THE MILTON FRIEDMAN INSTITUTE FOR RESEARCH IN ECONOMICS

MFI Working Paper Series
No. 2011-002

Existence of Optimal Mechanisms in Principal-Agent Problems

Ohad Kadan
Olin Business School, Washington University in St. Louis

Philip J. Reny
Department of Economics, University of Chicago

Jeroen M. Swinkels
Kellogg School of Management, Northwestern University

January 2011

MILTON FRIEDMAN INSTITUTE
THE UNIVERSITY OF CHICAGO

1126 East 59th Street
Chicago, Illinois 60637

T: 773.702.7587
F: 773.795.6891
mfi@uchicago.edu

Existence of Optimal Mechanisms in Principal-Agent Problems*

Ohad Kadan,†Philip J. Reny,‡and Jeroen M. Swinkels§

First Draft, October 2009
This Draft, January 2011

Abstract

We provide general conditions under which principal-agent problems admit mechanisms that are optimal for the principal. Our result covers as special cases those in which the agent has no private information – i.e., pure moral hazard – as well as those in which the agent's only action is a participation decision – i.e., pure adverse selection. We allow multi-dimensional actions and signals, as well as both financial and non-financial rewards. Beyond measurability, we require no *a priori* restrictions on the space of mechanisms. Consequently, our optimal mechanisms are optimal among all measurable mechanisms. A key to obtaining our result is to permit randomized mechanisms. We also provide conditions under which randomization is unnecessary.

1 Introduction

In his classic work on the principal-agent problem, Mirrlees (1976, 1999) identifies at least two important theoretical questions. The first is the question of the existence of an optimal incentive contract. Somewhat surprisingly, within even the most standard economic settings (e.g., logarithmic utility, normally distributed signals) Mirrlees shows that an optimal contract need not exist. The second question relates to the character of the optimal contract when one does exist. Here Mirrlees introduces the "first-order approach" (FOA) in which the agent's (possibly infinite number of) incentive constraints is replaced with the single constraint that the effort level targeted by the principal is a critical point of the agent's objective function. Much research has focused on providing conditions under which the FOA can be employed (see, e.g., Rogerson (1985), Jewitt (1988), Sinclair-Desgagne (1994), Conlon (2009)). But even the current state-of-the-art conditions are highly restrictive.

Our focus here is on the first of Mirrlees' questions, that of the existence of a solution to the principal's problem. However, our setting is more general. Specifically, an agent has a privately known type and a must choose an action. The principal presents the agent with a menu of contracts,

*We thank John Conlon and Wojciech Olszewski for helpful comments, and Max Stinchcombe for being especially generous with his time in discussing the paper with Swinkels during his Fall 2009 visit to UT Austin. We also thank Timothy Van Zandt for providing a number of very helpful comments in his thorough discussion of our paper at the 2011 ASSA/Econometric Society meetings in Denver. Reny gratefully acknowledges financial support from the National Science Foundation (SES-0922535), as well as the hospitality of MEDS at Northwestern University during his 2009-2010 visit.

†Olin Business School, Washington University in St. Louis

‡Department of Economics, University of Chicago

§Kellogg School of Management, Northwestern University

where each contract specifies a reward for the agent as a function of a signal whose distribution depends on the agent's action and type. After the agent chooses a contract from the menu and takes an action, the signal is generated and the contract is honored.[1] Thus, our setting includes as special cases pure moral hazard and pure adverse selection, but in general allows for combinations of the two.[2] The question addressed here is, "Under what conditions does a menu of contracts exist that maximizes the principal's ex-ante expected payoff subject to the agent's interim individual rationality constraints?"

We establish existence without imposing the restrictive conditions (e.g., one-dimensional signals and rewards, monotone likelihood ratio conditions, additively separable utility in rewards and actions, etc.) that are typical for the validity of the FOA. Similar to Holmström and Milgrom (1987, 1991) we allow multi-dimensional effort, signals, and rewards. Rewards may take the form of cash, promotions, restricted stock, or a nicer office. Unlike Holmström and Milgrom (1987, 1991), we do not restrict attention to a particular functional form of utility nor do we impose a particular statistical relationship between the agent's actions and the information received by the principal. We allow substantial flexibility in how the set of feasible rewards varies with the signal, as for example when a firm cannot contract to pay its manager more than some fraction of the total value of the firm. Finally, we allow the payoff of both the principal and agent to depend in a non-seperable way on the type of the agent, the action chosen, the signal, and the reward, with no order structure on either object. As such, our model incorporates, for example, a setting where the type of the agent is information about which of several projects is better, and the action of the agent and the "reward" of the principal consist of which project to focus their respective efforts on. In such an example, the payoffs of the agent and principal may well have a substantial component of common interest.

Even in the special case of pure moral hazard, establishing the existence of an optimal contract under general conditions has proven to be surprisingly difficult. The central issue is that the set of all possible contracts as typically conceived is not compact in a useful topology, i.e., one in which the participants' payoffs are continuous. For instance, Mirrlees' (1999) counterexample involves a sequence of contracts along which extreme punishments are inflicted at extremely low performance thresholds, and along which the principal's payoff approaches the full-information optimum. But, no contract can attain full-information optimum payoffs.

To restore compactness, Holmström (1979) begins by introducing upper and lower bounds on feasible payments. But, even with such bounds, the set of contracts is not usefully compact. Holmström thus also imposes a bounded variation restriction on contracts. While this delivers an existence result, the bounded variation restriction is ad hoc – there is no guarantee that otherwise reasonable contracts that do not meet the variation bound or the bounds on feasible payments

[1] No mechanism can improve upon a menu of contracts, which itself is equivalent to the principal conditioning the contract he offers the agent on the agent's report of his type. We will take this latter direct mechanism approach in the sequel. See Section 3.

[2] Examples of the latter sort of model can be found in Hellwig (2010), Laffont and Tirole (1986, 1993), Walsh (1995).

would not be better for the principal. In a model with substantively more general actions, signals, and preferences, Page (1987) takes a similar route to existence, again introducing upper and lower bounds on payments, and assuming that the set of feasible contracts is *a priori* restricted so as to be sequentially compact in the topology of pointwise convergence. Again, however, there is no guarantee that contracts outside the pre-specified set would not be even better.

In contrast, Grossman and Hart (1983) establish existence of an optimum in the moral hazard problem without ad hoc restrictions on contracts. They do this by restricting attention to a finite set of signals each of which occurs with probability bounded away from zero regardless of the agent's action. With a continuum of signals, Carlier and Dana (2005) and Jewitt, Kadan and Swinkels (2008) solve the existence problem while avoiding ad hoc restrictions on contracts by assuming that effort is one-dimensional, that likelihood ratios are monotone and bounded, and that the FOA is valid. All three papers require signals and rewards to be one-dimensional and the principal's losses to be additively separable in them, and none permits the agent to possess private information.

A contract is typically defined as a function from signals to rewards. This is too restrictive in our general setting, where randomization over rewards may be strictly optimal for the principal. Consequently, we allow contracts that randomize over rewards. But there is another–deeper–reason for allowing randomization. As our proof specialized to pure moral hazard reveals, randomization over rewards renders the set of all contracts whose costs to the prinicpal are uniformly bounded above, compact in a topology in which the principal's losses are lower semicontinuous. This insight builds on the ideas of Milgrom and Weber (1985), who introduce the concept of distributional strategies to compactify the set of behavioral strategies in Bayesian games with continuum type spaces.[3]

While permitting randomized contracts is crucial for our existence proof, the contracts in an optimal menu may or may not involve randomization. We establish conditions under which optimal menus can be composed of deterministic contracts and, in pure moral-hazard problems when, in addition, the cost minimizing contract for any given effort level is unique. We also provide examples in which randomization is necessary for optimality. The latter can occur when, for example, the sets of signals and rewards are finite.

Our results require two economically substantive assumptions, and several technical ones. The first substantive assumption is that the problem at hand can be formulated so that both the agent's utility and the principal's losses are bounded below (with the former ruling out the Mirrlees counterexample). Both restrictions appear to us to be exceedingly mild from an economic point of view, with the first being satisfied if, for example, contracts with arbitrarily draconian punishments are unavailable, and the second being satisfied if, for example, the gross benefits to the principal as a function of the agent's action have finite expectation and the gross costs to the principal are bounded below.

[3] To see how randomization can help with compactness in Bayesian games, consider an auction in which a bidder's value is drawn from $[0, 2]$, and suppose he bids zero if the n-th decimal place of his value is even, and bids 1 if it is odd. This pure strategy has no pointwise limit as $n \to \infty$. However, it does converge weakly to the randomized strategy in which the bidder is equally likely to bid zero or one regardless of his value.

Utility and losses are permitted to be unbounded above. This is important in particular because a priori upper bounds on payments are not easy to motivate. For example, it is easy to see why a CEO cannot be paid more than the value of the firm, but it is less obvious why the value of the firm should be bounded up front. Even if the problem at hand naturally leads to both utility and losses that are bounded, compactness of the space of contracts remains an issue that requires the techniques we introduce.

The second economically substantive assumption is that if the utility of the agent diverges along a sequence of rewards, the cost to the principal of providing that utility diverges even faster. When, for example, the wage is both the agent's reward and the firm's cost, this assumption is satisfied if the agent's marginal utility of wealth tends to zero as wealth increases without bound.

Our existence result is proven under two sets of hypotheses. In the first, we require that the support of the principal's signal is independent of the agent's action and private type. The resulting information structure is nonetheless quite general, and, in particular, it does not impose monotone likelihood ratio conditions or any other order structure on signals and/or actions.

The second set of hypotheses permits an even more general information structure, allowing, for example, signal supports that vary with the agent's action and type and that contain atoms. However, while we still permit the agent's preferences over rewards to depend on his action and type and the signal generated, we require here that for any given signal there is a worst reward for the agent that is independent of his action and type.

Finally, we require several purely technical assumptions. The agent's action space is compact metric, the reward, type, and signal spaces are complete, separable, and metric. As functions of the signal, action, reward, and type, the agent's utility is jointly continuous and the principal's loss is jointly lower semicontinuous.

Section 2 presents our assumptions on ambient spaces, payoffs, and information. Several example are provided that illustrate the model's scope. Section 3 defines the spaces of contracts, menus, and mechanisms, and introduces the principal's problem. Section 4 states the main result and provides a sketch of the proof. The formal proof is given in Section 5. Finally, Section 6 provides two results on when optimal contracts can be chosen to be deterministic, and provides examples in which optimality requires randomization.

2 The Model

2.1 Ambient Space and Payoff Assumptions

We maintain the following assumptions throughout.

Assumption 1 *The set of actions, A, is a non-empty compact metric space.*

Assumption 2 *The set of signals, S, is a Polish space, i.e., a complete separable metric space.*

Assumption 3 *The set of rewards is a Polish space, R. For each $s \in S$, $\phi(s) \subset R$ is a nonempty set of feasible rewards given the signal s. We assume that $\Phi = \{(s,r) \in S \times R : r \in \phi(s)\}$, the graph of the correspondence ϕ, is a Borel subset of $S \times R$.*

Assumption 4 *The set of types, T, is a Polish space, and the prior on T is a Borel probability measure H.*[4]

Assumption 5 *The agent has a continuous von Neumann-Morgenstern utility function, $u : \Phi \times A \times T \to \mathbb{R}$ that is bounded below, without loss of generality by 0.*

Assumption 6 *The principal has a lower semicontinuous von Neumann-Morgenstern loss (disutility) function $l : \Phi \times A \times T \to \mathbb{R}$ that is bounded below, again without loss of generality by 0.*

Assumption 7 *For every $\alpha \in \mathbb{R}$ and every compact subset Y of $S \times T$, $\{(s,r,a,t) \in \Phi \times A \times T : (s,t) \in Y$ and $l(s,r,a,t) \leq \alpha\}$ is compact.*[5]

Assumption 8 *If $u(s',r',a',t') \to \infty$ for some sequence (s',r',a',t') in $\Phi \times A \times T$, then $u(s',r',a',t')/l(s',r',a',t') \to 0$.*

Let $P(\cdot|a,t)$ denote the probability measure over the signal space S when the agent takes action $a \in A$ and his type is $t \in T$. We present and discuss our assumptions on P in section 2.3.

2.2 Examples and Discussion

The model reduces to pure moral hazard when the type space, T, is a singleton, and reduces to pure adverse selection when the action space, A, contains just two elements, "participate" and "don't participate," reflecting an interim IR constraint.

The model is intended to be as general as possible. Actions come from a compact space, but signals and rewards need not, and all three can be multi-dimensional (even infinite-dimensional) and perhaps discrete in some dimensions while continuous in others. No order structure is assumed on action, signal, or reward spaces. Even when a natural order exists, no monotonicity of payoffs is assumed. Our most economically significant assumption is that the agent's utility and the principal's losses are bounded below.

Each of the following examples satisfies all of the assumptions above as well as the informational assumptions in Section 2.3.

[4] Throughout the paper, whenever we refer to a probability measure over a metric space, we take the measurable sets to be the Borel sets. Spaces of probability measures are always endowed with the weak topology and are in fact metrizable in our setting because all ambient spaces are Polish.

