

# Knightian Uncertainty and Moral Hazard\*

Giuseppe Lopomo<sup>†</sup>  
Fuqua School of Business  
Duke University

Luca Rigotti<sup>‡</sup>  
Fuqua School of Business  
Duke University

Chris Shannon<sup>§</sup>  
Department of Economics  
UC Berkeley

September 2009

## Abstract

This paper presents a principal-agent model in which the agent has imprecise, or multiple, beliefs. This multiplicity can be interpreted as limited knowledge of the surrounding environment. We model this situation formally by assuming the agent's preferences are incomplete. In this setting, incentives must be robust to the agent's limited knowledge. We study whether robustness leads to simplicity of the resulting optimal contracts. Under mild conditions, we show that optimal contracts have a two-wage structure. That is, all output levels are divided into two groups, and the optimal incentive scheme pays the same amount for all output levels in each group. This can be interpreted as a flat payment plus bonus contract.

JEL Codes: D82, D81, D23

---

\*We thank Truman Bewley and Ben Polak for invaluable advice. Dirk Bergemann, Subir Bose, Simon Grant, Georg Kirchsteiger, June K. Lee, David Pearce, Aldo Rustichini, Matthew Ryan, and Dolf Talman gave useful comments at various stages of this work. We also thank seminar audiences at Arizona State, CORE, Universitat Autònoma de Barcelona, ITAM, Stanford, Tilburg, and Yale.

<sup>†</sup>E-mail: [glopomo@duke.edu](mailto:glopomo@duke.edu)

<sup>‡</sup>E-mail: [rigotti@duke.edu](mailto:rigotti@duke.edu)

<sup>§</sup>E-mail: [cshannon@econ.berkeley.edu](mailto:cshannon@econ.berkeley.edu)

# 1 Introduction

We study principal-agent problems in which the agent may have imprecise, or multiple, beliefs. This imprecision arises because the agent is not confident in assessing the possible consequences of his actions. We discuss the characteristics of an incentive scheme in such settings. Our main conclusion is that under mild conditions, optimal incentive schemes are simple, in particular, take only two values across all possible output states.

The moral hazard model sometimes generates extremely complex incentive structures. Optimal contracts often involve as many different payments as there are possible levels of output. In addition, small changes in the assumed distribution of outcomes can lead to large changes in the way an optimal scheme depends on output; that is, in its shape. Casual empiricism, on the other hand, suggests that many contracts are quite simple. For example, many labor contracts have a simple two-wage structure: a flat payment plus an “incentive bonus” at the end of the year. Why is this structure so common across different environments? The standard model’s answer is that all of them must share the same stochastic structure of output.

Some authors have speculated that contracts are simple because they need to be robust. Hart and Holmstrom (1987), for example, argue that real world incentives need to perform well across a wider range of circumstances than the ones accounted for in the standard model. Once this need for robustness is considered, simple optimal schemes might obtain. We follow this path by introducing a particular robustness requirement. Suppose the agent lacks confidence in judging the stochastic properties of the environment he operates in. This is reflected by non-unique beliefs about possible output levels. An incentive scheme which accounts for this problem is necessarily robust to the agent’s different beliefs. In this framework, we show the optimal scheme often has a two-wage structure. Thus, the need for robustness we examine generates simple contracts. Furthermore, these contracts have a shape we commonly observe in the real world.<sup>1</sup>

The situation we have in mind is the following. An entrepreneur is considering whether to hire

---

<sup>1</sup>Holmstrom and Milgrom (1987) provide conditions under which linear incentive schemes are optimal. These conditions include constant relative risk aversion and a specific dynamic property of stochastic output. Neither of these requirements is related to the idea of robustness stated above. We allow the agent to consider many stochastic structures of output. We also adopt a different notion of simplicity. A two-wage scheme is simple because it can be thought of as contingent on only two events; a linear contract is simple because it is contingent on an intercept and a slope for all events.

a worker. If she hires him, she cannot observe how much effort he puts into his job. He, however, cannot precisely evaluate the impact of what he does on the production process. Output depends on his effort, but also on many variables beyond his control, like what other workers in the firm do, or even pure luck. In particular, he has no previous experience working at her company. Thus, he may not feel confident evaluating the relation between the effort he devotes to his work and the output produced. She, however, as owner of the production technology can evaluate this relation more precisely.

In this relationship the agent is an outsider who is not familiar with all the details of the production process. The principal is an insider who is familiar with these details. The principal is at an informational disadvantage because she cannot observe the agent's effort choice, but the agent is entering a new environment and cannot evaluate precisely the consequences of his work. The agent's behavior may reflect the uncertainty he faces. The standard principal-agent model, however, neglects this possibility. It assumes both parties can precisely evaluate the stochastic consequences of the agent's action.

We model this situation formally by relaxing the assumption that an agent's preferences are complete. In this case, the agent may be unable to compare alternatives offered to him. Aumann (1962) and Bewley (1986) showed that incomplete preferences can be represented by a von Neumann-Morgenstern utility function with multiple probability distributions. This multiplicity can be thought of as imprecision of the agent's beliefs over uncertain outcomes. The agent computes an expected utility for each distribution, and they all matter in determining his behavior. One interpretation of this multiplicity follows Knight (1921). Individuals use a single distribution only when they regard events as risky; individuals use a set of distributions when they regard events as uncertain. The term Knightian uncertainty has been associated with the latter situation.

A question that arises with incomplete preference models is what do agents do when all alternatives are incomparable. Bewley (1986) adds a behavioral assumption, inertia, in part to address this issue. The inertia assumption states that, when faced with incomparable options, an individual sticks with his current behavior, the status quo, unless an alternative is strictly preferred. This 'uncertainty aversion' reflects reluctance to change behavior when the consequences of doing so are difficult to evaluate. Evidence of behavior under uncertainty that is consistent with such a status quo bias can

be found both in economics and psychology. A classic reference is Samuelson and Zeckhauser (1988). In an experimental setting, they find significant status quo bias in investment decisions regarding portfolio composition following a hypothetical inheritance. They also find similar evidence in field data on health plan choices and portfolio division in TIAA-CREF plans among Harvard employees. More recent evidence is given in Madrian and Shea (2001) and Ameriks and Zeldes (2000). See also Fox and Tversky (1995), Einhorn and Hogarth (1985) and Heath and Tversky (1991). In abstract settings, identifying a status quo is sometimes hard, making inertia less plausible. In our setting, however, there is a natural candidate for the status quo, the agent's outside option. We investigate features of optimal contracts both with and without the inertia assumption.

We consider the following moral hazard model. A risk-neutral principal has to design an incentive scheme for a risk-neutral agent who has imprecise beliefs over output outcomes. These beliefs are represented by sets of probability distributions, one set for each action. We assume risk-neutrality to focus on the impact of Knightian uncertainty alone. Therefore, multiplicity of beliefs is the only difference between our model and a risk-neutral version of Grossman and Hart (1983). The principal cannot observe the agent's action. Each action has a different disutility to the agent and induces different beliefs over output outcomes. For each action, the principal designs a contract which implements it at the lowest possible expected cost. Then, she selects the action that maximizes the difference between expected output and expected cost.

If agents have imprecise beliefs but do not satisfy the inertia assumption, it is often impossible for the principal to induce the agent to take a privately costly action such as hard work. For an incentive scheme to implement a particular action, the agent must prefer that action to his reservation utility and to all other actions. The agent, however, regards any two actions whose belief sets intersect as not comparable. Thus, an action can only be implemented if the agent's belief set corresponding to it does not intersect any of the belief sets corresponding to the other actions.

In many interesting situations, however, the agent's beliefs intersect. For example, if the agent chooses the lowest effort action, his beliefs may be extremely imprecise but the harder the agent works, the more precisely he evaluates his influence on the production process. In this case, all agent's belief sets intersect and, according to the above result, no action can be implemented.

Implementation is easier when the agent satisfies the inertia assumption. With inertia, an in-

centive scheme implements an action if this action is preferred to the reservation utility and, for each other action, either the first action is preferred, or the other action is not comparable to the reservation utility. Preferring one action to all others is no longer necessary. With the inertia assumption, implementing an action is sometimes possible even if the agent's belief sets intersect as in the example above.

Optimal contracts under imprecise beliefs are both robust and simple. Regardless of inertia, an optimal contract is robust because it provides incentives for the entire set of probability distributions the agent considers. Under mild conditions on the agent's imprecise beliefs, the unique optimal incentive scheme divides all the possible outputs levels into two groups and pays the same amount in all states belonging to the same group. First, we prove the result when the agent can choose between two actions. Then, we generalize it to the case in which many actions are available to him. Consider the following example. Suppose the number of events a contract is contingent upon is increased by one because the principal decides to make different payments in two output levels that previously corresponded to the same wage. In other words, one event is divided into two separate events. Risk-neutrality implies the agent does not place any premium on receiving different payments. Satisfying the constraints is now more difficult, however, because the new events have, in general, different probabilities for different elements of the agent's belief sets. Thus, dividing events makes it more difficult to provide incentives. Conversely, the formal proof shows that joining events is strictly profitable for the principal. Under mild assumptions, this result holds, regardless of the number of output levels, provided this number is finite. Additional restrictions guarantee it also holds for any finite number of actions available to the agent.

Recently, much attention has been devoted to incomplete contracts. For example, see Moore and Hart (1988), Tirole (1999), and Hart (1995). Tirole argues robustness (in the sense of limited knowledge of the surrounding environment) should be investigated as a source of incomplete contracts. Our main result does not deal with this problem explicitly, but suggests Knightian uncertainty as a useful tool. The main intuition of our result is that asymmetric confidence in beliefs introduces an uncertainty cost in contracts. This cost may depend positively on the number of events the contract is contingent upon. Therefore, it can be reduced by making the contract depend on fewer events. If this is the case, incomplete preference generate contracts that are simple, in the sense that they

depend on few contingencies.

Mukerji (1998a) and Ghirardato (1994) present moral hazard models similar in motivation to the one we describe, but unlike us, they get rather standard results about incentive schemes. The decision theoretic model they use, however, is different. In both cases, the principal and the agent are Choquet expected utility maximizers (see Schmeidler (1989)). Choquet expected utility is a model in which, loosely speaking, lack of confidence and uncertainty are reflected by the non-additivity of probabilities, not by the imprecision of beliefs. In that framework, there is no simple way to allow for asymmetric confidence, or lack thereof, across the parties involved, and optimal incentive schemes need not be simple in those settings. On the other hand, Mukerji (1998b) uses this framework to relate uncertainty to contract incompleteness.

