

# Robust Incentives for Information Acquisition

Gabriel Carroll

Department of Economics, Stanford University  
579 Serra Mall, Stanford, CA 94305, United States

`gdc@stanford.edu`

February 8, 2019

## Abstract

A principal needs to make a decision, and contracts with an expert, who can obtain information relevant to the decision by exerting costly effort. The principal can incentivize effort by paying a reward based on the expert's reported information and on the true state of nature, which is revealed *ex post*. Both parties are financially risk-neutral, and payments are constrained by limited liability. The principal is uncertain about the expert's information acquisition technology: she knows some actions (experiments) that he can take to obtain information, but there may also be other experiments available. The principal seeks robustness to this uncertainty, and so evaluates any incentive contract using a worst-case criterion. Under quite general conditions, we show that the optimal contract is a *restricted-investment contract*, in which the expert chooses from a subset of the decisions available to the principal, and is then rewarded proportionally to the value of his designated decision in the realized state.

JEL classification: D81; D82; D86

Keywords: information acquisition; principal-expert problem; restricted-investment contract; robustness; scoring rules; worst case

Thanks to (in random order) David Parkes, Sylvain Chassang, Bengt Holmström, Iván Werning, Alex Wolitzky, Drew Fudenberg, Ben Golub, Iván Marinovic, Yusuke Narita, Abhijit Banerjee, Daron Acemoglu, and Luis Zermeno for helpful discussions and advice, as well as several referees and many seminar and conference participants. The author is supported by a Sloan Foundation Fellowship. A one-page abstract for an earlier version of this paper was previously published in the EC 2013 conference proceedings.

# 1 Introduction

How should one pay for information that requires effort to produce? An extensive theoretical literature on proper scoring rules (e.g. Savage, 1971; Gneiting and Raftery, 2007) studies how one can incentivize agents to truthfully report pre-existing knowledge, such as beliefs about the probability of some event, under very general conditions. By contrast, the question of how best to give incentives to discover this information has not received correspondingly general attention. Yet there are plenty of situations where people are employed precisely to produce information: market researchers trying to forecast demand for new products; stock analysts, meteorologists, policy advisors, and so forth. We examine here the question of what shape and size of incentives are best for generating information, trading off the benefits of inducing effort against the fact that providing strong incentives can be expensive.

We adopt an agency model. A principal needs to make a decision that depends on an unknown state of nature, and so hires an expert, who can privately obtain information about the state by exerting effort. The expert has no intrinsic preferences over the decision being made, but the principal can incentivize him to exert effort by making his payment depend on how well the information he reports corresponds to the true state, which is publicly revealed *ex post*. We assume both parties are risk-neutral, and we impose limited liability — payments to the expert can never be less than zero.

We build on the work of Zermeño (2011, 2012), who gave an extremely general formulation of the principal-expert problem, in which the expert’s technology for acquiring information may take an arbitrary form. However, we depart from that work, and from most of the existing agency literature, by assuming that this technology is not common knowledge. Instead, as in this author’s previous work (Carroll, 2015), we take a robust-contracting approach: The principal knows some actions (here termed *experiments*) that the expert can perform to acquire information, but there may be other, unknown experiments available. The principal does not have a probabilistic belief about which experiments are and are not available. Rather, she wishes for a contract that can assure her a high expected net payoff without requiring any further knowledge about the available experiments. Accordingly, she evaluates incentive contracts based on their worst-case performance over all possible experiments the expert may have access to.

The above-mentioned previous work (Carroll, 2015) took this worst-case approach to a standard moral hazard problem, in which an agent’s action directly produces output for the principal, and the agent can be paid based on the observed output. In that simpler

setting, the concern for robustness results in linear contracts — in which the agent is paid a constant share of output — being optimal. Briefly, the intuition is as follows: the principal evaluates contracts based on the worst-case guarantee they provide on her expected payoff, and her limited knowledge provides only a worst-case guarantee on the agent’s expected payoff (via the actions she knows he has available); linearity provides a tight connection between these two.

In the present problem, a linear contract could be defined as follows: the expert recommends a decision, and is rewarded with some fixed fraction of the resulting payoff. The same forces leading to linearity in the simpler model apply here too. But the problem here is more complex, and it turns out the optimal contract is a variant which we call a *restricted-investment contract*. Roughly, instead of paying the expert a constant share of the principal’s own realized payoff, she allows him to choose an “investment” from a certain subset of her possible decisions, and pays him based on a fraction of the payoff that *would* have accrued (in the realized state) if she had made the decision thus designated by the expert. By excluding extreme decisions from the allowed investments, the principal can make the limited liability constraint slack, allowing her to pay less to the expert, while preserving enough of the linear connection between the expert’s payoff and her own to still provide the robust guarantee.

This paper aims to make two main points. One is descriptive: we give a recipe that can be used to write contracts in a particular agency setting, and the resulting contracts are optimally robust in a precise sense. As with most economic models, the model is not meant to be taken completely literally, but rather to deliver qualitative insights — the robustness of linear incentives, together with the benefits of excluding extreme decisions from the payment scheme — that seem relevant more generally. (The conclusion contains some further discussion on interpretation of the model.)

The other, broader point is methodological: we show how using a maxmin objective leads to a tidy and tractable model. By contrast, a traditional Bayesian approach, where the principal knows the expert’s information acquisition technology (or has a probabilistic belief about it), is unlikely to be tractable without much more specific functional form assumptions, e.g. binary states and one-dimensional effort choice by the expert. (Some discussion of previous related work appears below.) Here, we can give results on the shape of the optimal contract that are quite general, with essentially *no* structural assumptions needed on the set of known experiments.

Beyond the worst-case criterion, the other key assumption in our model is that the state is fully revealed ex post, regardless of the decision made. Thus, the appropriate

applications for our model are situations where the expert is predicting some event that will be publicly verifiable and is exogenous to the principal’s decision. For example, the expert may be a political analyst, hired by a firm to predict whether a government will enact some regulation that is currently under consideration; the firm’s business decisions depend on the predicted outcome, and afterwards it can be clearly observed whether the regulation was in fact enacted. Or the expert is a sports analyst, trying to predict a team’s chances of winning a game in order to help a betting house set the correct odds. Or, the expert is a public health researcher, predicting whether some new epidemic is likely to strike a particular country, in order to help the government decide how much to stock up on medicines; this fits the model as long as the disease’s arrival in the country is not causally affected by the government’s decision.

The full revelation assumption allows us to separate the problem of choosing the decision from that of providing incentives to the expert. This distinguishes our work from some prior literature on incentives for experts, including Zermeno (2011) as well as the computer science literature on decision markets (Othman and Sandholm, 2010; Chen, Kash, Ruberry, and Shnayder, 2011), in which optimal provision of incentives involves distorting decisions so as to reveal more information about the state. In the model here, the principal’s decision plays no role in the incentives faced by the expert. Nonetheless, we model the decision problem explicitly because it determines how information is valued by the principal.

In the next section, we present the formal model, and also present a stylized application that will serve as a running example. We then give a more detailed discussion of the intuition behind restricted-investment contracts, before proceeding to the formal analysis showing that such contracts are optimal. The main proof is essentially an application of the same linear separation techniques used in Carroll (2015). In Section 3, we discuss a number of extensions and variations. In particular, we discuss when an *unrestricted*-investment contract — which, up to a normalization, is just a linear contract — is or is not optimal; this generally can happen when the principal’s optimal decision is not too sensitive to the posterior, and in particular always holds if the decision problem is binary. The conclusion contains some additional discussion of interpretation and of the relation to realistic incentive schemes.

There is a variety of previous work on agency problems that considered incentives to acquire information, besides the work of Zermeno (2011, 2012) cited above. Early forerunners include Demski and Sappington (1987) (who introduced the term “expert”) and Malcolmson (2009); a recent variant by Lindbeck and Weibull (2017) is based on

rational inattention modeling. Unlike the present work, these models assume that the decision is delegated to the expert, and that only the realized payoff is observed, not the entire state. More recent years have seen much interest in dynamic contracts to incentivize experimentation (Bergemann and Hege, 1998; Manso, 2011; Gerardi and Maestri, 2012; Hörner and Samuelson, 2013; Halac, Kartik, and Liu, 2016). These works typically impose a very stylized structure (e.g. two states, binary signals, binary action choice for the agent), and study the dynamics of the optimal contract. Here, we consider one-shot information acquisition, but allow a rich state space and set of experiments, and impose no structural assumptions at all on the set of *known* experiments. Chassang (2013) considered a quite general model of information acquisition in a dynamic setting, also with very few structural assumptions, but with a different focus: the emphasis there was on showing that limited liability becomes approximately non-binding with a long time horizon, whereas here we focus on exactly optimal contracts in a one-shot problem. Finally, the scoring rules literature has also recognized at least informally that higher stakes give more incentives to acquire information (e.g. Karni, 2009), but without actually formulating the optimal contracting problem.

In addition to these strands of literature, the present paper also contributes to the growing literature on mechanism design using a maxmin objective in uncertain environments (e.g. Bergemann and Morris, 2005; Chung and Ely, 2007; Frankel, 2014; Garrett, 2014). The earlier paper (Carroll, 2015) contains more discussion of this literature, and of the interpretation of the worst-case objective.

There is other, less related literature on expert advice in economics, which focuses on other aspects of the problem. Most notably, a large literature beginning with Crawford and Sobel (1982) studies the incentive problems in reporting when the expert has intrinsic preferences over the decision being chosen; this issue is not part of our model.

## 2 Model and results

### 2.1 Notation

For  $X \subseteq \mathbb{R}^k$ ,  $\Delta(X)$  is the space of Borel probability distributions on  $X$ , with the weak topology. For  $x \in X$ , we write  $\delta_x$  for the degenerate distribution putting probability 1 on  $x$ . Expectations will often be written with the distribution subscripted; for example,  $E_F[g(x)]$  for the expectation of  $g(x)$  when  $x$  is drawn from distribution  $F$ .  $\mathbb{R}^+$  is the set of nonnegative real numbers.

## 2.2 The setup

As already indicated, the model is based on that of Zermeño (2011), with the main difference being the principal’s non-quantifiable uncertainty about the expert’s information acquisition technology.

The principal needs to choose a decision from some compact space  $D$ . The payoff to each decision depends on a state of nature, which will be realized in the future. We assume the set  $\Omega$  of possible states is finite. Payoffs are represented by a continuous function  $u : D \times \Omega \rightarrow \mathbb{R}$ . We will sometimes refer to  $u$  as the principal’s *gross* payoff (to distinguish it from the *net* payoff, which also reflects payments to the expert, below). We may refer to the triple  $(D, \Omega, u)$  as the principal’s *decision problem*. The principal also has a prior belief about the state,  $p_0 \in \Delta(\Omega)$ , with  $p_0(\omega) > 0$  for each  $\omega \in \Omega$ .

Before the principal makes her decision, she can hire the expert to obtain information about the state. The expert initially shares the prior  $p_0$ , but can obtain more information by performing an *experiment*. One can think of an experiment as producing a signal, observed by the expert; signals and states can follow any joint distribution such that the marginal distribution over states is  $p_0$ . However, the signal will matter only through the expert’s resulting posterior belief about the state, so we take the notational shortcut of representing experiments directly in terms of posteriors. Thus, we define an *experiment* to be a pair  $(F, c) \in \Delta(\Delta(\Omega)) \times \mathbb{R}^+$ , such that  $F$  has mean  $p_0$ . The interpretation is that the expert can, at a cost of  $c$ , perform the experiment, and obtain a posterior (an element of  $\Delta(\Omega)$ ) drawn from distribution  $F$ . The requirement that  $F$  should have mean  $p_0$  is simply the law of iterated expectations — the posterior should, in expectation, be equal to the prior. We will typically use the variable  $p$  for a posterior.

We give  $\Delta(\Delta(\Omega)) \times \mathbb{R}^+$  the natural product topology, and define an *information acquisition technology* (IAT) to be a nonempty, compact subset of  $\Delta(\Delta(\Omega)) \times \mathbb{R}^+$ , with every element  $(F, c)$  satisfying  $E_F[p] = p_0$ . An IAT then describes the set of experiments available to the expert. There is some exogenously given IAT,  $\mathcal{I}_0$ , consisting of the experiments which are mutually known to be available. From the principal’s point of view, the true set of experiments (known only to the expert) may potentially be any IAT  $\mathcal{I}$  such that  $\mathcal{I}_0 \subseteq \mathcal{I}$ . We will sometimes be specifically interested in  $\mathcal{I}_0$  that satisfy the following *full-support condition*: for every  $(F, c) \in \mathcal{I}_0$ , the support of  $F$  is all of  $\Delta(\Omega)$ , or else  $F = \delta_{p_0}$ .

After the contract is chosen, the expert can acquire information, and make a report to the principal, who then chooses a decision  $d \in D$ . After the decision is made, the

true state is revealed, and payments can be made contingent on all relevant observable information (the report and the realized state). Limited liability means that payments can never be negative. Thus, a *contract* is a pair  $(M, w)$ , consisting of a *message space*  $M$ , some nonempty compact space; and a *payment function*  $w : M \times \Omega \rightarrow \mathbb{R}^+$ , which must be continuous. (The topological assumptions ensure that the expert’s behavior is well-defined.)

We could potentially apply the revelation principle and assume that the expert’s message is simply his posterior belief (this approach is used in Zerméño (2011, 2012), and is also in keeping with the proper scoring rules literature). However, we will not do so, because it will be convenient during the analysis to study contracts formulated as indirect mechanisms.

On the other hand, we can safely assume that the principal chooses the decision that is optimal given the posterior  $p$  attained by the expert. This assumption requires a bit of justification. On its face, it seems inappropriate given the way we have modeled contracts thus far: if the message space  $M$  is small, then there may be many different posteriors  $p$  for which the expert chooses the same message  $m$ , and so this message may not be informative enough for the principal to infer what her optimal decision is. But any such “pooling” contract can be (weakly) improved upon as follows: expand the message space by having the expert also report his posterior  $p$ , in addition to the original message  $m$ ; but for purposes of calculating payment, the posterior report is ignored. Then the expert’s incentives for information acquisition, and for the  $m$  component of his report, are exactly the same as in the original contract; and he is indifferent about the  $p$  component of his report, so he is willing to report the true posterior, which clearly allows the principal to choose the best possible decision  $d$ .<sup>1</sup> Because we are concerned specifically with optimal contracts for the principal, we may as well assume she uses this modified contract.

To make this argument fully explicit would require some digression: we would need to specify the principal’s payoffs, which depend not only on the contract  $(M, w)$  as modeled above but also on how she maps messages she receives into decisions. These details are spelled out in Appendix A. For the rest of the main paper, we just assume that the message  $m$  is informative enough for the principal to identify the optimal decision given the expert’s posterior  $p$ . Whenever we discuss contracts that lack this property, they

---

<sup>1</sup>A referee observes that since this expansion may make the message space much larger, it may not be innocuous if communication is slightly costly: since the expert’s payment does not depend on what  $p$  he reports, he would claim whichever  $p$  is easiest to communicate, instead of the true  $p$ . Explicit modeling of communication costs would take us far afield, so for now, let us simply note that the indifference can be broken by adding  $\epsilon \cdot s(p, \omega)$  to the payment function, where  $s$  is any proper scoring rule.

should be understood as being shorthand for the more communicative contracts described above.

We can now summarize the timing of the game:

1. the principal offers a contract  $(M, w)$ ;
2. the expert, knowing  $\mathcal{I}$ , chooses experiment  $(F, c) \in \mathcal{I}$ ;
3. a posterior  $p \sim F$  is realized, and privately observed by the expert;
4. the expert chooses a message  $m \in M$  to send;
5. the principal chooses decision  $d$  (optimally given the posterior  $p$ );
6. the state  $\omega$  is revealed;
7. net payoffs accrue:  $u(d, \omega) - w(m, \omega)$  for the principal;  $w(m, \omega) - c$  for the expert.

It remains only to describe behavior. We give a brief summary here, and will introduce formal notation shortly. The expert knows  $\mathcal{I}$ , and he chooses  $(F, c)$  and then  $m$  to maximize his expected payoff. The above behavior by the expert gives rise to an expected payoff for the principal for each possible IAT  $\mathcal{I}$ ; we will notate this expected utility by  $V_P(M, w|\mathcal{I})$ . The principal evaluates each contract  $(M, w)$  in terms of the guarantee it can provide on  $V_P$ , across all possible IAT's  $\mathcal{I}$ , knowing only that  $\mathcal{I}_0 \subseteq \mathcal{I}$ . Finally, the question is how to design a contract to maximize this worst-case value.

Before developing the formal notation, we introduce some other useful objects. Suppose that the expert reaches a posterior  $p \in \Delta(\Omega)$ . Then, as above, the principal will choose  $d$  to maximize  $E_p[u(d, \omega)]$ . We denote this decision and the expected (gross) payoff by

$$d(p) = \arg \max_{d \in D} E_p[u(d, \omega)]; \quad U(p) = \max_{d \in D} E_p[u(d, \omega)].$$

(If there are several maximizers then  $d(p)$  can be chosen arbitrarily among them.) Note that  $U$  is a convex function on the simplex  $\Delta(\Omega)$ , since it is the maximum of the affine functions  $p \mapsto E_p[u(d, \omega)]$ , one such function for each  $d$ .

Similarly, when the expert has posterior  $p$ , he will choose message  $m$  so as to maximize  $E_p[w(m, \omega)]$ . We denote

$$W(p) = \max_{m \in M} E_p[w(m, \omega)],$$

and call this function  $W : \Delta(\Omega) \rightarrow \mathbb{R}^+$  the *reduced form* of the given contract  $(M, w)$ .  $W$  is likewise convex in  $p$ ; it is the upper envelope of the nonnegative-valued affine functions

given by  $p \mapsto E_p[w(m, \omega)]$ , one for each  $m \in M$ . As we shall see, the expert's incentives to acquire information depend only on the reduced form of the contract.<sup>2</sup>

Now we can describe behavior formally. The expert, given posterior  $p$ , chooses message  $m$  as above, leading to expected payment  $W(p)$ . At the earlier, experiment-choosing stage, he knows the IAT  $\mathcal{I}$  and so chooses  $(F, c) \in \mathcal{I}$  to maximize expected payoff  $E_F[W(p)] - c$ . We will write the value of the contract  $(M, w)$  to the expert as

$$V_E(M, w|\mathcal{I}) = \max_{(F, c) \in \mathcal{I}} (E_F[W(p)] - c)$$

and the expert's set of optimal experiments as

$$I^*(M, w|\mathcal{I}) = \arg \max_{(F, c) \in \mathcal{I}} (E_F[W(p)] - c).$$

When the principal learns the posterior  $p$ , she will make decision  $d(p)$ , gaining expected gross payoff  $U(p)$ . Thus, her expected net payoff from the contract under IAT  $\mathcal{I}$  is

$$V_P(M, w|\mathcal{I}) = \max_{(F, c) \in I^*(M, w|\mathcal{I})} (E_F[U(p)] - W(p)).$$

(The maximization reflects the fact that the expert may be indifferent between several optimal experiments. We assume he then chooses the one that is best for the principal. This is in line with the convention in contract theory whereby a contract specifies a “recommended action.”<sup>3</sup>)

Ex ante, the principal evaluates each possible contract by its worst-case guarantee on her expected payoff, over all IAT's  $\mathcal{I}$ :

$$V_P(M, w) = \inf_{\mathcal{I} \supseteq \mathcal{I}_0} V_P(M, w|\mathcal{I}).$$

The principal's problem, which we analyze, is then how to choose the contract  $(M, w)$  to maximize  $V_P$ .