[5] It follows from assumptions 3 and 7 that R is in fact a countable union of compact sets. That is, R is a sigma-compact Polish space. Every Euclidean space is a sigma-compact polish space, being the union of all closed balls with integer radii. Potentially useful infinite dimensional spaces, such as the space of bounded real-valued Lipschitz functions on a compact metric space, are also sigma-compact Polish spaces. Such function spaces can arise as reward spaces if, for example, the agent works in period 1 and is rewarded on the basis of the stock price in period 2, where rewards are stock options or other derivatives based on the stock price in period 3.

Example 1 ***Pure Moral Hazard.*** *Let $S = [\underline{s}, \bar{s}]$, $T = \{t\}$ (a singleton), $R = [0, \infty)$, $\phi(s) = R$ for all s, $A = [0, 1]$, and $l(s, r, a, t) = r - s$, where s is revenue, r is monetary compensation, and a is effort. Let $u(s, r, a, t) = w(r, a)$ where $w(\cdot, \cdot)$ is continuous, and assume that $P(\cdot|a)$ admits a positive density, $f(s|a)$, continuous in s and a. This yields the standard moral hazard problem, save that payments are bounded below and that there is no MLRP restriction on the information structure. When $w(r, a) = v(r) - c(a)$ we obtain the commonly employed case in which the agent's utility is separable in money, $v(\cdot)$, and cost of effort, $c(\cdot)$.*

The requirement that losses are bounded below may seem restrictive, as it appears to rule out arbitrarily large returns to the principal. In particular, Example 1 does not immediately admit cases in which the principal's revenue space is unbounded such as the exponential distribution example studied in Holmström (1979) or the case where $P(\cdot|a)$ represents the normal distribution with mean a and variance 1. However, as the next example illustrates, such cases are also often covered by our formulation, as long as the principal's *expected* revenue is bounded above.[6]

Example 2 ***Pure Moral Hazard with Unbounded Returns.*** *Assume that $S = (-\infty, \infty)$, $T = \{t\}$, $R = [0, \infty)$, $\phi(s) = R$ for all s. The principal receives revenue s, and pays compensation r. Let $u(s, r, a, t) = w(r, a)$ be continuous. If for some z, $\zeta(a) = \int s dP(s|a) \leq z$ for all $a \in A$, as would be the case if $\zeta(a)$ were finite and continuous in a on the compact set of actions A, then $l(s, r, a) = r + z - \zeta(a)$ is bounded below by zero (and happens to be independent of s).*

When all spaces are Euclidean, assumption 7 states that as rewards become unbounded, so does the loss to the principal. Finally, assumption 8 says that the losses to the principal per util provided are unbounded above when they provide the agent with arbitrarily high utility. Note that if the agent's utility is bounded, assumption 8 is satisfied trivially, regardless of the risk preferences of the principal and agent.[7] In the next two examples, there is a fixed $z \in \mathbb{R}$ such that $\zeta(a) = \int_S s dP(s|a) \leq z$ for each $a \in A$.

Example 3 ***Canonical Utility and Loss Functions.*** *Suppose that $S = (-\infty, \infty)$, $R = [0, \infty)$, $\phi(s) = R$ for all s, $l(r, s, a) = r + z - \zeta(a)$, and $u(s, r, a, t) = v(r) - c(a)$ where $c(\cdot)$ is continuous and v is differentiable with $\lim_{r \to \infty} v'(r) = 0$ (as is true for risk-averse utility functions typically used in practice). Then, assumption 8 is satisfied.*

If, in the previous example, $v(r) = r - \frac{1}{r}$, assumption 8 would fail. The next example illustrates some of the flexibility of our model.

Example 4 ***Multidimensional signals and rewards in a pure moral hazard problem.*** *A salesperson chooses his effort level in $[0, 1]$, and which of three styles of sales pitch to employ. Thus,*

[6] Situations in which the agent's utility is unbounded below in s can similarly often be brought into the confines of our model.

[7] Even if utility can diverge, assumption 8 imposes only a very weak form of asymptotic diminishing returns. The agent can, for example, be risk seeking over an arbitrarily large range of rewards.

$A = [0,1] \times \{1,2,3\}$. *The firm can observe whether the contract is or is not obtained* $\{yes, no\}$, *the contract price,* $p \in [0,\infty)$, *the delivery date,* $\tau \in [\underline{\tau}, \bar{\tau}]$, *and which of* 2 *basic platforms the customer chose,* $i \in \{1,2\}$. *Thus,* $S = \{yes, no\} \times [0,\infty) \times [\underline{\tau}, \bar{\tau}] \times \{1,2\}$. *The firm can give the salesperson a non-negative cash bonus, restricted by a liquidity constraint on the part of the firm to be no more than* p, *and decide whether or not to promote him. Hence,* $R = [0,\infty) \times \{$*promote,don't*$\}$, *with* $\phi(s) = [0,p] \times \{$*promote,don't*$\}$.

Example 5 ***Pure Adverse Selection.*** *Let* $S = \{s\}$ *be a singleton. The principal is a monopolist producing a good of quality* $q \in \left[\underline{q}, \bar{q}\right]$ *at a continuous cost* $\psi(q)$ *for a price* $p \in [0, \bar{p}]$. *The agent is a consumer with a private taste parameter for quality* $t \in T$, *where* T *is a subset of an Euclidean space, with prior* $H(\cdot)$. *A reward to the agent is a price-quality pair* $r = (p,q)$. *Thus,* $R = \phi(s) = [0,\bar{p}] \times \left[\underline{q}, \bar{q}\right]$. *The agent chooses whether or not to buy the good, i.e.,* $A = \{$*buy, don't buy*$\}$. *If the agent buys at price-quality pair* (p,q), *the loss to the principal is* $l(s,r,a,t) = \psi(q) - p$ *and the utility to the agent is* $u(s,r,a,t) = v(q,t) - p$, *where* v *is a continuous function representing the benefit to the agent from consuming a good with quality* q *given taste* t. *If the agent does not buy, both principal and agent receive a payoff of zero.*

The next example illustrates a mixed moral hazard/adverse selection setup.

Example 6 ***Moral Hazard and Adverse Selection.*** *The principal sells fire insurance and the agent is a homeowner. The value of the home is* $V_0 > 0$. *The homeowner can take preventive actions in* $[\underline{a}, \bar{a}]$ *and can also decide whether or not to purchase insurance. Hence,* $A = [\underline{a}, \bar{a}] \times \{$*insure, self-insure*$\}$. *The homeowner has private information about the probability of a fire indexed by* $t \in T = [\underline{t}, \bar{t}]$, *with prior* $H(\cdot)$. *The principal's signal set is* $S = [0, V_0]$, *reflecting possible damages, and* $P(\cdot|a,t)$ *is the distribution of damages given action* a *and type* t. *A reward is a pair* $r = (\pi, b)$, *where* $\pi \in [0, V_0]$ *is the insurance premium paid by the homeowner to the company, and* $b \in [0,s]$ *is the benefit paid by the company to the homeowner given damages of* s. *That is,* $\phi(s) = [0, V_0] \times [0,s]$ *for all* $s \in S$ *(for simplicity it is assumed that both the premium and the compensation are paid after the damages have been observed). If the agent chooses to insure, takes preventive action* a, *and the damage is* s *leading to reward* (π, b), *then the principal's loss is* $l(s,r,a,t) = b - \pi$ *and the agent's utility is* $u(s,r,a,t) = v(V_0 - s - \pi + b) - c(a)$, *where* $v(\cdot)$ *and* $c(\cdot)$ *are continuous. If the agent chooses to self-insure, taking preventive action* a, *then the principal's payoff is zero and the agent's utility in the event of damages* s *are* $v(V_0 - s) - c(a)$.

2.3 Information and Rewards

We now turn to our information structure. Our first assumption imposes a mild form of continuity of information in a and t.

Assumption 9 *If* $a_n \to a$, *and* $t_n \to t$, *then* $P(\cdot|a_n, t_n) \to P(\cdot|a,t)$, *where convergence of measures is in the weak topology.*

In general, we allow the supports of signals to vary with a and t. But, our next assumption requires that where supports overlap, they satisfy a form of absolute continuity. On the other hand, we do not require any form of monotone likelihood ratio property. Indeed, we impose no order structure on either A or S.

Assumption 10 *For every $a \in A$ and every $t \in T$ there is a Borel subset $S_{a,t}$ of S such that $P(S_{a,t}|a,t) = 1$ and $P(\cdot|a',t')$ is absolutely continuous with respect to $P(\cdot|a,t)$ on $S_{a,t}$ for every $a' \in A$ and $t' \in T$.*[8]

Assumption 10 is much weaker than assuming that $P(\cdot|a',t')$ is absolutely continuous with respect to $P(\cdot|a,t)$, because we ask for absolute continuity only on $S_{a,t}$. It is also weaker than assuming that $P(\cdot|a',t')$ is absolutely continuous with respect to $P(\cdot|a,t)$ on the support of $P(\cdot|a,t)$ because $S_{a,t}$ need not be closed and hence may be a strict subset of the support of $P(\cdot|a,t)$. Example 8 below illustrates the usefulness of this.

Assumption 10 implies, by the Radon-Nikodym theorem, that for all $a, a' \in A$ and all $t, t' \in T$, there is a measurable function, $\xi(\cdot, a', t', a, t) : S \to \mathbb{R}$ such that $P(B|a',t') = \int_B \xi(s,a',t',a,t) dP(s|a,t)$ for every measurable B contained in $S_{a,t}$. We require that the Radon-Nikodym derivative, ξ, can be chosen to be well-behaved in the following sense.

Assumption 11 *ξ is lower semicontinuous at (s,a',t',a,t) whenever $(a',t') \neq (a,t)$.*

Example 7 ***A Density.*** *S is Euclidean and $P(\cdot|a,t)$ can be represented by a density $f(s|a,t)$ where $f : S \times A \times T \to \mathbb{R}$ is continuous and everywhere strictly positive. In particular, here we can take $S_{a,t} = S$ for all a, t, and $\xi(s,a',t',a,t) = \frac{f(s|a',t')}{f(s|a,t)}$.*

Example 8 ***Moving support.*** *Let $A = [0,1]$, $S = [0,2]$, $T = \{t\}$, and let $P(\cdot|a,t)$ be uniform on $[a, a+1]$. Take $S_{a,t} = (a, a+1)$, and take $\xi(s,a',a) = 1$ for $s \in (a', a'+1) \cap (a, a+1)$, and 0 elsewhere.*[9,10]

Example 9 ***Moving atom.*** *Let $A = S = [0,1]$, $T = \{t\}$, and let $P(\cdot|a,t)$ be the Dirac measure placing mass one on $s = a$. That is, a is observable.*[11]

Example 10 ***Discrete signal distributions.*** *Suppose that S is a finite set and that $P(s|a,t)$ is continuous in (a,t) for each $s \in S$.*

Each of these satisfies our assumptions. An example that does not is the following.

[8] That is, $P(B|a',t') = 0$ for every Borel subset B of $S_{a,t}$ such that $P(B|a,t) = 0$.

[9] In examples of pure moral hazard, we will write simply $\xi(s,a',a)$.

[10] Note the usefulness, in terms of the lower semicontinuity of ξ, of allowing $S_{a,t}$ to be $(a, a+1)$ as opposed to insisting that it be the entire support $[a, a+1]$.

[11] This example illustrates why we only ask that ξ be lower semicontinuous where $(a',t') \neq (a,t)$. In particular, $\xi(s = a, a', a) = 0$ must hold for all $a' \neq a$, while $\xi(s = a, a, a) = 1$, violating lower semicontinuity at that point.

Example 11 ***A Non-Compact Set of Implementable Actions.*** *Let $R = [0,3]$, $A = S = [0,2]$, $T = \{t_0\}$, and $u(s,r,a,t) = a + r$. Suppose the principal sees the agent's action when it is strictly less than one, but not otherwise, i.e., the principal receives the signal $s = a$ if $a < 1$ and $s = 1$ if $a \geq 1$. Hence, assumption 11 is violated because $\xi(s,a',a) = 0$ for $(s,a',a) = (a,2,a)$ with $a < 1$ while $\xi(1,2,1) = 1$. Every action $a \in [0,1)$, being observable, is implementable. However, action $a = 1$ is not implementable because it is indistinguishable by the principal from the action $a = 2$, which is strictly preferred by the agent.*

One can see the role of the lower semicontinuity of ξ in this example. Effectively, as $a \to 1$, there is a discontinuous change in how much difficulty the principal has in rewarding action a without also giving the agent an incentive to deviate to $a = 2$. The role of lower semicontinuity is to ensure that, in a general sense, this does not occur.

Our main existence result requires that one of two additional assumptions is also satisfied. The first places more structure on information, the second more structure on payoffs.

Say that the environment has *uniform information* if the support of signals does not depend on a or t, i.e., if the following assumption holds.

Assumption 12 *The set of signals $S_{a,t}$ can be chosen to be S for every $a \in A$ and $t \in T$.*

Assumptions 10 and 12 are equivalent to saying that $P(\cdot|a,t)$ is absolutely continuous with respect to $P(\cdot|a',t')$ for every (a,t) and (a',t'). Example 7 satisfies this assumption. Examples 8 and 9 do not. When there is uniform information, no further structure is necessary for existence. But, when the support of the distribution of signals varies, we require instead some extra structure on rewards. In particular, say that the environment has *simple worst rewards* if the following assumption on payoffs holds.