In a related paper, Lopomo, Rigotti, and Shannon (2009), we study the implications of imprecise beliefs for general mechanism design problems. We show that mechanisms that are robust to uncertainty must be ex post incentive compatible in many standard mechanism design settings. This result can also be interpreted as an example of the theme we explore here: robustness to uncertainty might force designers to choose simple mechanisms.

The paper is organized as follows. The next section introduces some concepts of individual decision making when preferences are incomplete. Section 3 presents the basic framework and discusses the implementation rules. Section 4 shows that optimal incentive schemes are simple. Section 5 introduces a primitive model and establishes versions of our main results in this setting.

## 2 Incomplete Preferences and Inertia

We briefly describe individual behavior under incomplete preferences, and then introduce the inertia assumption. This approach to decision making was pioneered by Bewley (1986), and further developed in Bewley (1987).

### 2.1 Preference Representation without Completeness

If a preference relation satisfies the completeness axiom, the decision maker can compare any two alternatives. Von Neumann and Morgenstern were the first to observe that completeness is not an entirely satisfactory axiom in the presence of uncertainty.<sup>2</sup> If a preference relation is not complete, the

---

<sup>2</sup>They write:

decision maker cannot rank all pairs of alternatives. Bewley (1986) gives a set of axioms under which an incomplete preference relation can be characterized by multiplicity of beliefs and a corresponding family of expected utility functions, with alternatives compared one probability distribution at a time.

To formalize, let  $\mathcal{N} := \{1, \dots, N\}$  be a set of  $N$  states. Let  $\Delta(\mathcal{N}) := \{\pi \in \mathbf{R}^N : \pi_i \geq 0 \forall i, \sum_i \pi_i = 1\}$  be the set of probabilities on  $\mathcal{N}$ , and let  $x, y \in \mathbf{R}^N$  be random monetary payoffs defined on  $\mathcal{N}$ . In an Anscombe-Aumann setting, Bewley (1986) proves that under standard axioms on preferences except for completeness, a representation of the following form can be derived:

$$x \succ y \quad \text{if and only if} \quad E_\pi[u(x)] > E_\pi[u(y)] \text{ for all } \pi \in \Pi \quad (1)$$

where  $u$  is the von Neumann-Morgenstern utility derived from the payoffs,  $\Pi$  is a unique closed and convex set whose elements are subjective probability distributions, and  $E_\pi[z]$  denotes the expected value of  $z$  with respect to  $\pi$ .<sup>3</sup> The set  $\Pi$  has only one element if and only if the preference relation is complete, and in this case the usual subjective expected utility representation obtains. When the relevant state-space has  $N$  elements, (1) reduces to:

$$x \succ y \quad \text{if and only if} \quad \sum_{i=1}^N \pi_i u(x_i) > \sum_{i=1}^N \pi_i u(y_i) \text{ for all } \pi \in \Pi$$

Bewley argues that this model provides a possible formulation of Frank Knight's distinction between risk and uncertainty. A payoff is risky when the probabilities of different outcomes are known; if they are unknown, the payoff is uncertain. Hence, payoffs are risky when  $\Pi$  has only one element and uncertain otherwise. Informally, the size of  $\Pi$  measures the amount of uncertainty the individual perceives, and can be thought of as reflecting confidence in beliefs.

Figure 1 illustrates Bewley's representation. The axes measure consumption in each of the two possible states. Given a probability distribution over the states, an indifference curve through the bundle  $x$  represents all the bundles that have the same expected utility as  $x$  according to this

---

"It is conceivable - and may even in a way be more realistic - to allow for cases where the individual is neither able to state which of two alternatives he prefers nor that they are equally desirable...How real this possibility is, both for individuals and for organizations, seems to be an extremely interesting question...It certainly deserves further study." von Neumann and Morgenstern (1953), Section 3.3.4, pg. 19.

<sup>3</sup>This is Theorem 1.2 in Bewley (1986); a similar result is found in Aumann (1962).

distribution. As the distribution changes, many indifference curves arise.

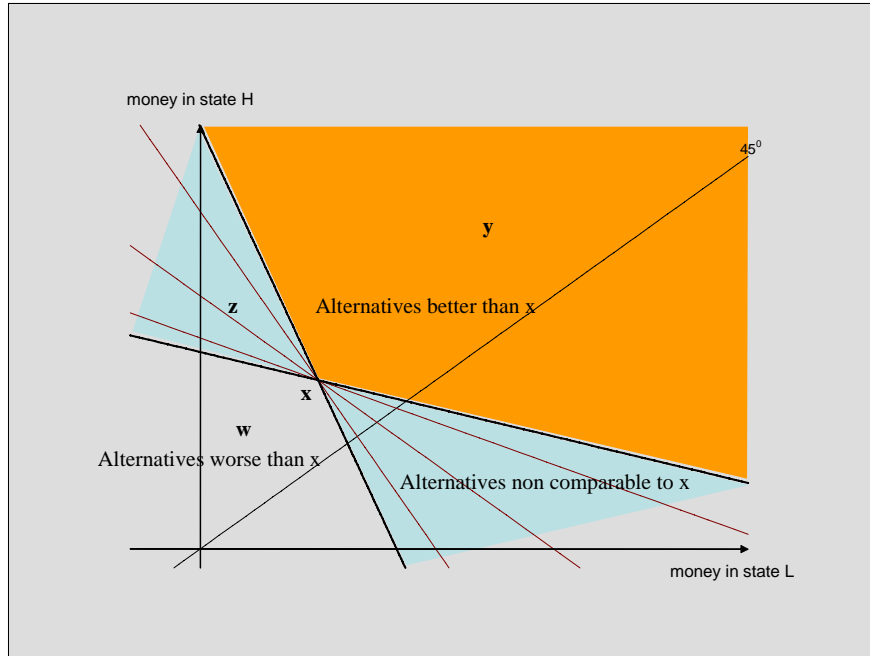


Figure 1: Incomplete Preferences

Figure 1 illustrates the three types of comparisons possible for incomplete preferences. First,  $y$  is preferred to  $x$  since it lies above all indifference curves through  $x$ . Second,  $x$  is preferred to  $w$  since  $w$  lies below all of the indifference curves through  $x$ . Third,  $z$  is not comparable to  $x$  since it lies above some indifference curves through  $x$  and below others. Thus corresponding to any bundle  $x$  there are three regions: bundles preferred to  $x$ , bundles worse than  $x$ , and bundles incomparable to  $x$ . This last region is empty only if there is a unique probability distribution over the two states and preferences are complete.

## 2.2 The Inertia Assumption

Revealed preference arguments must take incompleteness into account. If  $x$  is chosen when  $y$  is available, one cannot conclude that  $x$  is revealed preferred to  $y$ , but only that  $y$  is not revealed preferred to  $x$ . The concepts of status quo and inertia introduced in Bewley (1986) can sharpen revealed preference inferences when preferences are incomplete. Bewley's inertia assumption has two parts: the first posits the existence of a "status quo" or planned behavior that is taken as a reference

point, and the second assumes that the status quo is abandoned only for alternatives preferred to it. To illustrate, suppose that  $y$  is the status quo and the inertia assumption holds. Then if  $x$  is chosen when  $y$  is available,  $x$  must be preferred to  $y$ . In this case, when the inertia assumption holds, observing that  $x$  is chosen when  $y$  is available implies that  $x$  is revealed preferred to  $y$ . In Figure 1, for example, if  $x$  is the status quo and the inertia assumption holds, then alternatives like  $z$  will not be chosen since they are incomparable to  $x$ .

The inertia assumption is a behavioral assumption. Evidence consistent with this behavior can be found both in economics and psychology. A classic reference is Samuelson and Zeckhauser (1988), who find evidence of status quo biases in both field and experimental data. In an experimental setting, they find significant status quo bias in investment decisions regarding the portfolio composition following a hypothetical inheritance. Ameriks and Zeldes (2000) find similar evidence in data from TIAA-CREF and Surveys of Consumer Finance: almost half of their sample made no change in portfolio composition over the course of their 9 year sample, while the same period saw drastic changes in the returns to bond and equity holdings. Einhorn and Hogarth (1985) find evidence supporting a status quo bias in initial probability assessments in a number of experiments (see also Fox and Tversky (1995) and Heath and Tversky (1991)).

In many economic contexts, there is a natural candidate for the status quo. For example, consider a bargaining game in which each player has an outside option. If we interpret these options as the players' actions before entering the game, defining each player's outside option as his status quo seems natural. In the moral hazard model that follows, a natural candidate for a status quo is the action yielding the agent's reservation utility. This corresponds to the payoff the agent receives if he rejects the contract the principal offers. We interpret this as the reward corresponding to the agent's planned behavior before the possibility of contracting with the principal materialized. Below we explore the structure of optimal contracts both with and without the inertia assumption.

### 3 The Moral Hazard Setup

We present a moral hazard model where the agent perceives more uncertainty than the principal. Formally, the agent's preferences are not complete and the principal's preferences are complete. Because the inertia assumption is somewhat arbitrary, we establish conditions for a contract to

implement some action with and without inertia. We show this has different implications for the model; namely, the latter case imposes undesirable restrictions. Finally, we define an optimal incentive scheme, and derive some of its basic properties.

The principal owns resources that yield output. An agent has to perform some action for production to take place. The principal and the agent observe the realized output level, but the principal cannot observe the action performed by the agent. An incentive scheme is a contract that induces the agent to perform a particular action.

The  $N$  states of nature are distinguished by the amount of output produced. Let  $S := \{1, \dots, N\}$  denote the state space. Output is an  $N$ -vector denoted by  $y = (y_1, \dots, y_N)$ , with states labeled so that  $y_N > \dots > y_1$ . An incentive scheme is an  $N$ -vector denoted by  $w = (w_1, \dots, w_N)$ , where the payment from the principal to the agent in state  $j$  is  $w_j$ . The agent chooses an action  $a$  from a discrete set of available actions  $\mathcal{A} := \{1, 2, \dots, M\}$ , interpreted as effort levels. The agent's reservation utility, or outside option, is  $\bar{w}$ . This represents the reward the agent receives from his current behavior, which we will interpret as the status quo when we impose the inertia assumption. Each action has two consequences: it imposes a cost on the agent, measured by disutility of effort, and it generates beliefs about the likelihood of different output levels. Throughout, we use subscripts to denote states.