From here on, we will maintain the *non-triviality* assumption that there exists some contract  $(M, w)$  with  $V_P(M, w) > U(p_0)$ . That is, the principal benefits from hiring the

---

<sup>2</sup>In the language of Laffont and Tirole (1993, Section 1.4), the function  $w(m, \omega)$  “decentralizes” the incentives captured by the upper envelope  $W(p)$ .

<sup>3</sup>Under alternative assumptions, e.g. breaking ties in the way that is worst for the principal, the results would be substantively unchanged but involve more technicalities. In particular, an optimal contract may not exist, but the supremum would be approached by perturbations of the contract that is identified as optimal in the present model. This issue is also discussed in Carroll (2015).

expert. We shall shortly see conditions on primitives that ensure that this assumption is satisfied. (This assumption is not formally needed for our results. But if it is not satisfied, the problem is uninteresting — the optimal contract is clearly always to pay zero.)

Below is a stylized example to illustrate the environment. It will be analyzed further in later sections.

*Example 2.1.* Suppose that the principal runs a firm that operates in two countries,  $A$  and  $B$ . The firm has 1 unit of capital to divide between the two countries. Its earnings in each country is an increasing, concave function of the capital invested there. Specially, let us assume that the earnings from operating in a country after investing an amount  $x \in [0, 1]$  of capital there is given by  $\sin(\frac{\pi}{2}x)$ .

However, the government of  $B$  is considering enacting regulations that would make it impossible for the firm to operate there, thus dropping earnings from  $B$  to zero. The firm needs to plan its operations before the uncertainty is resolved. There are thus two possible states,  $\omega_R$  (regulation) and  $\omega_{NR}$  (no regulation). If we refer to the decision as the amount to invest in country  $A$ , then the decision is  $d \in [0, 1]$ , and the decision problem is described by  $\Omega = \{\omega_R, \omega_{NR}\}$  and

$$u(d, \omega_R) = \sin\left(\frac{\pi}{2}d\right), \quad u(d, \omega_{NR}) = \sin\left(\frac{\pi}{2}d\right) + \sin\left(\frac{\pi}{2}(1-d)\right).$$

These payoffs are shown in Figure 1, panel (a). Notice that decisions  $d < 1/2$  are dominated, so henceforth we regard  $D = [1/2, 1]$  as the relevant decision space.

The firm can hire a political analyst (the expert) to gather information about whether the regulation is likely, and thereby improve its investment decision. Suppose that the prior is a  $1/2$  probability of state  $\omega_R$ . (A belief is simply a probability placed on state  $\omega_R$ , so we can identify the space of beliefs with the unit interval.) As far as the firm is aware, the analyst can either do nothing (and learn nothing), or, by incurring a cost of  $c^* = 1/1200$ , can acquire a signal whose value is either  $R$  or  $NR$ . Conditional on the true state, the signal is correct with probability  $3/4$ . Thus, this generates a distribution  $F^*$  over posteriors that places mass points at  $1/4$  and  $3/4$ , of size  $1/2$  each. The analyst may also have other ways of acquiring information, which are unknown to the firm. So, in this example,  $\mathcal{I}_0 = \{(\delta_{p_0}, 0), (F^*, c^*)\}$ .

We can easily calculate the function  $U(p)$  in this example. Given any posterior  $p$ , the firm would choose  $d$  to maximize  $pu(d, \omega_R) + (1-p)u(d, \omega_{NR}) = \sin\left(\frac{\pi}{2}d\right) + (1-p)\sin\left(\frac{\pi}{2}(1-d)\right)$ . A quick calculation finds that the optimum is  $d(p) = (\cot^{-1}(1-p))/\left(\frac{\pi}{2}\right)$ ,

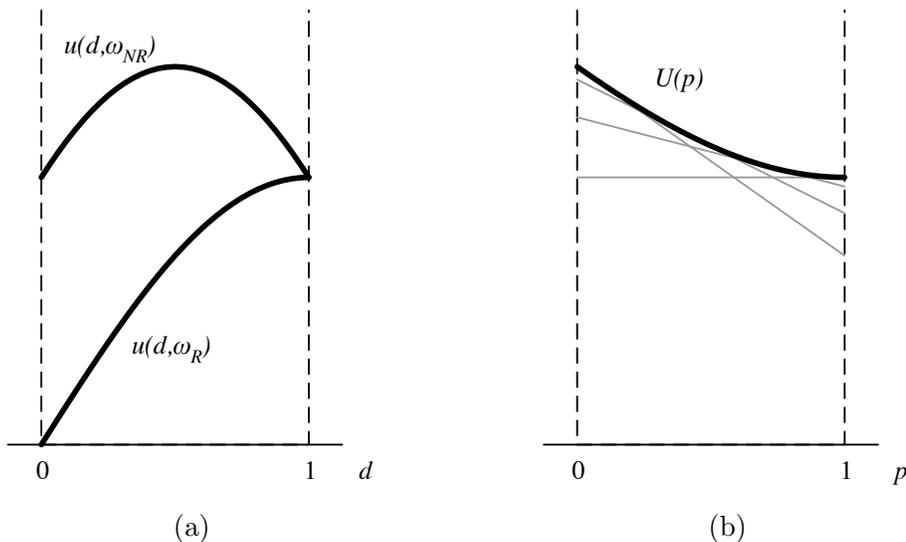


Figure 1: Decisions in Example 2.1. (a) Payoff to each decision in each state. (b) Sample decisions, and optimal payoff, plotted as a function of  $p$ .

and the resulting utility is

$$U(p) = \sin(\cot^{-1}(1-p)) + (1-p) \cos(\cot^{-1}(1-p)) = \sqrt{1 + (1-p)^2}.$$

This is shown in panel (b) of Figure 1. The horizontal axis is now the posterior  $p$ . Each possible decision  $d$  is illustrated by a thin line, plotting the firm's expected gross payoff as a function of  $p$ . The upper envelope of these lines is  $U(p)$ , shown by the dark curve.

As we will verify later, this example satisfies non-triviality. It does not satisfy the full-support condition on  $\mathcal{I}_0$ .  $\triangle$

### 2.3 Intuitions

How can our principal write a contract to guarantee herself a reasonably high expected payoff? A natural first try would be to use a linear contract: The expert recommends a decision; thus, the space of possible messages equals the decision space,  $M = D$ . The principal follows the recommendation and then pays the expert some fixed fraction  $\alpha \in (0, 1]$  of her gross payoff.

For such a contract, we could compute a payoff guarantee exactly as in Carroll (2015) (see also Chassang, 2013): For example, consider a contract in which the principal gives 1/3 of her gross payoff to the expert, and suppose the expert has some *known* experiment that would give him a total expected payoff of 100 under this contract. Then, the expert would not choose any other, unknown experiment unless it gave him at least 100 expected payoff. Since the principal gives the expert 1/3 of her gross payoff and keeps 2/3 for herself, her net payoff is at least twice his. (This inequality may be strict, since the expert also suffers the disutility of effort.) Hence, the principal is guaranteed at least 200 in expectation. More generally, for a linear contract with share  $\alpha$ , this logic shows the principal is guaranteed  $(1 - \alpha)/\alpha$  times the payoff that the expert gets from his best experiment in  $\mathcal{I}_0$ .

Actually, the above is not quite the right way to define linear contracts in our setting: Since the principal's gross payoffs  $u(m, \omega)$  may be negative (in which case the contract above violates limited liability) or may all be very high (in which case limited liability is slack), we should first add some constant adjustment  $\beta \in \mathbb{R}$ . Thus, a more proper definition of linear contracts in our formalism is given by

$$M = D, \quad w(m, \omega) = \alpha(u(m, \omega) + \beta). \quad (2.1)$$

When the expert has posterior  $p$ , she sends a message describing her recommended decision,  $m = d(p)$ . The resulting reduced form of the contract is

$$W(p) = \alpha(U(p) + \beta).$$

Explicitly, the optimal choice of  $\beta$  in such a contract would be such that limited liability just binds:  $\beta = -\min_{d, \omega} u(d, \omega)$ .

It is easy to incorporate this  $\beta$  adjustment into our calculation of the guarantee from such a contract. Let us write out the algebra explicitly.

If the expert chooses experiment  $(F, c)$ , his expected payoff will be  $E_F[W(p)] - c = \alpha(E_F[U(p)] + \beta) - c$ . Hence we have

$$\alpha(E_F[U(p)] + \beta) \geq \alpha(E_F[U(p)] + \beta) - c = V_E(M, w|\mathcal{I}) \geq V_E(M, w|\mathcal{I}_0).$$

Here the inequality at the end comes from the fact that  $\mathcal{I} \supseteq \mathcal{I}_0$ , so the expert certainly does at least as well with  $\mathcal{I}$  as  $\mathcal{I}_0$ .

Whenever the expert produces a gross payoff of  $y$ , he receives  $\alpha(y + \beta)$ , and the

principal receives  $(1 - \alpha)y - \alpha\beta$ . Hence, the principal's expected payoff will be

$$\begin{aligned}
(1 - \alpha)E_F[U(p)] - \alpha\beta &\geq \frac{1 - \alpha}{\alpha}V_E(M, w|\mathcal{I}_0) - \beta \\
&= \frac{1 - \alpha}{\alpha} \left[ \max_{(F, c) \in \mathcal{I}_0} \alpha(E_F[U(p)] + \beta) - c \right] - \beta \\
&= \max_{(F, c) \in \mathcal{I}_0} \left( (1 - \alpha)E_F[U(p)] - \frac{1 - \alpha}{\alpha}c \right) - \alpha\beta. \quad (2.2)
\end{aligned}$$

Thus the guarantee from the linear contract,  $V_P(M, w)$ , is at least the right-hand side of (2.2).

At this point, we can fulfill a promise made earlier: to give a sufficient condition on primitives to ensure the non-triviality assumption is satisfied. Let  $F$  be any distribution over posteriors such that  $E_F[U(p)] > U(p_0)$  — that is, any distribution that potentially provides useful information for the decision. One can check that, as long as

$$c < \left( \sqrt{E_F[U(p)] + \beta} - \sqrt{U(p_0) + \beta} \right)^2, \quad (2.3)$$

there exists  $\alpha$  such that the right side of (2.2) is greater than  $U(p_0)$ . Hence, if  $\mathcal{I}_0$  contains some such  $(F, c)$ , non-triviality holds.

Now, the logic so far is identical to that of the pure moral hazard setting in Carroll (2015). In that simpler setting, linear contracts were optimal: the only way to turn a guarantee on the agent's expected payoff (provided by the known technology) into a guarantee on the principal's payoff was to have a linear relationship between the two. Here, however, there is scope for improving on linear contracts, in two ways. To illustrate, we return to our firm-analyst example.

*Example 2.1, continued.* Recall the firm's decision problem as illustrated in Figure 1(b).

A starting point would be a linear contract. The analyst acquires information and then recommends an optimal decision  $d \in D = [1/2, 1]$ ; then he is paid, say,  $3/4$  of the firm's gross payoff. Actually, since all such decisions pay at least  $\sqrt{2}/2 \approx 0.707$  in each state, the firm can first subtract this quantity and then pay the analyst  $3/4$  of the difference. This is shown in Figure 2(a). The figure parallels the illustration of the firm's decision problem, but now the thin lines correspond to examples of messages the analyst can send — each line showing the analyst's expected payment from sending the given message, as a function of his posterior — and their upper envelope is the reduced form  $W(p)$ . ( $U(p)$  is shown in gray for comparison.)

Now, one simple way to improve on such a contract is to replace the single constant adjustment  $\beta = -\sqrt{2}/2$  by an adjustment that depends on the state,  $\beta(\omega)$ . Explicitly, instead of (2.1), we define the contract by

$$M = D, \quad w(m, \omega) = \alpha(u(m, \omega) + \beta(\omega)). \quad (2.4)$$

Note that this change can have no effect on the analyst's incentives as to which experiment to perform, because he has no control over which state arises and so which adjustment he receives. This generalization gives the firm more degrees of freedom, to reduce the amount paid to the analyst without changing incentives. The optimal choice of state-by-state adjustments is such that limited liability binds in each state separately:  $\beta(\omega) = -\min_d u(d, \omega)$ . In our example, we have  $\beta(\omega_R) = -\sqrt{2}/2$  and  $\beta(\omega_{NR}) = -1$ . This change in the contract is illustrated in Figure 2(b). The gray arrows emphasize that in each state, the lowest message now pays 0.

The second reason why linear contracts may not be optimal is more substantive: By cutting risky decisions out of the message space, the firm can relax the limited liability constraint, and thus further lower  $\beta$ . This is illustrated in Figure 2(c), which differs from (b) in that some of the messages corresponding to high and low choices of the decision  $d$  have been removed. Thus we keep only the messages representing an intermediate range of decisions,  $D' \subseteq D$ , whose upper envelope is shown by the black curve. (The light lines represent the extreme decisions within this range.) Notice that now there is once again slack in the limited liability constraint. Consequently, the firm can further adjust payments downward by a non-negligible amount in each state before limited liability binds, and thereby further save money.

Now, cutting out the extreme decisions makes  $W$  less convex, and so may weaken the analyst's incentives for information acquisition. However, if the known experiments rarely or never yield extreme posteriors, then this effect on incentives is small, and the effect of lowering  $\beta$  dominates. Note that under the new contract, the firm's *decisions* would not be restricted — if the analyst happened to reach an extremely high or low posterior, the firm could still choose a correspondingly extreme decision. But the analyst's payment would be based on the closest decision from the intermediate range  $D'$ .

In our example, we could describe the resulting contract verbally as follows. The contract is parameterized by share  $\alpha$  and decision space  $D' = [\underline{d}, \bar{d}] \subseteq [1/2, 1]$ . The analyst acquires information, and recommends an optimal decision  $d = d(p)$ . (In this example, we need not also include an explicit report of  $p$ , since  $d$  already tells the firm

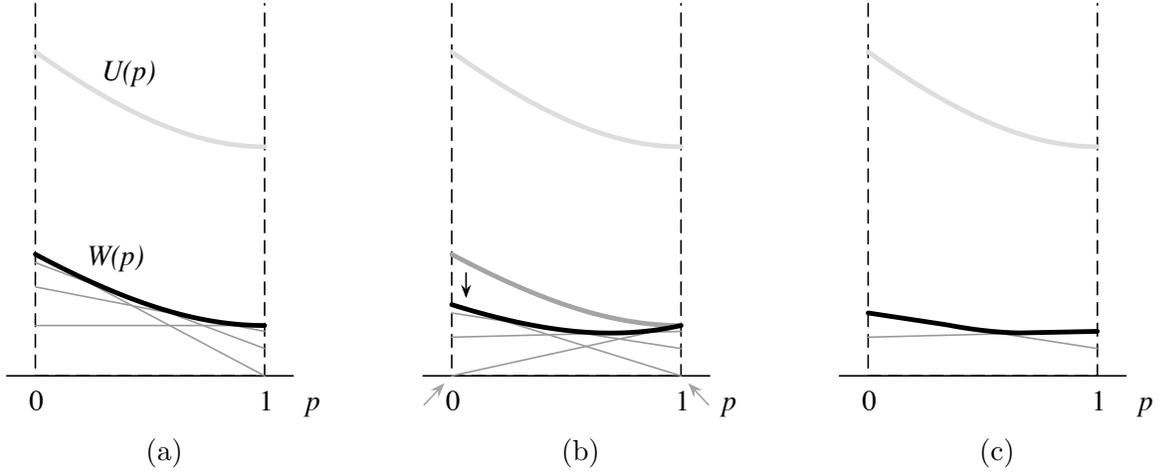


Figure 2: Intuition for improving on linear contracts. (a) A linear contract in Example 2.1. (b) State-by-state adjustment. (c) Trimming extreme messages.

the optimal decision.) If the analyst recommends  $d \in [\underline{d}, \bar{d}]$ , then if state  $\omega_R$  realizes, he is paid a share  $\alpha$  of the part of the firm's gross profit that exceeds a threshold; and if state  $\omega_{NR}$  realizes, he is paid a share  $\alpha$  of the profit exceeding a (different) threshold. These thresholds are set so that the analyst's payment from  $\underline{d}$  in state  $\omega_R$ , or  $\bar{d}$  in state  $\omega_{NR}$ , just hits zero. Finally, if he recommends a decision  $d < \underline{d}$ , the firm takes this decision but pays him as if he had recommended  $\underline{d}$ ; similarly, if  $d > \bar{d}$  then he is paid as if he had recommended  $\bar{d}$ .<sup>4</sup>  $\triangle$

Leaving the example and returning to the general setup, this idea leads to the notion of a *restricted-investment contract*, defined formally as follows: The message space is some compact subset of decisions,  $M \subseteq D$ , and the payment function is given by

$$w(m, \omega) = \alpha(u(m, \omega) + \beta(\omega)) \quad (2.5)$$

for some  $\alpha \in [0, 1]$  and some  $\beta : \Omega \rightarrow \mathbb{R}$ , such that the resulting payments are always nonnegative. The name emphasizes that the expert is allowed to “invest” in a restricted subset of decisions, and once the state is revealed, he is compensated according to the payoff that his chosen decision would have produced in that state.

<sup>4</sup>As a matter of practical implementation, it might be unrealistic to write a contract that refers to counterfactual profit levels. More realistically, the contract could just contain the explicit numerical values corresponding to these  $u(d, \omega)$ 's. Note that the principal must already have an idea what these values are in order to even formulate the optimal contracting problem.