Assumption 13 *There is a measurable function $r_* : S \to R$ with $r_*(s) \in \phi(s)$ for every $s \in S$ such that $u(s,r,a,t) \geq u(s,r_*(s),a,t)$ for every $(s,r,a,t) \in \Phi \times A \times T$, and such that $u(s,r_*(s),a,t)$ is lower semicontinuous in $(s,t) \in S \times T$ for each $a \in A$.*

That is, for each s, and independent of a and t, there is a worst reward for the agent.

Example 12 ***Worst Rewards.*** *If (i) R is partially ordered (Euclidean for example), (ii) a compact subset of R contains for each s a least element of $\phi(s)$, and (iii) u is continuous in its arguments and increasing in r, then assumption 13 holds.*

Thus, assumption 13 typically holds when r is a reward in the "usual" sense. The following is an example that fails to have simple worst rewards (and thus where our results require uniform information to establish existence).

Example 13 ***No Worst Reward.*** *Let $T = \{t\}$, let $A = R$ be a finite set, and for each $s \in S$ let $u(s,r,a,t) = 1$, if $a = r$, zero otherwise.*

3 Contracts, Menus, Mechanisms, and the Principal's Problem

In the introduction, we described the principal's problem as one of choosing a menu of contracts from which the agent chooses. It is well-known that this is without loss of generality by the revelation principle. Indeed, it is equivalent, and again without loss, for the principal to consider only direct mechanisms of the following form.

The agent reports his type and (i) the mechanism specifies a (possibly randomized) reward as a function of the signal the principal receives (i.e., the mechanism specifies a contract), and (ii) the mechanism suggests to the agent an action to take.[12] Moreover, the mechanism can be restricted to be incentive compatible in the sense that, regardless of his type, the agent can do no better than to report his type truthfully and to take the suggested action.[13]

Because menus of contracts and incentive compatible direct mechanisms are equivalent, we will, for convenience, conduct the analysis in terms of incentive compatible direct mechanisms. We now turn to the formal definitions, beginning with the notion of a contract.

Given the generality of the model, randomization over rewards will sometimes be strictly optimal.[14] However, even when randomization is not necessary for optimality, randomization nonetheless lies at the heart of establishing the *existence* of an optimal mechanism. Indeed, permitting randomized contracts is essential for establishing an appropriate compactness property for the space of contracts. This motivates the following definition of a contract, where $\Delta(Z)$ denotes the set probability measures on the Borel subsets of any topologized set Z.

Definition 1 *The set of contracts is* $K = \{\kappa(\cdot|\cdot)$ *s.t. (i)* $\kappa(\cdot|s) \in \Delta(\phi(s))$ *for every* $s \in S$*, and (ii)* $\kappa(B|s)$ *is a measurable function of* s *on* S *for every Borel subset* B *of* $R\}$.

That is, for each s, a contract specifies a lottery over rewards, with a basic measurability condition as s varies. Of course, randomization is permitted but is not required—for each s there may be a reward, r, such that $\kappa(\cdot|s)$ places probability one on r.

For $\kappa \in K$, $a \in A$, and $t \in T$, define

$$U(\kappa, a, t) = \int_{\Phi} u(s, r, a, t) d\kappa(r|s) dP(s|a, t)$$

[12] The mechanism need never randomize over the suggested action because the agent would have to be indifferent among all actions in the support of the randomization and one such action must yield the principal losses that are no greater than those expected under the randomization.

[13] Note that a mechanism induces a menu of contracts, namely the set of all contracts determined by the mechanism as the agent's report varies over all of his possible types. Moreover, the menu induced by an incentive compatible mechanism has the property that it is optimal for the agent to choose from the menu the contract the mechanism would have specified under truthful reporting and it is optimal for the agent to take the action that would have been suggested by the mechanism. Consequently, menus of contracts are equivalent to incentive compatible mechanisms from both the point of view of the principal and that of the agent.

[14] In Section 6 we examine conditions under which deterministic contracts are optimal, and present examples where they are not.

as the utility to the agent if he faces contract κ, takes action a, and has type t, and let

$$L(\kappa, a, t) = \int_{\Phi} l(s, r, a, t) d\kappa(r|s) dP(s|a, t)$$

denote the principal's expected losses. Because utility and losses are nonnegative, both integrals exist but could take on the value $+\infty$.

For each reported type of the agent, a mechanism determines a contract and suggests an action to the agent. Formally, we have the following.

Definition 2 *A* mechanism *is a pair of functions* $\boldsymbol{\kappa} : T \to K$, *and* $\mathbf{a} : T \to A$ *specifying the contract and suggesting an action for each reported type of the agent. For notational compactness, write* κ_t *for* $\boldsymbol{\kappa}(t)$, *and* a_t *for* $\mathbf{a}(t)$. *We require* $\mathbf{a} : T \to A$ *to be measurable and* $\int_B \kappa_t(C|s) dP(s|a_t, t)$ *to be measurable as a function from* T *into the reals, for all Borel subsets* B *of* S *and* C *of* R.

The measurability condition on $\int_B \kappa_t(C|s) dP(s|a_t, t)$ ensures that the principal can compute expected losses, and one can also compute the agent's ex-ante utility.[15]

Definition 3 *The mechanism* $(\boldsymbol{\kappa}, \mathbf{a})$ *is* incentive compatible *at* t *if it is optimal for the agent of type* t *to announce his true type and to take action* a_t. *That is,*

$$U(\kappa_t, a_t, t) \geq U(\kappa_{t'}, a', t) \text{ for all } a' \in A, t' \in T.$$

Finally, we have,

Definition 4 *A mechanism is* incentive compatible *if it is incentive compatible at* H*-a.e.* $t \in T$.

For any mechanism $(\boldsymbol{\kappa}, \mathbf{a})$, let

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\kappa}, \mathbf{a}) = \int_T L(\kappa_t, a_t, t) dH(t),$$

be the principal's expected loss assuming that the agent tells the truth and takes the recommended action for all t.

The principal's problem is

$$\min \boldsymbol{\mathcal{L}}(\boldsymbol{\kappa}, \mathbf{a}) \text{ s.t. } (\boldsymbol{\kappa}, \mathbf{a}) \text{ is incentive compatible.} \tag{1}$$

Although not explicitly modeled, our specification of the principal's problem easily allows the presence of interim individual rationality constraints for the agent.[16] To see this, suppose that the

[15] In particular, the condition ensures that for any measurable function $g : \Phi \times A \times T \to \mathbb{R}$, the integral $\int_T \int_\Phi g(s, r, a_t, t) d\kappa_t(r|s) dP(s|a_t, t) dH(t)$ exists. One might instead ask for the stronger condition that $\kappa_t(C|s)$ be jointly measurable in (s, t). Demonstrating the existence of such an optimal mechanism would require an extra step in an already lengthy proof.

[16] Such constraints arise when the agent must commit to participating in the mechanism only after learning his type.

agent has an outside option, $u_0(t)$, that may depend upon his type. The corresponding interim IR constraints can be incorporated by adding a distinct and isolated action a_0 to A and defining $u(r, s, a_0, t) = u_0(t)$ for all (r, s, t) and similarly defining $l(r, s, a_0, t)$ to be the loss to the principal when the agent exercises his outside option.[17]

4 The Main Existence Result

Theorem 1 *Suppose that assumptions 1-11 hold and that the environment has either uniform information or simple worst rewards, i.e., either assumption 12 or 13 also holds. If $\mathcal{L}(\boldsymbol{\kappa}, \mathbf{a})$ is finite for some incentive compatible mechanism $(\boldsymbol{\kappa}, \mathbf{a})$, then the principal's problem (1) possesses a solution.*

4.1 Proof Sketch for Pure Moral Hazard

A good deal of the effort in proving Theorem 1 is spent on the case of simple worst rewards without uniform information, and on establishing measurability of the mechanism in the agent's type when T is uncountable. Nonetheless, it is useful to illustrate some of the main ideas in the simpler environment of uniform information and pure moral hazard (T a singleton).

The challenge is to rule out examples along the lines of Mirrlees (1999), where the "optimum" can be arbitrarily closely approximated but never achieved. Let $c < \infty$ be the loss from some feasible contract and the action it implements. Given the compactness of A, it would be sufficient to establish that if κ_n is a sequence of contracts implementing actions $a_n \to a^*$ with $L(\kappa_n, a_n) \leq c$, then there is a contract $\kappa^* \in K$ implementing a^* with $L(\kappa^*, a^*) \leq \underline{\lim}_n L(\kappa_n, a_n)$.[18]

Note that κ_n and $P(\cdot|a_n)$ together induce a distribution μ_n on $S \times R$. The distribution μ_n is similar in spirit to a distributional strategy (Milgrom and Weber (1985)). Prohorov's theorem states that if the sequence μ_n is tight, then it converges along a subsequence to some μ^*.[19] The tightness of the sequence μ_n hinges on our assumption that $l(\cdot)$ is bounded from below. Together with the boundedness of the sequence $L(\kappa_n, a_n)$ of expected losses, this implies that rewards leading to large losses must be rare. Tightness follows because we also assume, roughly, that sets of rewards leading to bounded losses are compact. Our candidate contract will be κ^*, a regular conditional probability of μ^*. Several issues must be adressed.

[17] One might instead wish to satisfy an ex-ante individual rationality constraint (i.e., when the agent must commit to participating in the mechanism *before* learning his type). Our results can be extended to settings with an ex-ante IR constraint under somewhat stronger conditions, and we do not pursue the details here. Of course, in the case of pure moral hazard, the distinction between ex-ante and interim IR constraints vanishes and our model applies as is.

[18] For the case of pure moral hazard, our existence proof implicitly shows that there is a topology under which (i) the set of incentive compatible contracts with expected cost below a given amount is compact and (ii) the principal's expected loss function is lower semi-continuous. A similar statement holds when the type space is finite or countable. However, when the type space is a continuum, our proof, while still establishing the existence of an optimal mechanism, does not suggest a natural topology on the space of mechanisms (each of which specifies a contract and action for each type). See in particular step IV of the proof of Theorem 1.

[19] A set of probability measures on the Borel sets of a topological space is *tight* if, for any $\varepsilon > 0$, there is a compact C such that $P(C) > 1 - \varepsilon$ for all measures P in the set.

First, because $l(\cdot)$ and $u(\cdot)$ are lower semicontinuous and bounded below, but not necessarily bounded above, limits of their expectations under μ_n need not be equal to their expectations under μ^*. For example, suppose that $R = [0, \infty)$, and consider the sequence of contracts in which, regardless of the signal, the reward is n with probability $1/n$ and 0 with probability $(n-1)/n$. If $l(s, r, a) = r$, then, $L(\kappa_n, a_n) = 1$ for all n. However, under μ^*, rewards are zero with probability 1 and so $L(\kappa^*, a^*) = 0$. Consequently, $L(\kappa^*, a^*) < \lim_n L(\kappa_n, a_n)$ in this example, a result that generalizes. Indeed, by the portmanteau theorem,[20] our lower semicontinuity and lower bound assumptions imply that limits of expectations of $l(\cdot)$ and $u(\cdot)$ under μ_n are no less than their expectations under the weak limit μ^* so that, in particular, the desired inequality $L(\kappa^*, a^*) \leq \underline{\lim}_n L(\kappa_n, a_n)$ follows.

It remains to ensure that κ^* implements a^*. For this, the implication of the portmanteau theorem–that expectations cannot jump up in the limit–is not enough. Indeed, if the utility of the implemented action jumps down in the limit, incentive compatibility can fail. The role of assumption 8 is to ensure that the agent's utility does not jump down. To see how this works, return to the example in the previous paragraph and suppose that $u(s, r, a) = v(r) - a$, where v is utility over wealth. Then $U(\kappa_n, a_n) = \frac{1}{n} v(n) + \frac{n-1}{n} v(0) - a_n$. By assumption 8, $\frac{1}{n} v(n) \to 0$ as $n \to \infty$, and so it follows that $U(\kappa_n, a_n) \to v(0) - a^* = U(\kappa^*, a^*)$. Thus, in the example, while the loss to the principal jumps down in the limit, the utility of the agent converges. In general, assumption 8 permits us to show that

$$U(\kappa^*, a^*) = \lim_n U(\kappa_n, a_n). \tag{2}$$

Of course, what we must show is that $U(\kappa^*, a^*) \geq U(\kappa^*, a)$ for every $a \neq a^*$. Because κ_n implements a_n for every n, it suffices to show that for any $a \neq a^*$,

$$\int u(s, r, a_n)\, d\kappa_n(r|s)\, dP(s|a_n) \geq \int u(s, r, a)\, d\kappa_n(r|s)\, dP(s|a) \tag{3}$$

implies

$$\int u(s, r, a^*)\, d\kappa^*(r|s)\, dP(s|a^*) \geq \int u(s, r, a)\, d\kappa^*(r|s)\, dP(s|a). \tag{4}$$

By uniform information, we can write the right-hand side of (3) as

$$\int u(s, r, a)\, \xi(s, a, a_n)\, d\kappa_n(r|s)\, dP(s|a_n) = \int u(s, r, a)\, \xi(s, a, a_n)\, d\mu_n, \tag{5}$$

and the right-hand side of (4) as

$$\int u(s, r, a)\, \xi(s, a, a^*)\, d\kappa^*(r|s)\, dP(s|a^*) = \int u(s, r, a)\, \xi(s, a, a^*)\, d\mu^*. \tag{6}$$

Since utility is bounded from below, and since $u(s, r, a)\, \xi(s, a, a_n)$ is lower semicontinuous by assumption, the portmanteau theorem can be used to show that the liminf of the right-hand side of

[20] See, e.g., van der Vaart and Wellner (1996), Theorem 1.3.4.