The principal's beliefs are described by a function  $\pi^P : \mathcal{A} \rightarrow \Delta$ , while the agent's beliefs are described by a correspondence  $\Pi : \mathcal{A} \rightarrow 2^\Delta$  such that  $\Pi(a)$  is a closed and convex set for each  $a$ . Thus  $\pi^P(a)$  denotes the principal's beliefs regarding the distribution of output when the agent chooses action  $a$ , while  $\Pi(a)$  denotes the set of beliefs of the agent corresponding to the same action choice  $a$ . This description follows the parametrized distribution formulation of the principal-agent problem pioneered by Holmstrom (1987), in which  $\Pi$  is also a function.<sup>4</sup> We assume that both  $\pi^P$  and  $\Pi$  are common knowledge.

We assume that the principal's beliefs are consistent with the agents. Formally, for each  $a$  in  $\mathcal{A}$ ,  $\pi^P(a) \in \Pi(a)$ . This can be interpreted as implying there is asymmetric confidence in beliefs. The principal is more confident in evaluating the stochastic relationship between the agent's effort and output. Uniqueness of the principal's beliefs is assumed for analytical tractability. Most of the

---

<sup>4</sup>As we discuss in Section 5, the parametrized distribution set up may be less reasonable in the presence of Knightian uncertainty.

analysis carries over if  $\pi^P$  is also a correspondence, as long as the asymmetric confidence assumption, that beliefs are nested so that  $\pi^P(a) \subset \Pi(a)$ , is maintained. Finally, the assumption that  $\pi^P(a) \in \Pi(a)$  for each action rules out asymmetric information in the standard sense, since one of the agent's priors agrees with the principal's prior.

The cost of an action  $a$  is denoted  $c(a)$ . We assume actions are ordered such that  $c(a) > c(a')$  and  $\sum_{i=1}^N \pi_i^P(a)y_i \geq \sum_{i=1}^N \pi_i^P(a')y_i$  whenever  $a > a'$ . More expensive actions increase the expected value of output. Both assumptions are standard in the principal-agent literature.

The principal and the agent are risk-neutral. The agent's ex post utility is given by the payment from the contract minus the cost of the action he performs. To compare actions under a given contract, the agent computes many expected values, one for each element of the corresponding belief sets. His behavior depends on all of them. For each  $a$  and each  $\pi \in \Pi(a)$ ,

$$E_\pi[w] - c(a) = \sum_{j=1}^N \pi_j w_j - c(a).$$

The principal's utility function is given by the expected value of output minus the expected cost of the contract. These expectations are computed according to the probability distribution induced by the effort level chosen by the agent. The principal's utility is

$$E_{\pi^P(a)}[y - w] = \sum_{j=1}^N \pi_j^P(a)y_j - \sum_{j=1}^N \pi_j^P(a)w_j.$$

The set up parallels the standard principal-agent model. There, the principal is assumed risk-neutral while the agent is assumed risk-averse. Both parties have the same beliefs, and the same attitude toward uncertainty, but they evaluate risk in different ways. In our model, both parties have the same beliefs, and evaluate risk in the same way, but they have different attitudes toward uncertainty. Notice that even though the agent's utility function is linear, the corresponding marginal rate of substitution over money is not constant since it depends on which  $\pi \in \Pi(a)$  is used to evaluate the contract.

When the agent's action is observable (and/or verifiable), the principal can make the contract contingent on it. Denote the cost of this incentive scheme by  $C^{FB}(a)$ , where FB stands for first-best. As in the standard model,  $C^{FB}(a) = \bar{w} + c(a)$  for each action  $a$ . The principal induces the agent to pick  $a$  by offering a contract that in every state pays the agent  $C^{FB}(a)$  if action  $a$  is taken and  $-\infty$

otherwise. A contract that is not constant across all states of the world cannot be Pareto efficient: it is dominated by a “flatter” contract that leaves the principal’s expected utility the same but makes the agent strictly better off.<sup>5</sup>

Given this setup, the principal and the agent do not value the firm equally. In particular, the agent is more cautious about this value. Therefore, we cannot appeal to the standard way of dealing with moral hazard when the parties are risk-neutral, which would have the principal sell the firm to the agent. For a given action, the highest price the agent is willing to pay for the firm is the lowest expected value of output minus the cost of taking that action. The highest price the agent is willing to pay is thus lower than what the firm is worth to the principal.

### 3.1 Two Implementation Rules

The principal wants the agent to perform a specific action  $a^*$ . Because she cannot observe the agent, she must rely on the contract alone to provide the desired incentives. Obviously, different assumptions about the agent’s behavior will lead to different requirements the contract must satisfy. Here, we focus on contracts that leave no doubt the agent’s choice is  $a^*$ . A contract is a take-it-or-leave-it offer, and we say a contract implements  $a^*$  if the agent accepts the contract and chooses  $a^*$  among all possible actions.

#### 3.1.1 Implementing without inertia

Under our basic notion of implementation, a contract implements action  $a^*$  if it satisfies the standard participation and incentive compatibility conditions, modified to take multiplicity of beliefs into account. The participation constraint is satisfied if taking action  $a^*$  is preferred to the reservation utility, and incentive compatibility is satisfied if  $a^*$  is preferred to all other actions.<sup>6</sup>

**Definition 1.** *An incentive scheme  $w$  implements  $a^*$  if:*

$$\sum_{j=1}^N \pi_j(a^*) w_j - c(a^*) \geq \bar{w} \quad \forall \pi(a^*) \in \Pi(a^*) \quad (\text{P})$$

---

<sup>5</sup>Rigotti and Shannon (2005) give a general characterization of Pareto optimal allocations with incomplete preferences.

<sup>6</sup>In Lopomo, Rigotti, and Shannon (2009) we distinguish two notions of incentive compatibility in a general mechanism design framework: optimal and maximal. A mechanism is optimal incentive compatible if truth telling is preferred to all other reports. A mechanism is maximal incentive compatible if no report is preferred to truth telling. The notion of incentive compatibility we consider here corresponds to what we call optimal incentive compatibility in Lopomo, Rigotti, and Shannon (2009).

and for each  $a \in \mathcal{A}$  with  $a \neq a^*$

$$\sum_{j=1}^N \pi_j(a^*)w_j - c(a^*) \geq \sum_{j=1}^N \pi_j(a)w_j - c(a) \quad \forall \pi(a^*) \in \Pi(a^*) \quad \text{and} \quad \forall \pi(a) \in \Pi(a) \quad (\text{IC})$$

In words, for any probability distribution induced by  $a^*$ , the expected utility the agent receives from the contract must be weakly higher than the reservation utility, and weakly higher than the expected utility calculated according to all probability distributions induced by any other action.

The implementation requirements imposed by Definition 1 are restrictive. In particular, there could be many actions that cannot be implemented.

**Proposition 1.** *If  $\Pi(a) \cap \Pi(a') \neq \emptyset$  for some actions  $a \neq a'$ , then  $a$  and  $a'$  cannot both be implemented.*

*Proof.* Let the two actions be  $a$  and  $a'$  and, without loss of generality, assume  $a > a'$ . Suppose by way of contradiction that there exists  $\pi \in \Pi(a) \cap \Pi(a')$ , and there exists a contract  $w$  that implements  $a$ . Therefore, by (IC),

$$\sum_{j=1}^N \pi_j w_j - c(a) \geq \sum_{j=1}^N \pi_j w_j - c(a')$$

This implies  $c(a) \leq c(a')$ , and contradicts  $a > a'$ . ■

A possible implication of Proposition 1 is that, when the number of actions is large, only few could be implementable. In particular, no action whose belief set intersects the belief set of a cheaper action can be implemented. This result might have sharp implications in settings in which actions are naturally related to beliefs. For example, as an agent works harder he might gather valuable information regarding his influence on the productive process. Thus costlier actions might lead to a smaller set of beliefs. In such cases, the belief set for a cheaper action will always intersect the belief set of every more expensive action. Only the cheapest action could be implemented in this case.

### 3.1.2 Implementing with inertia

The inertia assumption says an alternative is chosen only if it is preferred to the status quo. In this case, an action not comparable to the reservation utility is not chosen. After an incentive scheme is proposed by the principal, the agent faces two basic decisions: does he accept the offer and if yes, which action does he choose? By rejecting the offer, the agent opts for the outside option, which has a natural interpretation as the status quo in this model. This corresponds to the agent's behavior

before the contractual offer was made available to him, and delivers the agent's reservation utility.

Under inertia, implementation becomes less restrictive. Any action that is not comparable to the outside option will not be chosen by the agent by assumption. That is, suppose  $a^*$  and  $a'$  are not comparable, but  $a^*$  is preferred to the status quo while  $a'$  is not. Then, the inertia assumption implies  $a'$  is not chosen. Thus to implement  $a^*$  when inertia holds, an incentive scheme does not need to make  $a^*$  preferred to all other actions, just those that are comparable to the status quo.

**Definition 2.** An incentive scheme  $w$  implements  $a^*$  with inertia in  $\mathcal{A}$  if

$$\sum_{j=1}^N \pi_j(a^*)w_j - c(a^*) \geq \bar{w} \quad \forall \pi(a^*) \in \Pi(a^*) \quad (\text{P})$$

and for each  $a \in \mathcal{A}$  with  $a \neq a^*$ , either

$$\sum_{j=1}^N \pi_j(a^*)w_j - c(a^*) \geq \sum_{j=1}^N \pi_j(a)w_j - c(a) \quad \forall \pi(a^*) \in \Pi(a^*) \quad \text{and} \quad \forall \pi(a) \in \Pi(a) \quad (\text{IC})$$

or

$$\sum_{j=1}^N \pi_j(a)w_j - c(a) \leq \bar{w} \quad \text{for some } \pi(a) \in \Pi(a) \quad (\text{NC})$$

*NC* is a non-comparability constraint. It says there exists at least one probability distribution induced by  $a$  such that the expected utility the agent derives from the contract is weakly lower than the reservation utility.

Implementation with inertia relaxes the incentive conditions. All incentive schemes that satisfy Definition 1 also satisfy Definition 2. Intuitively, to implement  $a^*$  with inertia, an incentive scheme must only make  $a^*$  the most preferred action among those that might be chosen over the status quo.

