Agency Theory with Maxmin Expected Utility Players

Edi Karni*

Johns Hopkins University

November 18, 2004

Abstract

This paper extends the analysis of incentive schemes, designed to mitigate the welfare loss associated with moral hazard, to the case in which the principal and the agent are maxmin expected utility players. Invoking the parametrized distribution approach to agency theory, the paper examines the axiomatic foundations of the principal’s and agent’s choice behaviors that are representable by the maximization of the minimum expected utility over action-dependent sets of priors. In the context of this model, the paper also discusses some implications of uncertainty aversion for the design of optimal incentive schemes.

*I am grateful to the National Science Foundation for financial support under grant SES-0314249.
1 Introduction

In recent years uncertainty aversion, displayed by a pattern of choice first noted by Ellsberg (1961), has become, a focal issue in the theory of decision making under uncertainty. This phenomenon is incompatible with representation of decision makers’ beliefs by additive (subjective) probability measures. Axiomatic models of decision making under uncertainty designed to accommodate uncertainty aversion include the Choquet expected utility model of Schmeidler (1989) and the maxmin expected utility model of Gilboa and Schmeidler (1989).

If uncertainty aversion is an important aspect of individual decision-making under uncertainty then, like risk aversion, it should manifest itself in shaping economic institutions and practices. In this paper I take a step toward incorporating uncertainty aversion into the theory of incentive contracts and the study of its implications.

By and large, the analysis of the principal-agent relationship in the presence of moral hazard is based on the assumption that the two parties are expected-utility maximizers. In particular, the parametrized distribution approach, pioneered by Mirrlees (1974, 1976), envisions a principal designing an incentive contract intended to induce an agent to choose an action that would maximize the principal’s expected utility over a class of action-dependent probability distributions on outcomes. Given the contract, the agent is supposed to choose the action that would maximize his own expected utility over the same class of action-dependent probability distributions on outcomes.
In view of the main concern of this work, I begin by developing axiomatic models of uncertainty-averse principals and agents, whose preferences are represented by maxmin expected utility over action-dependent sets of priors. These models are based on a synthesis, developed in Karni (2004a), of the maxmin expected utility model of Gilboa and Schmeidler (1989) and the analytical framework of Karni (2004). The analytical framework dispenses with the notion of states of nature which, for reasons explained in Karni (2004), I regard as unsatisfactory, introducing instead a choice space whose elements are action-bet pairs. Actions correspond to means by which agents may influence the likely realizations of different outcomes; bets represent outcome-contingent payoffs. The resulting representations are subjective versions of the parametrized distributions approach, lending themselves naturally to the investigation of incentive contracts in the presence of moral hazard.

Despite obvious similarities, the approach taken here is different from existing models in several important respects. First, unlike Mirrlees (1974, 1976), who assumes that the family of action-dependent distributions is given, I derive the relevant family of action-dependent (sets of) distributions from the principal’s and agent’s preferences. Second, the representations are different from Gilboa and Schmeidler (1989) in that, through their selection of actions, decision makers can choose among sets of distributions on outcomes. Moreover, the preference relation on the contingent payoff schemes, representing the contracts, may be outcome dependent.

To render this discussion more precise, I denote the set of actions by $\mathcal{A}$ and the set of contracts by $W$. Elements of $W$ are real-valued functions on a set, $\Theta$, of outcomes. Assume
that the principal and the agent have preference relations on the set, $A \times W$, of action-contract pairs that are denoted by $\succ^P$ and $\succ^A$, respectively. The principal’s problem may be stated as follows: Choose an action-contract pair $(a', w')$ such that

$$(a', w') \succ^P (a, w) \text{ for all } (a, w) \in A \times W$$

subject to the incentive compatibility constraints

$$(a', w') \succ^A (a, w') \text{ for all } a \in A$$

and the participation constraint

$$(a', w') \succ^A (a^0, w^0),$$

where $(a^0, w^0)$ denotes the agent’s “outside option” (that is, an action-contract pair equivalent to the best course of action available to the agent if he rejects the contract).

If the principal and the agent are maxmin expected utility maximizers and the agent’s utility function is additively separable in actions and payoffs, then the problem may be stated as follows: Choose an action-contract pair $(a', w')$,

$$(a', w') \in \arg \max_{A \times W} \left\{ \min_{\pi \in C^P(a)} \sum_{\theta \in \Theta} u^P (w (\theta); \theta) \pi (\theta) \right\}$$

subject to the incentive compatibility constraints, for all $a \in A$,

$$\min_{\pi \in C^A(a')} \sum_{\theta \in \Theta} u^A (w' (\theta), \theta) \pi (\theta) + v (a') \geq \min_{\pi \in C^A(a)} \sum_{\theta \in \Theta} u^A (w (\theta), \theta) \pi (\theta) + v (a),$$

and the participation constraint

$$\min_{\pi \in C^A(a')} \sum_{\theta \in \Theta} u^A (w' (\theta), \theta) \pi (\theta) + v (a') \geq \min_{\pi \in C^A(a^0)} \sum_{\theta \in \Theta} u^A (w^0 (\theta), \theta) \pi (\theta) + v (a^0),$$
where \( \{ u^P (\cdot; \theta) \mid \theta \in \Theta \} \) and \( \{ u^A (\cdot; \theta) \mid \theta \in \Theta \} \) are outcome-dependent utility functions of the principal and the agent, respectively; \( \{ C^P (a) \mid a \in \mathcal{A} \} \) and \( \{ C^A (a) \mid a \in \mathcal{A} \} \) are their respective sets of priors on \( \Theta \); and \( \nu \) denotes the agent’s disutility of actions.

In Karni (2004a) I axiomatized the principal’s choice behavior. Here, I extend this analysis to include the agent’s preferences. I assume that the payoffs of the bets are roulette lotteries (or lotteries, for short) in the sense of Anscombe and Aumann (1963). The key notion invoked in the axiomatization of the principal’s preferences is that of constant-valuation bets, that is, bets whose utility payoffs are the same across outcomes (see Karni [2004a]). Loosely speaking, constant-valuation bets are defined by the behavioral implication that, given such a bet, the decision-maker is indifferent among all actions. In the case of the agent, because the choice of action has a direct effect on his well-being, constant valuation bets may no longer be defined in this manner. Another approach is called for that permits the separation of the direct effect of the actions on the agent’s preferences from the indirect effect (that is, the impact of the actions on the distributions of outcomes).

In the following section I introduce the analytical framework. The principal’s preferences and their representation appear in Section 3. The agent’s preferences and their representations are taken up in Section 4. Additional results regarding the parties preferences, and the corresponding representations, are discussed in Section 5. In Section 6 I discuss issues pertaining to the application of the maxmin expected utility model to agency theory, and examine some welfare implications, within a simple model. Section 7 contains a brief summary and a discussion of related literature.
2 The Analytical Framework

2.1 Preliminaries

Let $\Theta$ be a finite set whose elements are referred to as outcomes, and let $\mathcal{A}$ be a set whose elements, called actions, describe activities by which a decision maker may influence the likely realization of different outcomes. Let $I(\theta)$ be an interval in $\mathbb{R}$ representing monetary prizes that are feasible if the outcome $\theta \in \Theta$ obtains. Denote by $\Delta(I(\theta))$ the set of all simple probability distributions (that is, distributions that have finite support) on $I(\theta)$. Elements of $\Delta(I(\theta))$ are referred to as lotteries. A bet, $b$, is a function on $\Theta$ such that $b(\theta) \in \Delta(I(\theta))$. Denote by $B$ the set of all bets (that is, $B := \prod_{\theta \in \Theta} \Delta(I(\theta))$). The choice set is the product set $\mathcal{C} := \mathcal{A} \times B$ whose generic element, $(a, b)$, is an action-bet pair.

