehat{a}}(w\big|Q^P)$, respectively; the intersection of the two associated half-spaces that both contain $\mathcal{S}_{\widehat{a}^*}(w\big|Q^A)$ is an affine convex cone. □

As an immediate implication of Step 3 we have the following relationship between the values from any given contract:

Lemma 9. *For any contract $w(y)$ that induces $\widehat{a}^*$ as the optimal action,*

$$V_P(w\big|Q) = \min_{q \in Q(\widehat{a}^*)} \mathbb{E}_q[y - w(y)] \quad \textit{subject to} \quad \min_{q \in Q(\widehat{a}^*)} [\mathbb{E}_q w(y) - g(a^*)] \geq V_A(w\big|Q)$$

Proof. By construction, for any contract that induces Agent to choose $\widehat{a}^*$ as the optimal action amongst the rationalizable actions $\widehat{\mathcal{A}}$, $\mathcal{S}_{\widehat{a}^*}(w\big|Q)$ contains all feasible $\{\mathbb{E}y,\, \mathbb{E}w(y)\}$ pairs. By definition, $V_P(w\big|Q)$ and $V_A(w\big|Q)$ are the worst-cases respectively for Principal and Agent over this feasible set with expectations taken over their respective worst-case distributions. □

We note that the existence of the solution to the problem specified in Lemma 9 follows from the analogous arguments given for the existence of the solution to the problem in Lemma 6.

Next we establish that a linear contract can be found that will induce a^* as a rationalizable action, while making Principal and Agent both as at least as well-off as under $w(y)$. In fact, within the family of linear contracts, there is an optimal one. We develop the analogous steps as in Theorem 1 to conclude.

(a$'$) Let a^* be the optimal action chosen by Agent under $w(y)$ according to Principal's coarse information Q^P, that provides a guarantee $V_P(w \big| Q^P) = \lambda$ to Principal. Let $\gamma(a^*)$ be the associated $\mathbb{E}_{q_A^*} w(y)$. First we establish that there is an affine contract that still rationalize a^* and guarantees Principal at least λ.

Construct the proposed affine contract as follows:

i.$'$ If $V_P(w \big| Q^P) < 0$, choose $w(y) = 0$

ii.$'$ If $V_P(w \big| Q^P) \geq 0$, let $w'(y) = \beta + \alpha y$, with $\alpha \in (0,1)$, and $\beta = (1-\alpha)\gamma - \alpha\lambda \geq 0$

It is straightforward to check that the line $\beta + \alpha y$ passes through the vertex of the affine convex cone generated by $\widehat{H_{A,w}}$ and $\widehat{H_{P,w}}$ hyperplanes. In light of the analysis in Step 3 the cone contains $\mathcal{S}_{a^*}(w \big| Q^P)$, and thereby has non-empty intersection with $\mathcal{S}_{a^*}(w \big| Q^P)$. In fact, the line segment defined by $gr(w'(y)) \cap \mathcal{S}_{a^*}(w \big| Q^P)$ represents the set $\{(\mathbb{E}_q y, \mathbb{E}_q w'(y)) \big| q \in Q^P(a^*)\}$. Recall also that any of these pairs $(\mathbb{E}_q y, \mathbb{E}_q w'(y))$ fall within the region representing both $y - w(y) \geq \lambda$ and $w(y) \geq \gamma$ [see Lemma 8]. In other words, a^* is still a rationalizable choice for Agent, while Principal is at least as well-off as before.

To finish this step of the argument, as in the symmetric case we consider Agent optimal choice $\widehat{a}^*$ from his rationalizable set of actions in $\widehat{\mathcal{A}}$, which might be different from a^* since Agent's precise information consistent with $Q^A \subset Q^P$ can take arbitrary forms.

To check for any possibly different optimal choice from a^* by Agent under the new contract $w'(y)$. By analogous 'revealed preference' argument given in the symmeric case, we first note that Agent would only choose an action in the set of rationalizable actions, $\widehat{\mathcal{A}}$: Agent would not choose any action a' with $\min_{q \in Q(a')} \mathbb{E}_q w'(y) < \min_{q \in Q(a^*)} \mathbb{E}_q w(y)$. We must have, if Agent were to choose $\widehat{a}^*$ in $\widehat{\mathcal{A}}$ under w', for some $Q \subset Q^P$

$$\min_{q \in Q(\widehat{a}^*)} \mathbb{E}_q w'(y) \geq \min_{q \in Q(a^*)} \mathbb{E}_q w'(y) \geq \min_{q \in Q(a^*)} \mathbb{E}_q w(y) \tag{24}$$

Hence we conclude that with this new affine contract w', Agent either still chooses a^*, or instead optimally chooses a different action $\widehat{a}^*$ with $\min_{q \in Q(\widehat{a}^*)} \mathbb{E}_q w'(y) \geq \min_{q \in Q(a^*)} \mathbb{E}_q w'(y)$. Given the affine nature of w', any such deviation would also improve Principal's worst-case. We have,

$$y - w'(y) = \frac{1-\alpha}{\alpha} w'(y) - \frac{1+\alpha}{\alpha}\beta \tag{25}$$

Taking expectations on both sides, and then taking minima, establishes the last claim. Since this is true for any a^*, we have found an affine contract that dominates $w(y)$.