### 3.2 The principal's problem

The principal maximizes her expected utility. We can divide her problem into two steps. First, for each action, find the cheapest contract which implements it. Second, decide which action to implement. For a given action  $a$ , let  $\mathcal{H}(a)$  be the (possibly empty) set of incentive schemes that implement  $a$ , and  $\mathcal{H}^I(a)$  be the (also possibly empty) set of incentive schemes that implement  $a$  with

inertia. We say  $\widehat{w}(a)$  is an *optimal incentive scheme to implement a* if it is a solution to

$$\min_{w \in \mathcal{H}(a)} \sum_{j=1}^N \pi_j^P(a) w_j \quad (2)$$

Let  $\widehat{W}(a)$  denote the set of solutions to (5). Similarly, we say that  $\widehat{w}(a)$  is an *optimal incentive scheme to implement a with inertia* if it is a solution to

$$\min_{w \in \mathcal{H}^I(a)} \sum_{j=1}^N \pi_j^P(a) w_j \quad (3)$$

Let  $\widehat{W}^I(a)$  denote the set of solutions to (6). For each  $a$ , define  $C(a) = \sum_{j=1}^N \pi_j^P(a) \widehat{w}_j(a)$  where  $\widehat{w}(a) \in \widehat{W}(a)$  solves (2) and  $\mathcal{H}(a)$  is not empty, and  $C(a) = \infty$  otherwise.  $C(a)$  is the expected cost of an optimal scheme to implement  $a$ . Similarly, define  $C^I(a) = \sum_{j=1}^N \pi_j^P(a) \widehat{w}_j(a)$  where  $\widehat{w}(a) \in \widehat{W}^I(a)$  solves (3) and  $\mathcal{H}^I(a)$  is not empty, and  $C^I(a) = \infty$  otherwise.

The second part of the principal's problem is to choose a target action to implement. Again there are two notions, depending on the notion of implementation used. An *optimal action* solves

$$\max_{a \in \mathcal{A}} \sum_{j=1}^N \pi_j^P(a) y_j - C(a).$$

while an *optimal action with inertia* solves

$$\max_{a \in \mathcal{A}} \sum_{j=1}^N \pi_j^P(a) y_j - C^I(a).$$

As noted before,  $\mathcal{H}(a)$  is a (possibly empty) subset of  $\mathcal{H}^I(a)$ . As a consequence, the principal cannot be worse off if the agent obeys the inertia assumption. Because of the linearity of the problem, existence of a solution is not an issue as long as some action can be implemented. Our main objective is to analyze the characteristics of optimal incentive schemes.

### 3.3 Some characteristics of the optimal incentive scheme

The agent's behavior depends on all the probability distributions in his belief set. Some of them, though, are more relevant for our analysis because they effectively determine whether incentive compatibility and participation constraints are satisfied. For any action  $a$  and incentive scheme  $w$ ,

let

$$\underline{\Pi}(a; w) := \arg \min_{\pi(a) \in \Pi(a)} \sum_{j=1}^N \pi_j(a) w_j$$

Given the contract  $w$ , any  $\underline{\pi} \in \underline{\Pi}(a; w)$  is a probability distribution yielding the minimum expected value for the agent when he chooses action  $a$ . A contract  $w$  satisfies the participation constraint for action  $a$  if and only if  $E_{\underline{\pi}}[w] - c(a) \geq \bar{w}$  for every  $\underline{\pi} \in \underline{\Pi}(a; w)$ . Similarly, given a contract  $w$ , action  $a'$  satisfies the non-comparability constraint if and only if  $E_{\underline{\pi}}[w] - c(a') < \bar{w}$  for some  $\underline{\pi} \in \underline{\Pi}(a'; w)$ . For any fixed action  $a$  and contract  $w$ , let

$$\bar{\Pi}(a; w) := \arg \max_{\pi(a) \in \Pi(a)} \sum_{j=1}^N \pi_j(a) w_j$$

Then  $\bar{\pi} \in \bar{\Pi}(a; w)$  is a probability yielding the maximum expected value of the contract  $w$  for the agent when he chooses action  $a$ . A contract  $w$  satisfies the incentive compatibility constraint for action  $a$  versus  $a'$  if and only if  $E_{\underline{\pi}}[w] - c(a) \geq E_{\bar{\pi}}[w] - c(a')$  for all  $\underline{\pi} \in \underline{\Pi}(a; w)$  and all  $\bar{\pi} \in \bar{\Pi}(a'; w)$ .

The following proposition specializes to our framework some results about the optimal contract which hold in the standard model (the proof is in the appendix).

**Proposition 2.** *Let  $a$  be the action the principal wants to implement.*

(i) *For any  $\hat{w}$  in  $\widehat{W}(a)$  or  $\widehat{W}^I(a)$ , the participation constraint binds when computed according to  $\underline{\pi}(a) \in \underline{\Pi}(a; \hat{w})$ ; formally,*

$$\sum_{j=1}^N \underline{\pi}_j(a) \hat{w}_j - c(a) = \bar{w}$$

*for any  $\hat{w} \in \widehat{W}(a) \cup \widehat{W}^I(a)$  and  $\underline{\pi}(a) \in \underline{\Pi}(a; \hat{w})$ .*

(ii) *If  $a$  is the least costly action, the scheme that pays  $\bar{w} + c(1)$  in all states implements  $a$  (with inertia); that is,  $\bar{w} + c(1) \in \widehat{W}(1)$ .*

(iii) *If  $\hat{w} \in \widehat{W}^I(a)$  and  $a$  is not the least costly action, then there exists an action less costly than  $a$  such that either (IC) or (NC) binds for this action; formally, if  $a > 1$  there exists an  $a' < a$  such that either*

$$\sum_{j=1}^N \underline{\pi}_j(a) \hat{w}_j - c(a) = \sum_{j=1}^N \bar{\pi}_j(a') \hat{w}_j - c(a') \quad \forall \underline{\pi}(a) \in \underline{\Pi}(a; \hat{w}) \text{ and } \forall \bar{\pi}(a') \in \bar{\Pi}(a'; \hat{w})$$

*or*

$$\sum_{j=1}^N \underline{\pi}_j(a') \hat{w}_j - c(a') = \bar{w} \quad \forall \underline{\pi}(a') \in \underline{\Pi}(a'; \hat{w})$$

for any  $\widehat{w} \in \widehat{W}^I(a)$ .

Similarly, if  $\widehat{w} \in \widehat{W}(a)$  and  $a$  is not the least costly action, then there exists an action less costly than  $a$  such that (IC) binds for this action.

## 4 Optimal Incentive Schemes Are Simple

This section contains the main result. We establish mild conditions for an optimal incentive scheme to be simple. Under these conditions, an optimal scheme divides the  $N$  possible states into two groups, and pays the same amount in all states belonging to the each group. The optimal contract distinguishes only between two events: it has a two-wage structure. First, we prove the result when the agent can choose only between two actions. Second, we introduce many actions and show how our result generalizes to this case.

Theorists have conjectured that contracts are simple because they need to be robust. Hart and Holmstrom (1987) argue that, in the real world, a contract must provide incentives across a wider range of circumstances than the ones the standard model considers. The idea of robustness we use represents one way to model this requirement. A different approach is used in Holmstrom and Milgrom (1987). They provide conditions for an optimal scheme to be linear. We also adopt a different notion of simplicity. Here simplicity is related to the number of different events on which the payments are contingent.

The main result depends on two characteristics of the beliefs of the parties involved in the contract.

**Assumption A-1:** For each action  $a$  in  $A$ ,  $\Pi(a)$  is the core of a convex capacity  $v^a$  on  $\mathcal{N}$ .<sup>7</sup> That is,

$$\Pi(a) = \{\pi \in \Delta(\mathcal{N}) : \pi(E) \geq v^a(E) \text{ for each } E \subset \mathcal{N}\} \quad (4)$$

Assumption A-1 places restrictions on the image of the correspondence  $\Pi$ .<sup>8</sup> Geometrically,  $\Pi(a)$  must be a polyhedron whose boundaries are determined by the linear inequalities in (4). Figure 2 displays some examples with three states. In the figure,  $\Pi^A$  and  $\Pi^B$  are cores of convex capacities,

<sup>7</sup>A convex capacity on  $\mathcal{N}$  is a function  $\nu : 2^{\mathcal{N}} \rightarrow [0, 1]$  such that (i)  $\nu(\emptyset) = 0$ , (ii)  $\nu(\mathcal{N}) = 1$ , (iii)  $\forall E, E' \subset \mathcal{N}$ ,  $E \subseteq E'$  implies  $\nu(E) \leq \nu(E')$ , and (iv)  $\forall E, E' \subset S$ ,  $\nu(E \cup E') \geq \nu(E) + \nu(E') - \nu(E \cap E')$ .

<sup>8</sup>In the literature on Choquet expected utility (see Schmeidler (1989)), convexity is assumed in almost all applications. For example, both Mukerji (1998a) and Ghirardato (1994) assume the beliefs of both parties involved in a moral hazard model are represented by convex capacities.

while  $\Pi^C$  is not.

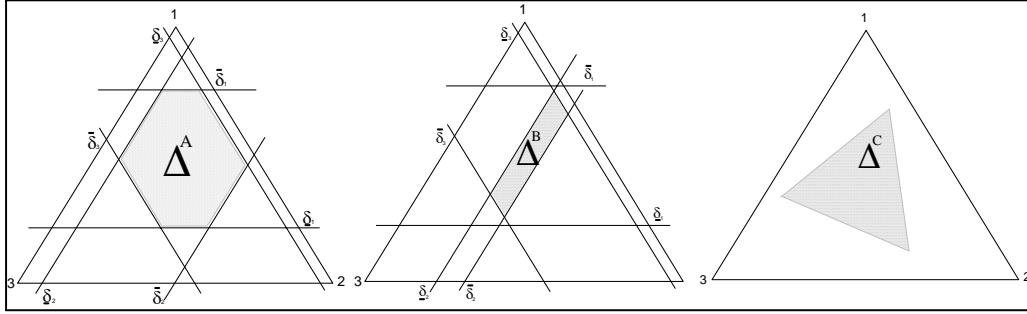


Figure 2: Agent's Beliefs with 3 Output States

**Assumption A-2:** For each action  $a \in \mathcal{A}$ ,  $\pi^P(a)$  is an element of the relative interior of  $\Pi(a)$ .

This makes more stringent the condition that the agent's belief sets contain an element corresponding to the principal's beliefs, and also requires that the agent perceives sufficiently rich uncertainty so that  $\Pi(a)$  has a nonempty relative interior for each action  $a$ .

#### 4.1 Simplicity with two actions

In this section, we specialize the framework to the case in which the agent can choose only among two actions,  $H$  and  $L$ . We interpret these as high and low effort respectively, hence  $c(H) > c(L)$ . As seen previously, the principal can implement the low effort action at the first best cost with a contract which promises the same payment in every state. A more interesting problem arises when the principal wants to implement the high effort action. Under our two basic assumptions, the optimal contract to implement high effort takes on only two values, regardless of the size of the state space.