Decision makers are characterized by preference relations, $\succ$, on $\mathcal{C}$ that have the usual interpretation.\footnote{A preference relation, $\succ$, is a binary relation on $\mathcal{C}$, and $(a, b) \succ (a', b')$ have the interpretation $(a, b)$ is at least as desirable as $(a', b')$.} In other words, decision makers are supposed to be able to choose among, or express preferences over, action-bet pairs, presumably taking into account the influence their choice of action may have on the likely realization of alternative outcomes and, consequently, on the desirability of the corresponding bets. The strict preference relation, $\succ$, and the indifference relation, $\sim$, are defined as usual. Henceforth I denote by $\succ^P$ and $\succ^A$ the preference relations of the principal and the agent, respectively.
For all $b, b' \in B$ and $\alpha \in [0, 1]$, define $\alpha b + (1 - \alpha) b' \in B$ by: $(\alpha b + (1 - \alpha) b')(\theta) = \alpha b(\theta) + (1 - \alpha) b'(\theta)$, for all $\theta \in \Theta$. I use the notation $b_{\theta \leftarrow p}$ to denote the bet that is obtained from the bet $b$ by replacing its $\theta-$coordinate with the lottery $p$.

Given $\succcurlyeq$, an outcome $\theta \in \Theta$ is null given the action $a$ if $(a, b_{\theta \leftarrow p}) \sim (a, b_{\theta \leftarrow q})$ for all $p, q \in \Delta(I(\theta))$, and $b \in B$, otherwise it is nonnull given the action $a$. In general, given $\succcurlyeq$, an outcome may be null under some actions and nonnull under others. Denote by $\Theta(a; \succcurlyeq)$ the subset of outcomes that are nonnull given $a$ and denote by $\Theta^c(a; \succcurlyeq)$ its complement in $\Theta$. It is customary in agency theory to suppose that the set of outcomes constituting the support of the distributions is invariable with respect to the actions. Formally,

**The full support assumption:** For all $a \in A$, $\Theta(a; \succcurlyeq) = \Theta$.

### 2.2 Axioms that apply to both the principal and the agent

In general, the principal and the agent are motivated by distinct considerations. In some applications the agent’s well-being may not be affected directly by the outcome; in most applications the principal’s well-being of is not directly affected by the agent’s actions. For example, in employer-employee relationships, the employer’s well-being is directly affected by the firm’s profits; while profits affect the employee’s well-being only to the extent that they determines his payoff under the employment contract. The employer’s sole concern is the action’s effect on the likely realization of profits, while the employee’s well-being is directly affected by the action (e.g., effort). Despite their different concerns, certain principles,
expressed by the following axioms, govern the choice behavior of the principal and the agent alike.\(^2\)

\[(A.1) \textbf{ (Weak order)} \triangleright \text{ is complete and transitive.}\]

\[(A.2) \textbf{ (Archimedean)} \text{ For all } a \in A \text{ and } b,b',b'' \in B, \text{ such that } (a, b) \triangleright (a, b') \triangleright (a, b''), \text{ there exist } \alpha, \beta \in (0, 1) \text{ satisfying } (a, \alpha b + (1 - \alpha) b') \triangleright (a, b') \triangleright (a, \beta b + (1 - \beta) b'').\]

\[(A.3) \textbf{ (Uncertainty Aversion)} \text{ For all } a \in A, b, b' \in B, \text{ and } \alpha \in (0, 1), \text{ if } (a, b) \sim (a, b') \text{ then } (a, \alpha b + (1 - \alpha) b') \succ (a, b).\]

\section{The Principal’s Preferences and their Representation}

\subsection{Preliminaries}

Given a preference relation, \(\succ^P\), constant valuation bets are bets whose contingent lottery-payoffs compensate for the variations in the principal’s well-being due to the outcomes themselves.\(^3\) Formally, let \(\mathcal{I}(b; a) = \{b' \in B \mid (a, b') \sim^P (a, b)\}\) and \(\mathcal{I}(p; \theta, b, a) = \{q \in \Delta(I(\theta)) \mid \)

\(^2\)This issue is taken up again in Section 5.

\(^3\)The concept of constant valuation bets was first defined in Karni (2004). In the context of the state space formulation, Drèze’s (1987) notion of “omnipotent” acts is analogous. Similar ideas appear in Karni (1993) and in Skiadas (1997).
\[(a, b_{-\theta q}) \sim^P (a, b_{-\theta p})\}. \quad \mathcal{I}(b; a) \text{ denotes the indifference class of } b \text{ given } a \text{ and, similarly, } \mathcal{I}(p; \theta, b, a) \text{ denotes the indifference class of } p \text{ given } \theta, b \text{ and } a.

**Definition 1:** A bet \(\bar{b} \in B\) is a constant valuation bet according to \(\succ\) if \((a, \bar{b}) \sim (a', \bar{b})\) for all \(a, a' \in \mathcal{A}\) and, for all \(b \in \cap_{a \in \mathcal{A}} \mathcal{I}(\bar{b}; a)\), \(b(\theta) \in \mathcal{I}(\bar{b}(\theta); \theta, \bar{b}, a)\) for all \(\theta \in \Theta\) and \(a \in \mathcal{A}\).

Definition 1 implies that the set of actions is sufficiently rich that there exist no additional actions that, if added to the set \(\mathcal{A}\), would reduce the set of bets that are indifferent to \(\bar{b}\) conditional on every action in \(\mathcal{A}\). Let \(B_{cv}^{\succ^P}\) denote the set of all constant valuation bets according to \(\succ^P\). In view of the definition of constant valuation bets, if \(\bar{b}', \bar{b} \in B_{cv}^{\succ^P}\) I write \(\bar{b}' \succ^P \bar{b}\) instead of \((a, \bar{b}') \succ^P (a, \bar{b})\).

To simplify the exposition, I assume that the choice set has a maximal and minimal elements whose bet components are constant valuation. Formally,

\[(A.0) \quad \text{There are } \bar{b}^{**}, \bar{b}^* \in B_{cv}^{\succ^P} \text{ such that } \bar{b}^{**} \succ^P (a, b) \succ^P \bar{b}^* \text{ for all } (a, b) \in \mathcal{C}, \text{ and } \bar{b}^{**} \succ^P \bar{b}^*.\]

### 3.2 Axioms

Stating the next axiom requires additional notation and definitions. Given \(p \in \Delta(I(\theta))\) denote by \(\bar{b}_{-\theta p}\) the constant valuation bet whose \(\theta-\) coordinate is \(p\). Define a preference
relation, $\succeq_\theta^P$ on $\Delta (I(\theta))$ by $p \succeq_\theta^P q$ if and only if $\overline{b - \theta p} \succ^P \overline{b - \theta q}$, for all $p, q \in \Delta (I(\theta))$. The symmetric and asymmetric parts of $\succeq_\theta^P$ are denoted by $\succ_\theta$ and $\sim_\theta$, respectively.