(5) is at least as large as the right-hand side of (6). Hence, the desired inequality follows because, by (2), the left-hand side of (3), which is simply $U(\kappa_n, a_n)$, converges to the left-hand side of (4), which is simply $U(\kappa^*, a^*)$.

Additional issues arise when the support of the signal distribution varies, for example, with the agent's action. In that case, uniform information fails and (5) need not hold. With uniform information, knowing the (randomized) reward that the limit contract specifies for almost every signal generated by the action a^* ties down the (randomized) reward specified for almost every signal generated by any other a in an incentive compatible manner. On the other hand, if some set of signals $\hat{S}$ has zero probability for all a in a neighborhood of a^*, then μ^* puts no weight on s, r pairs with $s \in \hat{S}$. Consequently, the conditional probability of μ^*, which forms the basis for defining the candidate limit contract, is arbitrary on $\hat{S}$, providing no guidance whatsoever in how the contract should be defined there. But, since $\hat{S}$ may receive positive probability when actions outside the neighborhood of a^* are taken, defining the contract appropriately on $\hat{S}$ is critical. It is here that the assumption of simple worst rewards is very helpful, as it allows us to specify a reward on $\hat{S}$ that, no matter what the agent's type and action, is as bad as it gets for the agent.

5 Proof of Theorem 1

For $\kappa \in K$ and $(a, t) \in A \times T$, let $\mu = \kappa * P(\cdot|a, t)$ denote the probability measure on $S \times R$ defined for all Borel subsets B of S and C of R by

$$\mu(B \times C) = \int_B \kappa(C|s) dP(s|a, t).$$

Note that $\mu \in \Delta(\Phi)$ because $\kappa(\phi(s)|s) = 1$ for every $s \in S$. We begin with a critical preliminary result.

Proposition 1 *Suppose that assumptions 1-11 hold and that the environment has either uniform information or simple worst rewards, i.e., either assumption 12 or 13 also holds. Let κ_n be a sequence of contracts in K, let a_n be a sequence of actions in A converging to $a^* \in A$, and let t_n be a sequence of types in T converging to $t^* \in T$. If $L(\kappa_n, a_n, t_n)$ is bounded above, then there is a subsequence, n_j of n, and a contract $\kappa^* \in K$ such that*

(i) *$\kappa_{n_j} * P(\cdot|a_{n_j}, t_{n_j})$ converges to $\kappa^* * P(\cdot|a^*, t^*)$*

(ii) *$\underline{\lim}_j L(\kappa_{n_j}, a_{n_j}, t_{n_j}) \geq L(\kappa^*, a^*, t^*)$,*

(iii) *$\lim_j U(\kappa_{n_j}, a_{n_j}, t_{n_j}) = U(\kappa^*, a^*, t^*) < \infty$, and*

(iv) *$\underline{\lim}_j U(\kappa_{n_j}, a', t'_j) \geq U(\kappa^*, a', t')$ for all $(a', t') \neq (a^*, t^*)$ and for all $t'_j \to t'$.*

That is, there is a contract, κ^*, such that as we pass from (κ_n, a_n, t_n) to (κ^*, a^*, t^*), the principal's losses can only fall, the agent's utility from taking the suggested action along the sequence of types

in question converges, and the agent's utility from taking any other action along a sequence of types converging to any other type can only fall.

Existence of an optimal contract for the case of pure moral hazard is immediate from this proposition, as in the sketch. It also follows for the case of pure moral hazard that if a can be implemented by some contract, then there exists a least-cost way of implementing a.

Corollary 1 *Suppose that T is a singleton and the environment has either uniform information or simple worst rewards. If an action a is optimal for the agent under some contract with finite expected losses to the principal then there is a contract that minimizes the principal's losses subject to implementing a.*

Proof of Proposition 1. By hypothesis,

$$L(\kappa_n, a_n, t_n) = \int_\Phi l(s, r, a_n, t_n) d\kappa_n(r|s) dP(s|a_n, t_n)$$

is bounded above by, say $c < \infty$.

Throughout the proof the ordered pairs (s, r) are restricted to the feasible set Φ and we omit explicit mention of this where convenient. The symbol δ_x denotes the Dirac measure placing probability one on x.

There are five steps. At Step 5, the proof differs significantly depending on which case we are in.

Step 1. We first show that the sequence $\mu_n = \kappa_n * P(\cdot|a_n, t_n)$ is tight. Fix $\varepsilon > 0$. Because S is Polish and $P(\cdot|a_n, t_n) \to P(\cdot|a^*, t^*)$, the sequence $P(\cdot|a_n, t_n)$ is a tight subset of $\Delta(S)$ by Prohorov's theorem. Consequently, there is a compact subset C of S such that $P(C|a_n, t_n) \geq 1 - \frac{\varepsilon}{2}$ for every n. Also, for every $\alpha > 0$,

$$\begin{aligned} c &\geq L(\kappa_n, a_n, t_n) \\ &= \int_\Phi l(s, r, a_n, t_n) d\kappa_n(r|s) dP(s|a_n, t_n) \\ &\geq \int_{l(s,r,a_n,t_n)>\alpha} l(s, r, a_n, t_n) d\kappa_n(r|s) dP(s|a_n, t_n) \\ &\geq \alpha \mu_n(l(s, r, a_n, t_n) > \alpha), \end{aligned}$$

for every n, where the second inequality follows because l is nonnegative. Therefore, setting $\alpha = 2c/\varepsilon$,

$$\mu_n(l(s, r, a_n, t_n) > \alpha) \leq \varepsilon/2, \text{ for every } n.$$

Letting $D_n = \{(s,r) \in \Phi : s \in C \text{ and } l(s,r,a_n,t_n) \leq \alpha\}$ and noting that $P(\cdot|a_n,t_n)$ is the marginal of μ_n on S, we have

$$\begin{aligned}\mu_n(D_n) &= 1 - \mu_n(\{(s,r) \in \Phi : s \in C \text{ and } l(s,r,a_n,t_n) > \alpha\}) - P(S\backslash C|a_n,t_n) \\ &\geq 1 - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} \\ &= 1 - \varepsilon.\end{aligned}$$

Let Y be the compact subset $C \times \{t^*, t_1, t_2, ...\}$ of $S \times T$. Clearly, the union $D_1 \cup D_2 \cup ...$ is contained in the projection onto Φ of $\{(s,r,a,t) \in \Phi \times A \times T : (s,t) \in Y \text{ and } l(s,r,a,t) \leq \alpha\}$, a compact set by assumption 7. Each μ_n therefore places weight at least $1-\varepsilon$ on the compact projection onto Φ, and so, since $\varepsilon > 0$ was arbitrary, $\{\mu_n\}$ is a tight subset of $\Delta(\Phi)$.

Step 2. We next define the requisite contract $\kappa^* \in K$ and establish (i). Since the sequence μ_n is tight, by Prohorov's theorem there is a subsequence, n_j, such that $\mu_{n_j} \to \mu^*$ for some $\mu^* \in \Delta(\Phi)$. Reindexing the subsequence, we have $\mu_n \to \mu^*$. Because $P(\cdot|a_n,t_n)$ is the marginal of μ_n on S and $P(\cdot|a_n,t_n) \to P(\cdot|a^*,t^*)$, the marginal of μ^* on S is $P(\cdot|a^*,t^*)$. Consequently, there is a regular conditional probability $\gamma(\cdot|\cdot)$ for μ^* (e.g. Dudley (2002), Theorem 10.2.2) such that for all Borel subsets B of R and D of S,

$$\mu^*(B \times D) = \int_D \gamma(B|s) dP(s|a^*,t^*),$$

where $\gamma(B|s)$ is measurable in s on S, and where (because $\mu^*(\Phi) = 1$) $\gamma(\cdot|s) \in \Delta(\phi(s))$ for all $s \in S_\gamma$, where S_γ is a measurable subset of S satisfying $P(S_\gamma|a^*,t^*) = 1$.

In the uniform information case, define $\kappa^* \in K$ as follows. For every Borel subset B of R and every $s \in S$,

$$\kappa^*(B|s) = \begin{cases} \gamma(B|s), & \text{if } s \in S_\gamma, \\ \tilde{\kappa}(B|s), & \text{otherwise,} \end{cases}$$

where $\tilde{\kappa}$ is any contract (as for example an element of $\{\kappa_n\}$).

For the simple worst reward case, let S_{a^*,t^*} be as given by assumption 10, and define $\kappa^* \in K$ as follows. For every Borel subset B of R and every $s \in S$,

$$\kappa^*(B|s) = \begin{cases} \gamma(B|s), & \text{if } s \in S_{a^*,t^*} \cap S_\gamma, \\ \delta_{r_*(s)}, & \text{otherwise.}^{21} \end{cases}$$

In either case, because $\kappa^*(B|s) = \gamma(B|s)$ for $P(\cdot|a^*,t^*)$ almost every s,

$$\mu^*(B \times D) = \int_D \kappa^*(B|s) dP(s|a^*,t^*),$$

[21] In the uniform information case, the support of P is independent of a and t, and so the only consideration there in defining the contract outside of S_γ was to preserve measurability. Here, the definition of the contract outside of $S_{a^*,t^*} \cap S_\gamma$ does not affect the payoff to type t^* from action a^*, but may well affect the payoff to other types and actions. Hence the assumption of simple worst rewards comes into play.

for all Borel subsets B of R and D of S.

Therefore, $\kappa_n * P(\cdot|a_n, t_n) = \mu_n \to \mu^* = \kappa^* * P(\cdot|a^*, t^*)$, proving (i).

Step 3. In this step we demonstrate (ii) as follows.

$$\begin{aligned}
\underline{\lim}_n L(\kappa_n, a_n, t_n) &= \underline{\lim}_n \int_\Phi l(s, r, a_n, t_n) d\mu_n \\
&= \underline{\lim}_n \int_{\Phi \times A \times T} l(s, r, a, t) d\left(\mu_n \times \delta_{(a_n, t_n)}\right) \\
&\geq \int_{\Phi \times A \times T} l(s, r, a, t) d\left(\mu^* \times \delta_{(a^*, t^*)}\right) \\
&= \int_\Phi l(s, r, a^*, t^*) d\mu^* \\
&= L(\kappa^*, a^*, t^*).
\end{aligned}$$

To see the inequality, note first that $\mu_n \times \delta_{(a_n, t_n)} \to \mu^* \times \delta_{(a^*, t^*)}$ by Billingsley (1999) Theorem 2.8 (ii), since $\mu_n \to \mu^*$ and $\delta_{(a_n, t_n)} \to \delta_{(a^*, t^*)}$. The inequality then follows from the portmanteau theorem since l is lower semi-continuous (van der Vaart and Wellner (1996), Theorem 1.3.4).

Step 4. In this step we demonstrate (iii), i.e., $\int u(s, r, a_n, t_n) d\mu_n \to \int u(s, r, a^*, t^*) d\mu^* < \infty$. Given our assumptions on u, this would follow directly from weak convergence if in addition u were bounded. However, because u need not be bounded, some extra care must be taken here.

Let $u_m(s, r, a, t) = \min(u(s, r, a, t), m)$. For every n,

$$\int u(s, r, a_n, t_n) d\mu_n = \int u_m(s, r, a_n, t_n) d\mu_n + \int_{u(s,r,a_n,t_n) > m} (u(s, r, a_n, t_n) - m) d\mu_n. \tag{7}$$

Consider the last term. Let x_m be the supremum of $u(s, r, a, t)/l(s, r, a, t)$ over all $(s, r, a, t) \in \Phi \times A \times T$ such that $u(s, r, a, t) > m$ (and where the sup of the empty set is zero). Then,

$$\begin{aligned}
\int_{u(s,r,a_n,t_n) > m} (u(s, r, a_n, t_n) - m) d\mu_n &\leq \int_{u(s,r,a_n,t_n) > m} u(s, r, a_n, t_n) d\mu_n \\
&= \int_{u(s,r,a_n,t_n) > m} \frac{u(s, r, a_n, t_n)}{l(s, r, a_n, t_n)} l(s, r, a_n, t_n) d\mu_n \\
&\leq x_m \int_\Phi l(s, r, a_n, t_n) d\mu_n \\
&= x_m L(\kappa_n, a_n, t_n) \\
&\leq x_m c.
\end{aligned}$$

Hence,

$$\int u_m(s, r, a_n, t_n) d\mu_n \leq \int u(s, r, a_n, t_n) d\mu_n \leq \int u_m(s, r, a_n, t_n) d\mu_n + x_m c, \tag{8}$$

where the second inequality follows from (7) .