In fact, in addition to restricted-investment contracts, it will be useful to define another class of contracts, which also expresses the idea of approximating a linear contract (plus state-by-state adjustments) but trimming some messages due to limited liability — and which will turn out to provide an alternative way of describing the optimum in our model. These contracts we call *affine-ceiling contracts*. Such a contract is parameterized by a value of  $\alpha \in [0, 1]$  and  $\beta : \Omega \rightarrow \mathbb{R}$ , and is defined as follows: First, define an affine-transformed version of the principal's indirect utility function,  $\widehat{U}_{\alpha, \beta} : \Delta(\Omega) \rightarrow \mathbb{R}$ , by

$$\widehat{U}_{\alpha, \beta}(p) = \alpha(U(p) + E_p[\beta(\omega)]).$$

Then, the expert is allowed to choose any nonnegative-valued affine function on the simplex that is bounded above by  $\widehat{U}_{\alpha, \beta}$ , and to be rewarded according to the function he chooses. Formally, the message space  $M$  of the affine-ceiling contract is the set of functions  $m : \Omega \rightarrow \mathbb{R}^+$  satisfying

$$E_p[m(\omega)] \leq \widehat{U}_{\alpha, \beta}(p) \quad \text{for all } p \in \Delta(\Omega),$$

and the payment function is simply

$$w(m, \omega) = m(\omega).$$

As long as the values of  $\beta$  are large enough so that  $U(p) + E_p[\beta(\omega)] \geq 0$  for all  $p$ , the resulting  $M$  is nonempty, and so we do obtain a contract. The concept is illustrated in Figure 3.

Panel (a) shows a possible function  $U$  (the medium-gray thick curve), a corresponding  $\widehat{U}_{\alpha, \beta}$  (the lighter curve), and some examples of affine functions bounded by  $\widehat{U}_{\alpha, \beta}$  (the thin lines). Panel (b) further shows, in reduced form, the corresponding affine-ceiling contract, which is the upper envelope of these thin lines. In general, there may be no contract whose reduced form equals  $\widehat{U}_{\alpha, \beta}$  itself because, if we tried to construct messages whose upper envelope equals  $\widehat{U}_{\alpha, \beta}$ , we might end up violating the limited liability constraint, as shown in panel (c). The affine-ceiling contract can thus be thought of as a best approximation to  $\widehat{U}_{\alpha, \beta}$ , subject to limited liability.

Affine-ceiling contracts are perhaps harder to describe verbally than restricted-investment contracts, but are more closely tied to the mathematics of the problem, and can be described with fewer parameters (an affine-ceiling contract is identified by  $|\Omega| + 1$  numbers, whereas a restricted-investment contract requires specifying the set  $M$ ). Moreover,

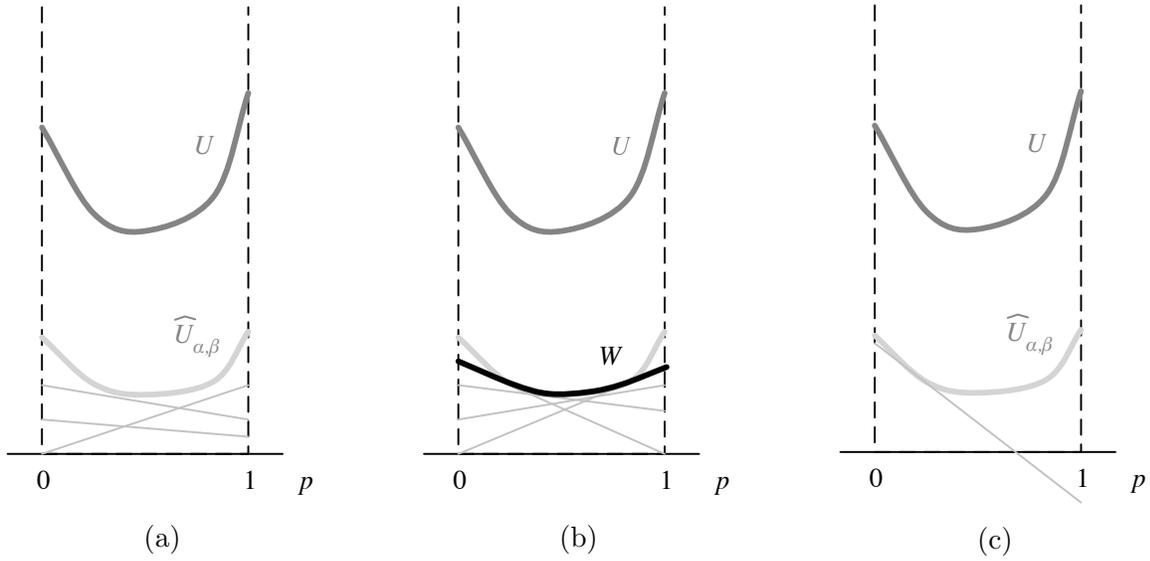


Figure 3: (a), (b) Construction of an affine-ceiling contract. (c) The affine-transformed utility function itself may not be the reduced form of a contract, due to limited liability.

restricted-investment contracts will be optimal under a convexity assumption on the principal’s decision problem (to be stated ahead); affine-ceiling contracts are optimal with or without such an assumption.

Thus, to summarize the ideas of this section: Linear contracts are a natural robust choice because they tightly align the principal’s interests with the expert’s, ensuring that the expert will never have the opportunity to produce a high payoff for himself without also producing a high value for the principal. In fact, this same alignment can be provided more cheaply by using a linear contract with state-by-state adjustments, as in (2.4). But the possible choices of parameters  $\alpha, \beta$  are constrained by limited liability. Using affine-ceiling contracts (or, under convexity, restricted-investment contracts) is the right tweak to satisfy the limited liability constraint for a wider range of parameters  $(\alpha, \beta)$  without losing the beneficial alignment between the expert’s and principal’s payoffs.

A natural question is: what makes our setting different from the pure moral hazard setting, in which linear contracts are optimal? In the pure moral hazard setting, the agent’s action produces (random) output, and the contract may be any nonnegative-valued function of output (aside from topological restrictions); this function determines the agent’s incentives. In the setting here, the expert’s action produces a random posterior  $p$  (with value  $U(p)$ ), and the function  $W(p)$  determines the expert’s incentives. But here,

$W(p)$  cannot be made an arbitrary nonnegative function of  $p$ , because it must be generated by state-by-state payments  $w(m, \omega)$  that are consistent with limited liability. Thus, there is now an additional layer of separation between the limited liability constraint and the incentive instrument. As a result, a departure from pure linearity can relax limited liability appreciably while having only a small effect on incentives.

## 2.4 Full analysis

Our results will use a notion of *equivalence* for contracts: two contracts are equivalent if they have the same reduced form. The values of  $V_E(M, w|\mathcal{I})$ ,  $I^*(M, w|\mathcal{I})$ ,  $V_P(M, w|\mathcal{I})$ , and therefore also  $V_P(M, w)$  all depend on the contract  $(M, w)$  only through its reduced form; hence, any two equivalent contracts give the same payoff both to the expert and to the principal. Our analysis will show that affine-ceiling contracts are optimal, and then that, under the convexity assumption below, affine-ceiling contracts are equivalent to restricted-investment contracts.

Thus, we state the first main step:

**Theorem 2.2.** *There is an optimal contract that is affine-ceiling. Moreover, if the known IAT  $\mathcal{I}_0$  satisfies the full-support condition, then every optimal contract is equivalent to an affine-ceiling contract.*

For the second portion of the analysis, we will say that the decision problem  $(D, \Omega, u)$  is *convex* if, for all  $d, d' \in D$  and all  $\lambda \in [0, 1]$ , there exists  $\hat{d} \in D$  such that

$$u(\hat{d}, \omega) \geq \lambda u(d, \omega) + (1 - \lambda)u(d', \omega) \quad \text{for all } \omega.$$

This is a richness condition that will ensure that the messages used in an affine-ceiling contract correspond to actual decisions available to the principal. Thus the following, purely technical result:

**Proposition 2.3.** *Suppose the decision problem is convex. Then every affine-ceiling contract is equivalent to a restricted-investment contract. In fact, the affine-ceiling contract with parameters  $\alpha, \beta$  is equivalent to the restricted-investment contract with the same parameters, and message space  $M$  equal to*

$$D_R(\beta) = \{d \in D \mid u(d, \omega) + \beta(\omega) \geq 0 \text{ for all } \omega\}, \quad (2.6)$$

*unless  $\alpha = 0$  in which case we can just take  $M = D$ .*

Combining these two results immediately gives our second characterization of optimal contracts:

**Corollary 2.4.** *Suppose the decision problem is convex. Then there is some restricted-investment contract that is optimal. If  $\mathcal{I}_0$  satisfies the full-support condition, then every optimal contract is equivalent to a restricted-investment contract.*

The convexity condition is satisfied, for example, whenever  $D$  is a convex subset of some vector space and  $u(d, \omega)$  is concave in  $d$  for each  $\omega$ . This is true in an investment application such as Example 2.1. It is also true, say, in an asset allocation application, where the principal is deciding how to invest a fixed amount of wealth across multiple assets, as long as the total return on each asset (conditional on the state) is weakly concave in the amount invested. However, if  $D$  is finite, then as observed by Clark (2016), convexity cannot be satisfied except in the trivial case where a dominant decision exists.

If convexity does not hold, then Proposition 2.3 and Corollary 2.4 do not apply directly. In this case, however, we can convexify the principal’s decision problem by allowing explicitly randomized decisions: Instead of  $D$ , have the decision space be the space of (Borel) probability distributions on  $D$ , with payoff in each state given by the expectation. It is clear that the indirect utility function  $U(p)$  is the same as before, and the new decision problem is trivially convex. Hence, the optimal contract can still be interpreted as a restricted-investment contract, where the expert may invest in a randomized decision. (Note that when the expert’s investment  $m$  is a randomization over  $d \in D$ , the payment to the expert  $w(m, \omega)$  is not actually randomized, but rather is based on the expected value of  $u(d, \omega)$ .)

Now we begin the proofs. The main steps will be described here, but a couple of lengthy details will be left to Appendix B.

The proof of Theorem 2.2 closely parallels the analysis in Carroll (2015). We take any given contract, and show that it can be improved to an affine-ceiling contract whose worst-case guarantee for the principal is at least as good.

To be more specific, we use a linear separation argument to show that, for any given contract, there is some linear relation between the principal’s and expert’s reduced-form payoffs that drives its worst-case guarantee: there exist parameters  $\alpha$  and  $\beta$  such that  $W(p) \leq \widehat{U}_{\alpha, \beta}(p)$  for all  $p$ , with equality on the support of the worst-case  $F$ . If we could replace this contract with another contract satisfying the relation with equality,  $W'(p) = \widehat{U}_{\alpha, \beta}(p)$ , then this new contract would be better for the expert. Then, it would quickly

follow that the new contract is better for the principal as well, using the linear relation between the two parties' payoffs and a calculation very similar to that used above for linear contracts (2.2). Unfortunately,  $\widehat{U}_{\alpha,\beta}$  may not be the reduced form of any contract, as depicted in Figure 3(c). Fortunately, we can use the corresponding affine-ceiling contract instead, and the adaptation of our argument (2.2) still works.

We begin the full argument. The first formal step is to get the existence question out of the way:

**Lemma 2.5.** *An optimal contract exists.*

The proof is topological — we show that we can restrict attention to a compact set of contracts (under an appropriate topology), and that  $V_P$  is upper semi-continuous on this set. (It is not continuous in general.) The details are in Appendix B.

The first substantive step is to characterize the worst-case payoff for any given contract. In doing this, we need to deal separately with *zero contracts*: those whose reduced form  $W$  satisfies  $W(p) = 0$  for all  $p$ . We denote the guarantee of a zero contract by  $V_P(0)$ . If there exists some experiment  $(F, c) \in \mathcal{I}_0$  with  $c = 0$ , then for any  $\mathcal{I}$ , the expert will choose whichever such experiment maximizes the principal's expected payoff  $E_F[U(p)]$ ; hence,  $V_P(0)$  is simply  $\max_{(F,0) \in \mathcal{I}_0} E_F[U(p)]$ . If there is no such experiment in  $\mathcal{I}_0$ , then when the IAT is  $\mathcal{I} = \mathcal{I}_0 \cup \{(\delta_{p_0}, 0)\}$ , the expert will choose the latter experiment, and we see that  $V_P(0) = U(p_0)$ . (The principal will not do worse than this under any other IAT, by convexity of  $U$ .)

For the rest of the analysis, it is useful to define *eligible* contracts as those satisfying both  $V_P(M, w) \geq V_P(0)$  and  $V_P(M, w) > U(p_0)$ . We know that some eligible contract exists — the contract guaranteed by the non-triviality assumption is eligible, unless it is worse than the zero contract, in which case the zero contract is eligible. Therefore, when searching for the optimal contract, we can focus our attention on eligible contracts.

Now we can describe the worst-case payoff in general:

**Lemma 2.6.** *Let  $(M, w)$  be any eligible contract, different from the zero contract. Let  $W$  be its reduced form. Then,*

$$V_P(M, w) = \min E_F[U(p) - W(p)] \quad \text{over } F \in \Delta(\Delta(\Omega)) \text{ such that} \quad (2.7)$$

$$E_F[p] = p_0 \text{ and } E_F[W(p)] \geq V_E(M, w|\mathcal{I}_0).$$

Moreover, for any  $F$  attaining the minimum, the constraint holds with equality:  $E_F[W(p)] = V_E(M, w|\mathcal{I}_0)$ .

Intuitively, the principal faces great uncertainty about the experiments available to the expert, and she knows only two things about the distribution  $F$  on posteriors that he will choose: she knows that  $E_F[p] = p_0$ , and she knows a lower bound on  $E_F[W(p)]$  due to the experiments that are known to be available. Lemma 2.6 shows that, indeed, these are the only constraints limiting how bad the principal's payoff might be. The actual proof requires a little more work than this brief explanation, due to subtleties introduced by the assumption of tie-breaking in favor of the principal. We leave the details to the full proof in Appendix B.

Next, we can give the key step in the proof of Theorem 2.2: a linear inequality relating the principal's and expert's payoffs under the given contract, with equality holding for the worst case. This follows from Lemma 2.6 by applying a linear separation theorem.

**Lemma 2.7.** *Let  $(M, w)$  be any eligible contract, different from the zero contract, and let  $W$  be its reduced form. Then there exist  $\alpha \in (0, 1)$  and  $\beta : \Omega \rightarrow \mathbb{R}$  such that*

$$W(p) \leq \alpha(U(p) + E_p[\beta(\omega)]) \quad \text{for all } p \in \Delta(\Omega); \quad (2.8)$$

$$V_E(M, w|\mathcal{I}_0) = \alpha(V_P(M, w) + V_E(M, w|\mathcal{I}_0) + E_{p_0}[\beta(\omega)]). \quad (2.9)$$

**Proof:** Consider the following two convex subsets of  $\Delta(\Omega) \times \mathbb{R} \times \mathbb{R}$ :

- $S$  is the convex hull of points  $(p, W(p), U(p) - W(p))$ ;
- $T$  is the set of all points  $(p_0, y, z)$  such that  $y \geq V_E(M, w|\mathcal{I}_0)$  and  $z < V_P(M, w)$ .

These two sets are disjoint — otherwise there would be some distribution  $F$  such that  $E_F[p] = p_0$ ,  $E_F[W(p)] \geq V_E(M, w|\mathcal{I}_0)$ , and  $E_F[U(p) - W(p)] < V_P(M, w)$ , contradicting Lemma 2.6. Applying a proper separating hyperplane theorem<sup>5</sup> to these sets gives us  $\lambda \in \mathbb{R}^\Omega$ ,  $\mu, \nu, \xi \in \mathbb{R}$  such that

$$\sum_{\omega} \lambda_{\omega} p(\omega) + \mu W(p) + \nu(U(p) - W(p)) \leq \xi \quad \text{for all } p \in \Delta(\Omega); \quad (2.10)$$

$$\sum_{\omega} \lambda_{\omega} p_0(\omega) + \mu V_E(M, w|\mathcal{I}_0) + \nu V_P(M, w) \geq \xi; \quad (2.11)$$

$\mu \geq 0$ ,  $\nu \leq 0$ ; and we do not simultaneously have  $\mu = \nu = 0$  and all the  $\lambda_{\omega}$  equal to each other.

---

<sup>5</sup>E.g. Rockafellar (1970, Theorem 11.3): given two disjoint convex sets in Euclidean space, there exists a hyperplane that weakly separates them *and* such that the two sets are not both contained in the hyperplane.

Moreover, we can let  $F^*$  be the distribution attaining the minimum in (2.7), and take the expectation over  $p \sim F^*$  in (2.10). Since  $E_{F^*}[W(p)] = V_E(M, w|\mathcal{I}_0)$  and  $E_{F^*}[U(p) - W(p)] = V_P(M, w)$ , we must have equality in (2.10) for all  $p$  in the support of  $F^*$ , and also have equality in (2.11).

At least one of the inequalities  $\mu \geq 0$ ,  $\nu \leq 0$  must hold strictly: otherwise  $\sum \lambda_\omega p(\omega) \leq \xi \leq \sum \lambda_\omega p_0(\omega)$  for all  $p \in \Delta(\Omega)$ , which would only be possible if all  $\lambda_\omega$  were equal (because  $p_0$  has full support), but proper separation means this cannot happen. Let's show that in fact both inequalities are strict:  $\mu > 0$  and  $\nu < 0$ .

- Suppose  $\nu = 0$ . Then  $\mu > 0$ , and (2.10) implies  $W(p)$  is bounded above by the affine function  $Z(p) = (\xi - \sum_\omega \lambda_\omega p(\omega))/\mu$ , with equality throughout the support of  $F^*$ . The last statement of Lemma 2.6 then assures us that  $V_E(M, w|\mathcal{I}_0) = E_{F^*}[W(p)] = E_{F^*}[Z(p)] = Z(p_0)$ . On the other hand, if  $(F, c)$  is the experiment the expert chooses under IAT  $\mathcal{I}_0$ , then

$$V_E(M, w|\mathcal{I}_0) = E_F[W(p)] - c \leq E_F[Z(p)] - c = Z(p_0) - c,$$

so it must be that  $c = 0$ . Thus, there is an experiment available for free in  $\mathcal{I}_0$  that gives the principal  $V_P(M, w|\mathcal{I}_0) \geq V_P(M, w)$ . But then the principal could have gotten a strictly higher payoff by using a zero contract, since she would get at least  $E_F[U(p)] > E_F[U(p) - W(p)] \geq V_P(M, w)$ . (The strict inequality holds because  $W$  is nonzero, so the expert can assure himself positive expected payment.) This contradicts eligibility. Therefore,  $\nu < 0$  strictly.