(A.4) **(Monotonicity)** For all $a \in A$, $b, b' \in B$, if $b(\theta) \succeq_\theta^P b'(\theta)$ for all $\theta \in \Theta$ then $(a, b) \succ^P (a, b')$.

The following axiom is analogous to the certainty-independence axiom of Gilboa and Schmeidler (1989). It requires that the independence axiom of expected utility theory applies when mixing bets with constant valuation bets.

(A.5) **(Constant-Valuation Independence)** For all $a \in A$, $b, b' \in B$, $\overline{b} \in B^{cv}$, and $\alpha \in (0, 1)$, $(a, b) \succ^P (a, b')$ if and only if $(a, \alpha \overline{b} + (1 - \alpha) \overline{b'}) \succ^P (a, \alpha b' + (1 - \alpha) \overline{b})$.

The next axiom - the only condition that has no analogue in Gilboa and Schmeidler - requires that if $\overline{b}, \overline{b'} \in B^{cv}$ then $(a, \alpha \overline{b} + (1 - \alpha) \overline{b'}) \sim^P (a', \alpha \overline{b} + (1 - \alpha) \overline{b'})$ for all $\alpha \in (0, 1)$. The intuition underlying this condition is that what the decision maker ultimately cares about are the consequences of his actions, that is, the outcome, $\theta$, that obtains and the payoff, $b(\theta)$, associated with it. Thus facing a choice between $(a, \alpha \overline{b} + (1 - \alpha) \overline{b'})$ and $(a', \alpha \overline{b} + (1 - \alpha) \overline{b'})$, the decision maker recognizes that the consequences are identical under the two actions. Whether he selects $a$ or $a'$, he is awarded a composite lottery $\alpha \overline{b} + (1 - \alpha) \overline{b'}$, whose consequences are $(\alpha \overline{b}(\theta) + (1 - \alpha) \overline{b'}(\theta), \theta)$, $\theta \in \Theta$. Because $\overline{b}$ and $\overline{b'}$ are constant-valuation bets, these consequences are equally preferred regardless of the outcome that obtains. Formally,
(A.6) **Convexity** For all \(a, a' \in \mathcal{A}, \bar{b}, \bar{b}' \in B^{cv},\) and \(\alpha \in (0, 1),\) \((a, \alpha \bar{b} + (1 - \alpha) \bar{b}') \sim^{P} (a', \alpha \bar{b} + (1 - \alpha) \bar{b}')\).

Axiom (A.6) renders the set of constant-valuation bets convex. Because what constitutes a constant valuation bet is a personal (subjective) matter, the convexity of the set \(B^{cv}\) is not implied by the structure of the choice space.\(^4\)

### 3.3 Representation

The representation of the principal’s preferences is analogous to the maxmin expected utility representation of Gilboa and Schmeidler (1989). The statement of the representation theorem involves the following terminology. A family \(\{f_\alpha\}\) of real-valued functions that figures in the representation of a preference relation is said to be **cardinally measurable and fully comparable** if, for the purpose of representation, it is equivalent to another set of functions \(\{g_\alpha\},\) where \(g_\alpha = a + cf_\alpha, c > 0\) for all \(\alpha.\)

**Theorem 1** (Karni [2004a]) Let \(\succeq^{P}\) be a binary relation on \(\mathbb{C}.\) Assume that (A.0) and the full support assumption hold. Then the following conditions are equivalent:

(i) \(\succeq^{P}\) satisfies (A.1)–(A.6).

\(^4\)Gilboa and Schmeidler (1989) implicitly assume that constant acts are constant-valuation acts. Because the set of constant functions is convex, the restriction imposed by Axiom (A.6) below has no bite.
(ii) There exist affine functions \( \{u^P(\cdot, \theta) : \Delta(I(\theta)) \to \mathbb{R} \}_{\theta \in \Theta} \), and a family of closed and convex sets, \( \{C^P(a)\}_{a \in A} \), of probability measures on \( \Theta \) such that, for all \((a, b), (a', b') \in C\),

\[
(a, b) \succ^P (a', b') \iff \min_{\pi \in C^P(a)} \sum_{\theta \in \Theta} u^P(b(\theta), \theta) \pi(\theta) \geq \min_{\pi \in C^P(a')} \sum_{\theta \in \Theta} u^P(b'(\theta), \theta) \pi(\theta).
\]

Moreover, the functions \( \{u^P(\cdot, \theta)\}_{\theta \in \Theta} \) are cardinally measurable and fully comparable, for every \( a \in A \), the set \( C^P(a) \) is unique, and each \( \pi \in C^P(a) \) has full support.

4 The Agent’s Preferences and their Representation

4.1 Preliminaries

Like the principal, the agent is characterized by a preference relation, \( \succ^A \), on \( C \). However, unlike the principal, for the agent the actions are costly, making it impossible to identify constant valuation bets my the method used in Definition 1.

In many situations the agent’s well-being is not directly affected by the outcomes. In such situations it is natural to suppose that, insofar as the agent is concerned, constant bets are constant valuation bets. The question is, what is the choice-theoretic expression of this supposition? In other words, what patterns of choice would validate, or invalidate, this supposition?
4.2 Axioms and Representations

Assume that \( I(\theta) = I \), for all \( \theta \in \Theta \), and denote by \( B^c \) the subset of constant bets (that is, constant functions) in \( B \). To simplify the notation I denote by \( r \) the constant bet whose value is \( r \).

The following axiom is the monotonicity axiom of Gilboa and Schmeidler (1989). Let \( b_c(p) \) be the constant bet whose payoff is \( p \) for every \( \theta \in \Theta \). Define a preference relation, \( \succeq^\Delta \) on \( \Delta(I(\theta)) \) by \( p \succeq^\Delta_q \) if and only if \( b_c(p) \succeq^A b'_c(q) \), for all \( p,q \in \Delta(I(\theta)) \).

(A.4') (Monotonicity) For all \( a \in \mathcal{A}, b,b' \in B, r \in B^c, \) and \( \alpha \in (0,1) \) \( (a,b) \succ^A (a,b') \) if and only if \( (a,\alpha b + (1-\alpha) r) \succ^A (a,\alpha b' + (1-\alpha) r) \).

The next axiom, due to Gilboa and Schmeidler (1989), plays a role analogous to that of constant-valuation independence in the preceding section.

(A.7) (Certainty Independence) For all \( a \in \mathcal{A}, b,b' \in B, r \in B^c, \) and \( \alpha \in (0,1) \) \( (a,b) \succ^A (a,b') \) if and only if \( (a,\alpha b + (1-\alpha) r) \succ^A (a,\alpha b' + (1-\alpha) r) \).

Theorem 2 below gives general representation of a minmax expected utility agent with action-dependent sets of priors and action-dependent utility. Unlike the representation of the principal’s preferences in Theorem 1, the representation of the agent’s preferences does not invoke the notion of constant valuation bets.

The statements of the theorems below require the following additional definitions.
Definition 2: Given $\succ^A$ on $\mathcal{A} \times \mathcal{B}$, the set $\mathcal{A}$ is **essential** if there exist $a, a' \in \mathcal{A}$ and $b \in \mathcal{B}$ such that $(a, b) \succ^A (a', b)$. Similarly, $\mathcal{B}$ is **essential** if there exist $a \in \mathcal{A}$ and $b, b' \in \mathcal{B}$ such that $(a, b) \succ^A (a, b')$.