To see that $\int u(s,r,a_n,t_n)d\mu_n$ is bounded, fix some m large enough that x_m is finite (such an m exists by assumption 8), and note that (8) implies that for all n,

$$\begin{aligned}\int u(s,r,a_n,t_n)d\mu_n &\leq \int u_m(s,r,a_n,t_n)d\mu_n + x_m c \\ &\leq m + x_m c.\end{aligned}$$

For any m, u_m is bounded and continuous, and so, as in step 3,

$$\int u_m(s,r,a_n,t_n)d\mu_n \to \int u_m(s,r,a^*,t^*)d\mu^*.$$

Hence, from (8) we have

$$\int u_m(s,r,a^*,t^*)d\mu^* \leq \underline{\lim}_n \int u(s,r,a_n,t_n)d\mu_n \leq \overline{\lim}_n \int u(s,r,a_n,t_n)d\mu_n \leq \int u_m(s,r,a^*,t^*)d\mu^* + x_m c.$$

Because $\lim_m x_m = 0$ by assumption 8, and because $\lim_m \int u_m(s,r,a^*,t^*)d\mu^* = \int u(s,r,a^*,t^*)d\mu^*$ by the monotone convergence theorem, we conclude that $\lim_n \int u(s,r,a_n,t_n)d\mu_n$ exists, is finite, and is equal to $\int u(s,r,a^*,t^*)d\mu^*$.

Step 5. In this step, we demonstrate (iv), i.e., $\underline{\lim}_n U(\kappa_n,a',t'_n) \geq U(\kappa^*,a',t')$ for all $(a',t') \neq (a^*,t^*)$ and for all $t'_n \to t'$. Here, the details are very different depending on the setting.

Uniform Information. Fix $(a',t') \neq (a^*,t^*)$. For every n

$$\begin{aligned}U(\kappa_n,a',t'_n) &= \int u(s,r,a',t'_n)d\kappa_n(r|s)dP\left(s|a',t'_n\right) \\ &= \int u(s,r,a',t'_n)\xi\left(s,a',t'_n,a_n,t_n\right)d\kappa_n(r|s)dP\left(s|a_n,t_n\right) \\ &= \int u(s,r,a',t'_n)\xi\left(s,a',t'_n,a_n,t_n\right)d\mu_n,\end{aligned}$$

where the second equality is valid because $S_{a',t_n} = S_{a_n,t_n} = S$ in the uniform information case. Hence,

$$\begin{aligned}\underline{\lim}_n U(\kappa_n,a',t'_n) &= \underline{\lim}_n \int u(s,r,a',t'_n)\xi\left(s,a',t'_n,a_n,t_n\right)d\mu_n \\ &\geq \int u(s,r,a',t')\xi\left(s,a',t',a^*,t^*\right)d\mu^* \\ &= \int u(s,r,a',t')\xi\left(s,a',t',a^*,t^*\right)d\kappa^*\left(r|s\right)dP\left(s|a^*,t^*\right) \\ &= \int u(s,r,a',t')d\kappa^*\left(r|s\right)dP\left(s|a',t'\right) \\ &= U\left(\kappa^*,a',t'\right),\end{aligned}$$

where the inequality follows as in step 3 using assumption 11.

Simple Worst Rewards. To keep the notation manageable in this step, for $X \subseteq S$ we shall write $\int_X$ to mean the integral over all $(r,s) \in \Phi$ with $s \in X$. For any set B and $\varepsilon > 0$, let B^ε denote the open set of points whose distance from B is strictly less than ε.

Fix $(a',t') \neq (a^*,t^*)$. Let F be a closed subset of S_{a^*,t^*}. By assumption 13,

$$\begin{aligned} U(\kappa_n,a',t'_n) &= \int u(s,r,a',t'_n)d\kappa_n(r|s)dP(s|a',t'_n) \\ &\geq \int_{S_{a_n,t_n}\cap F^\varepsilon} u(s,r,a',t'_n)d\kappa_n(r|s)dP(s|a',t'_n) + \int_{S^c_{a_n,t_n}\cup(F^\varepsilon)^c} u(s,r_*(s),a',t'_n)dP(s|a',t'_n). \end{aligned} \tag{9}$$

The second term on the right-hand side is at least

$$\int \left(1 - I_{\overline{F^\varepsilon}}(s)\right) u(s,r_*(s),a',t'_n)dP(s|a',t'_n), \tag{10}$$

since $u \geq 0$ and where $\overline{F^\varepsilon}$ denotes the closure of F^ε. By assumption 10, and the definition of ξ, the first term on the right-hand side is

$$\begin{aligned} &\int_{S_{a_n,t_n}\cap F^\varepsilon} u(s,r,a',t'_n)\xi(s,a',t'_n,a_n,t_n)d\kappa_n(r|s)dP(s|a_n,t_n) \\ =\ &\int_{F^\varepsilon} u(s,r,a',t'_n)\xi(s,a',t'_n,a_n,t_n)d\mu_n \\ =\ &\int u(s,r,a',t'_n)I_{F^\varepsilon}(s)\xi(s,a',t'_n,a_n,t_n)d\mu_n, \end{aligned} \tag{11}$$

where the first equality uses $P(S_{a_n,t_n}|a_n,t_n) = 1$.

Since F^ε is open, $I_{F^\varepsilon}(s)$ is lower semicontinuous. Consequently, since $(a',t') \neq (a^*,t^*)$, and by assumption 11, $u(s,r,a',t'_n)I_{F^\varepsilon}(s)\xi(s,a',t'_n,a_n,t_n)$ is lower semicontinuous at (s,r,a',t',a^*,t^*) for every $(s,r) \in \Phi$. Similarly, since $\overline{F^\varepsilon}$ is closed, $\left(1 - I_{\overline{F^\varepsilon}}(s)\right) u(s,r_*(s),a',t)$ is lower semicontinuous in (s,t) on $S \times T$ by assumption 13. Therefore, because $a_n \to a^*$, $t_n \to t^*$, $P(\cdot|a,t_n) \to P(\cdot|a,t^*)$, $P(\cdot|a,t'_n) \to P(\cdot|a,t')$ and $\mu_n \to \mu^*$, it follows from (10) and (11) as in step 3 that

$$\underline{\lim}_n U(\kappa_n,a',t'_n) \geq \int u(s,r,a',t')I_{F^\varepsilon}(s)\xi(s,a',t',a^*,t^*)d\mu^* + \int \left(1 - I_{\overline{F^\varepsilon}}(s)\right) u(s,r_*(s),a',t')dP(s|a',t').$$

Taking the limit of the right-hand side as $\varepsilon \to 0$, noting that both $I_{F^\varepsilon}(s)$ and $I_{\overline{F^\varepsilon}}(s) \searrow_\varepsilon I_F(s)$ for every $s \in S$, and applying the monotone convergence theorem,

$$\underline{\lim}_n U(\kappa_n,a',t'_n) \geq \int u(s,r,a',t')I_F(s)\xi(s,a',t',a^*,t^*)d\mu^* + \int (1 - I_F(s))\, u(s,r_*(s),a',t')dP(s|a',t').$$

Being a Borel probability measure on a Polish space, $P(\cdot|a^*,t^*)$ is regular. Consequently, there is an increasing sequence $F_1 \subset F_2 \subset F_3 ...$ of closed subsets of S_{a^*,t^*} such that $P(\cup_m F_m|a^*,t^*) = 1$. Substituting F_m into the right-hand side of the previous expression and noting that $I_{F_m}(s) \nearrow^m$

$I_{\cup_m F_m}(s)$ for every $s \in S$, the monotone convergence theorem implies that

$$\begin{aligned}
\underline{\lim}_n U(\kappa_n, a', t'_n) &\geq \int u(s,r,a',t') I_{\cup_m F_m}(s) \xi(s,a',t',a^*,t^*) d\kappa^*(r|s) dP(s|a^*,t^*) \\
&\quad + \int \left(1 - I_{\cup_m F_m}(s)\right) u(s, r_*(s), a', t') dP(s|a',t') \\
&\geq \int_{S_{a^*,t^*}} u(s,r,a',t') \xi(s,a',t',a^*,t^*) d\kappa^*(r|s) dP(s|a^*,t^*) \\
&\quad + \int_{S^c_{a^*,t^*}} u(s, r_*(s), a', t') dP(s|a',t') \\
&= \int_{S_{a^*,t^*}} u(s,r,a',t') d\kappa^*(r|s) dP(s|a',t') \\
&\quad + \int_{S^c_{a^*,t^*}} u(s, r_*(s), a', t') dP(s|a',t') \\
&= \int u(s,r,a,t^*) d\kappa^*(r|s) dP(s|a',t') \\
&= U\left(\kappa^*, a', t'\right),
\end{aligned}$$

where the second inequality follows because $P(S_{a^*,t^*} \setminus (\cup_m F_m))|a^*,t^*) = 0$ (first term) and because $S^c_{a^*,t^*} \subseteq (\cup_m F_m)^c$ and $u \geq 0$ (second term), where the first equality follows from the definition of ξ, and where the second equality follows because $\kappa^*(\cdot|s) = \delta_{r_*(s)}$ for $s \in S^c_{a^*,t^*}$. ∎

We now turn to the proof of Theorem 1.

Proof of Theorem 1. By hypothesis, the set of incentive compatible mechanisms $(\boldsymbol{\kappa}, \mathbf{a})$ such that $\boldsymbol{\mathcal{L}}(\boldsymbol{\kappa}, \mathbf{a})$ is finite is nonempty. We must show that $\boldsymbol{\mathcal{L}}(\boldsymbol{\kappa}, \mathbf{a})$ achieves a minimum on this set. Choose any sequence of incentive compatible mechanisms $(\boldsymbol{\kappa}^n, \mathbf{a}^n)$ such that $\boldsymbol{\mathcal{L}}(\boldsymbol{\kappa}^n, \mathbf{a}^n)$ converges to some $c < \infty$. It suffices to show that there is an incentive compatible mechanism $(\boldsymbol{\kappa}^*, \mathbf{a}^*)$ such that

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\kappa}^*, \mathbf{a}^*) \leq c. \tag{12}$$

The proof would be quite straightforward if there were a subsequence along which κ^n_t and a^n_t converged pointwise for H-a.e. $t \in T$.[22] But there need not exist any such subsequence when T is uncountable. Consequently, the proof strategy is to (step I) restrict attention to a carefully chosen countable dense subset of types and a carefully chosen subsequence for which κ^n_t and a^n_t have appropriate limits for each type t in the countable set. Then it is shown (step II) that the limit mechanism satisfies the incentive constraint when reports are restricted to the countable set. Next, it is shown (steps III and IV) that the limit mechanism can be extended to the entire set of types, T, in a measurable and incentive compatible manner, thereby producing an incentive compatible mechanism $(\boldsymbol{\kappa}^*, \mathbf{a}^*)$. Finally, owing to the carefully chosen subsequence from step I, it is shown (step V) that $(\boldsymbol{\kappa}^*, \mathbf{a}^*)$ satisfies (12).

[22] This discussion is informal. We do not define a topology on the space of contracts, K.

Step I. Because the countable intersection of measure one sets is measure one, there is a measurable subset T' of T with $H(T') = 1$ such that for all n, and for all $t \in T'$, $(\boldsymbol{\kappa}^n, \mathbf{a}^n)$ is incentive compatible at t.

For each n, and for each $t \in T$, define $L_n(t) = L(\kappa_t^n, a_t^n, t)$. Then each L_n is a $[0, +\infty]$-valued measurable function, due to the measurability properties of $\boldsymbol{\kappa}^n$ and $\mathbf{a}^n$, and because $l(\cdot)$ is lower semicontinuous and therefore measurable.

Because T' is a subset of the separable metric space T, it is itself separable metric. Hence, restricting each L_n to T', Lemma 1 (see Appendix) together with the boundedness of $\mathcal{L}(\boldsymbol{\kappa}^n, \mathbf{a}^n)$ and the compactness of A imply that there is a subsequence n_j and a countable dense subset T^0 of T' such that,

(a) $a_t^{n_j}$ converges, to $\hat{a}_t$ say, and $\lim_j L_{n_j}(t)$ exists and is finite for every $t \in T^0$, and

(b) For every $t \in T'$ there is a sequence t_m in T^0 converging to t such that

$$\lim_m \lim_j L_{n_j}(t_m) \leq \underline{\lim}_j L_{n_j}(t),$$

where the (possibly infinite) left-hand side limits in (b) exist.

By (a), the sequence $L_{n_j}(t)$ is bounded for each $t \in T^0$. Therefore, because T^0 is countable, Proposition 1 implies that there is a common subsequence n'_j of n_j such that for every $t \in T^0$,

(i′) $\mu_t^{n'_j} := \kappa_t^{n'_j} * P(\cdot|a_t^{n'_j}, t)$ converges to $\hat{\mu}_t := \hat{\kappa}_t * P(\cdot|\hat{a}_t, t)$

(ii′) $\lim_j L(\kappa_t^{n'_j}, a_t^{n'_j}, t) \geq L(\hat{\kappa}_t, \hat{a}_t, t)$,

(iii′) $\lim_j U(\kappa_t^{n'_j}, a_t^{n'_j}, t) = U(\hat{\kappa}_t, \hat{a}_t, t) < \infty$, and

(iv′) $\underline{\lim}_j U(\kappa_t^{n'_j}, a', t') \geq U(\hat{\kappa}_t, a', t')$ for all $(a', t') \neq (\hat{a}_t, t)$.

where the limit in (ii′) exists by (a).