The main result is the following.

**Proposition 3.** *Suppose Assumptions A-1 and A-2 hold, and the high action  $H$  is implementable (implementable with inertia) in  $\{L, H\}$ . Then, the unique optimal incentive scheme to implement  $H$  (with inertia) divides the  $N$  states into two subsets and is constant on each subset.*

The conclusion that the optimal contract must be simple does not depend on which notion of implementation is used. We provide the details of the proof for the inertia case. The other case can be proved similarly. Loosely, the following example drives the formal argument in the proof.

Consider a contract contingent on only two groups of output levels. Suppose the principal divides the group of output levels corresponding to one payment into two, and makes two different payments in each subset. As we show in the appendix, for a sufficiently small perturbation in the contract payments, a distribution yielding the agent’s extreme expectation of the old contract also yields an extreme expectation of the new one. Such an extreme distribution determines only the aggregate probability of the two new payments, however. Given this probability, in the minimum expectation of the new contract, the smallest weight is assigned to the largest of the two new payments. Therefore, the new contract is not as efficient as the old one in providing incentives.

Intuitively, this result follows from a simple consequence of assumption A-2. For any given state, there is one belief in the agent’s belief set such that the agent assigns lower probability than the principal to that state. Therefore, for any state there is a belief that makes the principal more optimistic about the likelihood of that state. Because all beliefs matter in Bewley’s model, the optimal contract must “accommodate” this probability. Since everyone is risk-neutral, the principal can always slightly modify a payment scheme to take advantage of the disagreement about any one particular state. In particular, this will be done by making the incentive scheme “flatter”, that is, more equal across states. Eventually, the process has to stop because the incentive scheme must have enough variation to distinguish among the two actions.

Note that assumption A-2 is crucial for the argument we developed. It is not, however, necessary to obtain the optimality of simple contracts. The following proposition provides an example where a simplicity result is obtained under different conditions.

**Proposition 4.** *Suppose there exists a state  $j$  such that*

$$\min_{\pi \in \Pi(L)} \pi_j = \pi_j^P(L) = \max_{\pi \in \Pi(L)} \pi_j > 0$$

and

$$\min_{\pi \in \Pi(H)} \pi_j = \pi_j^P(H) = \max_{\pi \in \Pi(H)} \pi_j > 0$$

*Then there exists an optimal incentive scheme (with inertia) that pays the same in all states but  $j$ .*

*Proof.* Such a scheme makes the contract conditional on two events which have precise, i.e. unique, probabilities. Hence, by writing a contract contingent on  $j$  on one hand and all states other than  $j$  on the other, the problem is reduced to the moral hazard model with a risk neutral agent. The solution to this yields a contract which achieves the first-best allocation. ■

Because there are  $N$  possible states and the optimal contract takes two values, Proposition 3

implies this contract can be found looking among  $2^N$  systems of two equations in two unknowns. Below we study features of this optimal contract. An interesting question for future research is how the principal selects among these  $2^N$  possibilities.

## 4.2 Simplicity with many actions

In this section we provide conditions under which optimal incentive schemes are simple when the agent chooses among many actions. We give two extensions of the results in the previous section. First, we illustrate how the main result of the previous section can be extended to more actions, without additional restrictions. Then, under additional restrictions on beliefs that are standard in the principal-agent literature, we show that the optimal contract may take only two values even when there are many actions.

The proof of Proposition 3 shows that the number of signals used in an optimal contract is the minimal number necessary to make a contract feasible. In the case of two actions, this is exactly two. If there are  $M$  possible actions, there will be at most  $M$  binding constraints. Using this observation, the argument used to prove Proposition 3 can be repeated, since it shows that reducing the number of states on which a contract depends is always profitable for the principal. Clearly, this becomes impossible when all constraints are binding. In the end, this shows that the optimal contract is contingent on at most  $M$  events.

A different way to proceed is to find conditions such that the same two-wage structure is optimal regardless of the number of actions. These conditions must reduce the many-action case to the two-action case. The exercise parallels what is done to obtain monotonicity (in output) of the optimal contract in the standard model, and uses analogous assumptions.

**Assumption A-3:** For each action  $a$ , every selection in

$$\{\pi^e : \mathcal{A} \rightarrow \Delta(\mathcal{N}) \mid \pi^e(a) \text{ is an extreme point of } \Pi(a) \text{ for each } a\}$$

satisfies the monotone likelihood ratio property and concavity of the distribution function. That is, if  $\pi^e : \mathcal{A} \rightarrow \Delta(\mathcal{N})$  is a selection such that  $\pi^e(a)$  is an extreme point of  $\Pi(a)$  for each  $a$ , then

- (i) (MLR) for any two actions  $a', a$ ,  $c(a) > c(a')$  implies that  $\frac{\pi_i^e(a')}{\pi_i^e(a)}$  is decreasing in  $i$
- (ii) (CDFC) for any three actions  $a'', a', a$  such that  $c(a) = \lambda c(a') + (1 - \lambda)c(a'')$  for some  $\lambda \in [0, 1]$ ,

$$\sum_{i=1}^j \pi_i^e(a) \leq \lambda \sum_{i=1}^j \pi_i^e(a') + (1 - \lambda) \sum_{i=1}^j \pi_i^e(a'') \text{ for all } j.$$

**Proposition 5.** *Suppose Assumptions A-1, A-2 and A-3 hold. Then for any action  $a^* > 1$ , the optimal contract to implement  $a^*$  (with inertia) has a two-wage structure.*

*Proof.* Again we prove the result for the case of implementation with inertia; the other case is analogous. Let  $\hat{w}$  be an optimal contract that implements  $a^*$  with inertia. Without loss of generality, label payments so that  $\hat{w}_N$  corresponds to the highest,  $\hat{w}_{N-1}$  the second highest, and so on. We claim there exists only one action  $a' \neq a^*$  for which the incentive compatibility constraint binds, and  $a' < a^*$ . Suppose not. Then, there exist two actions  $a'$  and  $a''$  different from  $a^*$  such that

$$\sum_{j=1}^N \pi_j(a'') \hat{w}_j - c(a'') = \sum_{j=1}^N \pi_j(a') \hat{w}_j - c(a') = \sum_{j=1}^N \pi_j(a^*) \hat{w}_j - c(a^*) = \bar{w}$$

for any  $\pi(a^*) \in \Pi(a^*; \hat{w})$ ,  $\pi(a') \in \Pi(a'; \hat{w})$ , and  $\pi(a'') \in \Pi(a''; \hat{w})$ . By construction  $\pi(a^*)$ ,  $\pi(a')$ , and  $\pi(a'')$  can be taken to be extreme points of the corresponding belief sets. From here on, following the argument in Grossman and Hart (1983) establishes the claim.

If only the constraint relative to one action  $a^*$  binds,  $\hat{w}$  must also be optimal in a problem where all other actions  $a \neq a^*$  are dropped from the constraints. Therefore, Proposition 3 applies to that problem and the optimal contract has a two-wage structure. ■

Interestingly, a sufficient condition to reduce the multi-action case to the two-action case is a generalized version of the requirement one needs in the standard model to obtain monotonicity in output of the optimal contract.

## 5 A Primitive Model

In this section, we consider a primitive version of the parameterized model. While the parametrized distribution model is standard in the contract theory literature, we argue that this primitive model is more natural when considering robustness to uncertainty. We show that the results derived above regarding simple contracts arising due to robustness to uncertainty have simple analogues involving analogous conditions on the primitive model.

Output is determined by a production function which depends on the agent's effort and on stochastic factors. Let  $\mathcal{S}$  be the state space, with elements  $s = 1, \dots, S$ . For each action  $a \in \mathcal{A}$ ,  $y(a) : \mathcal{S} \rightarrow \mathbf{R}$  denotes the stochastic output that results from the action  $a$ . For each action  $a, a'$ , we assume  $\text{supp } y(a) = \text{supp } y(a') := \{y_1, \dots, y_N\}$  (different output levels across different actions correspond to different permutations of the state space).

If the output is  $y$ , the principal's payoff is  $y - w$ , where  $w$  denotes the wage paid to the agent. The agent's payoff is  $w - c(a)$ , where as above  $c : \mathcal{A} \rightarrow \mathbf{R}$  denotes the agent's disutility of effort function. The wage received by the agent will depend on output  $y$  alone and not on the state. Since the agent knows which action he has taken, he also knows which state has occurred. The principal, on the other hand, does not know the realized state of the world since he can only observe realized output. We assume that the agent's beliefs over  $S$  are given by a closed and convex set  $\Pi \subset \Delta(\mathcal{S})$ , and that the principal's belief is  $\pi^P \in \Delta(\mathcal{S})$ .

Our notions of implementation can be reformulated as follows.

**Definition 3.** An incentive scheme  $w$  implements  $a^*$  if:

$$\sum_{s=1}^S \pi_s w(y_s(a^*)) - c(a^*) \geq \bar{w} \quad \forall \pi \in \Pi \quad (\text{P})$$

and for each  $a \in \mathcal{A}$  with  $a \neq a^*$

$$\sum_{s=1}^S \pi_s w(y_s(a^*)) - c(a^*) \geq \sum_{s=1}^S \pi_s w(y_s(a)) - c(a) \quad \forall \pi \in \Pi \quad (\text{IC})$$

**Definition 4.** An incentive scheme  $w$  implements  $a^*$  with inertia in  $\mathcal{A}$  if

$$\sum_{s=1}^S \pi_s w(y_s(a^*)) - c(a^*) \geq \bar{w} \quad \forall \pi \in \Pi \quad (\text{P})$$

and for each  $a \in \mathcal{A}$  with  $a \neq a^*$ , either

$$\sum_{s=1}^S \pi_s w(y_s(a^*)) - c(a^*) \geq \sum_{s=1}^S \pi_s w(y_s(a)) - c(a) \quad \forall \pi \in \Pi \quad (\text{IC})$$

or

$$\sum_{s=1}^S \pi_s w(y_s(a)) - c(a) \leq \bar{w} \text{ for some } \pi \in \Pi \quad (\text{NC})$$

To relate this model to the parameterized distribution model, given  $\pi \in \Pi$  and  $a \in \mathcal{A}$ , define  $\pi(a) \in \Delta(\mathcal{N})$  by

$$\pi_i(a) = \sum_{\{s \in \mathcal{S} : y_s(a) = y_i\}} \pi_s$$

For each  $a$ , set  $\Pi(a) := \{\hat{\pi} \in \Delta(\mathcal{N}) : \hat{\pi} = \pi(a) \text{ for some } \pi \in \Pi\}$ . Notice that  $\Pi(a)$  is closed and convex for each  $a$ . This means assumptions about the set  $\Pi$  can easily be translated into assumptions

about  $\Pi(a)$ . In particular, we show that Assumptions A-1 and A-2 can be derived from analogous assumptions on  $\Pi$  (the beliefs over the primitive state space).