Definition 3: Given $\succ^A$, two actions, $a$ and $a'$, are **directly linked** if there exist $r, r', \hat{r}, \hat{r}' \in \mathcal{B}^c$, $\hat{r} > r$ and $\hat{r}' > r'$ such that $(a, r) \sim^A (a', r')$ and $(a, \hat{r}) \sim^A (a', \hat{r}')$. Two actions, $a''$ and $a'''$, are **linked** if there is a sequence of actions $a'' = a_1, \ldots, a_k = a'''$ such that $a_i$ and $a_{i+1}, i = 1, \ldots, k - 1$ are directly linked. The set $\mathcal{A}$ is **linked** if all its elements are linked.

**Theorem 2** Let $\succ^A$ be a binary relation on $\mathcal{C}$. Assume that the full support assumption holds and the set $\mathcal{A}$ is linked. Then the following conditions are equivalent:

(i) $\succ^A$ satisfies (A.1)–(A.3), (A.4’) and (A.7).

(ii) There exist a family, $\{U^A (\cdot; a) : \Delta (I) \to \mathbb{R} \mid a \in \mathcal{A}\}$, of affine functions and a family, $\{C^A (a)\}_{a \in \mathcal{A}}$, of closed and convex sets of probability measures on $\Theta$ such that, for all $(a, b), (a', b') \in \mathcal{C}$,

$$(a, b) \succ^A (a', b') \iff \min_{\pi \in C^A (a)} \sum_{\theta \in \Theta} U^A (b (\theta), a) \pi (\theta) \geq \min_{\pi \in C^A (a')} \sum_{\theta \in \Theta} U^A (b' (\theta), a') \pi (\theta).$$

Moreover, the functions $\{U^A (\cdot, a)\}_{a \in \mathcal{A}}$ are cardinally measurable and fully comparable, for every $a \in \mathcal{A}$, the set $C^A (a)$ is unique if and only if the set $\mathcal{B}$ is essential, and each $\pi \in C^A (a)$ has full support.
**Proof.** (i) $\rightarrow$ (ii). For each $a \in \mathcal{A}$, let $\succsim^A_a$ denote the agent’s action-dependent preferences on $B$, defined by, $b \succsim^A_a b'$ if and only if $(a, b) \succsim^A (a, b')$. Then (A.1)–(A.3), (A.4’) and (A.7) imply that, for each $a \in \mathcal{A}$, $\succsim^A_a$ satisfies axioms (A.1)–(A.5) of Gilboa and Schmeidler (1989). Hence their Theorem 1 implies that, for each $a \in \mathcal{A}$, there exist affine function, $U^A(\cdot, a) : \Delta(I) \rightarrow \mathbb{R}$, unique up to positive linear transformation, and a non-empty, closed and convex set, $C^A(a)$, of probability measures on $\Theta$, such that $\succsim^A_a$ on $B$ is represented by

$$\min_{\pi \in C(a)} \sum_{\theta \in \Theta} U^A(b(\theta), a) \pi(\theta).$$

Moreover, $C^A(a)$ is unique if and only if $B$ is essential. (The essentiality of $B$ implies that $\succsim^A_a$ satisfies Gilboa and Schmeidler’s non-degeneracy axiom.)

Fix $a, a' \in \mathcal{A}$ that are directly linked. Let $\bar{r}(a), \overline{r}(a), \bar{r}(a')$, $\overline{r}(a') \in B^c$ be such that $U^A(\bar{r}(a), a) > U^A(\overline{r}(a), a)$, $(a, \bar{r}(a)) \sim^A (a', \bar{r}(a'))$, and $(a, \overline{r}(a)) \sim^A (a', \overline{r}(a'))$. Invoking the uniqueness of the functions $U(\cdot, a), a \in \mathcal{A}$, normalize $U^A(\cdot, a)$ and $U^A(\cdot, a')$ so that $U^A(\bar{r}(a), a) = U^A(\bar{r}(a'), a') = 1$ and $U^A(\overline{r}(a), a) = U^A(\overline{r}(a'), a') = 0$. Because $\mathcal{A}$ is linked, this normalization may be extended to link all the action-dependent utility functions $U^A(\cdot, a), a \in \mathcal{A}$. Then, for all $(a, r), (a', r') \in \mathcal{A} \times B^c$,

$$\quad (a, r) \succsim^A (a', r') \iff U^A(r, a) \geq U^A(r', a'). \quad (7)$$

For all $(a, b) \in C$, let $r(b, a) \in B^c$ satisfy $(a, b) \sim^A (a, r(b, a))$ (that is, $r(b, a)$ is a constant-bet equivalent of $b$ given $a$). That such $r(b, a)$ exist follows from (A.3). But

$$\sum_{\theta \in \Theta} U^A(r(b, a), a) \pi(\theta) = U^A(r(b, a), a) \quad \text{for all } \pi \in C(a).$$

Thus, by the representation,

$$\quad (a, b) \sim^A (a, r(b, a)) \iff \min_{\pi \in C^A(a)} \sum_{\theta \in \Theta} U^A(b(\theta), a) \pi(\theta) = U^A(r(b, a), a). \quad (8)$$

15
Equations (7) and (8) and axiom (A.1) imply that, for all \((a, b), (a', b') \in \mathbb{C}\),

\[
(a, b) \succ^{A} (a', b') \iff \min_{\pi \in C^{A}(a)} \sum_{\theta \in \Theta} U^{A}(b(\theta), a) \pi(\theta) \geq \min_{\pi \in C(a')} \sum_{\theta \in \Theta} U^{A}(b'(\theta), a') \pi(\theta).
\]  

(9)

Hence \((i) \rightarrow (ii)\). The proof that \((ii) \rightarrow (i)\) is straightforward. The uniqueness of \(U^{A}(\cdot, a)\) and of \(C^{A}(a)\) follows from the Gilboa and Schmeidler (1969). That the fact that each \(\pi \in C^{A}(a)\) has full support is an implication of the full support assumption.

\[
\ast \quad \ast
\]

The action-dependent utility in Theorem 2 is rather general. Consider a more restrictive model in which the \(U^{A}(\cdot, a) = \lambda^{A}(a) u^{A}(\cdot) + \kappa^{A}(a), \lambda^{A}(a) > 0\). This type of action-dependent utility figures in Grossman and Hart (1983). The axiomatic underpinning of this utility function is given by the next axiom, which is analogous to the sure thing principle of Savage (1954).\(^5\)

\[(A.8) \text{ (Coordinate independence)} \quad \text{For all } a, a' \in \mathcal{A} \text{ and } r, r' \in B^{c}, (a, r) \succ^{A} (a', r') \text{ if and only if } (a', r) \succ^{A} (a', r') \text{ and } (a, r) \succ^{A} (a', r) \text{ if and only if } (a, r') \succ^{A} (a', r'). \]

**Theorem 3** Let \(\succ^{A}\) be a binary relation on \(\mathbb{C}\). Assume that the full support assumption holds, and the set \(\mathcal{A}\) is linked. Then the following conditions are equivalent:

\[(i) \succ^{A} \text{ satisfies (A.1)--(A.3), (A.4'), (A.7) and (A.8)}.\]

