Only Valuable Experts Can Be Valued

Moshe Babaioff* Liad Blumrosen† Nicolas S. Lambert‡ Omer Reingold§

March 30, 2011

Abstract

Suppose a principal Alice wishes to reduce her uncertainty regarding some future payoff. Consider a self-proclaimed expert Bob that may either be an informed expert knowing an exact (or approximate) distribution of a future random outcome that may affect Alice's utility, or an uninformed expert who knows nothing more than Alice does. Alice would like to hire Bob and solicit his signal. Her goal is to incentivize an informed expert to accept the contract and reveal his knowledge while deterring an uninformed expert from accepting the contract altogether. The starting point of this work is a powerful negative result (Olszewski and Sandroni, 2007), which tells us that in the general case for any contract which guarantees an informed expert some positive payoff an uninformed expert (with no extra knowledge) has a strategy which guarantees him a positive payoff as well.

At the face of this negative result, we reexamine the notion of an expert and conclude that knowing some hidden variable (i.e., the description of the aforementioned distribution), does not make Bob an expert, or at least not a "valuable expert". The premise of our paper is that if Alice only tries to incentivize experts which are valuable to her decision making then she can indeed screen them from uninformed experts.

On a more technical level, we consider the case where Bob's signal about the distribution of a future event cannot be an arbitrary distribution but rather comes from some subset $\mathcal{P}$ of all possible distributions. We give rather tight conditions on $\mathcal{P}$ (which relate to its convexity), under which screening is possible. We formalize our intuition that if these conditions are not met then an expert is not guaranteed to be valuable. We give natural and arguably useful scenarios where indeed such a restriction on the distribution arise.

1 Introduction

Consider a principal Alice and a self-proclaimed expert Bob, claiming to have information that may affect Alice's future payoff. Alice is interested in reducing her uncertainty about the future payoff. To do so she would like to offer Bob a contract where the following conditions hold: (1) If Bob is indeed an informed expert then Bob will have an incentive to accept the contract and reveal his true signal, and (2) If Bob is uninformed (meaning that Bob does not know more than Alice) then Bob will reject the contract. In this paper we consider conditions that make such a contract possible.

More concretely, assume that p^* is the probability that some future Boolean event X will take the value one.[1] Assume that if Bob is informed then he knows the probability p^* exactly[2], while if

*Microsoft Research - Silicon Valley, moshe@microsoft.com

†Department of Economics, The Hebrew University, blumrosen@huji.ac.il. Part of this research was performed while at Microsoft Research - Silicon Valley

‡Stanford University Graduate School of Business, nlambert@gsb.stanford.edu. Part of this research was performed while visiting Microsoft Research - Silicon Valley

§Microsoft Research - Silicon Valley, omer.reingold@microsoft.com

[1] While our results apply to the case where the set of outcomes is an arbitrary finite set, to simplify the exposition in the introduction we concentrate on the Boolean case.

[2] In the paper we also consider a much more subtle and general scenario where informed experts are not always

uninformed he does not know the value of p^*. A contract Π determines the reward that Bob will receive after reporting his signal q (which supposedly equals the true distribution p^*) and when the outcome of X is later observed to be ω. In such a case, Bob's reward would be $\Pi(q, \omega)$. (Note that $\Pi(q, \omega)$ may be negative and in such a case, as the contract is binding, Bob will have to pay Alice $-\Pi(q, \omega)$). Bob's expected reward when reporting q and ω is distributed according to p^* is $\Pi(q, p^*)$.

A rich literature, which deals with scoring rules (see the survey [21] and more references in Section 1.3), gives methods of defining a contract Π that addresses Alice's first goal. This means that $\Pi(p^*, p^*)$ can be made larger than some predefined threshold δ_A, *for every value of* p^* (recall that ω is one with probability p^*). In other words, a risk neutral informed Bob which knows p^* and has an acceptance threshold δ_A, has an incentive to accept the contract. Furthermore, once Bob accepts the contract it is the unique dominant strategy for Bob to report his true signal (we then say that the contract is *strictly proper*). In fact, contracts with very useful properties are known. For example, based on the Brier scoring rule (also known as Quadratic scoring rule, see [4]), one can get a contract such that the reward of Bob improves as the Euclidean distance of his report q from the true probability p^* decreases (in this paper we refer to such contracts as "distance-rewarding contracts", see more details in Section 2).