For each  $E \subset \{1, \dots, N\}$  and each  $a$ , let  $E(a) := \{s \in \mathcal{S} : y_s(a) \in E\}$ . Then notice that  $\hat{\pi} \in \Pi(a) \iff \exists \pi \in \Pi$  such that  $\hat{\pi}(E) = \pi(E(a))$  for each  $E \subset \{1, \dots, N\}$ .

**Proposition 6.** *Let  $\nu : 2^{\mathcal{S}} \rightarrow [0, 1]$  be a convex capacity on  $\mathcal{S}$  and let  $\Pi$  be the core of  $\nu$ . For each  $a \in \mathcal{A}$  define  $\nu^a : 2^{\mathcal{N}} \rightarrow [0, 1]$  by  $\nu^a(E) = \nu(E(a))$  for each  $E \subset \{1, \dots, N\}$ . Then for each  $a \in \mathcal{A}$ ,  $\nu^a$  is a convex capacity on  $\{1, \dots, N\}$  with core  $\Pi(a)$ .*

*Proof.* Fix  $a \in \mathcal{A}$ . Let  $E, F \subset \{1, \dots, N\}$ . Then

$$\begin{aligned} \nu^a(E) + \nu^a(F) &= \nu(E(a)) + \nu(F(a)) \\ &\leq \nu(E(a) \cap F(a)) + \nu(E(a) \cup F(a)) \\ &= \nu((E \cap F)(a)) + \nu((E \cup F)(a)) \\ &= \nu^a(E \cap F) + \nu^a(E \cup F) \end{aligned}$$

Thus  $\nu^a$  is a convex capacity. Moreover, if  $\hat{\pi} \in \Pi(a)$ , then  $\hat{\pi} = \pi(a)$  for some  $\pi \in \Pi$ . For each  $\hat{\pi} \in \Pi(a)$ ,  $\hat{\pi}(E) = \pi(E(a)) \geq \nu(E(a)) = \nu^a(E)$ . Thus  $\Pi(a)$  is a subset of the core of  $\nu^a$ . Since  $\Pi(a)$  is closed and convex, it suffices to show that  $\Pi(a)$  contains all of the ‘‘marginal contribution’’ vectors for  $\nu^a$ , that is, any vector  $\pi$  of the form  $\pi_j = \nu^a(\{\sigma(1), \dots, \sigma(j)\}) - \nu^a(\{\sigma(1), \dots, \sigma(j-1)\})$  where  $\sigma$  is a permutation on  $\{1, \dots, N\}$ . Without loss of generality, consider the identity permutation and corresponding vector  $\pi$  in which  $\pi_j = \nu^a(\{1, \dots, j\}) - \nu^a(\{1, \dots, j-1\})$  for each  $j$ . Thus for each  $j$ ,  $\pi_j = \nu(\{y_1, \dots, y_j\}(a)) - \nu(\{y_1, \dots, y_{j-1}\}(a))$ . For each  $j$ , set  $\{s_1^j, \dots, s_{k_j}^j\} := \{y_j\}(a) = \{y_1, \dots, y_j\}(a) \setminus \{y_1, \dots, y_{j-1}\}(a)$ . Define  $\bar{\pi} \in \Delta(\mathcal{S})$  as follows. Set

$$\bar{\pi}(s_1^1) = \nu(\{s_1^1\})$$

and for each  $k \neq 2, \dots, k_1$ , set

$$\bar{\pi}(s_k^1) = \nu(\{s_1^1, \dots, s_k^1\}) - \nu(\{s_1^1, \dots, s_{k-1}^1\})$$

For  $j = 2, \dots, N$ , similarly define

$$\pi(s_1^j) = \nu(\{y_1, \dots, y_{j-1}\}(a) \cup \{s_1^j\}) - \nu(\{y_1, \dots, y_{j-1}\}(a))$$

and for  $k = 2, \dots, k_j$ ,

$$\bar{\pi}(s_k^j) = \nu(\{y_1, \dots, y_{j-1}\}(a) \cup \{s_1^j, \dots, s_k^j\}) - \nu(\{y_1, \dots, y_{j-1}\}(a) \cup \{s_1^j, \dots, s_{k-1}^j\})$$

Then  $\bar{\pi}$  is an element of the core of  $\nu$ , since it is the marginal contribution vector corresponding to the permutation  $(s_1^1, \dots, s_{k_1}^1, \dots, s_1^N, \dots, s_{k_N}^N)$ . Thus  $\bar{\pi} \in \Pi$ . By construction, for each  $j$ ,

$$\bar{\pi}(\{y_j\}(a)) = \sum_{k=1}^{k_j} \bar{\pi}(s_k^j) = \nu(\{y_1, \dots, y_j\}(a)) - \nu(\{y_1, \dots, y_{j-1}\}(a)) = \pi_j$$

Thus  $\pi \in \Pi(a)$ , and the claim is established. ■

We say that a subset  $B \subset \mathcal{S}$  is *unambiguous* if  $\pi(B) = \pi'(B)$  for all  $\pi, \pi' \in \Pi$ . While the empty set and the entire state space  $\mathcal{S}$  are always unambiguous, every nonempty, proper subset of  $\mathcal{S}$  is ambiguous under the assumption that  $\Pi$  has nonempty relative interior. Here we assume that  $\Pi$  has nonempty relative interior, and that the principal's unique prior  $\pi^P \in \Delta(\mathcal{S})$  is an element of  $\text{rel int } \Pi$ . Under these assumptions, together with the assumption that the support of stochastic output  $y(a)$  is the same for all actions  $a$ , no action eliminates ambiguity. That is,  $\Pi(a)$  also has a nonempty relative interior, and  $\pi^P(a) \in \text{rel int } \Pi(a)$  for each action  $a$ .

**Proposition 7.** *If  $\pi^P \in \text{rel int } \Pi$ , then  $\pi^P(a) \in \text{rel int } \Pi(a)$  for each action  $a$ .*

*Proof.* Fix an action  $a$ . For each  $j = 1, \dots, N$ ,  $\{y_j\}(a) \subset \mathcal{S}$  is a nonempty and proper subset of  $\mathcal{S}$ , since  $y(a)$  has support  $\{y_1, \dots, y_N\}$ . Thus for a nonempty, proper subset  $E \subset \{1, \dots, N\}$ ,  $E(a) \subset \mathcal{S}$  must also be nonempty and proper. By assumption,

$$\min_{\pi \in \Pi} \pi(E(a)) < \pi^P(E(a)) < \max_{\pi \in \Pi} \pi(E(a))$$

From this we conclude

$$\min_{\pi \in \Pi(a)} \pi(E) < \pi_j^P(a)(E) < \max_{\pi \in \Pi(a)} \pi(E)$$

Thus the claim is established. ■

These two results allow us to identify assumptions in this primitive model that will suffice to ensure that optimal contracts robust to uncertainty will be simple, as in the previous section.

**Assumption A-4:**  $\Pi$  is the core of a convex capacity  $\nu$  on  $\mathcal{S}$ .

**Assumption A-5:**  $\pi^P$  is an element of the relative interior of  $\Pi$ .

**Proposition 8.** *Suppose Assumptions A-4 and A-5 hold, and the high action  $H$  is implementable (implementable with inertia) in  $\{L, H\}$ . Then, the unique optimal incentive scheme to implement  $H$  (with inertia) divides the  $\mathcal{S}$  states into two subsets and is constant on each subset.*

## 6 Appendix

### 6.1 Proof of Proposition 2

We prove each statement separately.

*Proof of (i):* Suppose not. Then  $\widehat{w}(a)$  is optimal and  $\sum_{j=1}^N \pi_j(a) \widehat{w}_j(a) - c(a) > \bar{w}$ . Reduce the payment in each state by  $\varepsilon > 0$ ; that is, for all  $j$ , let  $\tilde{w}_j = \widehat{w}_j^a - \varepsilon$ . For  $\varepsilon$  small enough,  $\tilde{w}_j$  implements  $a$  according to both definitions because it satisfies (P), (IC), and (NC). Furthermore,  $\sum_{j=1}^N \pi_j^P(a) \tilde{w}_j < \sum_{j=1}^N \pi_j^P(a) \widehat{w}_j(a) - \varepsilon$ , contradicting the optimality of  $\widehat{w}(a)$ .

*Proof of (ii):* Because  $\pi^P(1) \in \Pi(1)$ , for any payment scheme  $w$  and any  $\pi(1) \in \underline{\Pi}(1; w)$ , by definition  $\sum_{j=1}^N \pi_j(1) w_j \leq \sum_{j=1}^N \pi_j^P(1) w_j$ . The constant contract  $\widehat{w}_j = \bar{w} + c^1$  for all  $j$  is feasible: it satisfies (P), and it satisfies (IC) because alternative actions are more costly for the agent. It is a solution to (2) because  $\bar{w} + c^1 = \sum_{j=1}^N \pi_j(1) \widehat{w}_j(1) = \sum_{j=1}^N \pi_j^P(1) \widehat{w}_j(1)$ .

*Proof of (iii):* Suppose the claim does not hold. That is,  $\widehat{w}(a)$  is optimal and for each  $a' < a$ , either

$$\sum_{j=1}^N \pi_j(a) \widehat{w}_j(a) - c(a) > \sum_{j=1}^N \bar{\pi}_j(a') \widehat{w}_j(a) - c(a')$$

for all  $\pi(a) \in \underline{\Pi}(a; \widehat{w})$  and  $\bar{\pi}(a') \in \bar{\Pi}(a'; \widehat{w})$ , or

$$\sum_{j=1}^N \pi_j(a') \widehat{w}_j(a) - c(a') < \bar{w}$$

for some  $\pi(a') \in \Pi(a')$ . The latter inequality implies  $\sum_{j=1}^N \pi_j(a') \widehat{w}_j(a) - c(a') < \bar{w}$  for any  $\pi(a') \in \underline{\Pi}(a'; \widehat{w})$ . Because none of the respective constraints binds,  $\widehat{w}(a)$  is a solution for a problem like (2) where all actions like  $a'$  have been dropped from the constraints. In this new problem,  $a$  is the least costly action and a contract that pays  $\bar{w} + c(a)$  in all states is optimal. That is,  $\widehat{w}_j = \bar{w} + c(a)$  for each  $j$ . Thus,  $\bar{w} + c(a) = \sum_{j=1}^N \pi_j(a) \widehat{w}_j(a) = \sum_{j=1}^N \bar{\pi}_j(a') \widehat{w}_j(a) = \sum_{j=1}^N \pi_j(a') \widehat{w}_j(a)$ . Hence  $c(a') > c(a)$ , contradicting  $a' < a$ . ■

### 6.2 Proof of the Main Result

In the proof, we use the following results adapted from Chateauneuf and Jaffray (1989).