- Suppose  $\mu = 0$ . Then (2.10) for  $p_0$  and the equality in (2.11) imply

$$U(p_0) - W(p_0) \geq \frac{\sum_\omega \lambda_\omega p_0(\omega) - \xi}{-\nu} = V_P(M, w),$$

so  $V_P(M, w) \leq U(p_0)$ , again contradicting eligibility. Thus  $\mu > 0$  strictly.

Now we can finish the proof of the lemma. We can rearrange (2.10) to obtain

$$W(p) \leq \frac{-\sum_\omega \lambda_\omega p(\omega) - \nu U(p) + \xi}{\mu - \nu} \tag{2.12}$$

for all  $p$ . Put

$$\alpha = \frac{-\nu}{\mu - \nu}, \quad \beta(\omega) = \frac{-\lambda_\omega + \xi}{-\nu} \text{ for each } \omega.$$

Then (2.8) is just (2.12), and (2.9) follows from equality in (2.11). Note also that indeed  $\alpha \in (0, 1)$ . So the lemma is proven.  $\square$

Now we can complete the proof of Theorem 2.2, showing that an affine-ceiling contract is optimal. We can start with any given contract, and use the values of  $\alpha$  and  $\beta$  given by Lemma 2.7 to construct an affine-ceiling contract that improves on it. In particular, if we start from an optimal contract, then the affine-ceiling contract we construct must again be optimal.

**Proof of Theorem 2.2:** By Lemma 2.5, there exists an optimal contract, call it  $(M, w)$ ; and we know it must be eligible. Write  $W$  for the reduced form. We can assume it is nonzero, since otherwise it is already equivalent to an affine-ceiling contract.

Let  $\alpha, \beta$  be the values given by Lemma 2.7. Let  $(M', w')$  be the affine-ceiling contract with these parameters, and let  $W'$  be its reduced form.

Let  $m$  be any message in the original contract. Then for all  $p$ , we have

$$E_p[w(m, \omega)] \leq W(p) \leq \alpha(U(p) + E_p[\beta(\omega)])$$

where the first inequality is by definition of  $W(p)$  and the second is from (2.8). Thus, the map  $\omega \mapsto w(m, \omega)$  is in  $M'$ . In particular, for all  $p$ ,  $W'(p) \geq E_p[w(m, \omega)]$ . Taking the maximum over  $m$  gives  $W'(p) \geq W(p)$ : the new contract dominates the original one (for the expert).

On the other hand, we still have

$$W'(p) \leq \alpha(U(p) + E_p[\beta(\omega)]) \tag{2.13}$$

by definition of the affine-ceiling contract. This can be rearranged to give

$$U(p) - W'(p) \geq \frac{1 - \alpha}{\alpha} W'(p) - E_p[\beta(\omega)]. \tag{2.14}$$

The relation  $W'(p) \geq W(p)$  for all  $p$  implies immediately that  $V_E(M', w' | \mathcal{I}_0) \geq V_E(M, w | \mathcal{I}_0)$ , since whichever experiment the expert chooses under  $(M, w)$  and  $\mathcal{I}_0$  will pay at least as much under  $(M', w')$  as under  $(M, w)$ . Moreover, under the new contract, for any IAT  $\mathcal{I}$ , the expert will choose an experiment  $(F, c)$  such that  $E_F[W'(p)] \geq$

$E_F[W'(p)] - c = V_E(M', w'|\mathcal{I}) \geq V_E(M', w'|\mathcal{I}_0)$ . And so

$$\begin{aligned} E_F[U(p) - W'(p)] &\geq E_F \left[ \frac{1-\alpha}{\alpha} W'(p) - E_p[\beta(\omega)] \right] && \text{(by (2.14))} \\ &\geq \frac{1-\alpha}{\alpha} V_E(M', w'|\mathcal{I}_0) - E_{p_0}[\beta(\omega)] \\ &= V_P(M, w) + \frac{1-\alpha}{\alpha} (V_E(M', w'|\mathcal{I}_0) - V_E(M, w|\mathcal{I}_0)) && \text{(by (2.9)).} \end{aligned}$$

The left-hand side equals  $V_P(M', w'|\mathcal{I})$ . Taking the infimum over all  $\mathcal{I}$ , we have

$$V_P(M', w') \geq V_P(M, w) + \frac{1-\alpha}{\alpha} (V_E(M', w'|\mathcal{I}_0) - V_E(M, w|\mathcal{I}_0)) \geq V_P(M, w). \quad (2.15)$$

Since  $(M, w)$  was assumed to be an optimal contract, this must be an equality, and the affine-ceiling contract  $(M', w')$  is again optimal.

It remains to prove the final statement: we will show that under the full-support condition, the initial optimal contract is equivalent to our affine-ceiling contract, that is,  $W'$  is identically equal to  $W$ . Suppose not. Then, we have  $V_E(M', w'|\mathcal{I}_0) > V_E(M, w|\mathcal{I}_0)$ , since the experiment chosen by the expert under  $(M, w)$  and  $\mathcal{I}_0$  has full support and so gives the expert strictly higher expected payoff under  $(M', w')$ . (The expert's chosen experiment cannot be distribution  $\delta_{p_0}$ , because then  $V_P(M, w) \leq U(p_0)$  contradicting eligibility.) Together with  $\alpha \in (0, 1)$ , this implies that (2.15) is a strict inequality. But this contradicts the optimality of  $(M, w)$ . Thus,  $W'$  must be equal to  $W$  after all, which completes the proof.  $\square$

Now, to complete our main analysis (aside from the details that we have left to Appendix B), it remains only to make the leap from affine-ceiling contracts to restricted-investment contracts, Proposition 2.3. We need to show that, under convexity, for any given affine-ceiling contract, the message optimally chosen by the expert at any given posterior  $p$  actually represents some decision available in the corresponding restricted decision space.

**Proof of Proposition 2.3:** Suppose that the decision problem is convex. Let  $(M, w)$  be the affine-ceiling contract with parameters  $\alpha, \beta$ , and  $(M', w')$  the corresponding restricted-investment contract identified in the proposition statement. Write  $W, W'$  for the corresponding reduced forms. We wish to show that  $W$  and  $W'$  are identical. The  $\alpha = 0$  case is trivial (both  $W$  and  $W'$  are zero), so assume  $\alpha > 0$ .

First the easy direction: Every message available in the restricted-investment contract is also available in the affine-ceiling contract. More explicitly, take any decision  $d \in M'$ ,

and consider the function  $m_d : \Omega \rightarrow \mathbb{R}^+$  defined by  $m_d(\omega) = w'(d, \omega) = \alpha(u(d, \omega) + \beta(\omega))$ . This is nonnegative, and its expectation over any posterior  $p$  is bounded above by  $\alpha(U(p) + E_p[\beta(\omega)])$  (by definition of  $U$ ), so  $m_d \in M$ . Therefore, for every posterior  $p$ ,  $W(p) \geq E_p[w'(d, \omega)]$  for any  $d \in M'$ . And so  $W(p) \geq \max_{d \in M'} E_p[w'(d, \omega)] = W'(p)$ .

Now for the reverse inequality. Consider the following subset  $S$  of  $\mathbb{R}^\Omega$ :  $S$  is the set of all functions  $m : \Omega \rightarrow \mathbb{R}$  such that there exists some  $d \in D$  with  $\alpha(u(d, \omega) + \beta(\omega)) \geq m(\omega)$  for all  $\omega$ .  $S$  is closed, and the convexity assumption on the decision problem implies  $S$  is also convex. Now fix any posterior  $p$ . By definition of the reduced form  $W$ , there is some  $m^* \in M$  such that  $E_p[m^*(\omega)] = W(p)$ .

We will show that  $m^* \in S$ . Suppose not. Then we can apply a strict separating hyperplane theorem to conclude the existence of nonzero  $\lambda \in \mathbb{R}^\Omega$  and  $\xi$  such that

$$\sum_{\omega} \lambda_{\omega} m(\omega) < \xi \quad \text{for all } m \in S, \quad (2.16)$$

$$\sum_{\omega} \lambda_{\omega} m^*(\omega) > \xi. \quad (2.17)$$

(2.16) implies that  $\lambda_{\omega} \geq 0$  for all  $\omega$ . Then, we can normalize to assume that  $\sum_{\omega} \lambda_{\omega} = 1$ , so that  $\lambda$  equals some probability distribution  $q \in \Delta(\Omega)$ . Consider the decision  $d = d(q)$ , and  $m \in S$  given by  $m(\omega) = \alpha(u(d, \omega) + \beta(\omega))$ ; then (2.16) and (2.17) give us

$$E_q[\alpha(u(d, \omega) + \beta(\omega))] < \xi < E_q[m^*(\omega)].$$

But the definition of the affine-ceiling contract requires that

$$E_q[m^*(\omega)] \leq \alpha(U(q) + E_q[\beta(\omega)]) = E_q[\alpha(u(d, \omega) + \beta(\omega))],$$

and so we have a contradiction.

Thus,  $m^* \in S$ , which means that there is some  $d \in D$  satisfying

$$\alpha(u(d, \omega) + \beta(\omega)) \geq m^*(\omega) \geq 0$$

for all  $\omega$ . Dividing through by  $\alpha > 0$ , we see that this decision  $d$  lies in the restricted decision space  $M'$  as well. Then the value of the restricted-investment contract at posterior  $p$  satisfies

$$W'(p) \geq E_p[\alpha(u(d, \omega) + \beta(\omega))] \geq E_p[m^*(\omega)] = W(p).$$

Now we have  $W(p) \geq W'(p)$  and  $W'(p) \geq W(p)$  for every  $p$ , so we are finished.  $\square$

At this point we have reached our main goal: the optimum can be expressed either as an affine-ceiling contract with some parameters  $\alpha, \beta$ , or (under convexity) the corresponding restricted-investment contract.

To close up the analysis, we should comment a bit further on how to identify the optimal parameters  $\alpha, \beta$ . Although it is not possible to express them in fully closed form, we can give an implicit characterization that will be useful for further investigation in Subsection 3.6.

Given  $\beta : \Omega \rightarrow \mathbb{R}$ , put

$$M(\beta) = \{m : \Omega \rightarrow \mathbb{R}^+ \mid E_p[m(\omega)] \leq U(p) + E_p[\beta(\omega)] \text{ for all } p \in \Delta(\Omega)\}$$

and then put

$$U_R(p; \beta) = \max_{m \in M(\beta)} (E_p[m(\omega)] - E_p[\beta(\omega)]).$$

(If  $M(\beta) = \emptyset$  we put  $U_R(p; \beta) = -\infty$ .) Clearly  $U_R(p; \beta) \leq U(p)$ . In the convex case, we have an especially straightforward interpretation for  $U_R$ : Proposition 2.3 implies that  $U_R(p; \beta) = \max_{d \in D_R(\beta)} E_p[u(d, \omega)]$ , where  $D_R(\beta)$  is the restricted decision space defined in (2.6).

**Proposition 2.8.** *The payoff guarantee from the optimal contract is equal to*

$$\max_{\substack{\alpha \in [0,1] \\ \beta : \Omega \rightarrow \mathbb{R}}} \left( \max_{(F,c) \in \mathcal{I}_0} \left( (1-\alpha)E_F[U_R(p; \beta)] - \frac{1-\alpha}{\alpha}c \right) - \alpha \cdot E_{p_0}[\beta(\omega)] \right), \quad (2.18)$$

where, if  $\alpha = 0$ , we interpret the  $-((1-\alpha)/\alpha)c$  term as 0 for  $c = 0$  and  $-\infty$  otherwise. (In particular, the maximum in (2.18) exists.)

Moreover, when  $\alpha, \beta$  attain the maximum, the corresponding affine-ceiling contract is optimal.

The derivation follows our earlier calculation (2.2) that gave a payoff guarantee for linear contracts. A very similar calculation applies for any affine-ceiling contract, and shows that for any  $\alpha, \beta$ , the corresponding affine-ceiling contract guarantees at least the value in (2.18). In general, this is only a lower bound for the guarantee, but by retracing the steps of the proof of Theorem 2.2, we can show that it holds with equality at the maximum. The full proof is in Appendix B.

Proposition 2.8 is potentially helpful if one wants to actually compute the optimal contract. In particular, as long as  $\mathcal{I}_0$  consists of a manageable number of experiments,

one can try out a grid of values of  $\alpha$  and  $\beta$ , and calculate the value of (2.18) in each case by considering each  $(F, c) \in \mathcal{I}_0$  in turn, and then pick the best one. Since the dimension of the vector  $\beta$  equals the number of states, this approach is viable when the number of states is small. Even with a large number of states, the problem can be computationally tractable in some cases. For example, if the decision space is an interval in  $\mathbb{R}$  and the payoff  $u(d, \omega)$  is concave in  $d$  for each  $\omega$  (as in Example 2.1), then for every possible  $\beta$ , the restricted decision space identified in (2.6) is actually an interval; so we actually only need to optimize over the endpoints of this interval and the share  $\alpha$ .

*Example 2.1, continued.* We illustrate the application of Proposition 2.8 to calculate an optimal restricted-investment contract in our firm-analyst example. The relevant range of values for  $\beta(\omega_R)$  is  $[-1, -\frac{\sqrt{2}}{2}]$ : any lower  $\beta(\omega_R)$  would leave no decisions satisfying (2.6), so that the restricted decision space would be empty, while any higher value would needlessly make limited liability slack for all decisions. Similarly, the range of values for  $\beta(\omega_{NR})$  is  $[-\sqrt{2}, -1]$ .

For any such vector  $\beta$ , the restricted decision range given by (2.6) is  $D' = [\underline{d}, \bar{d}]$ , whose endpoints are given by equations  $u(\underline{d}, \omega_R) + \beta(\omega_R) = u(\bar{d}, \omega_{NR}) + \beta(\omega_{NR}) = 0$ . Plugging in the definition of  $u(d, \omega)$  and solving gives

$$\underline{d} = \frac{\sin^{-1}(-\beta(\omega_R))}{\pi/2}, \quad \bar{d} = \frac{3\pi/4 - \sin^{-1}\left(-\frac{\sqrt{2}}{2}\beta(\omega_{NR})\right)}{\pi/2}.$$

(If  $\underline{d} > \bar{d}$  then we do not get a valid contract.) We can then evaluate the  $U_R(p, \beta)$  term in (2.18) as  $E_p[u(d'(p), \omega)]$ , where  $d'(p)$  is the point in  $[\underline{d}, \bar{d}]$  closest to the ideal decision  $d(p) = (\cot^{-1}(1-p))/(\frac{\pi}{2})$ .

Recall our description of the IAT  $\mathcal{I}_0$ : it consisted of the null experiment available for free, and a distribution  $F^*$  achieving posterior 1/4 or 3/4 with probability 1/2 each, at cost  $c^* = 1/1200$ .

We now numerically compute the optimum in (2.18) over parameters  $\alpha, \beta$ . We can focus on the experiment  $(F^*, c^*)$ , since the alternative (the null experiment) can never yield a payoff higher than  $U(p_0)$ . The optimal guarantee is 1.1215, given by parameters  $\alpha = 0.0928$ ,  $\beta(\omega_R) = -0.8268$ ,  $\beta(\omega_{NR}) = -1.2570$ . These values of  $\beta$  correspond to a restricted decision space  $D' = [\underline{d}, \bar{d}] = [0.6196, 0.8031]$ .

We comment also that for the posteriors achieved by the known experiment  $(F^*, c^*)$ , the ideal decisions lie outside the restricted decision space:  $d(1/4) = 0.5903 < \underline{d}$  and  $d(3/4) = 0.8440 > \bar{d}$ . Thus, if the expert chooses this experiment, the restriction on

messages is binding.

Finally, observe that the payoff guarantee of 1.1215 is indeed higher than the payoff of  $U(p_0) = 1.1180$  that the principal would get without hiring the expert (so non-triviality does hold as claimed earlier).  $\triangle$

## 2.5 On uniqueness of the optimum

We have shown that affine-ceiling contracts (or restricted-investment contracts) are worst-case optimal. However, an issue with the worst-case criterion is that perhaps there are many choices that do equally well in the worst case, but one can distinguish among them by looking outside the worst case. In particular, as suggested by Börgers (2017), we may wish to refine the worst-case criterion by considering only undominated contracts. Specifically, say that a contract  $(M, w)$  is (*weakly*) *dominated* by  $(M', w')$  if  $V_P(M', w'|\mathcal{I}) \geq V_P(M, w|\mathcal{I})$  for all  $\mathcal{I} \supseteq \mathcal{I}_0$ , with strict inequality for some  $\mathcal{I}$ . If the optimal contracts we have focused on are weakly dominated by some other contract, this would suggest we should not really view them as optimal.

If  $\mathcal{I}_0$  satisfies the full-support condition, then Theorem 2.2 says that only affine-ceiling contracts (up to equivalence) can be optimal, so there is no worry that some other form of contract actually does better in this sense. However, without full support, there can be other contracts attaining the worst-case optimum. Indeed, let  $\alpha, \beta$  be parameters attaining the max in (2.18), and  $W$  the reduced form for the resulting affine-ceiling contract. Suppose some other contract has a reduced form  $W'$  that satisfies  $W' \leq W$  everywhere, with equality on the support of the experiment that is optimal under  $\mathcal{I}_0$ . Then  $W'$  gives the expert the same payoff as  $W$  under  $\mathcal{I}_0$ , and so the argument behind Proposition 2.8 shows that  $W'$  guarantees at least (2.18) for the principal, i.e.  $W'$  is also optimal.

However, even without full support, we can at least be assured that our affine-ceiling contracts survive the weak-dominance criterion:

**Proposition 2.9.** *Let  $(M, w)$  be any eligible contract, different from the zero contract, and let  $\alpha, \beta$  be as given by Lemma 2.7. Let  $(M', w')$  be the affine-ceiling contract with parameters  $\alpha, \beta$ . Then,  $(M', w')$  is not weakly dominated by  $(M, w)$ .*

*Consequently, if the parameters attaining the max in (2.18) are unique, then the corresponding affine-ceiling contract is not weakly dominated by any other contract.*

The proof is in Appendix B.

### 3 Extensions and variations

We now consider extending the model in several directions. These extensions are independent of each other; we do not attempt to incorporate them all into a single general model. Several of these extensions are analogous to extensions of the basic robust moral hazard model from Carroll (2015), and the arguments are identical to arguments in that paper. In these cases, we only describe the ideas briefly; the reader is referred to the earlier paper for details.

#### 3.1 Participation constraint

The original model assumes that the expert will accept whatever contract is offered. Suppose, instead, that the principal needs to guarantee the expert some minimum expected payoff  $\bar{U}_E > 0$  in order to hire him. Thus, the principal is restricted to contracts that satisfy  $E_F[W(p)] - c \geq \bar{U}_E$  for some  $(F, c) \in \mathcal{I}_0$ . We maintain non-triviality: assume there exists some such contract  $(M, w)$  with  $V_P(M, w) > U(p_0)$ .