\(^5\)See Wakker (1989) for a detailed discussion of coordinate independence.
(ii) There exists an affine function \( u^A(\cdot): \Delta(I) \to \mathbb{R} \), \( \lambda^A \in \mathbb{R}^{\vert A \vert}_{++}, \kappa^A \in \mathbb{R}^{\vert A \vert} \), and a family of closed and convex sets, \( \{C^A(a)\}_{a \in A} \), of probability measures on \( \Theta \) such that, for all \( (a,b),(a',b') \in \mathcal{C}, (a,b) \succ^A (a',b') \) if and only if

\[
\min_{\pi \in C^A(a)} \sum_{\theta \in \Theta} \lambda^A(a) u^A(b(\theta)) \pi(\theta) + \kappa^A(a) \geq \min_{\pi \in C^A(a')} \sum_{\theta \in \Theta} \lambda^A(a') u^A(b(\theta)) \pi(\theta) + \kappa^A(a').
\]

Moreover, the functions \( \{u^A(\cdot)\}_{a \in A} \) are cardinally measurable and fully comparable, for each \( a \in A \), the set \( C^A(a) \) is unique if and only if \( B \) is essential, and each \( \pi \in C^A(a) \) has full support.

**Proof.** \((i) \to (ii)\). By Theorem 2 there exist affine functions, \( U^A(\cdot,a): \Delta(I) \to \mathbb{R}, a \in A \), unique up to positive linear transformation, and non-empty, closed and convex set, \( C^A(a) \), of probability measures on \( \Theta \) such that \( \succ^A \) on \( B \) is represented by \( \min_{\pi \in C(a)} \sum_{\theta \in \Theta} U^A(b(\theta),a) \pi(\theta) \).

For all \( r \in B^c \) and \( a \in A \), \( \min_{\pi \in C(a)} \sum_{\theta \in \Theta} U^A(r,a) \pi(\theta) = U^A(r,a) \). By coordinate independence, \( (A.8) \), for all \( r, r' \in B^c \) and \( a, a' \in A \), \( U^A(r,a) \geq U^A(r',a) \) if and only if \( U^A(r,a') \geq U^A(r',a') \). Hence, by the von Neumann-Morgenstern theorem, \( U^A(\cdot,a) \) and \( U^A(\cdot,a') \) are linear transformations of one another.

Fix \( a \in A \) and define \( U^A(\cdot,a) := u^A(\cdot) \). Note that \( u^A \) is unique up to positive linear transformation. Let \( \bar{r}, \bar{r} \in B^c \) such that \( (a,\bar{r}) \succ^A (a,\bar{r}) \). Invoking the uniqueness of \( u^A \) and \( U^A(\cdot,a) \), set \( u^A(\bar{r}) = 1 \) and \( u^A(\bar{r}) = 0 \). For all \( a' \in A \) such that \( a' \) and \( a \) are directly linked, let \( \lambda^A(a') \) and \( \kappa^A(a') \) be the solution to the equations

\[
\lambda^A(a') u^A(\bar{r}) + \kappa^A(a') = 1
\]
\[ \lambda^A(a') u^A(r') + \kappa^A(a') = 0, \]

where \((a, r) \sim^A (a', r')\) and \((a, r) \sim^A (a', r')\). (That is, \(\lambda^A(a') = \left[ u^A(r') - u^A(r') \right]^{-1} > 0 \)

and \(\kappa^A(a') = -u^A(r') / \left[ u^A(r') - u^A(r') \right]\). Since \(\mathcal{A}\) is linked this way of parameterizing the utility functions can be extended to include all actions. Then, for all \((a, r)\), \((a', r')\) \(\in \mathcal{A} \times \mathcal{B}^c\),

\[(a, r) \succ^A (a', r') \iff \lambda^A(a) u^A(r) + \kappa^A(a) \geq \lambda^A(a') u^A(r') + \kappa^A(a'). \quad (10)\]

Hence, by Theorem 2, for all \((a, b)\), \((a', b')\) \(\in \mathcal{C}\), \((a, b) \succ^A (a', b')\) if and only if

\[
\min_{\pi \in \mathcal{C}^A(a)} \lambda^A(a) \sum_{\theta \in \Theta} u^A(b(\theta)) \pi(\theta) + \kappa^A(a) \geq \min_{\pi \in \mathcal{C}(a')} \lambda^A(a') \sum_{\theta \in \Theta} u^A(b'(\theta)) \pi(\theta) + \kappa^A(a') \quad (11)
\]

That \((ii) \rightarrow (i)\) is immediate. The uniqueness of the follow from Theorem 2 and the normalization. 

\[\star \quad \star\]

In many applications it is assumed that the agent’s utility function is additively separable over payoff and actions (see Shavel [1979] and Holmstrom [1979]). Next I introduce a well-known axiom that entails a separately additive representation of the agent’s preferences over actions and bets. Notice, however, that, like coordinate independence, it holds only on the Cartesian product of actions and constant-valuation bets.

\[(A.9) \quad \textbf{(Hexagon condition)} \quad \text{For all } a, a', a'' \in \mathcal{A} \text{ and } r, r', r'' \in \mathcal{B}^c, \text{ if } (a, r') \sim^A (a', r) \text{ and } (a, r'') \sim^A (a', r') \sim^A (a'', r) \text{ then } (a', r'') \sim^A (a'', r').\]
An array of real-valued functions \((v_s)_{s \in S}\) is said to be a \textit{jointly cardinal additive representation} of a binary relation \(\succeq\) on a product set \(D = \Pi_{s \in S} D_s\) if, for all \(d, d' \in D\), \(d \succeq d'\) if and only if \(\sum_{s \in S} v_s(d_s) \geq \sum_{s \in S} v_s(d'_s)\), and the class of all functions that constitute an additive representation of \(\succeq\) consists of those arrays of functions, \((\hat{v}_s)_{s \in S}\), for which \(\hat{v}_s = \lambda v_s + \zeta_s\), \(\lambda > 0\) for all \(s \in S\). The functions \((v_s)_{s \in S}\) that have the uniqueness property cited above are \textit{jointly cardinal}.

**Theorem 4** Let \(\succ^A\) be a binary relation on \(C\). Assume that the full support assumption holds and the set \(A\) is linked. Then the following conditions are equivalent:

(i) \(\succ^A\) satisfies \((A.1)-(A.3), (A.4'), (A.7)-(A.9),\) and \(A\) and \(B\) are essential.

(ii) There exist an affine function \(u^A : \Delta(I) \to \mathbb{R}\), a function \(v^A : A \to \mathbb{R}\), and a family of closed and convex sets, \(\{C^A(a)\}_{a \in A}\), of probability measures on \(\Theta\) such that, for all \((a, b), (a', b') \in C\), \((a, b) \succ^A (a', b')\) if and only if

\[
\min_{\pi \in C^A(a)} \sum_{\theta \in \Theta} u^A(b(\theta)) \pi(\theta) + v^A(a) \geq \min_{\pi \in C^A(a')} \sum_{\theta \in \Theta} u^A(b'(\theta)) \pi(\theta) + v^A(a').
\]

Moreover, the functions \(u^A\) and \(v^A\) are jointly cardinal, \(v^A\) is non-constant if and only if \(A\) is essential and, for each \(a \in A\), the set \(C^A(a)\) is unique if and only if \(B\) is essential and each \(\pi \in C^A(a)\) has full support.