**Lemma 1.** *Let  $v$  be a convex capacity on  $S$  with core  $\Pi$ , and let  $\Pi^e$  be the set of extreme points of*

II. Then, for any  $N$ -vector  $z$  such that  $z_1 \leq z_2 \leq \dots \leq z_N$ ,

$$\min_{\pi \in \Pi} \sum_{j=1}^N \pi_j z_j$$

is attained at  $\pi \in \Pi^e$  such that  $\sum_{s=j}^N \pi_s = \sum_{s=j}^N v(\{s\})$  for each  $j = 2, \dots, N$ . In particular, if  $z$  and  $z'$  are two  $N$ -vectors such that  $z_1 \leq \dots \leq z_N$  and  $z'_1 \leq \dots \leq z'_N$ , then

$$\arg \min_{\pi \in \Pi} \sum_{j=1}^N \pi_j z_j = \arg \min_{\pi \in \Pi} \sum_{j=1}^N \pi_j z'_j$$

The first result follows from Chateauneuf and Jaffray (1989), Propositions 10 and 13. The second follows from the first, which shows that the set of minimizing distributions depends only on the order of the elements of  $z$  and not on the values of the components.

When Assumption A-1 holds, the lemma provides a characterization of the probability distributions that, among a set, yield the smallest and largest expected values of a fixed contract. In particular, this lemma can be used to show that for each action  $a$ , the smallest expected value for a given  $z$  is attained at an extreme point of  $\Pi(a)$  such that  $\pi_K(a) < \pi_K^P(a)$  and  $\pi_1(a) \geq \pi_1^P(a)$ , where  $K$  is an index at which the maximum value of  $z_k$  is attained. Moreover, when two incentive schemes are “ordered” in the same way, their minimum or maximum expectations are attained using the same distributions.

### 6.2.1 Proof of Proposition 3

The main step in the proof is to show, by contradiction, that an optimal contract cannot be contingent on more than two subsets of output levels.

Define a partition of the state space such that each element in this partition corresponds to different payments in the optimal contract  $\hat{w}$ . Let  $k$  be an element of this partition; we say  $k$  is an event. Any probability distribution over the original space defines a probability distribution over this partition. Label events so that  $K$  corresponds to the largest  $\hat{w}_k$ ,  $K - 1$  to the second largest, and so on until event 1 corresponds to the smallest value of  $\hat{w}_k$ . By construction,  $K$  is the number of events the optimal contract is contingent upon, and  $\hat{w}_1 < \hat{w}_2 < \dots < \hat{w}_K$ . We need to show  $K$  equals 2.

By Proposition 2,  $\widehat{w}$  must satisfy

$$\sum_{k=1}^K \underline{\pi}_k(H) \widehat{w}_k = \bar{w} + c(H) \quad (5)$$

for any  $\underline{\pi}(H) \in \underline{\Pi}(H; \widehat{w})$ . We claim  $\widehat{w}$  must also satisfy

$$\sum_{k=1}^K \underline{\pi}_k(L) \widehat{w}_k = \bar{w} + c(L) \quad (6)$$

for any  $\underline{\pi}(L) \in \underline{\Pi}(L; \widehat{w})$ . Suppose not. Then,  $\sum_{k=1}^K \underline{\pi}_k(L) \widehat{w}_k < \bar{w} + c(L)$ . Let

$$\widetilde{w} = \left( \widehat{w}_1 - \varepsilon \frac{\underline{\pi}_K(H)}{\underline{\pi}_1(H)}, \widehat{w}_2, \dots, \widehat{w}_{K-1}, \widehat{w}_K + \varepsilon \right)$$

where  $\varepsilon$  is positive and small enough so that the ranking of the payments for  $\widetilde{w}$  and  $\widehat{w}$  is the same. By Lemma 1,  $\underline{\pi}(H)$  and  $\underline{\pi}(L)$  minimize the expected value of  $\widetilde{w}$  for the agent. For any distribution  $\pi$ , the expected values of  $\widetilde{w}$  and  $\widehat{w}$  are related by the following:

$$\sum_{k=1}^K \pi_k \widetilde{w}_k - \sum_{k=1}^K \pi_k \widehat{w}_k = \frac{\pi_1 \underline{\pi}_K(H) - \pi_K \underline{\pi}_1(H)}{\underline{\pi}_1(H)} \varepsilon \quad (7)$$

$\pi = \underline{\pi}(H)$  implies the right hand side of (7) is equal to 0; thus,  $\widetilde{w}$  satisfies (5). For some  $\varepsilon$  close enough to zero and  $\pi = \underline{\pi}(L)$  the right hand side of (7) is very small; thus  $\widetilde{w}$  satisfies (NC) because  $\widehat{w}$  satisfies it strictly. By Lemma 1 and Assumption A-1,  $\underline{\pi}_K(H) < \pi_K^P(H)$  and  $\underline{\pi}_1(H) \geq \pi_1^P(H)$ ; thus, the right hand side of (7) is negative when  $\pi = \pi^P(H)$ . Summarizing,  $\widetilde{w}$  is feasible and cheaper than  $\widehat{w}$ , contradicting the optimality of  $\widehat{w}$ . Hence (6) must hold for  $\widehat{w}$  to be an optimum.

We claim  $K$  must be strictly larger than 1. Suppose not, i.e.,  $K = 1$ . If this is the case, all payments are the same, and the left hand sides of equations (5) and (6) are the same. Thus, we have  $\bar{w} + c(L) = \bar{w} + c(H)$ , contradicting  $c(H) > c(L)$ .

We claim  $K$  is not larger than 2. Suppose not. Then  $\widehat{w}$  is optimal, so satisfies equations (5) and (6), and  $K > 2$ . Equations (5) and (6) constitute a system of two equations which can be solved for  $\widehat{w}_K$  and some  $\widehat{w}_{k'}$ , yielding:

$$\widehat{w}_K = \frac{\underline{\pi}_{k'}(L) (\bar{w} + c(H)) - \underline{\pi}_{k'}(H) (\bar{w} + c(L))}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} + \sum_{k \neq K, k'} \frac{\underline{\pi}_{k'}(H) \underline{\pi}_k(L) - \underline{\pi}_k(H) \underline{\pi}_{k'}(L)}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} \widehat{w}_k \quad (8)$$

$$\widehat{w}_{k'} = \frac{\pi_K(H)(\bar{w} + c(L)) - \pi_K(L)(\bar{w} + c(H))}{\pi_{k'}(L)\pi_K(H) - \pi_K(L)\pi_{k'}(H)} + \sum_{k \neq K, k'} \frac{\pi_K(L)\pi_k(H) - \pi_K(H)\pi_k(L)}{\pi_{k'}(L)\pi_K(H) - \pi_K(L)\pi_{k'}(H)} \widehat{w}_k \quad (9)$$

These are well defined unless

$$\pi_K(L)\pi_{k'}(H) - \pi_{k'}(L)\pi_K(H) = 0 \text{ for all } k' \neq K \quad (10)$$

In that case:

$$\begin{aligned} 0 &= \pi_K(L) \sum_{k=1}^{K-1} \pi_k(H) - \pi_K(H) \sum_{k=1}^{K-1} \pi_k(L) \\ &= \pi_K(L)(1 - \pi_K(H)) - \pi_K(H)(1 - \pi_K(L)) \\ &= \pi_K(L) - \pi_K(H) \\ &\Rightarrow \pi_K(H) = \pi_K(L) \end{aligned}$$

Using this result in (10):

$$\pi_K(H)(\pi_{k'}(H) - \pi_{k'}(L)) = 0$$

Thus, either  $\pi_K(H) = \pi_K(L) = 0$ , or  $\pi_{k'}(H) = \pi_{k'}(L)$  for all  $k' \neq K$ . If  $\pi_K(H) = \pi_K(L) = 0$ ,  $\widehat{w}$  cannot be optimal because it makes the largest payment in a state that does not affect the constraints and, by Assumption A-2, has positive probability for the principal. If  $\pi_{k'}(H) = \pi_{k'}(L)$  for all  $k' \neq K$ , then  $\pi(L) = \pi(H)$ . In this case,  $\sum_{k=1}^K \pi_k(H)\widehat{w}_k = \sum_{k=1}^K \pi_k(L)\widehat{w}_k$  and  $c(H) = c(L)$ , a contradiction.

Using equations (8) and (9), we can write the expected cost of the optimal incentive scheme as follows:

$$\sum_{k=1}^K \pi_k^P(H)\widehat{w}_k = \widehat{\alpha} + \sum_{k \neq K, k'} \widehat{\beta}_k \widehat{w}_k$$

where

$$\begin{aligned} \widehat{\alpha} &= \frac{\pi_K^P(H)\pi_{k'}(L) - \pi_{k'}^P(H)\pi_K(L) + \pi_{k'}^P(H)\pi_K(H) - \pi_K^P(H)\pi_{k'}(H)}{\pi_{k'}(L)\pi_K(H) - \pi_K(L)\pi_{k'}(H)} \bar{w} \\ &+ \frac{\pi_K^P(H)\pi_{k'}(L) - \pi_{k'}^P(H)\pi_K(L)}{\pi_{k'}(L)\pi_K(H) - \pi_K(L)\pi_{k'}(H)} c(H) \\ &+ \frac{\pi_{k'}^P(H)\pi_K(H) - \pi_K^P(H)\pi_{k'}(H)}{\pi_{k'}(L)\pi_K(H) - \pi_K(L)\pi_{k'}(H)} c(L) \end{aligned}$$

and

$$\widehat{\beta}_k = \pi_k^P(H) - \underline{\pi}_k(H) \frac{\pi_K^P(H) \underline{\pi}_{k'}(L) - \pi_{k'}^P(H) \underline{\pi}_K(L)}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} - \underline{\pi}_k(L) \frac{\pi_{k'}^P(H) \underline{\pi}_K(H) - \pi_K^P(H) \underline{\pi}_{k'}(H)}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} \quad (11)$$