As in Carroll (2015), adding a participation constraint does not change our main results. The compactness argument that ensures existence of an optimal contract still holds when we add the participation constraint; and since the main step of the proof of Theorem 2.2 replaced the given contract  $(M, w)$  by an affine-ceiling contract that is weakly better for both the expert and the principal, we see that if the original contract satisfies the participation constraint, so does the new one. Then the main results (Theorem 2.2, Corollary 2.4) still hold.

#### 3.2 Smaller sets of experiments

We have written the model so that the principal evaluates contracts by their worst case over all possible IAT's  $\mathcal{I} \supseteq \mathcal{I}_0$ . This is perhaps an unrealistically large class of IAT's. In fact, we do not need such drastic uncertainty: As in Carroll (2015), the results hold as long as every IAT of the form  $\mathcal{I}_0 \cup \{(F, c)\}$ , for some  $(F, c)$ , is considered possible. Moreover, we can restrict to distributions  $F$  whose support consists of at most  $2|\Omega| + 2$  posteriors. This holds because the minimum in (2.7) is attained by some distribution whose support has size at most  $|\Omega| + 2$ , by Carathéodory's Theorem (e.g. Rockafellar, 1970, Theorem 17.1). The extra  $|\Omega|$  support points are needed to construct the auxiliary distributions used in the proof of Lemma 2.6.

### 3.3 Screening on technology

In general, the principal’s minmax payoff typically is strictly greater than her maxmin payoff: that is, there is some payoff  $\bar{V}_P$  strictly greater than the guarantee of the maxmin-optimal contract, such that, if she knew the IAT  $\mathcal{I}$  with certainty when choosing the contract, she could achieve an expected payoff at least  $\bar{V}_P$ , no matter what  $\mathcal{I} \supseteq \mathcal{I}_0$  she faced. (See Carroll (2015, Proposition 1); a similar argument applies here.) This suggests the possibility of screening experts according to their IATs, by offering a *menu* of contracts  $(M, w)$ , which the expert chooses from before performing any experiment, so that experts with different IATs may choose different contracts. It is natural to imagine that the principal could do better with this additional tool, compared to the basic model where she can only offer one contract.

It turns out, however, that the principal cannot actually achieve a better worst-case guarantee with screening than she can without screening. The argument is exactly the same as that given for Theorem 4 of Carroll (2015). If the principal can guarantee herself some payoff  $V_P^*$  using a menu of contracts, we show she can also do so using just the contract  $(M_0, w_0)$  that the expert with IAT  $\mathcal{I}_0$  would choose from this menu. If this were not the case, then there would be some experiment  $(F', c')$  such that, under  $\mathcal{I}_1 = \mathcal{I}_0 \cup \{(F', c')\}$  and contract  $(M_0, w_0)$ , the expert strictly prefers to perform experiment  $(F', c')$ , and the principal’s resulting expected payoff is less than  $V_P^*$ . Then, given the choice from the menu, the expert with IAT  $\mathcal{I}_1$  might choose a different contract, but he would again perform experiment  $(F', c')$ , since we know his payoff under  $\mathcal{I}_1$  is strictly higher than the best he could do using experiments in  $\mathcal{I}_0$ . This means the principal’s resulting payoff can only be even worse than under  $(F', c')$  and  $(M_0, w_0)$  (since the principal gets the same distribution over posteriors but now has to pay the expert more). This means that the menu of contracts does not guarantee  $V_P^*$ , a contradiction.

### 3.4 Public information

Our model has assumed that information acquisition is private: when the expert performs an experiment, only he observes the outcome. What if instead we assume that the information acquired becomes public (and verifiable), so that the principal updates to the same posterior  $p$  as the expert, and  $p$  can be contracted on? What contract provides the optimal worst-case guarantee in this setting?

In this version of the model, there is no strategic choice of message, so no reason to have contracts depend on the realized state. Instead, we can have them depend directly

on the posterior: a contract is now any continuous function  $W : \Delta(\Omega) \rightarrow \mathbb{R}^+$ . The model now looks like the simple robust moral hazard model of Carroll (2015), since we can treat the expected gross payoff  $U(p)$  as being the observed output variable, and the expert’s experiment  $(F, c)$  directly determines the distribution of output via  $p \sim F$ . This suggests that the optimal contract should simply be affine in  $U(p)$  — that is, of the form  $W(p) = \alpha(U(p) + \beta)$ .

This is almost correct, except that the actions potentially available to the expert do not produce all possible distributions over posteriors, only those distributions with mean  $p_0$ . Hence, our observable-posteriors model does not collapse to the main model of Carroll (2015), but rather to the extension described in Section II.A of that paper, with a lower bound on costs. Here all of  $p$  is observable, and the lower bound on the cost of an experiment is 0 if it has mean  $p_0$  and  $\infty$  otherwise. The results from that model then show that the optimal contract is of the form  $W(p) = \alpha(U(p) + E_p[\beta(\omega)])$  for some function  $\beta : \Omega \rightarrow \mathbb{R}$ .

Since the choice of  $\beta$  can have no effect on incentives (as discussed in Subsection 2.3), its only role can be to relax the limited liability constraint. Therefore the optimal  $\beta$  is chosen by minimizing  $E_{p_0}[\beta(\omega)]$  subject to  $U(p) + E_p[\beta(\omega)] \geq 0$  for all  $p$ . Clearly, if we can make equality hold at  $p_0$ , then this  $\beta$  is optimal; this is achieved by taking  $\beta(\omega) = -u(d(p_0), \omega)$ . The resulting contract has a simple interpretation: The expert is paid a share  $\alpha$  of the principal’s expected gain, conditional on the newly discovered information, from choosing the new optimal decision rather than the decision she would have chosen in his absence.

### 3.5 Influencing states

What if we adopt the main model, except that we allow the expert’s actions to influence the distribution of states, in addition to providing information? That is, we drop the requirement  $E_F[p] = p_0$  from the definition of an experiment, so an experiment is now any element of  $\Delta(\Delta(\Omega)) \times \mathbb{R}^+$ ; where we originally imposed a full-support common prior  $p_0$ , we now simply require that for each  $(F, c) \in \mathcal{I}_0$ ,  $E_F[p]$  should have full support. The rest of the model is left unchanged. (Zermeño (2012) also considers allowing the expert to influence the distribution. See also Shi, Conitzer, and Guo (2009).)

We can repeat the analysis, and find that there is an optimal contract that is affine-ceiling with the adjustment term  $\beta(\omega)$  constant across all states  $\omega$ . In the convex case, this means likewise that a restricted-investment contract with  $\beta$  constant across all states

is optimal. (For an overview of what changes in the analysis: Lemma 2.6 is the same as before, except that the constraint  $E_F[p] = p_0$  is dropped from (2.7). Lemma 2.7 is the same but with  $\beta$  constant across states. The separation argument used to prove Lemma 2.7 is now simply done in  $\mathbb{R} \times \mathbb{R}$ , so that the  $\lambda_\omega$  terms in inequalities (2.10) and (2.11) are absent.)

### 3.6 Restricted versus unrestricted investment

In this subsection we briefly consider the question of when the optimal contract is an *unrestricted-investment* contract, with  $M = D$ . Such contracts have a natural interpretation, as delegating the decision directly to the expert, who is then paid a fraction  $\alpha$  of the (gross) payoff plus the state-by-state adjustment  $\beta$ . If it also happens that the worst possible payoff in each state is zero ( $\min_d u(d, \omega) = 0$  for each  $\omega$ ), then the interpretation is even simpler: the optimum is a linear contract. In this latter case, we also do not need our assumption that the full state  $\omega$  is observed ex post; it is enough for the realized payoff  $u(d, \omega)$  to be observed.

Identifying situations where unrestricted-investment contracts are optimal helps to bridge the gap with some of the earlier principal-expert literature (Demski and Sappington, 1987; Malcolmson, 2009) which assumed that the decision must be delegated to the expert and compensation could depend only on the realized payoff. These contracts also seem more in line with incentives observed in reality. This comparison will be discussed further in the concluding section.

We cannot give a complete characterization of when restriction is or is not optimal, but we can give some partial results. Roughly, if the principal's optimal decision is extremely sensitive to her posterior belief, then unrestricted-investment contracts are generally not optimal: There are some extreme decisions that are very unlikely to come into play, and trimming them out of the message space is useful in relaxing limited liability. Conversely, if the optimal decision is not so sensitive, then unrestricted-investment contracts may be optimal. In the special case where the principal's decision problem is binary, unrestricted-investment contracts are always optimal.

For formal statements, we maintain several assumptions. We assume throughout this subsection that the zero contract is not optimal. We also assume that in every state, not all decisions give the same payoff. Furthermore, for Proposition 3.1 below, we assume that the decision problem is convex; thus Proposition 2.3 applies, and we can describe the optimal contract either as an affine-ceiling contract or as a restricted-investment contract

with the same parameters  $\alpha, \beta$ . Finally, we assume that dominated decisions have been eliminated a priori: For any distinct  $d, d' \in D$ , there exists  $\omega$  such that  $u(d, \omega) > u(d', \omega)$ . This eliminates uninteresting cases where there is some decision giving extremely low payoffs in every state, which would make an unrestricted-investment contract extremely costly (via limited liability forcing  $\beta(\omega)$  to be high) even though the expert would never invest in that decision.

Also, we need a little more terminology. Say that a decision  $d$  is  $\omega$ -*extremal* if it minimizes  $u(d, \omega)$  among all decisions; and it is *extremal* if it is  $\omega$ -extremal for some  $\omega$ .

Now we can state our main result relating the structure of the principal's decision problem to the optimality of unrestricted investment:

**Proposition 3.1.** *Fix the decision problem  $(D, \Omega, u)$ , satisfying convexity.*

- (a) *Suppose that, for every decision  $d$ , there is only one posterior  $p$  for which  $d$  is optimal for the principal. Then, for any  $p_0$  and  $\mathcal{I}_0$ , there is a properly-restricted-investment contract (i.e.  $M \subset D$  strictly) that is optimal.*
- (b) *Suppose that, for every state  $\omega$ , there is an  $\omega$ -extremal decision  $d \in D$ , and a nonempty open set of posteriors  $P \subseteq \Delta(\Omega)$ , such that  $d$  is optimal for any posterior  $p \in P$ . Then, for any prior  $p_0$ , it is possible to choose  $\mathcal{I}_0$  so that an unrestricted-investment contract is optimal. We can do this while satisfying the non-triviality and full-support assumptions.*
- (c) *Suppose that there exists a nonempty open set of posteriors  $P \subseteq \Delta(\Omega)$  such that there is no one decision that is optimal for all  $p \in P$ , and such that extremal decisions are not optimal for any  $p \in P$ . Then, it is possible to choose  $p_0 \in P$  and  $\mathcal{I}_0$ , satisfying non-triviality and full support, such that a properly-restricted-investment contract is optimal.*

To aid in understanding this proposition, consider Figure 4, which shows a couple of possible decision problems. Each panel shows a decision problem with state space  $\Omega = \{\omega_1, \omega_2\}$ ; each decision is represented by the pair of payoffs  $(u(d, \omega_1), u(d, \omega_2))$ . (Note that this differs from the illustration style used in Figure 1 where one axis was the decision itself.) The thick curve depicts the set of available decisions  $D$ . Within each state, payoffs have been normalized so that the payoff of the worst decision is 0. A posterior  $p$  can be represented as a vector  $(p(\omega_1), p(\omega_2))$ ; the best decision for this posterior is then the one where the tangent to the decision set is orthogonal to the vector.

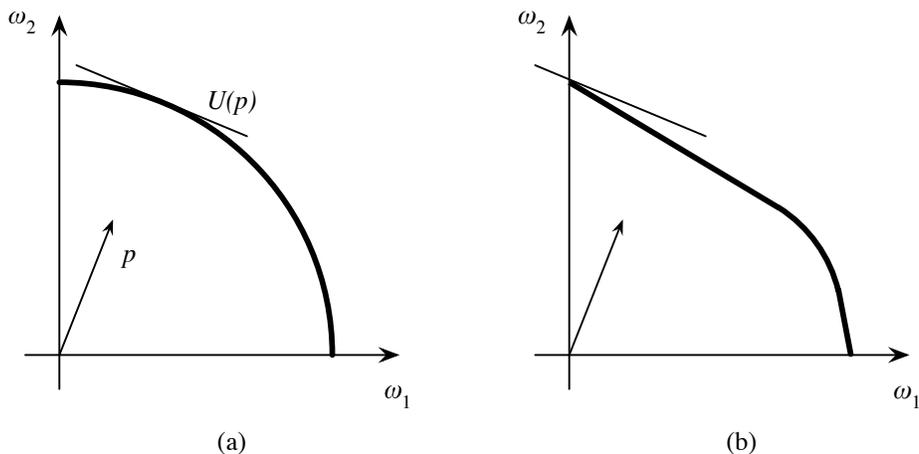


Figure 4: Example decision problems. (a) Properly-restricted-investment contracts are always optimal. (b) Unrestricted-investment contracts may be optimal, depending on  $p_0$  and  $\mathcal{I}_0$ .

Panel (a) of the figure depicts part (a) of the proposition. In this case, for every possible posterior there is a different optimal decision, and the proposition tells us that properly-restricted-investment contracts are always optimal.

In panel (b), by contrast, the leftmost decision (where the tangent line is illustrated) is an  $\omega_1$ -extremal decision that is optimal for all posteriors that put sufficiently low weight on  $\omega_1$ . Similarly, the rightmost decision is  $\omega_2$ -extremal and is optimal for all posteriors that put sufficiently low weight on  $\omega_2$ . Hence, part (b) of the proposition tells us that unrestricted-investment contracts may be optimal in this case.

For a more general example in the same spirit, consider a decision space given by starting with a finite set of decisions, and then allowing randomization as mentioned in Subsection 2.4. More precisely, suppose we begin with some finite set of decisions  $D_0$ ; we identify each decision  $d$  with the corresponding vector of payoffs in  $\mathbb{R}^\Omega$ , and form the convex hull  $D_1$ ; and then  $D$  consists of all  $d \in D_1$  that are not weakly dominated by some other vector in  $D_1$  (in order to adhere to our assumption of no weakly dominated decisions). In this case, one can show, using a separating hyperplane argument, that the condition in (b) is satisfied. So at least for some choices of  $\mathcal{I}_0$ , an unrestricted-investment contract is optimal. This, in turn, is the same as an unrestricted-investment contract without allowing mixed decisions — simply letting the expert choose from among all

decisions in  $D_0$  — as long as no such decision is weakly dominated by some mixture. We omit the details here.

Finally, panel (b) of the figure also illustrates part (c) of the proposition: The smooth bump on the right side represents a set of non-extremal decisions, which correspond to an open interval of posteriors, with the optimal decision varying as the posterior moves. Thus, by part (c), properly-restricted-investment contracts may also be optimal for this decision problem.

Are there simple conditions on the decision problem sufficient to ensure that unrestricted-investment contracts are optimal? The logic of part (c) is quite general, and shows that this cannot happen unless “most” decisions are extremal. But there is one important special case where the conclusion does hold: when the principal’s decision problem is binary. (Note this is true even though the problem is not convex in this case.)

**Proposition 3.2.** *Suppose that  $D$  consists of just two decisions. Then, there is an unrestricted-investment contract that is optimal.*

The proofs of the two propositions from this subsection are in Appendix C. Proposition 3.1 builds on Proposition 2.8, which gave a formula identifying the parameters  $\alpha, \beta$  of an optimal contract; the proof of Proposition 3.1 involves more detailed analysis of this formula. Proposition 3.2 holds because, in the binary-decision case, restricted- and unrestricted-investment contracts are actually the same under a suitable change of parameters.

## 4 Discussion

We wrap up by briefly recapitulating our results and then discussing interpretation. We considered a principal-expert model, with risk-neutrality and limited liability, and ex-post revelation of the state of nature. We assumed unquantifiable uncertainty about the expert’s information acquisition technology, and sought out contracts that are robust to this uncertainty, as expressed by a maxmin objective. This led quite generally to a novel form of contracts being optimal: affine-ceiling contracts, or equivalently (under a convexity assumption) restricted-investment contracts, in which the expert chooses an investment from a restricted subset of less-extreme decisions, and is paid proportionally to the payoff that his investment yields in the realized state, even though the decision that the principal actually makes may be different. The result reflects the intuition that linear payment rules are robust because they tightly align the principal’s and expert’s payoffs,

turning an expected-payoff guarantee for the expert into a guarantee for the principal. But it also reflects the added wrinkle that by prohibiting investment in extreme decisions, the principal can relax the limited liability constraints and so pay the expert less.

Some readers might find jarring the combination of fully-Bayesian modeling of the state space  $\Omega$  with extreme maxmin uncertainty about the expert’s set of experiments. An interpretation is that the relevant states are relatively easy to specify (does a certain regulation get enacted or not?) and so easy to formulate a prior over; whereas the nature of information that the expert might obtain is abstract and hard for the principal to think about. Indeed, the principal-expert situation is one where the assumption of ignorance seems relatively natural: it is inherent in the nature of expertise that the principal who hires an expert does not fully understand how he does his job. Of course, in reality, contract writers usually have neither fully-specified priors nor complete uncertainty. The model here simply is one tractable (albeit imperfect) way of representing the idea that the principal understands the state space better than she understands the expert’s job.

There are two ways to view the contribution of this paper. The more direct interpretation is that our results show how one can go about providing incentives to acquire information in uncertain environments. They offer qualitative insights into the structure of a contract that is optimally robust, in a sense that the model makes precise. They also show that this optimal contract can be characterized by a small number of parameters (equal to the number of states, or less in some cases), which can be helpful in actually computing it.

On this view, it is natural to ask how our findings compare to incentive schemes seen in reality. Linear contracts based on realized payoffs are quite widespread (see e.g. the references in Bhattacharyya and Lafontaine (1995, pp. 763–4)), and this includes some settings where the agency problem involves acquiring information: consider a portfolio manager paid via a fixed percentage of the returns he earns, and thereby motivated to exert effort to find good stocks. One can also argue that some such situations involve restricting the possible decisions; a portfolio manager may be given limits on how much to allocate to each asset class. But one key finding in our model — that the expert is paid based on a decision that differs from the one actually taken — does not seem to have obvious analogues in practice.<sup>6</sup> This can be viewed either as an innovation in

---

<sup>6</sup>The academic literature seems to offer very little empirical documentation on explicit, contractual incentives for information acquisition. Even in the case of financial analysts, where one might expect such contracts to be common, the literature instead emphasizes indirect, competitive, or reputational incentives (e.g. Cowen, Groysberg, and Healy, 2006; Marinovic, Ottaviani, and Sørensen, 2013). Witkowski, Freeman, Vaughan, Pennock, and Krause (2018) documents several uses of forecasting competitions with

contract design, or as a descriptive weakness of the model. If the latter, then the next question is what is missing from the model to account for the discrepancy. Perhaps, without strict incentives over some part of the posterior space, actual experts might be motivated by other factors not captured in the model (e.g. overconfidence, or a desire to appear informed), leading to exaggerated reports and therefore excessively risky decisions. Or perhaps, in realistic applications, it is simply not possible to specify the prior beliefs and the payoff function  $u(d, \omega)$  precisely enough to write out the restricted-investment contract explicitly, and paying based on the observed outcomes seems a good enough approximation.