(b′) We next show that, for any affine contract $w'(y)$, there is a linear contract $\ell(y)$ that does at least as well, or strictly better, for Principal, and still rationalize the same action as $w'(y)$.

If $\beta = 0$, then $w'(y)$ is already linear and we are done. If $\beta > 0$, then note that $w'(0) = \beta$. Define $\ell(y) = \alpha y = w'(y) - \beta$. This improves Principal's payoff and leaves Agent's incentives unchanged.

(c′) The previous steps, taken together, imply that any action that is rationalized by an arbitrary contract, $w(y)$, can be also rationalized by an appropriately specified linear contract, that makes Principal and Agent both at least as well off as under $w(y)$. In other words, it is enough for Principal to only optimize within the class of linear contracts.

From (a′), we can strengthen (19) to equality:

$$\mathbb{E}_{q_P^*}[y - \ell(y)] = \frac{1-\alpha}{\alpha}\mathbb{E}_{q_P^*}\ell(y) = \frac{1-\alpha}{\alpha}\mathbb{E}_{q_A^*}\ell(y) \tag{26}$$

This follows from noting that with a linear contract, for any action a, Principal and Agent agree on the worst-case $q_P^* = q_A^*$, even if Principal is not able to perfectly infer $\widehat{a}^*$ the linear contract ensures that Agent's optimal action choice maximizes Principal's guarantee.

In other words, with a linear contract, the best worst-case guarantees for Princial and Agent are related via a linear relationship described in (26). Now using this linear relationship we establish the existence of optimal contract in linear form $\ell(y) = \alpha y$. From the linear contract αy Principal's guarantee value is $V_P(\alpha) = \max_{a\in\mathcal{A}}[(1-\alpha)E_{q^*(a)}[y] - \frac{1-\alpha}{\alpha}g(a)]$. Continuity of V_P in α follows from the Theorem of Maximum by taking α in the role of the parameter, and an $\alpha^*(a^*)$ exists in $[0,1]$ that maximizes it.[13]

□

4.2. **Optimal Linear Contract.** We have established the optimality of linear contracts. Next we identify the optimal linear contract. Using Theorem 2 and especially the linear relationship between the values in (26) the optimal share maximizes

$$\max_{a\in\mathcal{A}}[(1-\alpha)E_{q^*(a)}[y] - \frac{1-\alpha}{\alpha}g(a)] \tag{27}$$

jointly over $a \in \mathcal{A}$ and $\alpha \in [0,1]$. Here we use the observation that $q_p^*(a)$, the lower envelope of $Q^P(a)$ is the common worst-case that minimizes the expected output in that

[13] There are some subtleties regarding $\alpha = 0$, for which we provide more details below in Section 4.2.

set. Our approach is to find the optimal α for any given action a, then find the maximum over $\mathcal{A}$. If the worst-case expected output from an action a is not high enough to cover its effort cost, that is if $E_{q_p^*(a)}[y] < g(a)$, then the optimum α is zero, corresponding to zero contract. Otherwise, the optimal share α for action a that solves (27) is equal to $\alpha(a) = \sqrt{g(a)/E_{q_p^*(a)}[y]}$.

Using $\alpha(a)$ in (27) and by continuity of the objective function in a over a compact $\mathcal{A}$ we solve for the optimal action a^* and pick as the share $\alpha^* = \alpha(a^*) = \sqrt{g(a^*)/E_{q^*(a^*)}[y]}$. Notice that as Principal's perception of technology becomes more ambiguous $Q^P(a) \subset \hat{Q}^P(a)$ the expected wort-case becomes worse $E_{\hat{q}^*(a^*)}[y] \leq E_{q^*(a^*)}[y]$ and the share of output paid to Agent increases (this observation is true for *uniform* increase in ambiguity, i.e. when $Q^P(a)$ becomes larger for each a).

5. Discussion

We illustrate optimality of linear contracts in a moral hazard problem when both parties have ambiguous perceptions of technology, and Principal faces more ambiguity than Agent. The main result derives from the ability of linear contracts to align the worst-case scenarios for both parties, thereby reducing the impact of the endogenous uncertainty of Agent's choice of action on Principal's payoff.

We conclude with a few remarks about our assumptions. It is important that Principal has coarser understanding of technology relative to Agent's. This assumption ensures that Principal's worst-case value, relative to her own coarser information is well-defined (it lies on the boundary of a certain convex sudset of the convex hull of the graph of $w(\cdot)$. Our assumption on the coarseness of Principal's understanding allows for richness in Agent's action choice as perceived by Principal. In this sense, it generates in an endogenous manner Carroll (2015) richness condition on the set of actions available to Agent.