**Proof.** \((i) \rightarrow (ii)\). Consider the restriction of \(\succ^A\) to \(A \times B^c\). By \((A.1)\) and \((A.2)\) it is a continuous weak order. Moreover, it satisfies coordinate independence, the hexagon condition,
and both $\mathcal{A}$ and $B$ are essential. Thus, by Theorem III.4.1 of Wakker (1989) and Theorem 3, there exist jointly cardinal, continuous, additive-valued functions $u^A : B^c \to \mathbb{R}$ and $v^A : \mathcal{A} \to \mathbb{R}$, where $u^A$ is affine and $v^A$ is non-constant, such that $\succeq^A$ on $\mathcal{A} \times B^c$ is represented by $U^A(r,a) = u^A(r) + v^A(a)$. (If $\mathcal{A}$ is not essential, that is, $(a,r) \sim^A (a',r)$ for all $r \in B^c$ and $a,a' \in \mathcal{A}$, then $U^A(r,a) = U^A(r,a') = u^A(r)$ and $v^A$ is a constant function.)

Fix $\bar{r} \in B^c$ and $a \in \mathcal{A}$. For all $a' \in \mathcal{A}$ such that $a$ and $a'$ are linked, let $\bar{r}(a') \in B^c$ satisfy $(a,\bar{r}) \sim^A (a',\bar{r}(a'))$. Then, invoking the uniqueness of $u^A$ and $U^A(\cdot,a)$, set $u^A(\bar{r}) = 1$ and, for all $a' \in \mathcal{A}$, let $v^A(a')$ be the solution of $u^A(\bar{r}') + v^A(a') = 1$. (That is, $v^A(a') = u^A(\bar{r}) - u^A(\bar{r}')$.) Then, for all $(a,r),(a',r') \in \mathcal{A} \times B^c$,

\[(a,r) \succeq^A (a',r') \Leftrightarrow u^A(r) + v^A(a) \geq u^A(r') + v^A(a') .\]  

Hence, by Theorem 3, for all $(a,b),(a',b') \in C$,

\[(a,b) \succeq^A (a',b') \Leftrightarrow \min_{\pi \in C^A(a)} \sum_{\theta \in \Theta} u^A(b(\theta)) \pi(\theta) + v^A(a) \geq \min_{\pi \in C^A(a')} \sum_{\theta \in \Theta} u^A(b'(\theta)) \pi(\theta) + v^A(a') .\]  

Hence $(i) \rightarrow (ii)$. The proof that $(ii) \rightarrow (i)$ is straightforward.

The uniqueness of $u^A$ and $v^A$ is an implication of the jointly cardinal additive in Theorem III.4.1 of Wakker (1989). The uniqueness of $C^A(a)$ follows from the uniqueness of the set of measures in Gilboa and Schmeidler (1969). The fact that each $\pi \in C^A(a)$ has full support is an implication of the full support assumption. ■
5 Two Additional Corollaries

There are situations, involving moral hazard, in which the agent’s well-being depends on the outcome directly, while the principal’s sole concern is the outcome-contingent payoffs. For example, in the case of health insurance, the agent’s well-being is directly affected by his state of health (that is, the outcome), while that of the principal (that is, the insurer) is affected by the patient’s health only through the associated indemnities. In this case, the objective function of the principal, in Theorem 1, and that of the agent, in Theorems 2, 3, and 4, no longer depicts the parties’ concerns. In this section I modify the theory to accommodate situations in which the principal’s utility function is outcome-independent and that of the agent is outcome-dependent.

5.1 Outcome-independent utility functions of the principal

Consider a principal-agent relationship in which the outcomes and the agent’s payoffs are denominated the same. For example, in sharecropping agreements, the outcomes are levels of output, to be divided between the principal and the agent. Thus the outcome and the agent’s payoff enter the principal’s utility function additively (that is, they are perfect substitutes). Moreover, insofar as the principal is concerned, constant bets are constant-valuation bets (that is, $B_{cv} = B^c$). Since the set of constant-valuation bets is convex, Axiom (A.6) is satisfied. Furthermore, the constant-valuation independence axiom, (A.7), is reduced to Gilboa and Schmeidler’s constant independence axiom and the monotonicity axiom (A.4) is
equivalent to (A.4'). Applying (A.8) and assuming that the set of actions is not essential we obtain the following result:

**Corollary 5** Let $\succ^P$ be a binary relation on $\mathbb{C}$. Assume that the full support assumption holds, the set $\mathcal{A}$ is linked and not essential and $B$ is essential. Then the following conditions are equivalent:

(i) $\succ^P$ satisfies (A.1)–(A.5), (A.7), and (A.8).

(ii) There exist an affine, real-valued, function $u^P(\cdot)$ on $\Delta(I)$ and a family of closed and convex sets, $\{C^P(a)\}_{a \in \mathcal{A}}$, of probability measures on $\Theta$, such that, for all $(a,b), \ (a',b') \in \mathbb{C}$,

$$(a,b) \succ^P (a',b') \iff \min_{\pi \in C^P(a)} \sum_{\theta \in \Theta} u^P(b(\theta)) \pi(\theta) \geq \min_{\pi \in C^P(a')} \sum_{\theta \in \Theta} u^P(b'(\theta)) \pi(\theta).$$

Moreover, the function $u^P(\cdot)$ is unique up to positive linear transformation, for every $a \in \mathcal{A}$, $C^P(a)$ is unique and each $\pi \in C^P(a)$ has full support.

The proof follows from Theorem 3 and the assumption that $\mathcal{A}$ is not essential. Specifically, the fact that $\mathcal{A}$ is not essential and the representation in Theorem 3 imply that, for all $a,a' \in \mathcal{A}$ and $r \in B^c$, $(a,r) \sim^P (a',r)$ if and only if

$$u^P(r) \lambda^P(a) + \kappa^P(a) = u^P(r) \lambda^P(a') + \kappa^P(a').$$

But, if $B$ is essential then $\lambda^P(a) = \lambda^P(a') = \lambda$ and $\kappa^P(a) = \kappa^P(a') = \kappa$ for all $a,a' \in \mathcal{A}$. The representation is obtained after the normalization of $u^P$ so that $\lambda = 1$ and $\kappa = 0$. 

22
5.2 Outcome-dependent utility functions for the agent

Consider again the health insurance example at the beginning of this section. The insurer’s behavior in this case has a representation, as in Corollary 5. However, insofar as the agent is concerned, constant bets are not constant-valuation bets. Thus Theorems 2, 3, and 4, do not apply. Nevertheless, the approach to modeling the agent’s choice behavior that was taken there is still valid.

Let \( B_0 \) be a subset of bets such that the restriction of \( \succeq A \) to the set \( \mathcal{A} \times B' \) satisfies Axioms (A.7)-(A.9) with \( B' \) instead of \( B^c \) and also Axiom (A.6) with \( B' \) instead of \( B^{cv} \). Suppose also that both \( \mathcal{A} \) and \( B' \) are essential. Since \( \succeq A \) is a continuous weak order, Theorem III.4.1 of Wakker (1989) implies that there exist jointly cardinal, continuous, additive valued functions \( u^A : B' \rightarrow \mathbb{R} \) and \( v^A : \mathcal{A} \rightarrow \mathbb{R} \) such that, for all \( \succeq ^A \) on \( \mathcal{A} \times B' \) is represented by \( u^A (b) + v^A (a) \).