Note that (a) holds when the sequence n_j is replaced by the subsequence n'_j because the limits remain the same. As a consequence of this, the left-hand side of (b) remains unchanged when n_j is replaced by n'_j. Therefore, because the right-hand of (b) can only increase when the sequence n_j is replaced by a subsequence, both (a) and (b) hold when n_j is replaced by n'_j. To simplify notation, reindex the subsequence n'_j as n for the remainder of the proof. Along this reindexed subsequence, $(\mu_t^n, a_t^n) \to (\hat{\mu}_t, \hat{a}_t)$ for all $t \in T^0$ and $\mathcal{L}(\boldsymbol{\kappa}^n, \mathbf{a}^n) \to c$.

For each $t \in T$, let $\underline{L}(t) = \underline{\lim}_n L_n(t)$. Then $\underline{L}$ is a $[0, +\infty]$-valued measurable function, being the liminf of a sequence of such measurable functions. By Fatou's lemma,

$$\int_T \underline{L}(t)\, dH(t) \leq \underline{\lim}_n \int_T L_n(t) dH(t) = c < \infty. \tag{13}$$

Consequently, $\underline{L}(t) < \infty$ for H-a.e. $t \in T$ and so removing from T' all of those t with $\underline{L}(t) = \infty$, we have $\underline{L}(t) < \infty$ for every $t \in T'$ and it remains the case that $H(T') = 1$. Moreover, by (a), no $t \in T^0$ is removed from T' and so $T^0 \subset T'$ continues to hold.

By (b) and (ii′),

(c) For every $t \in T'$ there is a sequence t_m in T^0 converging to t such that,

$$\underline{\lim}_m L\left(\hat{\kappa}_{t_m}, \hat{a}_{t_m}, t_m\right) \leq \underline{L}(t) < \infty.$$

Step II. In this step, we show that $((\hat{\kappa}_t, \hat{a}_t))_{t \in T^0}$ satisfies the incentive constraints for each $t \in T^0$ when reports are restricted to T^0. For all $t, t' \in T^0$ and all $a' \in A$,

$$\begin{aligned} U\left(\hat{\kappa}_t, \hat{a}_t, t\right) &= \lim_n U(\kappa_t^n, a_t^n, t) \\ &\geq \underline{\lim}_n U(\kappa_{t'}^n, a', t), \end{aligned}$$

where the equality follows from (iii′) and the inequality follows because each mechanism $(\boldsymbol{\kappa}^n, \mathbf{a}^n)$ is incentive compatible at every $t \in T^0 \subset T'$. Consequently, reversing the roles of t and t' in (iv′), we have

$$U\left(\hat{\kappa}_t, \hat{a}_t, t\right) \geq U\left(\hat{\kappa}_{t'}, a', t\right), \tag{14}$$

whenever $(a', t) \neq (\hat{a}_{t'}, t')$. But because (14) clearly holds when $(a', t') = (\hat{a}_t, t)$, (14) holds for all $t', t \in T^0$ and all $a' \in A$, as desired.

Step III. Because T is Polish H is regular. Hence, the measure of T' can be approximated arbitrarily well by the measure of compact subsets of T'. Since each such compact subset is itself Polish, we may apply Lusin's theorem to obtain a sequence $T_1, T_2, ...$ of compact subsets of T' whose union, $T^* \in \mathcal{B}(T)$, has measure one under H, and on each of which $\underline{L}$ is continuous.

For each $t \in T^*$, let $F(t)$ denote the set of (μ, a) in $\Delta(\Phi) \times A$ such that for some sequence t_j in T^0 converging to t,

(A) $\left(\hat{\mu}_{t_j}, \hat{a}_{t_j}\right) \to (\mu, a)$, and

(B) $\underline{\lim}_j L\left(\hat{\kappa}_{t_j}, \hat{a}_{t_j}, t_j\right) \leq \underline{L}(t)$.

That is, $F(t)$ is the set of all (μ, a) that are limits of pairs $\left(\hat{\mu}_{t_j}, \hat{a}_{t_j}\right)$ for types $t_j \in T^0$ near t whose associated contracts $\hat{\kappa}_{t_j}$ yield expected losses to the principal that, in the limit, are no more than $\underline{L}(t)$ when the agent takes action $\hat{a}_{t_j}$ and is type t_j. We claim that the correspondence $F : T^* \twoheadrightarrow \Delta(\Phi) \times A$ is nonempty-valued, closed-valued, and measurable.

Nonempty-valued. Fix $t \in T^*$. By (c), there is a sequence t_j in T^0 converging to t such that

$$\underline{\lim}_j L\left(\hat{\kappa}_{t_j}, \hat{a}_{t_j}, t_j\right) \leq \underline{L}(t) < \infty.$$

Therefore, by the compactness of A and by (i) of Proposition 1, there is a subsequence along which $(\hat{\mu}_{t_j}, \hat{a}_{t_j})$ converges to some $(\mu_0, a_0) \in \Delta(\Phi) \times A$. Hence, $(\mu_0, a_0) \in F(t)$.

Closed-valued. Follows similarly from a straightforward diagonal argument.

Measurable. Let C be a closed set in $\Delta(\Phi) \times A$. We wish to show that

$$F^{-1}(C) = \{t \in T^* : F(t) \cap C \neq \emptyset\}$$

is a member of $\mathcal{B}(T)$. Since $T^* = \cup T_i$,

$$F^{-1}(C) = \bigcup_{i=1}^{\infty} F_i^{-1}(C),$$

where $F_i^{-1}(C) = \{t \in T_i : F(t) \cap C \neq \emptyset\}$. It therefore suffices to show that $F_i^{-1}(C)$ is a closed subset of T for each i.

So, fix i, and let t_m be a sequence of elements of $F_i^{-1}(C)$ converging to $t_0 \in T$. We wish to show that $t_0 \in F_i^{-1}(C)$. Since T_i is compact, $t_0 \in T_i$. It remains to show that $F(t_0) \cap C$ is nonempty. For each m, since $t_m \in F_i^{-1}(C)$, there is a sequence t_{mj} in T^0 converging to t_m such that $(\hat{\mu}_{t_{mj}}, \hat{a}_{t_{mj}})$ converges to some $(\mu_m, a_m) \in C$ and $\underline{\lim}_j L(\hat{\kappa}_{t_{mj}}, \hat{a}_{t_{mj}}, t_{mj}) \leq \underline{L}(t_m)$.

For each m, we may choose j large enough so that defining $t'_m = t_{mj}$, we have t'_m within distance $1/m$ of t_m, $(\hat{\mu}_{t'_m}, \hat{a}_{t'_m})$ within distance $1/m$ of (μ_m, a_m), and $L(\hat{\kappa}_{t'_m}, \hat{a}_{t'_m}, t'_m) \leq \underline{L}(t_m) + \frac{1}{m}$.[23] Consequently, t'_m converges to t_0 and $\underline{\lim}_m L(\hat{\kappa}_{t'_m}, \hat{a}_{t'_m}, t'_m) \leq \underline{L}(t_0)$, where the inequality follows from the continuity of $\underline{L}$ on T_i. Therefore because $\underline{L}(t_0) < \infty$, we may extract a subsequence of m along which $L(\hat{\kappa}_{t'_m}, \hat{a}_{t'_m}, t'_m)$ is bounded and, by (i) of Proposition 1, also along which $(\hat{\mu}_{t'_m}, \hat{a}_{t'_m})$ converges to some (μ_0, a_0). Consequently, $(\mu_0, a_0) \in F(t_0)$. But (μ_m, a_m), being within distance $1/m$ of $(\hat{\mu}_{t'_m}, \hat{a}_{t'_m})$, therefore also converges to (μ_0, a_0) along this subsequence. Hence, because C is closed, $(\mu_0, a_0) \in C$ and we are done.

Step IV. We now define the mechanism $(\boldsymbol{\kappa}^*, \mathbf{a}^*)$ and show that it is incentive compatible. Since F is nonempty-valued, closed-valued and measurable, it follows from Wagner (1977, Theorem 4.1) and the discussion therein (p.863) that measurability of F implies weak measurability, that F has a measurable selection. That is, for each $t \in T^*$ there exists $(\mu_t^*, a_t^*) \in F(t)$ such that (μ_t^*, a_t^*), as a function from T^* into $\Delta(\Phi) \times A$, is measurable.

For each $t \in T^*$, the definition of $F(t)$ implies that there exists a sequence t_j in T^0 converging to t such that $(\hat{\mu}_{t_j}, \hat{a}_{t_j}) \to (\mu_t^*, a_t^*)$. For typical elements $t, t' \in T^*$, we denote their corresponding sequences by t_j and t'_j. By Proposition 1, for each $t \in T^*$, there is a contract $\kappa_t^* \in K$ such that,

(i″) $\hat{\mu}_{t_j} = \hat{\kappa}_{t_j} * P(\cdot|\hat{a}_{t_j}, t_j)$ converges to $\kappa_t^* * P(\cdot|a_t^*, t)$,

(iii″) $\lim_j U(\hat{\kappa}_{t_j}, \hat{a}_{t_j}, t_j) = U(\kappa_t^*, a_t^*, t) < \infty$, and

[23] The distance between $\mu_{t'_m}$ and μ_m is measured by the Prohorov metric.

(iv″) $\underline{\lim}_j U(\hat{\kappa}_{t_j}, a', t'_j) \geq U(\kappa^*_t, a', t')$ for every $(a', t') \in A \times T^* \setminus \{(a^*_t, t)\}$.

Consequently, by (i″), $\mu^*_t = \kappa^*_t * P(\cdot|a^*_t, t)$ for every $t \in T^*$. Having defined (κ^*_t, a^*_t) for each $t \in T^*$, we now extend to all of T. Fix some $t_0 \in T^*$ and let $(\kappa_0, a_0) = (\kappa^*_{t_0}, a^*_{t_0})$. For $t \in T \setminus T^*$, define $(\kappa^*_t, a^*_t) = (\kappa_0, a_0)$. Consequently, a^*_t is measurable as a function from T into A, and for every measurable subset B of R and D of S, $\int_D \kappa^*_t(B|s) dP(s|a^*_t, t)$ is measurable as a function from T into $\mathbb{R}$, because $\mu^*_t = \kappa^*_t * P(\cdot|a^*_t, t)$ is a measurable function of t into $\Delta(\Phi)$.[24] Given these measurability properties, we have therefore defined a mechanism $(\boldsymbol{\kappa}^*, \mathbf{a}^*)$. It remains to show that it is incentive compatible. Indeed, it suffices to show incentive compatibility at each $t \in T^*$ because $H(T^*) = 1$.

Reporting a type in $T \setminus T^*$ is equivalent to reporting $t_0 \in T^*$. Consequently, it suffices to show that for all $t, t' \in T^*$, and all $a' \in A$ with $(a', t') \neq (a^*_t, t)$,

$$\begin{aligned} U(\kappa^*_t, a^*_t, t) &= \lim_j U(\hat{\kappa}_{t_j}, \hat{a}_{t_j}, t_j) \\ &\geq \underline{\lim}_j U(\hat{\kappa}_{t'_j}, a', t_j) \\ &\geq U\left(\kappa^*_{t'}, a', t\right). \end{aligned}$$

But the equality follows from (iii″), the first inequality follows from (14), and the final inequality follows by reversing the roles of t and t' and the roles of t_j and t'_j in (iv″), so long as $(a', t) \neq (a^*_{t'}, t')$, or equivalently, so long as $(a', t') \neq (a^*_t, t)$, as desired.

Step V. For $t \in T^*$, $(\mu^*_t, a^*_t) \in F(t)$ implies that $(\hat{\mu}_{t_j}, \hat{a}_{t_j}) \to (\mu^*_t, a^*_t)$ and $\underline{\lim}_j L\left(\hat{\kappa}_{t_j}, \hat{a}_{t_j}, t_j\right) \leq \underline{L}(t)$ for some sequence t_j in T^0 converging to t. Consequently,

$$\begin{aligned} L(\kappa^*_t, a^*_t, t) &= \int l(s, r, a^*_t, t) d\mu^*_t \\ &\leq \underline{\lim}_j \int_{\Phi \times A \times T} l(s, r, a, t) d[\hat{\mu}_{t_j} \times \delta_{(\hat{a}_{t_j}, t_j)}] \\ &= \underline{\lim}_j L\left(\hat{\kappa}_{t_j}, \hat{a}_{t_j}, t_j\right) \\ &\leq \underline{L}(t), \end{aligned}$$

where the second line follows from the portmanteau theorem. Therefore, since $H(T^*) = 1$,

$$\begin{aligned} \mathcal{L}(\boldsymbol{\kappa}^*, \mathbf{a}^*) &= \int_T L(\kappa^*_t, a^*_t, t)\, dH(t) \\ &\leq \int_T \underline{L}(t)\, dH(t) \\ &\leq c, \end{aligned}$$

[24] To see measurability, define for each measurable $Z \in \Phi$, $g_Z : \Delta(\Phi) \to \mathbb{R}$ by $g_Z(\gamma) = \int I_Z(s, r) d\gamma(s, r)$, and define $h : T \to \Delta(\Phi)$ by $h(t) = \mu^*_t$. Then, as a real-valued function of t, $\int_D \kappa^*_t(B|s) dP(s|a^*_t, t)$ is the composition of $g_{B \times D}$ and h. By construction, h, our selection from F, is measurable. Hence, it suffices to show that g_Z is measurable. But this follows from the fact that g_Z is lower semicontinuous (and therefore measurable) for each open set Z and that the collection of Borel sets Z for which g_Z is measurable is a Dynkin class, which, containing the open sets must therefore consist of all the Borel sets as desired.

where the last inequality follows from (13). ■

6 Randomization and Optimal Contracts

In much of this section, we specialize to the case of pure moral hazard (and so drop the notation t) and provide several results relating to the use of randomization in optimal contracts. Our final result returns to the setting including adverse selection.