This latter line of thinking suggests that a model with even less information might lead to contracts where the payment rule is purely linear, perhaps with the *actual* decision being restricted. For example, what if the state were not fully observed, but only the realized payoff? In such a model, a contract would specify a decision rule as a function of the expert’s report (or equivalently, delegate the decision to the expert), and payment could depend only on the report and the realized gross payoff. However, this simple change would not be enough to generate linear payment rules. In an application such as Example 2.1, the restricted-investment contract can be replicated to within  $\epsilon$ , by limiting decisions to an arbitrarily dense grid within which  $u(d, \omega) \neq u(d', \omega')$  whenever  $(d, \omega) \neq (d', \omega')$ . Then the state can be inferred from the realized payoff, and the assumption that the state is not observed has no bite. It may well be that some model with additional uncertainty would lead to linear (or otherwise more intuitive) contracts, but more work will be needed to formulate such a model. In the meantime, the model we have studied here can be viewed as exploring the consequences of a relatively modest departure from the traditional approach of fully-Bayesian modeling.

This segues into the second, broader perspective on our contribution: as an illustrative case study of the worst-case methodology in contract theory. The traditional Bayesian version of the principal-expert problem, even under risk-neutrality and limited liability as we have assumed here, seems to be intractable without making strong functional form assumptions. The maxmin version offers traction quite generally — without any assumptions on the structure of the known information acquisition technology — by extending the linear separation methods of the earlier paper (Carroll, 2015). As in that earlier paper, we also saw that the model can be extended in various directions without changing ranking-based prizes. Proper scoring rules are used in many fields to measure accuracy of forecasts (see Carvalho, 2016, for a review), but apparently these measures are not used as a basis for explicit incentives, except in experiments.

the underlying technical machinery. And the analysis has led us to a new qualitative insight — the benefit of cutting extreme messages out of a contract, so as to relax limited liability — that is not knife-edge sensitive to the worst-case formulation of the objective. Regardless of whether one finds this insight realistic, being able to solve the model helps in understanding how and why it generates this conclusion.

It seems likely that the methods used here generalize to other kinds of robust moral hazard contracting problems. A natural task for future work is to identify the general class of models that are amenable to the same proof approach — writing a linear program to identify the worst case for a given contract, using duality to infer a linear bound relating the principal’s and agent’s payoffs, and from there producing a new contract that improves on the original. At the same time, the worst-case modeling framework may also bring traction to difficult problems by other routes. For example, recent works by the author (Carroll, 2017; Carroll and Segal, 2018) take the worst-case approach to multidimensional screening problems and auctions with resale (respectively), to give formulations where the optimal mechanism has a simple form. The proofs there also work by duality, but rather than starting with any given mechanism and finding an improvement, they involve direct construction of an adversarial environment for the principal.

## A On direct reporting of posteriors

Here we elaborate formally the revelation-principle-style argument outlined in Subsection 2.2: when looking for optimal contracts, we may assume that the expert communicates enough to the principal so that she can choose the optimal decision  $d(p)$  given his posterior  $p$ .<sup>7</sup>

As mentioned in the text, to formalize this argument, we need to specify the principal’s payoffs in the game with the expert when she may potentially use “pooling” contracts. Properly speaking, the principal’s payoffs depend both on the contract and on the rule she uses to choose a decision given the expert’s message. Note that we cannot simply appeal to some kind of Bayesian updating to pin down the decision rule, since the principal in our model is fundamentally non-Bayesian. We also need to be more careful about how the expert responds to any given contract; we cannot simply assume he breaks ties in favor of the principal, when the principal’s payoffs have not yet been defined.

In this appendix, we will show that any contract, together with *any* decision strategy

---

<sup>7</sup>The notation  $d(p)$  was defined in the main text, as  $\arg \max_{d \in D} E_p[u(d, \omega)]$ .

for the principal, is weakly outperformed by a contract in which the expert reports his posterior truthfully as part of his message, and the principal takes the corresponding optimal decision. And to avoid the ambiguity in the expert's tie-breaking, we define contracts differently than in the main text: here we follow the standard approach of requiring a contract to explicitly specify a recommended strategy for the expert, which must be incentive-compatible, in addition to specifying the compensation rule. We then think of the model formulated in the main paper as expressing the same concept, but in a more compact notation; we elaborate at the end of this appendix.

Let  $\mathcal{IAT}$  denote the set of all possible IAT's  $\mathcal{I}$  with  $\mathcal{I}_0 \subseteq \mathcal{I}$ . We properly define a *contract* to consist of four parts:

- a compact message space  $M$ ;
- a continuous payment function  $w : M \times \Omega \rightarrow \mathbb{R}^+$ ;
- an experiment strategy  $\sigma : \mathcal{IAT} \rightarrow \Delta(\Delta(\Omega)) \times \mathbb{R}^+$ , such that  $\sigma(\mathcal{I}) \in \mathcal{I}$  for each  $\mathcal{I}$ ;
- a reporting strategy  $\rho : \mathcal{IAT} \times \Delta(\Omega) \rightarrow M$ , which should be measurable in its second argument.

Thus  $(\sigma, \rho)$  describes the strategy recommended for the expert. (We could also allow for mixed strategies; this would change nothing below.)

We want to impose that the specified strategy for the expert be incentive-compatible. This requires defining payoffs. So given  $(M, w, \sigma, \rho)$ , define

$$W(p) = \max_{m \in M} E_p[w(m, \omega)], \quad V_E(M, w|\mathcal{I}) = \max_{(F, c) \in \mathcal{I}} (E_F[W(p)] - c)$$

just as in the main text. We now say that the contract  $(M, w, \sigma, \rho)$  is *incentive-compatible* if

- for all  $\mathcal{I}$ , the experiment  $(F, c) = \sigma(\mathcal{I})$  satisfies  $E_F[W(p)] - c = V_E(M, w|\mathcal{I})$ ; and moreover
- for all  $\mathcal{I}$  and all  $p$ ,

$$E_p[w(\rho(\mathcal{I}, p), \omega)] = W(p).$$

Now define a *decision rule* for the principal to be any measurable  $r : M \rightarrow D$ . Given the contract and the decision rule, we define the principal's payoff from a given  $\mathcal{I}$  to be

$$V_P(M, w, \sigma, \rho; r|\mathcal{I}) = E_F[E_p[u(r(\rho(\mathcal{I}, p)), \omega)] - W(p)]$$

where  $F$  is the distribution chosen by  $\sigma(\mathcal{I})$ .

Now we can give the formal justification for the modeling of contracts in the main text, including the assumption that the principal is able to choose decision  $d(p)$ .

**Proposition A.1.** *Let  $(M, w, \sigma, \rho)$  be any incentive-compatible contract, and  $r : M \rightarrow D$  a decision rule. Then define a new contract  $(M', w', \sigma', \rho')$  and decision rule  $r'$  by taking:*

- $M' = M \times \Delta(\Omega)$ ;
- $w'((m, p), \omega) = w(m, \omega)$ ;
- $\rho'(\mathcal{I}, p) = (\rho(\mathcal{I}, p), p)$ ;
- $r'(m, p) = d(p)$ ; and
- $\sigma'(\mathcal{I})$  is taken to be any  $(F, c) \in \mathcal{I}$  that lexicographically maximizes  $E_F[W(p)] - c$  and then  $E_F[U(p) - W(p)]$ , where  $W$  is the reduced form of the original contract.

*Then the new contract is also incentive-compatible, and (with  $r'$ ) weakly dominates the old contract (with  $r$ ): for any  $\mathcal{I}$ ,*

$$V_P(M', w', \sigma', \rho'; r'|\mathcal{I}) \geq V_P(M, w, \sigma, \rho; r|\mathcal{I}).$$

**Proof:** At any posterior  $p$ ,  $(m, p)$  is an optimal report under the new contract iff  $m$  was optimal under the old contract; and the reduced form of the new contract is the same as the old,  $W'(p) = W(p)$ . This in turn implies that the new contract has the same  $V_E$  as the old one. Incentive-compatibility for the new contract immediately follows.

As for the comparison of payoffs, consider any  $\mathcal{I}$ . Let  $(F, c) = \sigma(\mathcal{I})$  and  $(F', c') = \sigma'(\mathcal{I})$ . Since  $(F', c')$  is an optimal experiment, as is  $(F, c)$  (by incentive-compatibility of the original contract), lexicographic maximization in the definition of  $\sigma'$  implies that  $E_{F'}[U(p) - W(p)] \geq E_F[U(p) - W(p)]$ . Thus,

$$\begin{aligned} V_P(M', w', \sigma', \rho'; r'|\mathcal{I}) &= E_{F'} [E_p[u(r'(\rho'(\mathcal{I}, p)), \omega)] - W'(p)] \\ &= E_{F'} [E_p[u(d(p), \omega)] - W(p)] \\ &= E_{F'} [U(p) - W(p)] \\ &\geq E_F [U(p) - W(p)] \\ &\geq E_F [E_p[u(r(\rho(\mathcal{I}, p)), \omega)] - W(p)] \\ &= V_P(M, w, \sigma, \rho; r|\mathcal{I}). \end{aligned}$$

□

Now we can think of the modeling approach taken in the main text as a shorthand, where if the contract  $(M, w)$  does not fully distinguish posteriors at which different decisions should be taken, then it really refers to the full contract  $(M', w', \sigma', \rho')$  (and decision rule  $r'$ ) above. Indeed, the choice of experiment and the principal's payoff as defined in the main text agree with the definitions given here for that full contract. Since we are concerned specifically with identifying optimal contracts for the principal, it is sufficient to look at these contracts that are constructed in Proposition A.1 (with the appropriate decision rule).

## B Omitted proofs for main optimal-contract results

### Proof of Lemma 2.5:

Let  $\bar{U} = \max_p U(p)$ , and  $\bar{W} = (\bar{U} - U(p_0))/(\min_\omega p_0(\omega))$ . We may restrict attention to contracts whose reduced form satisfies  $W(\delta_\omega) \leq \bar{W}$  for all  $\omega$ . To see this, note that if there were some message guaranteeing the expert payoff higher than  $(\bar{U} - U(p_0))/p_0(\omega)$  in some state  $\omega$ , then no matter what experiment the expert performs, he can always force the principal to pay more than  $\bar{U} - U(p_0)$  in expectation (by just always sending this message), so the principal's expected payoff must be less than  $U(p_0)$ , and the principal is worse off than by not hiring the expert.

Since  $W$  is convex, this restriction implies  $W(p) \leq \bar{W}$  for all  $p$ . Now say that a function  $W : \Delta(\Omega) \rightarrow \mathbb{R}^+$  is a *reduced-form contract* if it is the reduced form of some contract. We make two claims:

- Claim 1: The set of reduced-form contracts  $W : \Delta(\Omega) \rightarrow [0, \bar{W}]$  is compact in the sup-norm topology.
- Claim 2:  $V_P$  is upper semi-continuous on the set of reduced-form contracts (with respect to the sup-norm topology).

Together, these claims imply that  $V_P$  attains a maximum over the reduced-form contracts whose values never exceed  $\bar{W}$ , which is then a global maximum, as needed.

To prove Claim 1, let  $W_1, W_2, \dots$  be a sequence of reduced-form contracts taking values in  $[0, \bar{W}]$ . Note that each  $W_k$  must be a Lipschitz function with constant  $\bar{W}$  (relative to the  $L_1$  metric on  $\Delta(\Omega)$ ). By passing to a subsequence, we may assume that  $W_k$  converges pointwise at each *rational* point  $p \in \Delta(\Omega)$ . Let  $W_\infty(p) = \lim_k W_k(p)$  for each such  $p$ .

Define  $M$  as the set of all functions  $m : \Omega \rightarrow [0, \overline{W}]$  such that  $E_p[m(\omega)] \leq W_\infty(p)$  for each rational  $p$ . Notice that  $M$  is a nonempty, compact subset of  $\mathbb{R}^\Omega$ . Define  $w : M \times \Omega \rightarrow [0, \overline{W}]$  by  $w(m, \omega) = m(\omega)$ . Then  $(M, w)$  is a contract, with reduced form  $W(p) = \max_{m \in M} E_p[m(\omega)]$ .

For each rational  $p$ , we will show  $W(p) = W_\infty(p)$ . The direction  $W(p) \leq W_\infty(p)$  is immediate from the definition of  $W$ . For the reverse inequality, fix  $p$ . For each  $k$ , since  $W_k$  is a reduced-form contract, there is some affine  $m_k : \Delta(\Omega) \rightarrow [0, \overline{W}]$  such that  $m_k(p) = W_k(p)$  and  $m_k(p') \leq W_k(p')$  for all other rational  $p'$ . There is some subsequence along which the  $m_k$  converge to some  $m_\infty$ . Then,  $m_\infty(p) = W_\infty(p)$ , and  $m_\infty(p') \leq W_\infty(p')$  for each other rational  $p'$ , so that  $m_\infty \in M$ . Therefore  $W(p) \geq W_\infty(p)$ , and equality follows.

Now we claim that  $W_k \rightarrow W$  in sup norm. If not, there exists some  $\epsilon > 0$  and a subsequence of  $k$ 's and points  $p_k$  along which  $|W_k(p_k) - W(p_k)| > \epsilon$ . Again by taking a subsequence, we may assume the  $p_k$  converge to some point  $p$ . Now, pick a rational point  $q$  such that the  $L_1$  distance between  $p$  and  $q$  is less than  $\epsilon/4\overline{W}$ . Then, for  $k$  high enough,  $|W_k(q) - W(q)| \leq \epsilon/4$  (because  $W(q) = W_\infty(q)$  by the previous paragraph), and for any reduced-form contract with values at most  $\overline{W}$ , its values at  $p_k$  and  $q$  differ by at most  $\epsilon/4$ . Therefore,

$$|W_k(p_k) - W(p_k)| \leq |W_k(p_k) - W_k(q)| + |W_k(q) - W(q)| + |W(q) - W(p_k)| \leq \frac{3\epsilon}{4},$$

a contradiction.

This shows that the set of reduced-form contracts taking values in  $[0, \overline{W}]$  is sequentially compact, proving Claim 1.

For Claim 2, suppose we have a convergent sequence of reduced-form contracts  $W_k \rightarrow W$ . We need to show that  $V_P(W) \geq \limsup_k V_P(W_k)$ . By passing to a subsequence, we may assume  $V_P(W_k)$  converges (to the lim sup of the original sequence). Fix any IAT  $\mathcal{I}$ , and let  $(F_k, c_k)$  be the expert's chosen experiment under contract  $W_k$ . Then, again by taking a subsequence, we may assume that  $(F_k, c_k)$  converges to some limit  $(F, c)$ . It follows that  $(F, c) \in \mathcal{I}$ , and since

$$E_F[W(p)] - E_{F_k}[W_k(p)] = (E_F[W(p)] - E_{F_k}[W(p)]) + (E_{F_k}[W(p)] - E_{F_k}[W_k(p)]) \rightarrow 0$$

(the first parenthesized expression goes to 0 by weak convergence of  $F_k$ , and the second by sup-norm convergence of  $W_k$ ), we can conclude that  $(F, c)$  is an optimal experiment

for the expert under  $W$  and  $\mathcal{I}$ : If there were some strictly better experiment  $(F', c')$ , then this experiment would also be preferred to  $(F_k, c_k)$  under  $W_k$  for high enough  $k$ , a contradiction.

This same double-convergence argument also implies that

$$E_F[U(p) - W(p)] - E_{F_k}[U(p) - W_k(p)] \rightarrow 0$$

from which

$$V_P(W|\mathcal{I}) \geq E_F[U(p) - W(p)] = \lim_k E_{F_k}[U(p) - W_k(p)] = \lim_k V_P(W_k|\mathcal{I}) \geq \limsup_k V_P(W_k).$$

Since this holds for all  $\mathcal{I}$ , we have

$$V_P(W) \geq \limsup_k V_P(W_k)$$

which proves Claim 2. □

**Proof of Lemma 2.6:**

First, for any  $\mathcal{I} \supseteq \mathcal{I}_0$ , and any experiment  $(F, c)$  chosen by the expert,  $E_F[W(p)] \geq E_F[W(p)] - c \geq V_E(M, w|\mathcal{I}_0)$ , so  $F$  satisfies the constraints in (2.7), and hence  $V_P(M, w|\mathcal{I})$  is at least the indicated minimum. This holds for any  $\mathcal{I}$ , so  $V_P(M, w)$  is at least the minimum in (2.7). (Note that this minimum is well-defined.)

To prove the reverse inequality, we begin by defining the affine function  $Z : \Delta(\Omega) \rightarrow \mathbb{R}^+$  by  $Z(p) = \sum_{\omega} p(\omega)W(\delta_{\omega})$ . By convexity,  $W(p) \leq Z(p)$  for every  $p$ . Now let  $F$  be a distribution attaining the minimum in (2.7). Suppose  $F$  places positive probability on posteriors  $p$  such that  $W(p) < Z(p)$ . Then we have

$$E_F[W(p)] < E_F[Z(p)] = Z(p_0)$$

where the equality holds because  $Z$  is affine. For small  $\epsilon > 0$ , define a distribution  $F'$  as follows: with probability  $1 - \epsilon$ ,  $F'$  chooses a posterior according to  $F$ ; with the remaining probability  $\epsilon$ ,  $F'$  picks a state  $\omega \sim p_0$  and gives posterior  $\delta_{\omega}$ . Evidently,  $E_{F'}[p] = p_0$ , and

$$E_{F'}[W(p)] = (1 - \epsilon)E_F[W(p)] + \epsilon Z(p_0) > E_F[W(p)] \geq V_E(M, w|\mathcal{I}_0).$$

So if the IAT is  $\mathcal{I} = \mathcal{I}_0 \cup \{(F', 0)\}$ , the expert's unique optimal choice of experiment is

$(F', 0)$ . The principal's expected payoff  $V_P(M, w|\mathcal{I})$  is then

$$E_{F'}[U(p) - W(p)] = (1 - \epsilon)E_F[U(p) - W(p)] + \epsilon E_{p_0}[U(\delta_\omega) - W(\delta_\omega)].$$

By taking  $\epsilon \rightarrow 0$ , we see that the principal cannot be guaranteed a payoff higher than  $E_F[U(p) - W(p)]$ , which is exactly the amount in (2.7).