Elements of the set \( B' \) are constant-valuation bets. Applying Theorem 1 with \( B' \) instead of \( B^{cv} \), to obtain the following:

**Corollary 6** Let \( \succeq ^A \) be a binary relation on \( C \), and assume that the full support assumption hold, that the set \( \mathcal{A} \) is linked, and both \( \mathcal{A} \) and \( B' \) are essential. Then the following conditions are equivalent:

(i) \( \succeq ^A \) satisfies (A.1)-(A.4), (A.6)-(A.9) with \( B' \) instead of \( B^{cv} \) in (A.6) and instead of \( B^c \) in (A.7)-(A.9).

(ii) There exist jointly cardinal, continuous, functions \( \{ u^A (\cdot ; \theta) : \Delta (I (\theta)) \rightarrow \mathbb{R} \mid \theta \in \Theta \} \),

23
and \( v^A : \mathcal{A} \rightarrow \mathbb{R} \), and a family of closed, convex sets, \( \{C^A(a)\}_{a \in \mathcal{A}} \), of probability measures on \( \Theta \), such that for all \( (a, b), (a', b') \in \mathcal{C} \), \( (a, b) \succ^A (a', b') \) if and only if

\[
\min_{\pi \in C^A(a)} \sum_{\theta \in \Theta} u^A(b(\theta), \theta) \pi(\theta) + v^A(a) \geq \min_{\pi \in C^A(a')} \sum_{\theta \in \Theta} u^A(b'(\theta), \theta) \pi(\theta) + v^A(a').
\]

Moreover, \( v^A \) is non-constant, for every \( a \in \mathcal{A} \) the set \( C^A(a) \) is unique, and each \( \pi \in C^A(a) \) has full support.

The proof follows from Theorems 4 and 1.

### 6 A Simple Application

To explore some rudimentary agency-theory implications of uncertainty aversion, I consider a principal who engages an agent to perform a task. The task must be performed in one of two ways, involving distinct actions, \( a_0 \) and \( a_1 \); either way may result in success or failure. Assume that the agent’s action is his private information and the outcome is publicly observable. If the agent performs the task successfully, the principal stands to gain \( x_s \) dollars; if the agent fails, the principal stands to gain \( x_f \) dollars, where \( x_s > x_f \). Assume that the action \( a_1 \) is more costly to the agent and yields a more favorable prospect of success. Assume also that the principal is risk neutral and the agent is risk averse. To explore the implications of uncertainty aversion for the design of incentive contracts, I consider next alternative attitudes toward uncertainty.
6.1 The benchmark case: Expected utility-maximizing behavior

Consider the standard problem in which both the agent and the principal are uncertainty neutral (that is, they are expected utility maximizers). As is customary in agency theory, assume that the principal and the agent agree on the likelihood of success, conditional on the alternative actions. Denote by $p_s(a_i)$ the probability of success under $a_i$, $i = 0, 1$, and assume that $p_s(a_1) > p_s(a_0)$. Suppose that the principal wants to implement action $a_1$. His problem is to design a contract, $w = (w_s, w_f) \in \mathbb{R}^2$, so as to maximize

$$p_s(a_1)(x_s - w_s) + (1 - p_s(a_1))(x_f - w_f)$$

subject to the incentive compatibility (IC) constraint

$$(p_s(a_1) - p_s(a_0))(u(w_s) - u(w_f)) \geq v(a_1) - v(a_0)$$

and the individual rationality (IR) constraint

$$p_s(a_1)u(w_s) - (1 - p_s(a_1))u(w_f) - v(a_1) \geq u,$$

where $u$, the agent’s utility function, is increasing and strictly concave, $v(a_1) > v(a_0)$, and $u$ is the agent’s utility of his, uncertainty free, “outside option.”

It is well known that in this case both constraints are binding and the solution to the problem is given by\(^6\)

$$u(w^*_s) = u + v(a_0) + \frac{1 - p_s(a_0)}{p_s(a_1) - p_s(a_0)}(v(a_1) - v(a_0))$$

$$u(w^*_f) = u + v(a_0) - \frac{p_s(a_0)}{p_s(a_1) - p_s(a_0)}(v(a_1) - v(a_0)).$$

Clearly, $w^*_s > w^*_f$.

\(^6\)See Salanié (1997).
6.2 Uncertainty neutral principal and uncertainty averse agent

Analogous to the assumption that the principal is risk neutral and the agent is risk averse, I assume next that the principal is uncertainty neutral and the agent uncertainty averse. The principal’s objective function is as in the previous case problem. To describe the constraints he faces, let \( C(a_i) = [\alpha_i, \beta_i], \ i = 0, 1 \), be the agent’s set of probabilities of success under \( a_i \).

In the standard analysis of moral hazard problems, it is customary to suppose that costlier actions yield a higher probability of success. In the present context, the analogous assumption requires that \( \alpha_1 > \alpha_0 \) and \( \beta_1 > \beta_0 \). Furthermore, since the agent’s preferences are represented by minmax expected utility functional with action-dependent sets of priors, while the principal is an expected utility maximizer, the “common priors” assumption must be modified. Assumed that the principal is aware of the agent’s uncertainty aversion, that he knows \( C(a_i), i = 0, 1 \), and that his own conditional probabilities of success satisfy \( p_s(a_i) \in C(a_i), i = 0, 1 \).

As in the benchmark case, suppose that the principal would like to implement the action \( a_1 \). Note that under these assumptions, if \( w = (w_s, w_f) \) is the optimal incentive contract, \( w_s > w_f \). To see this, observe that because the IC constraint is binding, if \( w_s \leq w_f \), then the IC constraint implies that \( (\beta_1 - \beta_0)(u(w_s) - u(w_f)) = v(a_1) - v(a_0) > 0 \). If \( \beta_1 > \beta_0 \) then \( u(w_s) - u(w_f) > 0 \) and, since \( u \) is strictly monotonic increasing, \( w_s > w_f \), a contradiction. If \( w_s \leq w_f \) and \( \beta_1 = \beta_0 \), then the agent cannot be induced to choose \( a_1 \).

Let \( \hat{w} = (\hat{w}_s, \hat{w}_f) \) be the optimal incentive contract designed to induce the agent to choose
the action $a_1$. Suppose that $\alpha_1 > \alpha_0$ (otherwise, it is impossible to implement $a_1$). Because $\hat{w}_s > \hat{w}_f$ the IC constraint is

$$ (\alpha_1 - \alpha_0) (u(\hat{w}_s) - u(\hat{w}_f)) \geq v(a_1) - v(a_0), \quad (15) $$

and the corresponding IR constraint is

$$ \alpha_1 u(\hat{w}_s) + (1 - \alpha_1) u(\hat{w}_f) - v(a_1) \geq u. \quad (16) $$

It is easy to verify that, since the two constraints are binding, $\hat{w}$ is given by

$$ u(\hat{w}_s) = u + v(a_0) + \frac{1 - \alpha_1}{\alpha_1 - \alpha_0} (v(a_1) - v(a_0)) $$

$$ u(\hat{w}_f) = u + v(a_0) - \frac{\alpha_1}{\alpha_1 - \alpha_0} (v(a_1) - v(a_0)). \quad (17) $$

Comparing this solution to the benchmark case, if $p_s(a_1) - p_s(a_0) \geq \alpha_1 - \alpha_0$ then

$$ \frac{1 - p_s(a_1)}{p_s(a_1) - p_s(a_0)} < \frac{1 - \alpha_1}{\alpha_1 - \alpha_0} \quad (18) $$

and

$$ \frac{p_s(a_1)}{p_s(a_1) - p_s(a_0)} \geq \frac{\alpha_1}{\alpha_1 - \alpha_0}. \quad (19) $$

Thus $\hat{w}_s > w^*_s$ and $\hat{w}_f \leq w^*_f$.