It is well known that contracts involving randomization over rewards may or may not be necessary for creating optimal incentives. For example, Holmström's (1979) sufficient statistic result yields as a corollary that randomization is undesirable when the agent's utility function displays risk aversion and is additively separable in effort and one-dimensional rewards, and the principal's profit function displays risk aversion and strictly decreases when higher rewards are given to the agent. See also Grossman and Hart (1983). On the other hand, Gjesdal (1982), and Arnott and Stiglitz (1988) provide examples in which randomization is desirable. The following result is an important stepping stone to a generalization (Corollary 2) of Holmström's corollary on the optimality of deterministic contracts and provides some insight into the nature of contracts when randomization is necessary for the principal to achieve optimality.

Theorem 2 *Suppose assumptions 1-6 are satisfied and that* $u(s,r,a) = v(s,r) - c(s,a)$, *so that for each* s, *utility is additively separable in rewards and actions. Suppose also that* $\kappa \in K$ *achieves a minimum expected loss for the principal subject to implementing* $a \in A$, *and that this loss is finite. If* Φ *is closed, then for* $P(\cdot|a)$*-a.e.* $s \in S$, $\kappa(\cdot|s)$ *solves*

$$\min_{\eta \in \Delta(\phi(s))} \int_R l(s,r,a)\, d\eta \tag{15}$$

subject to

$$\int_R u(s,r,a)\, d\eta = \int_R u(s,r,a)\, d\kappa(r|s).$$

That is, the principal, for each signal, chooses a least cost means of delivering the utility in question. This is intuitively obvious, as otherwise, changing the contract to deliver the same utility at each s more cheaply leaves the incentives of the agent unaffected (for each action, given the additive separability assumed) and the principal better off. The only issue is in the technicality of showing that such changes can be made in a measurable fashion as s varies. The proof is in the Appendix.

Theorem 2 permits a rather general result on the redundancy of randomization. We'll need the following definition.

Definition 5 *Say that the principal and the agent have weakly conflicting preferences over rewards if whenever*

$$u(s,r,a) > u(s,r',a)$$

for $r, r' \in \phi(s)$, *then there exists* $r'' \in \phi(s)$ *such that*

$$u(s, r'', a) = u(s, r', a) \text{ and } l(s, r'', a) < l(s, r, a).$$

Thus the principal and agent have weakly conflicting preferences over rewards if whenever the principal can reduce the agent's utility to a given level by changing the reward, he can reduce the agent's utility to the same level and decrease his losses. Weakly conflicting preferences are present in the classic moral hazard problem with one-dimensional rewards because the principal's losses strictly fall when the (monetary) reward to the agent falls. Indeed, in the classic problem the principal and the agent have diametrically opposed preferences over rewards. When the reward space is multidimensional, it would be far too restrictive to assume that their preferences are diametrically opposed, i.e., that the principal's losses are an increasing function of the agent's utility. On the other hand, weakly conflicting preferences can easily arise even when the reward space is multidimensional as the following example illustrates.

Example 14 ***Weakly Conflicting Preferences.*** *For each* $s \in S$, $\phi(s) = R = \mathbb{R}^n_+$. *For each* $a \in A$ *and each* $s \in S$, $u(s, \underline{0}, a) \leq u(s, r, a)$ *for all* $r \in R$, *and* $l(s, \cdot, a)$ *is strictly increasing coordinatewise. Weakly conflicting preferences holds because for any* $r, r' \in \mathbb{R}^n_+$ *with* $u(s, r, a) > u(s, r', a)$, *either* $l(s, r', a) < l(s, r, a)$, *and we are done, or set* r'' *as an appropriate convex combination of* r *and* $\underline{0}$.

Corollary 2 *Suppose that (i)* $R \subset \mathbb{R}^N$, *(ii)* $\phi(\cdot)$ *is convex-valued, (iii)* $u(s, r, a) = v(s, r) - c(s, a)$ *is concave in* $r \in \phi(s)$, *(iv)* $l(s, r, a)$ *is convex in* $r \in \phi(s)$, *and (v) the principal and agent have weakly conflicting preferences over rewards. Under the hypotheses of Theorem 2, if some contract* $\kappa \in K$ *achieves a finite minimum expected loss for the principal subject to implementing* $a \in A$, *then some deterministic contract does so as well. Moreover, if for* $P(\cdot|a)$*-a.e.* $s \in S$, *either* $u(s, \cdot, a)$ *is strictly concave or* $l(s, \cdot, a)$ *is strictly convex,*[25] *then the loss minimizing contract subject to implementing* a *is unique and deterministic, each up to events of* $P(\cdot|a)$ *measure zero.*

Proof of Corollary 2. For each $s \in S$ let $\bar{r}(s) = \int_R r d\kappa(r|s) \in \phi(s)$, and choose $\hat{r}(s) \in \phi(s)$ to be any certainty equivalent for the agent, i.e., such that $u(s, \hat{r}(s), a) = \int_R u(s, r, a) d\kappa(r|s)$. Because $\phi(s)$ is closed and convex, such an $\hat{r}(s)$ always exists.

Because u is concave in r,

$$u(s, \bar{r}(s), a) \geq \int_R u(s, r, a) d\kappa(r|s) = u(s, \hat{r}(s), a). \tag{16}$$

If the inequality in (16) is strict, then because preferences are weakly conflicting, there exists $\tilde{r}(s) \in \phi(s)$ such that

$$u(s, \tilde{r}(s), a) = u(s, \hat{r}(s), a)$$

[25] Which of the two alternatives occurs is allowed to depend upon $s \in S$.

and

$$l(s,\tilde{r}(s),a) < l(s,\bar{r}(s),a) \leq \int_R l(s,r,a)d\kappa(r|s),$$

where the last weak inequality follows because l is convex in r. Theorem 2 therefore implies that the inequality in (16) can be strict for at most a $P(\cdot|a)$-measure zero subset of S.[26] Consequently, $u(s,\bar{r}(s),a) = \int_R u(s,r,a)d\kappa(r|s)$ for $P(\cdot|a)$-a.e. $s \in S$ and therefore, again by Theorem 2, $l(s,\bar{r}(s),a) = \int_R l(s,r,a)d\kappa(r|s)$ for $P(\cdot|a)$-a.e. $s \in S$. These two equalities imply that the deterministic contract $\bar{r}(\cdot)$ implements a, proving the first part of the result, and Jensen's inequality applied to the two equalities proves the deterministic piece of the second part. For the uniqueness piece of the second part, simply note that if two deterministic contracts are optimal and differ on a set of $P(\cdot|a)$ positive measure, the randomized contract where after any signal a fair coin determines which of the two specified rewards is given to the agent, is also optimal, contradicting the deterministic piece of the second part. ■

Randomization can be necessary for optimality for example when risk-aversion fails over some range of rewards or when the reward and signal spaces display some discreteness. In the latter case, randomization convexifies the set of rewards.[27]

Example 15 ***Promotion.*** *The action space (effort) is $A = [0, 0.5]$, the reward space is $R = \{P, D\}$, i.e., "promote (P)" and "don't promote (D)," and the signal space is $S = \{succeed, fail\}$. Suppose $u(s,r,a) = v_r - \frac{1}{2}a^2$ and $l(s,r,a) = c_r - \rho_s$, where $\frac{1}{2} \geq v_P - v_D > 0$, $c_P > c_D = 0$, $\rho_{succeed} = 1$ and $\rho_{fail} = 0$. Finally, suppose that $\Pr(succeed|a) = a$. Any cost minimizing contract for effort level a uses D in the event of 'fail', and P with some probability τ given 'succeed'. The effort level implemented is then*

$$a = \tau(v_P - v_D), \tag{17}$$

with expected cost to the principal

$$a\tau c_P = a^2 \frac{c_P}{v_P - v_D}.$$

Hence, the principal chooses $a \in [0, 0.5]$ to maximize

$$a - a^2 \frac{c_P}{v_P - v_D},$$

so that the optimal effort level is

$$a^* = \frac{v_P - v_D}{2c_P}.$$

Comparing, with (17),

$$\tau = \frac{1}{2c_P}.$$

So, for $c_P > \frac{1}{2}$, optimization requires randomization.

[26] Note that the functions on the right-hand and left-hand sides of (16) are measurable.

[27] When preferences are not additively separable, the agent may be risk averse after one action but risk neutral after another, so that randomization is helpful in dissuading the first action. In a more general setting with a non-trivial type space, randomization can easily be optimal as a way of inducing separation between types.

Even when there is randomization, Theorem 2 has strong implications. For example, assume that $R = \mathbb{R}^N$, that rewards on the n-th coordinate can be continuously varied, and that the principal's losses are linear in that coordinate. Then, except in so much as the feasibility constraint $\phi(s)$ precludes it, the principal will equalize the marginal utility of this reward across any randomization taking place among other coordinates. The following example illustrates.

Example 16 *Modify Example 15 so that the reward space is* $\{P, D\} \times [0, \infty]$ *(for example, the principal can also pay cash,* m*). Assume that* $u(s, P, m) = v_P + 2\sqrt{m}$*, and* $u(s, D, m) = v_D + \sqrt{m}$*, where* $v_P, v_D \in \mathbb{R}$*. Let* $l(s, P, m) = c_P + m$ *and* $l(s, D, m) = c_D + m$*, where* $c_P, c_D \in \mathbb{R}$*. An optimal contract consists of the probability* τ *of promotion given success, along with a specification,* (m_P, m_D)*, of how much cash the agent receives in the event of success conditional on being promoted or not. Given* (τ, m_P, m_D) *the agent chooses*

$$a = \tau(v_P - v_D) + \tau 2\sqrt{m_P} + (1 - \tau)\sqrt{m_D},$$

and, as before, for appropriate c_P*, the principal's optimum will involve* $\tau \in (0, 1)$*. But, the concavity of utility in the monetary coordinate of utility together with Theorem 2 also imply*

$$\frac{1}{\sqrt{m_P}} = \frac{1}{2}\frac{1}{\sqrt{m_D}},$$

so that the agent's marginal utility of income, in the event of success, is equalized across whether he is promoted or not. Indeed, if this were not the case then increasing payments where marginal utility is low and decreasing them where marginal utility is high can be done in a way that maintains the same expected utility for the agent,and yet reduces the principal's losses.

Our second result on the redundancy of randomization is based upon a purification result due to Dvoretsky et. al. (1951). The version of their result used here is distinct from, although related to, the version used in the purification literature in game theory (see Milgrom and Weber (1985)).[28] We do not impose any of the assumptions from previous sections here. We require only that $u(\cdot)$ and $l(\cdot)$ are real-valued measurable functions on $\Phi \times A \times T$ and that $\phi(s) = R$ for all $s \in S$. However, in contrast to the previous results, we require that A and T are finite, and that R is compact.

Theorem 3 *If* A *and* T *are finite,* R *is a compact metric space,* $\phi(s) = R$ *for all* $s \in S$*, and* $P(\cdot|a, t)$ *is atomless for every* $(a, t) \in A \times T$*, then either the principal's problem, (1), has a solution that is deterministic, or it has no solution at all. Consequently, under the additional hypotheses of Theorem 1, the principal's problem has a deterministic solution.*

Proof: Suppose that $(\boldsymbol{\kappa}^0, \mathbf{a}^0)$ is an incentive compatible mechanism. Let $w_1(\cdot) = u(\cdot)$ and $w_2(\cdot) = l(\cdot)$. Consider first the case in which R is finite. Following Dvoretsky, Wald, and Wolfowitz (1951),

[28] For other techniques toward purification in games, see Aumann et. al. (1983).

henceforth DWW, for each $r \in R$, $a \in A$, $t' \in T$, and $n \in \{1,2\}$, let

$$\nu_{(r,a,t',n)}(B) = \int_B w_n(s,r,a,t')dP(s|a,t'),$$

for every Borel subset B of S. Then $\{\nu_{(r,a,t',n)}\}$ is a finite collection of atomless measures. Consequently, by Theorem 2.1 in DWW, for each $t \in T$ there is a deterministic contract κ_t^* such that

$$\int_S \kappa_t^*(r|s)d\nu_{(r,a,t',n)}(s) = \int_S \kappa_t^0(r|s)d\nu_{(r,a,t',n)}(s),$$

for each $r \in R$, $a \in A$, $t' \in T$, and $n \in \{1,2\}$.