We claim that  $\widehat{\beta}_{k''} \neq 0$  for some  $k'' \neq K, k'$ . Suppose not. Then,  $\widehat{\beta}_k = 0$  for each  $k \neq K, k'$ . Using equation (11),

$$\begin{aligned} \pi_k^P(H) (\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)) &= \underline{\pi}_k(H) (\pi_K^P(H) \underline{\pi}_{k'}(L) - \pi_{k'}^P(H) \underline{\pi}_K(L)) \\ &\quad + \underline{\pi}_k(L) (\pi_{k'}^P(H) \underline{\pi}_K(H) - \pi_K^P(H) \underline{\pi}_{k'}(H)) \end{aligned}$$

Summing over  $k$ , rearranging, and solving for  $\pi_{k'}^P(H)$ :

$$\pi_{k'}^P(H) = - \frac{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H) + \pi_K^P(H) (\underline{\pi}_{k'}(H) - \underline{\pi}_{k'}(L))}{\underline{\pi}_K(L) - \underline{\pi}_K(H)}$$

This implies:

$$\frac{\pi_K^P(H) \underline{\pi}_{k'}(L) - \pi_{k'}^P(H) \underline{\pi}_K(L)}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} = \frac{\underline{\pi}_K(L) - \pi_K^P(H)}{\underline{\pi}_K(L) - \underline{\pi}_K(H)}$$

and

$$\frac{\pi_{k'}^P(H) \underline{\pi}_K(H) - \pi_K^P(H) \underline{\pi}_{k'}(H)}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} = \frac{\pi_K^P(H) - \underline{\pi}_K(H)}{\underline{\pi}_K(L) - \underline{\pi}_K(H)}$$

We know that  $\underline{\pi}_K(H) < \pi_K^P(H)$ . Thus,  $\frac{\underline{\pi}_K(L) - \pi_K^P(H)}{\underline{\pi}_K(L) - \underline{\pi}_K(H)} < 1$  and  $\frac{\pi_K^P(H) - \underline{\pi}_K(H)}{\underline{\pi}_K(L) - \underline{\pi}_K(H)} > 0$ . Hence:

$$\begin{aligned} \sum_{k=1}^K \pi_k^P(H) \widehat{w}_k &= \widehat{\alpha} = \overline{w} + c(H) + \frac{\pi_K^P(H) - \underline{\pi}_K(H)}{\underline{\pi}_K(L) - \underline{\pi}_K(H)} (c(L) - c(H)) \\ &< \overline{w} + c(H) \end{aligned}$$

a contradiction.

Because  $\widehat{\beta}_{k''} \neq 0$  for some  $k'' \neq K, k'$ , we find a feasible contract which is cheaper than  $\widehat{w}$ . Let  $\widetilde{w}$  be defined as follows:

$$\begin{aligned} \widetilde{w}_k &= \widehat{w}_k && \text{when } k \neq K, k', k'' \\ \widetilde{w}_K &= \widehat{w}_K + \frac{\underline{\pi}_{k'}(H) \underline{\pi}_{k''}(L) - \underline{\pi}_{k''}(H) \underline{\pi}_{k'}(L)}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} \varepsilon \\ \widetilde{w}_{k'} &= \widehat{w}_{k'} + \frac{\underline{\pi}_K(L) \underline{\pi}_{k''}(H) - \underline{\pi}_K(H) \underline{\pi}_{k''}(L)}{\underline{\pi}_{k'}(L) \underline{\pi}_K(H) - \underline{\pi}_K(L) \underline{\pi}_{k'}(H)} \varepsilon \\ \widetilde{w}_{k''} &= \widehat{w}_{k''} + \varepsilon \end{aligned}$$

where

$$|\varepsilon| < \min \{ \widehat{w}_{k''} - \widehat{w}_{k''-1}, \widehat{w}_{k''+1} - \widehat{w}_{k''}, \widehat{w}_K - \widehat{w}_{K-1}, \widehat{w}_{k'} - \widehat{w}_{k'-1}, \widehat{w}_{k'+1} - \widehat{w}_{k'} \}$$

By construction, the payments in  $\widetilde{w}$  and  $\widehat{w}$  are ranked in the same order. Lemma 1 applies, and  $\underline{\pi}(H)$  and  $\underline{\pi}(L)$  yield the minimum expected values of  $\widetilde{w}$  under actions  $H$  and  $L$ . Moreover,

$$\begin{aligned} 0 &= \underline{\pi}_K(L) \frac{\underline{\pi}_{k'}(H)\underline{\pi}_{k''}(L) - \underline{\pi}_{k''}(H)\underline{\pi}_{k'}(L)}{\underline{\pi}_{k'}(L)\underline{\pi}_K(H) - \underline{\pi}_K(L)\underline{\pi}_{k'}(H)} + \underline{\pi}_{k'}(L) \frac{\underline{\pi}_K(L)\underline{\pi}_{k''}(H) - \underline{\pi}_K(H)\underline{\pi}_{k''}(L)}{\underline{\pi}_{k'}(L)\underline{\pi}_K(H) - \underline{\pi}_K(L)\underline{\pi}_{k'}(H)} + \underline{\pi}_{k''}(L) \\ 0 &= \underline{\pi}_K(L) \frac{\underline{\pi}_{k'}(H)\underline{\pi}_{k''}(L) - \underline{\pi}_{k''}(H)\underline{\pi}_{k'}(L)}{\underline{\pi}_{k'}(L)\underline{\pi}_K(H) - \underline{\pi}_K(L)\underline{\pi}_{k'}(H)} + \underline{\pi}_{k'}(L) \frac{\underline{\pi}_K(L)\underline{\pi}_{k''}(H) - \underline{\pi}_K(H)\underline{\pi}_{k''}(L)}{\underline{\pi}_{k'}(L)\underline{\pi}_K(H) - \underline{\pi}_K(L)\underline{\pi}_{k'}(H)} + \underline{\pi}_{k''}(L) \end{aligned}$$

Hence,  $\widetilde{w}$  is feasible because  $\widehat{w}$  is. The expected cost of  $\widetilde{w}$  is given by

$$\begin{aligned} \sum_{k=1}^K \pi_k^P(H) \widetilde{w}_k &= \widehat{\alpha} + \sum_{k \neq K, k'} \widehat{\beta}_k \widetilde{w}_k \\ &= \widehat{\alpha} + \sum_{k \neq K, k'} \widehat{\beta}_k \widehat{w}_k + \widehat{\beta}_{k''} \varepsilon \end{aligned}$$

Thus, we can choose  $\varepsilon > 0$  whenever  $\widehat{\beta}_{k''} < 0$  and  $\varepsilon < 0$  whenever  $\widehat{\beta}_{k''} > 0$ . In either case,  $\widetilde{w}$  is feasible and cheaper than  $\widehat{w}$ , contradicting the optimality of  $\widehat{w}$ . Summarizing, if an optimal contract is contingent on  $K > 2$  events, we can find a feasible contract which is cheaper. Therefore, because we already proved  $K < 2$  is impossible, a contract can be optimal only if  $K = 2$ . ■

## References

- AMERIKS, J., AND S. P. ZELDES (2000): “How Do Household Portfolio Shares Vary with Age?,” Discussion paper, Columbia University.
- AUMANN, R. J. (1962): “Utility Theory without the Completeness Axiom,” *Econometrica*, 30, 445–462.
- BEWLEY, T. F. (1986): “Knightian Decision Theory: Part I,” Discussion paper, Cowles Foundation.
- (1987): “Knightian Decision Theory: Part II,” Discussion paper, Cowles Foundation.
- CHATEAUNEUF, A., AND J.-Y. JAFFRAY (1989): “Some Characterizations of Lower Probabilities and other Monotone Capacities through the Use of Möbius Inversion,” *Mathematical Social Sciences*, 17, 263–283.
- EINHORN, H. J., AND R. M. HOGARTH (1985): “Ambiguity and Uncertainty in Probabilistic Inference,” *Psychological Review*, 92, 433–461.
- FOX, C. R., AND A. TVERSKY (1995): “Ambiguity Aversion and Comparative Ignorance,” *Quarterly Journal of Economics*, 110, 586–603.
- GHIRARDATO, P. (1994): “Agency Theory with Uncertainty Aversion,” Discussion paper, Caltech.
- GROSSMAN, S. J., AND O. D. HART (1983): “An Analysis of the Principal-Agent Problem,” *Econometrica*, 51, 7–45.
- HART, O. D. (1995): *Firms, Contracts, and Financial Structure*. Oxford: Oxford University Press.
- HART, O. D., AND B. HOLMSTROM (1987): “The Theory of Contracts,” in *Advances in Economic Theory*, ed. by T. F. Bewley.
- HEATH, C., AND A. TVERSKY (1991): “Preference and Belief: Ambiguity and Competence in Choice Under Uncertainty,” *Journal of Risk and Uncertainty*, 4, 5–28.
- HOLMSTROM, B. (1987): “Moral Hazard and Observability,” *Bell Journal of Economics*, 10, 74–91.

- HOLMSTROM, B., AND P. MILGROM (1987): “Aggregation and Linearity in the Provision of Intertemporal Incentives,” *Econometrica*, 55, 303–328.
- KNIGHT, F. H. (1921): *Uncertainty and Profit*. Boston: Houghton Mifflin.
- LOPOMO, G., L. RIGOTTI, AND C. SHANNON (2009): “Uncertainty in Mechanism Design,” *mimeo*.
- MADRIAN, B. C., AND D. F. SHEA (2001): “The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior,” *Quarterly Journal of Economics*, 107, 1149–1187.
- MOORE, J., AND O. D. HART (1988): “Incomplete Contracts and Renegotiation,” *Econometrica*, 56, 755–785.
- MUKERJI, S. (1998a): “Ambiguity and Contractual Forms,” Discussion paper, Yale University.
- (1998b): “Ambiguity Aversion and Incompleteness of Contractual Form,” *American Economic Review*, 88, 1207–1231.
- RIGOTTI, L., AND C. SHANNON (2005): “Uncertainty and Risk in Financial Markets,” *Econometrica*, 73, 203–243.
- SAMUELSON, W., AND R. ZECKHAUSER (1988): “Status Quo Bias in Decision Making,” *Journal of Risk and Uncertainty*, 1, 7–59.
- SCHMEIDLER, D. (1989): “Subjective Probability and Expected Utility without Additivity,” *Econometrica*, 57(3), 571–587.
- TIROLE, J. (1999): “Incomplete Contracts: Where Do We Stand?,” *Econometrica*, 67, 741–781.
- VON NEUMANN, J., AND O. MORGENSTERN (1953): *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.