(a) If $\hat{w}_f > w^*_f$, then the cost to the principal of implementing $a_1$ when the agent is uncertainty averse is higher than in the benchmark case.

(b) If $\hat{w}_f \leq w^*_f$, then implementation of $a_1$ requires exposing the agent to greater risk. In addition, since $\alpha_1 u^A(w_s) + (1 - \alpha_1) u^A(w_f) < p_s(a_1) u^A(w_s) - (1 - p_s(a_1)) u^A(w_f)$ for all $w$
such that \( w_s > w_f \), the cost to the principal of meeting the individual rationality constraint, necessary to implement \( a_1 \), is higher.

Thus \textit{compared with the benchmark case, when the principal is uncertainty neutral and the agent is uncertainty averse, the implementation of the costly action, \( a_1 \), results in a Pareto inferior allocation.}

A different question concerns the welfare effects of increasing the agent’s uncertainty aversion. One agent is said to display greater uncertainty aversion than another if, for any given action, the set of priors of the former contains that of the latter. Inspection of equations (17) reveals that, ceteris paribus, the effect of greater uncertainty aversion depends on the coefficient \( \alpha_1 / (\alpha_1 - \alpha_0) \) and \( (1 - \alpha_1) / (\alpha_1 - \alpha_0) \). Specifically, increasing uncertainty aversion corresponds to a decline in the values of \( \alpha_i, i = 0, 1 \). Hence if the shift, \( \alpha_1 - \alpha_0 \), due to the change of action, is independent of the agent’s uncertainty aversion, then greater uncertainty aversion implies a larger spread between the payoffs associated with success and failure. Because the agent is risk averse, the mean payoff required to induce the agent to implement \( a_1 \) must increase, and the principal is made worse off. Since the utility of the outside option of the agent remains the same, the overall result is Pareto inferior.

\section{6.3 Uncertainty-averse principal and agent}

Suppose that the principal and the agent are both uncertainty averse and that \( C^P(a_i) = C^A(a_i) := C(a_i), i = 0, 1 \). As in the preceding case, denote by \( \hat{w} = (\hat{w}_s, \hat{w}_f) \) the contract
required to implement $a_1$ (the solution to equation (17) and let $\bar{w} = (\bar{w}, \bar{w})$ be the contract required to implement $a_0$. What are the consequences of the uncertainty aversion of the principal? In particular, is it possible that, ceteris paribus, an uncertainty-neutral principal would choose to implement an action different from that of an uncertainty-averse principal?

Recall that $\hat{w}_s > \hat{w}_f$ and suppose that $(x_s - \hat{w}_s) > (x_f - \hat{w}_f)$. If, in addition, $p_s(a_1) - p_s(a_0) > \alpha_1 - \alpha_0$, it is possible that

$$p_s(a_1)(x_s - \hat{w}_s) + (1 - p_s(a_1))(x_f - \hat{w}_f) > p_s(a_0)x_s + (1 - p_s(a_0))x_f - \bar{w},$$

and

$$\alpha_0x_s + (1 - \alpha_0)x_f - \bar{w} > \alpha_1(x_s - \hat{w}_s) + (1 - \alpha_1)(x_f - \hat{w}_f).$$

In other words, an uncertainty-neutral principal would induce the agent to implement $a_1$ while an uncertainty-averse principal, in the same situation, would prefer that the agent implement $a_0$.

7 Summary and Related Literature

This paper axiomatizes the parametrized distributions approach to agency theory, in which the preferences of the principal and the agent are representable by the maxmin expected utility functionals, displaying uncertainty aversion. The two features that distinguish maxmin expected utility representations in this paper from the original maxmin expected utility representation of Gilboa and Schmeidler (1989) are the action-dependent sets of priors and
outcome-dependent utility functions of money. These features reflect the distinct analytical frameworks more than the characteristics of the individual preferences.

Using a simple set-up, this paper also discusses some considerations that need to be addressed when modeling principal-agent relations in which either party or both are minmax expected utility players. Within this framework the paper examines some implications of uncertainty aversion for the design of incentive contracts, illustrating complications and welfare implications that arise due to uncertainty aversion.

Little has been written on agency theory with uncertainty-averse players. Ghirardato (1994) examines the validity of some properties of optimal incentive schemes, in the context of Schmeidler’s (1989) Choquet expected utility model. While the motivation is similar, the present work differs from that of Ghirardato in terms of the analytical framework and the representations of the players’ preferences. Like Schmeidler’s Choquet expected utility model, on which it is based, Ghirardato’s analysis does not allow for outcome-dependent preferences over monetary payoffs. This aspect of the preference relations, which is relevant in many applications, is accommodated by the approach taken here.

Drèze (1961, 1987) was the first to axiomatize subjective expected utility theory with state-dependent preferences and moral hazard. He argues that the “reversal of order” axiom of Anscombe and Aumann (1963) embodies the decision-maker’s belief that he cannot influence the likely realization of alternative states. In situations in which a decision maker may take actions that would tilt the probabilities of the states in his favor, he would pre-
fer knowing the outcome of a lottery (and, consequently, his payoff contingent on the state that obtains) before rather than after the state of nature is revealed. In Drèze’s theory the decision maker’s actions are tacit.

Like the maxmin expected utility model, the representation in Drèze’s theory entails the maximization of subjective expected utility over a convex set of subjective probability measures. Yet Drèze’s theory is different from the maxmin expected utility model in its motivation and interpretation. Drèze’s motivation is to model decision making in the presence of moral hazard, not uncertainty aversion. Consequently, the set of probability measures in his theory represents the leverage a decision maker believes he has over the likely realization of the states, rather than his uncertainty aversion.

The present work differs from that of Drèze in several respects, the most important of which is the analytical framework. In this paper the agent’s actions and the relation that the principal and the agent believe to exist between these actions and probability distributions of outcomes are explicit aspects of the representation. This difference in the treatment of the actions emanates from distinct methodological outlooks. Drèze is looking at the problem of modeling the agent’s behavior from the perspective of an observer (or the principal) who may “see” only the decision-makers’ preferences over acts and over the timing of resolution of risk. In contrast, here the principal and the agent have preferences over the set of action-contract pairs that are known to them and, in principle, observable.

Also related to the present work is Karni (2004b), where I explore axiomatic foundations
of agency theory based on an analytical framework similar to the one used here. In that paper I avoided the use of probabilities (or lotteries) as a primitive concept and was concerned only with principal and agent who are subjective expected utility maximizers.
References