Also, if  $W(p) = Z(p)$  throughout the support of  $F$ , but  $E_F[W(p)] > V_E(M, w|\mathcal{I}_0)$  strictly, then we can give a similar argument by simply taking  $\mathcal{I} = \mathcal{I}_0 \cup \{(F, 0)\}$ .

This leaves us with the case where

$$V_E(M, w|\mathcal{I}_0) = E_F[W(p)] = E_F[Z(p)] = Z(p_0).$$

For this to happen, it must be that whatever experiment  $(F_0, c_0)$  is chosen under  $\mathcal{I}_0$  satisfies  $W(p) = Z(p)$  throughout the support of  $F_0$ , and  $c_0 = 0$ . However, in this case, the expert would be willing to perform the same experiment under the zero contract, so

$$V_P(0) \geq E_{F_0}[U(p)] > E_{F_0}[U(p) - W(p)] = V_P(M, w|\mathcal{I}_0) \geq V_P(M, w).$$

(The strict inequality follows from the assumptions that  $W$  is nonzero, and  $F_0$  has mean  $p_0$  which has full support.) This contradicts eligibility, so this case cannot happen.

Finally, let  $F$  be a distribution attaining the minimum in (2.7). Then certainly

$$E_F[U(p) - W(p)] = V_P(M, w) > U(p_0) \geq U(p_0) - W(p_0).$$

If  $E_F[W(p)] > V_E(M, w|\mathcal{I}_0)$  strictly, then define another distribution  $F'$  by drawing a posterior from  $F$  with probability  $1 - \epsilon$ , and placing the remaining probability mass  $\epsilon$  on  $p_0$ . For small  $\epsilon$ ,  $F'$  still satisfies the constraints of (2.7), and

$$E_{F'}[U(p) - W(p)] = (1 - \epsilon)E_F[U(p) - W(p)] + \epsilon(U(p_0) - W(p_0)) < E_F[U(p) - W(p)],$$

contradicting minimality for  $F$ . □

**Proof of Proposition 2.8:** Let  $(M, w)$  be the affine-ceiling contract with parameters  $\alpha, \beta$ . We will show that its payoff guarantee satisfies

$$V_P(M, w) \geq \max_{(F, c) \in \mathcal{I}_0} \left( (1 - \alpha)E_F[U_R(p; \beta)] - \frac{1 - \alpha}{\alpha}c \right) - \alpha \cdot E_{p_0}[\beta(\omega)]. \quad (\text{B.1})$$

We will also show that there are some  $\alpha, \beta$  for which the corresponding affine-ceiling

contract is optimal and (B.1) is an equality. This will imply the result. (Note that once we show this, if there exist additional maximizers  $\alpha', \beta'$  for the right side of (B.1), then (B.1) must be an equality for those parameters as well: otherwise the contract with parameters  $\alpha', \beta'$  is strictly better than the one with parameters  $\alpha, \beta$ , contradicting optimality.)

First suppose  $\alpha = 0$ . If there does not exist any  $(F, 0) \in \mathcal{I}_0$ , the right side of (B.1) is  $-\infty$ ; and if there is such an experiment, the right side is  $\max_{(F,0) \in \mathcal{I}_0} E_F[U_R(p; \beta)] \leq \max_{(F,0) \in \mathcal{I}_0} E_F[U(p)]$  which is exactly the payoff guarantee from that contract. Also, if a zero contract is optimal, we can obtain equality in (B.1) by taking  $\alpha = 0$  and  $\beta$  large enough so that  $U_R(p; \beta) = U(p)$ .

Now suppose  $\alpha > 0$ . The reduced form of the affine-ceiling contract is precisely

$$W(p) = \alpha(U_R(p; \beta) + E_p[\beta(\omega)]). \quad (\text{B.2})$$

Once the expert attains posterior  $p$ , his expected payoff is  $W(p) \leq \alpha(U(p) + E_p[\beta(\omega)])$ , while the principal's is

$$\begin{aligned} U(p) - W(p) &\geq (1 - \alpha)U_R(p; \beta) - \alpha E_p[\beta(\omega)] \\ &= \frac{1 - \alpha}{\alpha} W(p) - E_p[\beta(\omega)]. \end{aligned}$$

Therefore, whatever experiment  $(F, c)$  the expert performs, the principal's expected payoff  $V_P(M, w | \mathcal{I})$  is

$$\begin{aligned} E_F[U(p) - W(p)] &\geq E_F \left[ \frac{1 - \alpha}{\alpha} W(p) - E_p[\beta(\omega)] \right] \\ &= \frac{1 - \alpha}{\alpha} E_F[W(p)] - E_{p_0}[\beta(\omega)]. \end{aligned}$$

Since

$$E_F[W(p)] \geq E_F[W(p)] - c \geq V_E(M, w | \mathcal{I}_0) = \max_{(F,c) \in \mathcal{I}_0} (E_F[W(p)] - c),$$

we can plug in to obtain

$$V_P(M, w | \mathcal{I}) \geq \frac{1 - \alpha}{\alpha} \cdot \max_{(F,c) \in \mathcal{I}_0} (E_F[W(p)] - c) - E_{p_0}[\beta(\omega)].$$

Now plugging in from (B.2) and rearranging gives exactly the right side of (B.1). Since this applies for all  $\mathcal{I}$ , (B.1) follows.

Now consider an optimal contract  $(M, w)$ . Assume it is nonzero, since we already dealt

with the case where a zero contract is optimal. Lemma 2.6 identifies its payoff guarantee to the principal; let  $F^*$  be the worst-case distribution given by that lemma. In the proof of Theorem 2.2, we obtained parameters  $\alpha, \beta$ , such that the corresponding affine-ceiling contract  $(M', w')$ , with reduced form  $W'$ , satisfies:

- $W'(p) \geq W(p)$  for all  $p$ , hence  $V_E(M', w'|\mathcal{I}_0) \geq V_E(M, w|\mathcal{I}_0)$ ;
- $V_P(M', w') \geq V_P(M, w)$  (from which  $(M', w')$  is again an optimal contract); and
- $\alpha(U(p) + E_p[\beta(\omega)]) \geq W'(p) \geq W(p)$ , with equality on the support of  $F^*$  (this holds because in the proof of Lemma 2.7 we showed equality in (2.8) throughout the support of the worst-case distribution).

The proof of Theorem 2.2 also showed that we cannot have  $V_E(M', w'|\mathcal{I}_0) > V_E(M, w|\mathcal{I}_0)$  strictly, because that would imply  $V_P(M', w') > V_P(M, w)$ , contradicting optimality of the original contract. Hence  $V_E(M', w'|\mathcal{I}_0) = V_E(M, w|\mathcal{I}_0)$ , and so  $F^*$  satisfies the constraints of (2.7) for the new contract  $(M', w')$ . Since  $E_{F^*}[U(p) - W'(p)] = E_{F^*}[U(p) - W(p)]$ ,  $F^*$  must actually attain the minimum in (2.7) (for the new contract) — otherwise the minimum would be strictly lower, and we would have  $V_P(M', w') < V_P(M, w)$ , a contradiction.

Now we show that (B.1) is an equality for the new contract  $(M', w')$ . From (2.7) we have

$$\begin{aligned}
V_P(M', w') &= E_{F^*}[U(p) - W'(p)] \\
&= E_{F^*}[(1 - \alpha)U(p) - \alpha E_p[\beta(\omega)]] \\
&= \frac{1 - \alpha}{\alpha} E_{F^*}[W'(p)] - E_{p_0}[\beta(\omega)] \\
&= \frac{1 - \alpha}{\alpha} V_E(M', w'|\mathcal{I}_0) - E_{p_0}[\beta(\omega)] \\
&= \frac{1 - \alpha}{\alpha} \max_{(F, c) \in \mathcal{I}_0} (E_F[W'(p)] - c) - E_{p_0}[\beta(\omega)].
\end{aligned}$$

Now plugging in  $W'(p) = \alpha(U_R(p; \beta) + E_p[\beta(\omega)])$  gives us equality in (B.1) for the new contract.  $\square$

**Proof of Proposition 2.9:** For the first statement: Let  $W$  and  $W'$  be the reduced forms of the two contracts respectively, so  $W(p) \leq W'(p)$  everywhere. If  $W = W'$  then we are done, so let  $p^*$  be a posterior at which  $W(p^*) < W'(p^*)$ .

Let  $(F, c)$  be an optimal experiment for the expert under contract  $(M, w)$  and technology  $\mathcal{I}_0$ . The same calculation as in the proof of Theorem 2.2 shows that inequality (2.15) holds here. We cannot have  $(M', w')$  weakly dominated by  $(M, w)$  unless

$V_P(M, w) \geq V_P(M', w')$ , which requires  $V_E(M', w'|\mathcal{I}_0) = V_E(M, w|\mathcal{I}_0)$ . In particular, we must have  $W'(p) = W(p)$  throughout the support of  $F$ , and  $(F, c)$  must be an optimal experiment under  $(M', w')$  and  $\mathcal{I}_0$  as well.

Let us show that  $c > 0$ . If  $c = 0$ , and  $E_F[W(p)] > 0$ , then  $V_P(0) \geq E_F[U(p)] > E_F[U(p) - W(p)] = V_P(M, w|\mathcal{I}_0) \geq V_P(M, w)$ , contradicting the assumption that  $(M, w)$  was eligible. And if  $E_F[W(p)] = 0$  then  $W$  is identically zero, contradicting the assumption that  $(M, w)$  was not the zero contract.

So now we know that  $c > 0$ , so  $V_E(M, w|\mathcal{I}_0) < E_F[W(p)]$ . For some small  $\epsilon$ , let  $F' = (1 - \epsilon)F + \epsilon\delta_{p^*}$ . Then  $E_{F'}[W(p)] < E_{F'}[W'(p)]$ . Choose  $c' > 0$  such that

$$E_{F'}[W(p)] - V_E(M, w|\mathcal{I}_0) < c' < E_{F'}[W'(p)] - V_E(M, w|\mathcal{I}_0).$$

This is doable since the left side is positive for small  $\epsilon$ .

Also, consider the payoffs from  $F$  under contract  $(M, w)$ . Since  $F$  does not achieve the minimum in Lemma 2.6 (due to the final assertion of the lemma), we have

$$V_P(M, w) < E_F[U(p) - W(p)]. \tag{B.3}$$

Let  $\underline{F}$  be a distribution that does achieve the minimum in Lemma 2.6, and let  $\underline{F}' = (1 - \zeta)\underline{F} + \zeta F$ , where  $\zeta > 0$  is small compared to  $\epsilon$ . In the proof of Proposition 2.8, we observed that the affine-ceiling contract  $W'$  must coincide with  $W$  on the support of  $\underline{F}$ . Consequently,  $W$  and  $W'$  agree on the support of  $\underline{F}'$ .

Now suppose the true technology is

$$\mathcal{I} = \mathcal{I}_0 \cup \{(F', c'), (\underline{F}', 0)\}.$$

Under contract  $(M, w)$ , we have

$$E_{F'}[W(p)] - c' < V_E(M, w|\mathcal{I}_0) = E_F[W(p)] - c < E_{\underline{F}'}[W(p)]$$

which means that  $(\underline{F}', 0)$  is the unique optimal experiment in  $\mathcal{I}$ . The resulting payoff for the principal is

$$V_P(M, w|\mathcal{I}) = (1 - \zeta)V_P(M, w) + \zeta E_F[U(p) - W(p)]. \tag{B.4}$$

Whereas under contract  $(M', w')$ , we have

$$V_E(M', w' | \mathcal{I}_0) = E_F[W'(p)] - c < E_{F'}[W'(p)] - c'$$

and also

$$E_{\underline{F}'}[W'(p)] = E_{\underline{F}'}[W(p)] = (1 - \zeta)V_E(M, w | \mathcal{I}_0) + \zeta E_F[W(p)] < E_{F'}[W'(p)] - c'$$

when  $\zeta$  is small. Thus  $(F', c')$  is the unique optimal experiment in  $\mathcal{I}$ . The resulting payoff for the principal is then

$$\begin{aligned} V_P(M', w' | \mathcal{I}) &= (1 - \epsilon)E_F[U(p) - W'(p)] + \epsilon(U(p^*) - W'(p^*)) \\ &= (1 - \epsilon)E_F[U(p) - W(p)] + \epsilon(U(p^*) - W'(p^*)). \end{aligned} \quad (\text{B.5})$$

Comparing (B.4) and (B.5) using (B.3), we see that when  $\epsilon$  and  $\zeta$  are small,  $V_P(M, w | \mathcal{I}) < V_P(M', w' | \mathcal{I})$ . So  $(M, w)$  cannot weakly dominate  $(M', w')$ , as desired.

Now we prove the final statement of the proposition. Let  $(M, w)$  be an arbitrary contract; we prove that it does not weakly dominate the maxmin-optimal affine-ceiling contract with parameters  $(\alpha, \beta)$  given by (2.18). We can assume  $(M, w)$  is also maxmin-optimal (otherwise the conclusion is obvious). If  $(M, w)$  is the zero contract, then its guarantee equals expression (2.18) with parameters  $(0, \beta')$ , for suitable  $\beta'$ . Uniqueness of the maximum then implies  $\alpha = 0$ , so the two contracts being compared are identical (both are zero). Thus, assume  $(M, w)$  is not the zero contract. The proof of Proposition 2.8 showed that if  $(\alpha', \beta')$  are the parameters given by Lemma 2.7, the resulting affine-transformed contract gives at least as high a guarantee as  $(M, w)$ , and this guarantee equals the maximand in (2.18). So  $(\alpha', \beta')$  also attain the maximum in (2.18). Uniqueness of the maximum implies  $(\alpha', \beta') = (\alpha, \beta)$ . Now this second statement of Proposition 2.9 follows from the first.  $\square$

## C Proofs for restricted versus unrestricted investment

Here we prove Proposition 3.1. We first develop some preliminary tools. Recall Proposition 2.8, which gives the formula (2.18) for the parameters of the optimal contract, and solve out for  $\alpha$ : given  $\beta$  and  $(F, c)$ , the  $\alpha$  that attains the maximum is  $\alpha =$

$\sqrt{c/(E_F[U_R(p; \beta)] + E_{p_0}[\beta(\omega)])}$ . Plugging in this value of  $\alpha$ , (2.18) turns into

$$\max_{\substack{\beta: \Omega \rightarrow \mathbb{R} \\ (F, c) \in \mathcal{I}_0}} \left( E_F[U_R(p; \beta)] + c - 2\sqrt{c(E_F[U_R(p; \beta)] + E_{p_0}[\beta(\omega)])} \right). \quad (\text{C.1})$$

Moreover, the optimal contract uses the corresponding value of  $\beta$ .

Let  $\beta_0$  be given by  $\beta_0(\omega) = -\min_{d \in D} u(d, \omega)$ , for each  $\omega$ . The optimal  $\beta$  certainly satisfies  $\beta(\omega) \leq \beta_0(\omega)$  for each  $\omega$ , since otherwise  $\beta$  can be decreased without changing the restricted decision space  $M$  in the corresponding affine-ceiling contract, thus saving money without affecting the expert's incentives for information acquisition. Thus, in the maximization (C.1), we can restrict the domain to  $\beta \leq \beta_0$  componentwise.

The resulting contract is unrestricted-investment iff  $\beta = \beta_0$ . Thus, if the maximum in (C.1) is not attained at  $\beta = \beta_0$ , then there is a properly-restricted-investment contract that is optimal; if it is attained at  $\beta = \beta_0$ , then an unrestricted-investment contract must be optimal.

**Lemma C.1.**  $U_R(p; \beta)$  is concave and increasing in  $\beta$ .

**Proof:** For values  $\beta$  and  $\beta'$ , let  $m \in M(\beta)$  and  $m' \in M(\beta')$  attain the respective maxima in the definitions of  $U_R(p; \beta), U_R(p; \beta')$ . Suppose  $\beta'' = \lambda\beta + (1 - \lambda)\beta'$  with  $\lambda \in [0, 1]$ . It is immediate that  $m'' = \lambda m + (1 - \lambda)m' \in M(\beta'')$ , from which

$$U_R(p; \beta'') \geq E_p[m''(\omega)] - E_p[\beta''(\omega)] = \lambda U_R(p; \beta) + (1 - \lambda)U_R(p; \beta').$$

And to see that  $U_R$  is increasing in  $\beta$ : If  $\beta \leq \beta'$  componentwise, and  $m$  attains the maximum for  $\beta$ , then  $m'(\omega) = m(\omega) + \beta'(\omega) - \beta(\omega)$  is in  $M(\beta')$  and gives the same value for the maximand in defining  $U_R(p; \beta')$ ; thus  $U_R(p; \beta') \geq U_R(p; \beta)$ .  $\square$

It follows that, for each posterior  $p$  and each state  $\omega$ , the one-sided partial derivative

$$\psi(p; \omega) = \left. \frac{\partial^- U_R}{\partial \beta(\omega)} \right|_{(p; \beta_0)}$$

is well-defined and nonnegative. We will use this throughout the proof below.

**Proof of Proposition 3.1:**

*Part (a).* Fix any state  $\omega^*$ . We first show that, under the assumptions of this part,  $\psi(p, \omega^*) = 0$  for all  $p$ . Suppose not; pick  $p^*$  with  $\psi(p^*, \omega^*) > 0$ . Write  $\kappa = \psi(p^*, \omega^*)$ . Then, for each  $\epsilon \geq 0$ , if  $\beta_\epsilon : \Omega \rightarrow \mathbb{R}$  is defined by  $\beta_\epsilon(\omega) = \beta_0(\omega)$  for  $\omega \neq \omega^*$  and  $\beta_\epsilon(\omega^*) = \beta_0(\omega^*) - \epsilon$ , we get  $U_R(p^*; \beta_\epsilon) \leq U_R(p^*; \beta_0) - \kappa\epsilon$ . (This follows by concavity of

$U_R$ .) That is, for any decision  $d$  satisfying  $u(d, \omega^*) \geq -\beta_0(\omega^*) + \epsilon$ , we have  $E_{p^*}[u(d, \omega)] \leq E_{p^*}[u(d(p^*), \omega)] - \kappa\epsilon$ . So for any  $d \in D$ , we can take  $\epsilon = u(d, \omega^*) + \beta_0(\omega^*)$  and obtain

$$E_{p^*}[u(d, \omega)] \leq E_{p^*}[u(d(p^*), \omega)] - \kappa(u(d, \omega^*) + \beta_0(\omega^*)). \quad (\text{C.2})$$

Plugging in  $d = d(p^*)$ , transposing, and dividing by  $\kappa > 0$  gives  $u(d(p^*), \omega^*) + \beta_0(\omega^*) \leq 0$ . But since  $u(d, \omega^*) + \beta_0(\omega^*) \geq 0$  for all  $d$  by definition of  $\beta_0$ , we must have equality:  $\beta_0(\omega^*) = -u(d(p^*), \omega^*)$ . Then, (C.2) rewrites as

$$E_{p^*}[u(d, \omega)] + \kappa u(d, \omega^*) \leq E_{p^*}[u(d(p^*), \omega)] + \kappa u(d(p^*), \omega^*) \quad (\text{C.3})$$

for all  $d \in D$ .