Thus, our analysis provides intuition as to why a similar argument for linearity may not work if the asymmetry of uncertainty is reversed, i.e., if Principal is better informed about technology than Agent (for example, when an experienced Principal hires a rookie Agent).[14]. In that case, Principal is not able to uniquely identify a lower bound on Agent's value for an arbitrary contract. Characterizing the optimal contract in such a case is our next step in this research project.

[14]See Lopomo et al. (2011) for a treatment of such a problem in a different formal setting; a two-part contract with a flat payment and a bonus is proposed.

References

Charalambos Aliprantis and Kim Border. *Infinite dimensional analysis*. Springer, 2006.

Nemanja Antic. Contracting with unknown technologies. *unpublished paper, Princeton University*, 2014.

Lawrence M Ausubel and Raymond J Deneckere. A generalized theorem of the maximum. *Economic Theory*, 3(1):99–107, 1993.

Dirk Bergemann and Stephen Morris. Robust mechanism design. *Econometrica*, 73(6): 1771–1813, 2005.

Dirk Bergemann and Karl Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146(6):2527–2543, 2011.

Gabriel Carroll. Robustness and linear contracts. *The American Economic Review*, 105(2): 536–563, 2015.

Sylvain Chassang. Calibrated incentive contracts. *Econometrica*, 81(5):1935–1971, 2013.

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

Alexander Frankel. Aligned delegation. *The American Economic Review*, 104(1):66–83, 2014.

Daniel F Garrett. Robustness of simple menus of contracts in cost-based procurement. *Games and Economic Behavior*, 87:631–641, 2014.

Paolo Ghirardato. Agency theory with non-additive uncertainty. *Unpublished working paper, Cal. Tech*, 1994.

Sanford J Grossman and Oliver D Hart. An analysis of the principal-agent problem. *Econometrica: Journal of the Econometric Society*, pages 7–45, 1983.

Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Springer Science & Business Media, 2012.

Bengt Holmstrom and Paul Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328, 1987.

Leonid Hurwicz and Leonard Shapiro. Incentive structures maximizing residual gain under incomplete information. *The Bell Journal of Economics*, pages 180–191, 1978.

Edi Karni. Subjective expected utility theory without states of the world. *Journal of Mathematical Economics*, 42(3):325–342, 2006.

Edi Karni. A reformulation of the maxmin expected utility model with application to agency theory. *Journal of Mathematical Economics*, 45(1):97–112, 2009.

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

Sujoy Mukerji. Ambiguity aversion and incompleteness of contractual form. *American Economic Review*, pages 1207–1231, 1998.

Appendix A. Existence of Solutions to Agent's Optimization Problem

Assume that the correspondence that maps a in compact set $\mathcal{A}$ to $Q(a)$ in compact convex set of probabilities over output levels $\mathbb{K}_{\Delta(\mathcal{Y})}$ is upper hemi-continuous.

We want to show that for a given continuous contract w, the maxmin action $a^* \in \mathcal{A}$ is well-defined.

We start with characterizing Agent's decision rule for the effort choice for a given contract. Notice that for each action a Agent effort choice solves:

$$V_A(a; w) = \min_{q \in Q(a)} E_q[w(y)] - g(a)$$

and the value function is continuous in a. To see this, consider effort level a in the role of a parameter in the constrained optimization problem. In particular, the objective function is continuous in effort a and the feasible set correspondence $Q(a)$ is uhc. Therefore, by the (generalized) Maximum Theorem from Ausubel and Deneckere (1993), that uses upper-hemicontinuity of the feasible set of alternatives, the worst-case value $V_A(a; w)$ is continuous in the parameter, or in this case the action a.

For the contract w Agent's optimal effort $a^*(w)$ which maximizes his continuous value function $V_A(\cdot; w)$ over a compact set $\mathcal{A}$ is well-defined by the Extreme Value Theorem (Aliprantis and Border (2006)).

By symmetry, Principal infers Agent's decision rule a^* and if the optimal action is unique Principal uses it to compute his worst-value from the contract:

$$V_p(w) = \min_{q \in Q(a^*(w))} E_q[y - w(y)]$$

If Agent's decision rule assigns multiple actions as optimal to a contract w, from such a contract Principal evaluates his guarantee according to the worst case. That is, if $A^*(w)$ is Agent's set of optimizers, Principal's guarantee considers the worst of these actions from her perspective:

$$V_p(w) = \min_{a \in A^*(w)} \min_{q \in Q(a)} E_q[y - w(y)]$$

Since $A^*(w)$ compact, being a closed subset of compact $\mathcal{A}$, Principal's guarantee is again well-defined. An alternative approach to the decision rules with multiple actions is that Agent when indifferent picks the best action for Principal. This alternative approach which is typically used in standard moral hazard problems is easy to adopt in our guarantee specification above by taking max over $A^*(w)$ instead of min, and without affecting the analysis and the results qualitatively.

Department of Economics, Universidad Carlos III de Madrid, e-mail: mdumav@gmail.com, and Department of Economics, University of California at Riverside, e-mail: urmeek@ucr.edu