Hence, for each $a \in A$, $t,t' \in T$ and $n \in \{1,2\}$,

$$\sum_{r\in R}\int_S \kappa_t^*(r|s)w_n(s,r,a,t')dP(s|a,t') = \sum_{r\in R}\int_S \kappa_t^0(r|s)w_n(s,r,a,t')dP(s|a,t').$$

But this is equivalent to

$$U(\kappa_t^*,a,t') = U(\kappa_t^0,a,t') \text{ and } L(\kappa_t^*,a,t') = L(\kappa_t^0,a,t'),$$

for each $a \in A$, and $t,t' \in T$, from which it follows that the deterministic mechanism $(\boldsymbol{\kappa}^*, \mathbf{a}^0)$ is incentive compatible and yields the same losses to the principal as the incentive compatible mechanism $(\boldsymbol{\kappa}^0, \mathbf{a}^0)$. The desired conclusions follow.

The remaining case, in which R is a compact metric space, is handled as in Section 4 of DWW.

■

7 Appendix.

Lemma 1 *Suppose that X is a separable metric space, that H is a probability measure on the Borel subsets of X whose support is X, and that g_n is a sequence of integrable $[0,+\infty]$-valued functions on X such that $\int_X g_n dH$ is bounded. Then there is a subsequence n_j and a countable dense subset X^0 of X such that,*

(i) $\lim_j g_{n_j}(x)$ *is finite for every $x \in X^0$, and*

(ii) *For every $x \in X$ there is a sequence x_m in X^0 converging to x such that*

$$\lim_m \lim_j g_{n_j}(x_m) \leq \underline{\lim}_j g_{n_j}(x).$$

It is implicit in (i) and (ii) when writing "lim" that the associated limit (which may be $+\infty$) exists.

Proof. Let τ be an upper bound for the sequence $\int_X g_n dH$. Because the support of H is X,

$H(U) > 0$ for any open subset U of X. Hence, setting $\alpha = 1/H(U)$ gives

$$\inf_{x \in U} \underline{\lim}_n g_n(x) \leq \alpha \underline{\lim}_n \int_U g_n dH \leq \alpha \underline{\lim}_n \int_X g_n dH \leq \alpha\tau < \infty, \tag{18}$$

where the first inequality uses Fatou's Lemma.

Fix a countable dense subset of X and let $U_1, U_2, ...$ be a list of the countably many open subsets of X that are open balls with rational radii centered around points in the countable dense set. We now construct X^0 inductively. First choose $x_1 \in U_1$ such that,

$$\underline{\lim}_n g_n(x_1) - 1 \leq \inf_{x \in U_1} \underline{\lim}_n g_n(x),$$

and choose $\{n_1(j)\}_j$, a subsequence of $\{n\}$, such that $\lim_j g_{n_1(j)}(x_1) = \underline{\lim}_n g_n(x_1)$.

Next, choose $x_2 \in U_2$ such that,

$$\underline{\lim}_j g_{n_1(j)}(x_2) - \frac{1}{2} \leq \inf_{x \in U_2} \underline{\lim}_j g_{n_1(j)}(x),$$

and choose $\{n_2(j)\}_j$, a subsequence of $\{n_1(j)\}_j$, such that $\lim_j g_{n_2(j)}(x_2) = \underline{\lim}_j g_{n_1(j)}(x_2)$.

Continuing this process, for each k, choose $x_k \in U_k$ such that,

$$\underline{\lim}_j g_{n_{k-1}(j)}(x_k) - \frac{1}{k} \leq \inf_{x \in U_k} \underline{\lim}_j g_{n_{k-1}(j)}(x), \tag{19}$$

and choose $\{n_k(j)\}_j$, a subsequence of $\{n_{k-1}(j)\}_j$, such that $\lim_j g_{n_k(j)}(x_k) = \underline{\lim}_j g_{n_{k-1}(j)}(x_k)$.

We claim that (i) and (ii) are satisfied by setting $X^0 = \{x_1, x_2, ...\}$ and defining $n_j = n_j(j)$. Indeed, because for each k, $\{n_j\}_{j \geq k}$ is a subsequence of $\{n_k(j)\}_{j \geq 1}$, $\lim_j g_{n_j}(x_k)$ exists and is equal to $\lim_j g_{n_k(j)}(x_k)$. Moreover, both limits are finite in view of their construction and (18). Hence, (i) is satisfied. To see that (ii) is also satisfied, consider any $x \in X$. Choose U_k containing x. Then, $\lim_j g_{n_j}(x_k) = \underline{\lim}_j g_{n_{k-1}(j)}(x_k)$ and (19) imply that,

$$\lim_j g_{n_j}(x_k) - \frac{1}{k} \leq \inf_{x' \in U_k} \underline{\lim}_j g_{n_{k-1}(j)}(x') \leq \underline{\lim}_j g_{n_j}(x), \tag{20}$$

where the second inequality follows because $\{n_j\}_{j \geq k-1}$ is a subsequence of $\{n_{k-1}(j)\}_{j \geq 1}$ and because $x \in U_k$. By considering a sequence of U_k containing x such that the associated sequence of x_k converges to x as $k \to \infty$, (ii) follows by taking the limit over an appropriate subsequence of $\{k\}$ in the left-hand side of (20).

Finally, because $\{x_1, x_2, ...\}$ is, by construction, dense in X, the result follows. ∎

Proof of Theorem 2 If (15) fails, there is a measurable subset $\hat{S}$ of S, $\delta > 0$, and (because expected losses are finite under (κ, a)) $M < \infty$, such that $P(\hat{S}|a) > 0$ and for every $s \in \hat{S}$, some $\eta_s \in \Delta(\phi(s))$ satisfies

$$\int l(s, r, a)\, d\eta_s \leq \int l(s, r, a)\, d\kappa(r|s) - \delta < M, \tag{21}$$

and

$$\int u(s,r,a)\,d\eta_s = \int u(s,r,a)\,d\kappa(r|s). \tag{22}$$

By Lusin's theorem (see, e.g., Aliprantis and Border, Theorem 12.8, p. 438), we may further suppose that $\hat{S}$ is compact and that both $\int l(s,r,a)\,d\kappa(r|s)$ and $\int u(s,r,a)\,d\kappa(r|s)$ are continuous in s on $\hat{S}$.

Let $\psi_u(s) = \int u(s,r,a)\,d\kappa(r|s)$ and let $\psi_l(s) = \int l(s,r,a)\,d\kappa(r|s) - \delta$. Define the correspondence $F : \hat{S} \twoheadrightarrow \Delta(R)$ by,

$$F(s) = \{\eta \in \Delta(\phi(s)) : \int u(s,r,a)\,d\eta = \psi_u(s) \text{ and } \int l(s,r,a)\,d\eta \leq \psi_l(s)\}.$$

If F admits a measurable selection $\hat{\eta}(\cdot|s)$, where the sigma algebra on $\hat{S}$ is $\hat{S} \cap \mathcal{B}(S)$, then we may define

$$\hat{\kappa}(r|s) = \begin{cases} \kappa(\cdot|s), & \text{if } s \in S \backslash \hat{S} \\ \hat{\eta}(\cdot|s), & \text{otherwise.} \end{cases}$$

Then $\hat{\kappa} \in K$.[29] But then, (22) and the additive separability of u imply that $\hat{\kappa}$ implements a. Yet, (21) implies that $\hat{\kappa}$ yields the principal strictly lower losses than κ - contradicting that κ is optimal. Hence, it suffices to show that F admits a selection that is $\hat{S} \cap \mathcal{B}(S)$-measurable.

Both $\hat{S}$ and $\Delta(R)$ are Polish, the latter because R is Polish (see Billingsley, (1999) Theorem 6.8, p. 73). By Wagner (1977) Theorem 4.1 and the discussion therein (p.863) that measurability of F implies weak measurability, it suffices to show that F has a closed graph.

So, suppose $\nu_n \in F(s_n)$, $s_n \to s^*$, and $\nu_n \to \nu^*$. We must show that $\nu^* \in F(s^*)$. Because $\hat{S}$ is closed, $s^* \in \hat{S}$. Because Φ is closed, $\nu^* \in \Delta(\phi(s^*))$.[30] Similar to steps 3 and 4 in the proof of Proposition 1 (but treating s_n here as a_n there) we have that (step 3) $\int l(s^*,r,a)d\nu^* \leq \underline{\lim}_n \int l(s_n,r,a)d\nu_n$ and (step 4) $\lim_n \int u(s_n,r,a)d\nu_n = \int u(s^*,r,a)d\nu^*$. Hence, the continuity of ψ_u and ψ_l on $\hat{S}$ imply that $\nu^* \in F(s^*)$. ■

References

[1] Aliprantis, C.D., and K.C. Border (2006), *Infinite Dimensional Analysis: A Hitchhiker's Guide,* 3rd Edition, Springer-Verlag.

[2] Arnott, R., and J. E. Stiglitz (1988): "Randomization with Asymmetric Information," *The Rand Journal of Economics,* 19, 344-362.

[29] The only issue is whether, for each Borel subset B of R, $\hat{\eta}(B|s)$ is a measurable real-valued function of s on $\hat{S}$. To see that it is measurable, define $g_B : \Delta(R) \to \mathbb{R}$ by $g_B(\gamma) = \gamma(B)$, and define $h : \hat{S} \to \Delta(R)$ by $h(s) = \hat{\eta}(\cdot|s)$. Then $\hat{\eta}(B|\cdot)$ is the composition of g_B and h. By construction, h, our selection from F, is measurable. Hence, it suffices to show that g_B is measurable. But this follows much as in footnote 24.

[30] Because Φ is closed, for every $\varepsilon > 0$, $\phi(s_n) \subset (\phi(s^*))^\varepsilon$ for n large enough, where $(\phi(s^*))^\varepsilon$ is the set of rewards within distance ε from $\phi(s^*)$. Hence, $\nu^*((\phi(s^*))^\varepsilon) = 1$ for every $\varepsilon > 0$, implying $\nu^*(\phi(s^*)) = 1$ since $\phi(s^*)$ is closed.

[3] Aumann, R. J., Y. Katznelson, R. Radner, R. W. Rosenthal, and B. Weiss (1983): "Approximate Purification of Mixed Strategies, *Mathematics of Operations Research*, 8, 327-341.

[4] Billingsley, P., (1999): *Convergence of Probability Measures*, 2nd Edition, Wiley Series in Probability and Statistics.

[5] Carlier, G., and R. A. Dana (2005): "Existence and Monotonicity of Solutions to Moral Hazard Problems," *Journal of Mathematical Economics,* 41, 826–843.

[6] Conlon, J. (2009): "Two New Conditions Supporting the First-Order Approach to Multisignal Principal-Agent Problems," *Econometrica,* 77, 249–278.

[7] Dudley, R. M. (2002): *Real Analysis and Probability*, Cambridge University Press.

[8] Dvoretzky, A., A. Wald, and J. Wolfowitz (1951): "Elimination of Randomization in Certain Statistical Decision Procedures and Zero-Sum Two-Person Games," *The Annals of Mathematical Statistics*, 22, 1-21.

[9] Gjesdal, F. (1982): "Information and Incentives: The Agency Information Problem," *The Review of Economic Studies*, 49, 373-390.

[10] Grossman, S., and O. Hart (1983): "An Analysis of the Principal Agent Problem," *Econometrica*, 51, 7-45.

[11] Econ. Theor)] 20 (April 1979), 231-259.

[12] Holmström, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics,* 10, 74-91.

[13] Holmström, B., and P. R. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica,* 55, 303-328.

[14] Holmström, B., and P. R. Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 7, 24-52.

[15] Jewitt, I. (1988): "Justifying the First Order Approach to Principal-Agent Problems," *Econometrica,* 56, 1177-1190.

[16] Jewitt I., O. Kadan, and J. M. Swinkels (2008): "Moral Hazard with Bounded Payments," *Journal of Economic Theory,* 143, 59-82.

[17] Laffont, J.-J., and J. Tirole (1986): "Using cost observation to regulate firms," *Journal of Political Economy,* 94, 614-641.

[18] Laffont, J.-J., and J. Tirole (1993): "A Theory of Incentives in Procurement and Regulation," MIT Press, Cambridge.

[19] Milgrom, P. R., and R. J. Weber (1985): "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research,* 10, 619-632.

[20] Mirrlees, J. (1976): "The Optimal Structure of Incentives and Authority within an Organization," *Bell Journal of Economics,* 7, 105-131.

[21] Mirrlees, J. (1999): "The Theory of Moral Hazard and Unobservable Behavior: Part I," *Review of Economic Studies,* 66, 3-21.

[22] Page, F. H. (1987): "The Existence of Optimal Contracts in the Principal-Agent Model," *Journal of Mathematical Economics,* 16, 157-167.

[23] Rogerson, W. P. (1985): "The First-Order Approach to Principal-Agent Problems," *Econometrica,* 53, 1357-1367.

[24] Sinclair-Desgagne, B. (1994): "The First-Order Approach to Multi-Signal Principal-Agent Problems," *Econometrica,* 62, 459-465.

[25] Van der Vaart, A. W., and J. A. Wellner (1996): *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer Series in Statistics.

[26] Wagner, D. H. (1977): "Survey of Measurable Selection Theorems," *Siam Journal of Control and Optimization,* 15, 859-903.

[27] Walsh, C. E. (1995), "Optimal Contracts for Central Bankers," *The American Economic Review*, 85, 150-167.