Hence, if we define  $q(\omega) = p^*(\omega)$  for each  $\omega \neq \omega^*$  and  $q(\omega^*) = p^*(\omega^*) + \kappa$ , (C.3) tells us that  $d(p^*)$  maximizes  $\sum_{\omega} q(\omega)u(d, \omega)$  over all  $d \in D$ . Hence, scaling down  $q$  by a factor of  $1 + \kappa$  gives another posterior belief, distinct from  $p^*$ , for which  $d(p^*)$  is optimal. This contradicts the assumption that each decision can only be optimal for one belief. (This argument would fail if  $p^*$  puts weight 1 on  $\omega^*$ , since then  $q/(1 + \kappa)$  is actually equal to  $p^*$ . But this cannot be the case:  $d(p^*)$  is the best decision for  $p^*$ , but we saw above that it is the worst decision in state  $\omega^*$ .)

This contradiction completes the proof that  $\psi(p, \omega) = 0$  for all  $p$  and  $\omega$ .

Now, the concavity of  $U_R$  implies that  $(U_R(p; \beta_0) - U_R(p; \beta_\epsilon))/\epsilon$  decreases as  $\epsilon \rightarrow 0$ , so we can invoke the monotone convergence theorem to conclude that

$$\left. \frac{d^+}{d\epsilon} E_F[U_R(p; \beta_\epsilon)] \right|_{\epsilon=0} = E_F \left[ \left. \frac{d^+}{d\epsilon} U_R(p; \beta_\epsilon) \right|_{\epsilon=0} \right] = 0 \quad \text{for any } F \in \Delta(\Delta(\Omega)).$$

Now consider any  $\mathcal{I}_0$ , and consider the maximand in (C.1), corresponding to the optimal contract. The maximizing choice of  $(F, c)$  must have  $c > 0$ , since otherwise the zero contract is optimal, contrary to assumption. Suppose that the maximizing choice of  $\beta$  is  $\beta_0$ . Consider replacing  $\beta$  by  $\beta_\epsilon$ . Then the derivative of the maximand in (C.1) with respect to  $\epsilon$  (at 0) is equal to

$$p_0(\omega) \sqrt{\frac{c}{E_F[U_R(p; \beta_0)] + E_{p_0}[\beta_0(\omega)]}} > 0.$$

This contradicts maximality at  $\beta_0$ . So the maximum is not attained at  $\beta_0$  after all, and so our optimal contract is a properly-restricted-investment contract.

*Part (b).* In this case, we first choose, for each state  $\omega$ , a posterior  $p_\omega$  lying in the corresponding open set  $P$ . We claim that  $\psi(p_\omega, \omega) > 0$  for each state  $\omega$ .

This is essentially the mirror image of the argument at the beginning of (a). Fix any state  $\omega^*$ , and put  $p^* = p_{\omega^*}$ . By assumption, there is an  $\omega^*$ -extremal decision  $d^* = d(p^*)$  that is optimal throughout the corresponding  $P$ . In particular, if we take  $\kappa > 0$  sufficiently small, and define a belief  $q$  by  $q(\omega) = p^*(\omega)/(1 + \kappa)$  for all  $\omega \neq \omega^*$  and  $q(\omega^*) = (p^*(\omega^*) + \kappa)/(1 + \kappa)$ , then  $d^*$  will be optimal for each such belief  $q$ . Multiplying through by  $1 + \kappa$  tells us that, for every  $d \in D$ , we have

$$\sum_{\omega} p^*(\omega)u(d, \omega) + \kappa u(d, \omega^*) \leq \sum_{\omega} p^*(\omega)u(d^*, \omega) + \kappa u(d^*, \omega^*)$$

or

$$\begin{aligned} E_{p^*}[u(d, \omega)] &\leq E_{p^*}[u(d^*, \omega)] - \kappa(u(d, \omega^*) - u(d^*, \omega^*)) \\ &= E_{p^*}[u(d^*, \omega)] - \kappa(u(d, \omega^*) + \beta_0(\omega^*)) \end{aligned}$$

where the equality holds by assumption that  $d^*$  is  $\omega^*$ -extremal. In particular, if  $d$  is a decision satisfying  $u(d, \omega^*) \geq -\beta_0(\omega^*) + \epsilon$ , then we have  $E_{p^*}[u(d, \omega)] \leq E_{p^*}[u(d^*, \omega)] - \kappa\epsilon$ . It follows that  $\psi(p^*, \omega^*) \geq \kappa$ , proving the claim.

So now let  $\eta = \min_{\omega} \psi(p_\omega, \omega) > 0$ . Let  $F$  be any distribution with full support, and mass at least  $\epsilon$  on each  $p_\omega$ , whose mean equals the given  $p_0$ . We will show that an unrestricted-investment contract is optimal under  $\mathcal{I}_0 = \{(\delta_{p_0}, 0), (F, c)\}$ , for sufficiently small  $c > 0$ . Note that full support and non-triviality are satisfied.

We claim that for any  $\beta$  that is  $\leq \beta_0$  componentwise,

$$E_F[U_R(p; \beta)] \leq E_F[U_R(p; \beta_0)] - \frac{\epsilon\eta}{|\Omega|} \sum_{\omega} (\beta_0(\omega) - \beta(\omega)). \quad (\text{C.4})$$

Indeed, choose the state  $\omega^*$  for which  $\beta_0(\omega) - \beta(\omega)$  is largest. Put  $\beta_1(\omega^*) = \beta(\omega^*)$  and  $\beta_1(\omega) = \beta_0(\omega)$  for other states  $\omega$ . Then increasingness and concavity imply

$$U_R(p; \beta) \leq U_R(p; \beta_1) \leq U_R(p; \beta_0) - (\beta_0(\omega^*) - \beta(\omega^*))\psi(p, \omega^*).$$

Applying expectations under  $F$ , noting that  $p = p_{\omega^*}$  arises with probability at least  $\epsilon$ , and  $\beta_0(\omega^*) - \beta(\omega^*) \geq \sum_{\omega} (\beta_0(\omega) - \beta(\omega))/|\Omega|$  by choice of  $\omega^*$ , leads to (C.4).

Thus, on the domain  $\beta \leq \beta_0$ , the function  $E_F[U_R(p; \beta)]$  is bounded above by an affine

function of  $\beta$  that is uniquely maximized at  $\beta_0$ , and this upper bound (C.4) holds with equality at  $\beta = \beta_0$ . Since the quantity  $\sqrt{E_F[U_R(p; \beta)] + E_{p_0}[\beta(\omega)]}$  is locally Lipschitz in  $\beta$  near  $\beta_0$ , adding a small multiple of it will not change this fact. Hence, for  $c$  sufficiently small, the maximand in (C.1) is still uniquely maximized over  $\beta$  by taking  $\beta = \beta_0$ . (The optimum with respect to choice of experiment must indeed be given by  $(F, c)$ , not  $(\delta_{p_0}, 0)$ , by non-triviality.) It follows that an unrestricted-investment contract is optimal.

*Part (c).* We may shrink  $P$  if necessary to assume it is an open ball. Let  $F_0$  be any continuous distribution over posteriors, whose support is the closure of  $P$ . The assumption that no decision is optimal throughout  $P$  ensures that  $E_{F_0}[U(p)] > U(E_{F_0}[p])$ . By continuity, this remains true when we replace  $P$  by a sufficiently large closed sub-ball  $\bar{P}$ , and replace  $F_0$  with the distribution conditioned on  $p \in \bar{P}$ ; call this distribution  $\bar{F}$ . Let the prior be  $p_0 = E_{\bar{F}}[p]$ . Then  $p_0 \in P$  by convexity.

As noted at the beginning of Subsection 2.3, if  $c > 0$  satisfies the bound (2.3), then  $\mathcal{I}_0 = \{(\delta_{p_0}, 0), (\bar{F}, c)\}$  satisfies the non-triviality assumption. Let  $F'$  be any distribution with full support on  $\Delta(\Omega)$  and with mean  $p_0$ . By continuity of the formula (2.18), we can take  $F = (1 - \epsilon)\bar{F} + \epsilon F'$  for sufficiently small  $\epsilon > 0$ , and then  $\mathcal{I}_0 = \{(\delta_{p_0}, 0), (F, c)\}$  still satisfies non-triviality. It clearly satisfies full support as well.

We will show that for  $\epsilon$  sufficiently small, under this  $\mathcal{I}_0$ , a properly-restricted-investment contract is optimal. Compactness implies that for all  $p \in \bar{P}$ , optimal decisions are uniformly bounded away from extremal decisions; that is, one can choose  $\beta_1 : \Omega \rightarrow \mathbb{R}$  with  $\beta_1(\omega) < \beta_0(\omega)$ , such that for all  $p \in \bar{P}$ , any optimal decision  $d$  satisfies  $u(d, \omega) + \beta_1(\omega) \geq 0$  for each  $\omega$ .

In particular,  $U_R(p; \beta_1) = U(p) = U_R(p; \beta_0)$  for each such  $p$ .

Now, as long as  $\epsilon$  is small enough, we know that the optimum in (C.1) is attained with  $(F, c)$  rather than  $(\delta_{p_0}, 0)$  (by non-triviality). On the other hand, it is immediate that when  $\epsilon = 0$ , the value of (C.1) is strictly higher with  $\beta = \beta_1$  than  $\beta_0$ , since the  $U_R(p; \beta)$  terms are unchanged and the  $E_{p_0}[\beta(\omega)]$  term is smaller. By continuity, this is still true for sufficiently small  $\epsilon$ . So in this case, the value of the maximum in (C.1) is attained at some  $\beta \neq \beta_0$ , which means that the corresponding optimal contract is a properly-restricted-investment contract. □

**Proof of Proposition 3.2:** Let  $D = \{d_0, d_1\}$ , and let  $D'$  be the space of all randomized decisions, which we notate as  $\{(1 - \gamma)d_0 + \gamma d_1 \mid \gamma \in [0, 1]\}$ , thereby making the problem convex. So we know that there are some  $\alpha$  and  $\beta$  such that the corresponding restricted-investment contract, described in Proposition 2.3, is optimal.

Note that the set of  $\gamma$ 's corresponding to decisions satisfying (2.6) is closed and convex, so it is an interval, say  $[\underline{\gamma}, \bar{\gamma}]$ . Also, if  $\bar{\gamma} = \underline{\gamma}$  then the restricted-investment contract allows only a single message, hence there is zero incentive for information acquisition and the optimal contract is the zero contract, contrary to assumption. Hence,  $\bar{\gamma} > \underline{\gamma}$  (and  $\alpha > 0$ ).

Now, notice that

$$\alpha \left( u((1 - \gamma)d_0 + \gamma d_1, \omega) + \beta(\omega) \right) \quad \text{with } \gamma \in [\underline{\gamma}, \bar{\gamma}]$$

is equal to

$$\alpha' \left( u((1 - \delta)d_0 + \delta d_1, \omega) + \beta'(\omega) \right)$$

where

$$\alpha' = \alpha(\bar{\gamma} - \underline{\gamma}), \quad \beta'(\omega) = \frac{1}{\bar{\gamma} - \underline{\gamma}} \left( (1 - \bar{\gamma})u(d_0, \omega) + \underline{\gamma}u(d_1, \omega) + \beta(\omega) \right)$$

and  $\delta$  is linearly related to  $\gamma$  by

$$\delta = \frac{\gamma - \underline{\gamma}}{\bar{\gamma} - \underline{\gamma}} \in [0, 1].$$

Therefore, the restricted-investment contract with parameters  $\alpha, \beta$ , and restricted decision space given by  $\{(1 - \gamma)d_0 + \gamma d_1 \mid \gamma \in [\underline{\gamma}, \bar{\gamma}]\}$ , is equivalent to the restricted-investment contract with parameters  $\alpha', \beta'$ , and the full decision space  $D'$  — that is, the unrestricted-investment contract with these parameters  $\alpha', \beta'$ .

Finally, since the optimal decision in  $D'$  is always either  $d_0$  or  $d_1$ , this is equivalent to the unrestricted-investment contract with parameters  $\alpha', \beta'$  over the original pure decision space  $D$ .  $\square$

## References

- [1] Dirk Bergemann and Ulrich Hege (1998), “Venture Capital Financing, Moral Hazard, and Learning,” *Journal of Banking & Finance* 22 (6–8), 703–735.
- [2] Dirk Bergemann and Stephen Morris (2005), “Robust Mechanism Design,” *Econometrica* 73 (6), 1771–1813.

- [3] Sugato Bhattacharyya and Francine Lafontaine (1995), “Double-Sided Moral Hazard and the Nature of Share Contracts,” *RAND Journal of Economics* 26 (4), 761–781.
- [4] Tilman Börgers (2017), “(No) Foundations of Dominant-Strategy Mechanisms: A Comment on Chung and Ely (2007),” *Review of Economic Design* 21 (2), 73–82.
- [5] Benjamin Brooks and Songzi Du (2018), “Optimal Auction Design with Common Values: An Informationally-Robust Approach,” unpublished paper, University of Chicago.
- [6] Gabriel Carroll (2015), “Robustness and Linear Contracts,” *American Economic Review* 105 (2), 536–563.
- [7] Gabriel Carroll (2017), “Robustness and Separation in Multidimensional Screening,” *Econometrica* 85 (2), 453–488.
- [8] Gabriel Carroll and Ilya Segal (2018), “Robustly Optimal Auctions with Unknown Resale Opportunities,” *Review of Economic Studies*, forthcoming.
- [9] Arthur Carvalho (2016), “An Overview of Applications of Proper Scoring Rules,” *Decision Analysis* 13 (4), 223–242.
- [10] Sylvain Chassang (2013), “Calibrated Incentive Contracts,” *Econometrica* 81 (5), 1935–1971.
- [11] Yiling Chen, Ian Kash, Mike Ruberry, and Victor Shnayder (2011), “Decision Markets with Good Incentives,” in *Internet and Network Economics* (WINE-11), 72–83.
- [12] Kim-Sau Chung and J. C. Ely (2007), “Foundations of Dominant-Strategy Mechanisms,” *Review of Economic Studies* 74 (2), 447–476.
- [13] Aubrey Clark (2016), “Contracts for Information Acquisition,” unpublished paper, Harvard University.
- [14] Amanda Cowen, Boris Groyberg, and Paul Healy (2006), “Which Types of Analyst Firms are More Optimistic?” *Journal of Accounting and Economics* 41 (1–2), 119–146.
- [15] Vincent P. Crawford and Joel Sobel (1982), “Strategic Information Transmission,” *Econometrica* 50 (6), 1431–1451.

- [16] Joel S. Demski and David E. M. Sappington (1987), “Delegated Expertise,” *Journal of Accounting Research* 25 (1), 68-89.
- [17] Alexander Frankel (2014), “Aligned Delegation,” *American Economic Review* 104 (1), 66–83.
- [18] Daniel Garrett (2014), “Robustness of Simple Menus of Contracts in Cost-Based Procurement,” *Games and Economic Behavior* 87, 631–641.
- [19] Dino Gerardi and Lucas Maestri (2012), “A Principal-Agent Model of Sequential Testing,” *Theoretical Economics* 7 (3), 425–463.
- [20] Tilmann Gneiting and Adrian E. Raftery (2007), “Strictly Proper Scoring Rules, Prediction, and Estimation,” *Journal of the American Statistical Association* 102 (477), 359–378.
- [21] Marina Halac, Navin Kartik, and Qingmin Liu (2016), “Optimal Contracts for Experimentation,” *Review of Economic Studies* 83 (3), 1040–1091.
- [22] Johannes Hörner and Larry Samuelson (2013), “Incentives for Experimenting Agents,” *RAND Journal of Economics* 44 (4), 632–663.
- [23] Edi Karni (2009), “A Mechanism for Eliciting Probabilities,” *Econometrica* 77 (2), 603–606.
- [24] Jean-Jacques Laffont and Jean Tirole (1993), *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.
- [25] Assar Lindbeck and Jörgen Weibull (2017), “Investment, Rational Inattention, and Delegation,” unpublished paper, Stockholm University.
- [26] James M. Malcolmson (2009), “Principal and Expert Agent,” *B.E. Journal of Theoretical Economics*, 9 (1) (Contributions), Article 17.
- [27] Gustavo Manso (2011), “Motivating Innovation,” *Journal of Finance* 66 (5), 1823–1860.
- [28] Iván Marinovic, Marco Ottaviani, and Peter Sørensen (2013), “Forecasters’ Objectives and Strategies,” in Graham Elliott and Allan Timmermann, eds., *Handbook of Economic Forecasting* vol. 2B, 691–720.

- [29] Abraham Othman and Tuomas Sandholm (2010), “Decision Rules and Decision Markets,” in *Autonomous Agents and Multi-Agent Systems* (AAMAS ’10), 625–632.
- [30] R. Tyrrell Rockafellar (1970), *Convex Analysis*. Princeton, NJ: Princeton University Press.
- [31] Leonard J. Savage (1971), “Elicitation of Personal Probabilities and Expectations,” *Journal of the American Statistical Association* 66 (336), 783–801.
- [32] Peng Shi, Vincent Conitzer, and Mingyu Guo (2009), “Prediction Mechanisms That Do Not Incentivize Undesirable Actions,” in *Internet and Network Economics* (WINE-09), 89-100.
- [33] Jens Witkowski, Rupert Freeman, Jennifer Wortman Vaughan, David M. Pennock, and Andreas Krause (2018), “Incentive-Compatible Forecasting Competitions,” in *AAAI Conference on Artificial Intelligence* (AAAI ’18), 1282–1289.
- [34] Luis Zermelo (2011), “A Principal-Expert Model and the Value of Menus,” unpublished paper, Massachusetts Institute of Technology.
- [35] Luis Zermelo (2012), “The Role of Authority in a General Principal-Expert Model,” unpublished paper, Massachusetts Institute of Technology.