Now that informed experts can be incentivized to reveal their true signal, can we simultaneously deter uninformed experts from accepting the contract? In other words, can we screen informed experts from uninformed experts? Unfortunately, it was shown by Olszewski and Sandroni [16], that Alice's two goals are conflicting: If every informed expert has an incentive to accept the contract (meaning for example that $\Pi(p^*, p^*)$ is positive for every p^*), then an uninformed Bob (knowing nothing but the description of Π) has a randomized report q such that both $\Pi(q, 1)$ and $\Pi(q, 0)$ are positive in expectation (where the expectation is now only over Bob's randomized report q).

Our Approach In this paper we attempt to explain the difficulty in screening experts and to characterize the cases where one can motivate informed experts without rewarding uninformed experts. We consider a scenario where the probability p^* belongs to a set of distributions $\mathcal{P}$, where $\mathcal{P}$ is not necessarily the set of all distributions but rather possibly a subset of it.[3] We give rather tight conditions on $\mathcal{P}$ under which screening is possible.

In the following we will try to describe these conditions and to give natural examples where such conditions hold. Furthermore, we will try to argue that when these conditions do not hold then some experts are not useful to Alice (in some formal sense). The premise of this paper is thus that the quality of experts can be valued only when experts are guaranteed to be valuable (justifying the title of this paper). Indeed, we show that even without any restriction on p^*, Alice can offer Bob a contract such that if Bob is uninformed he will reject the contract whereas if Bob is a "valuable" expert then he will have an incentive to accept the contract and be truthful.

Our results come in two flavors. In the first, we assume that Alice has some prior P over $\mathcal{P}$. This prior is also known to Bob (even when Bob is an uninformed expert). In the second, neither Alice nor an uninformed Bob have any prior over $\mathcal{P}$ (this is the setting of [16]). We will start discussing the first scenario, where indeed our intuition that experts can be screened if and only if they are sufficiently valuable is much cleaner.

perfectly informed. Such an expert knows p^* up to some small error $\epsilon > 0$. For simplicity of the exposition in the introduction, unless explicitly stated otherwise, we assume that an informed expert is perfectly informed.

[3]In the context of forecast testing, restrictions on the set of distributions were recently considered by Olszewski and Sandroni [18] and Al-Najjar et al. [1], see Section 1.3 for discussion of related work.

1.1 Screening at the Presence of a Prior

Consider the case where the prior P (known to Alice as well as to an uninformed expert) is that p^* is distributed uniformly between 0 and 1. From the point of view of Alice, the probability that $\omega = 1$ is $\frac{1}{2}$. An informed expert which knows that $p^* \neq \frac{1}{2}$ has an advantage in predicting ω and thus we consider such an expert valuable (and the further p^* is from $\frac{1}{2}$ the more valuable the expert is). On the other hand, if $p^* = \frac{1}{2}$ then although the informed expert knows that $p^* = \frac{1}{2}$ he has no advantage in predicting ω over an uninformed expert. In other words, being an expert should be determined not by knowledge of some hidden variable (the value of p^* in our scenario) but rather by better ability to predict the event that will eventually be measured (the value of ω).

More generally, consider some prior P and let $\bar{p}$ denote the expected value of p^* according to the prior P. It is easy to argue (by the linearity of expectation), that if $\bar{p} \in \mathcal{P}$ (that is, the expected value is a valid report) and if for every value of $p^* \in \mathcal{P}$ it holds that $\Pi(p^*, p^*) > \delta_A$ (which means that every informed expert accepts the contract) then an uninformed expert has an incentive to accept the contract and report $\bar{p}$, as his expected reward (when p^* is sampled according to P) is $E[\Pi(\bar{p}, p^*)] = \Pi(\bar{p}, \bar{p}) > \delta_A$.

In the presence of a prior, we therefore observe that if $\bar{p} \in \mathcal{P}$ then screening experts is impossible. We also have the intuition that in this case, some informed experts (specifically, those who know that $p^* = \bar{p}$) are not valuable. We complement this with a positive result that screening *is possible* whenever $\bar{p} \notin \mathcal{P}$.[4] We will demonstrate that this condition is not only tight but it is also quite applicable by discussing natural examples where $\bar{p} \notin \mathcal{P}$.

We further observe that our contract can also differentiate between uninformed experts and "partially informed" experts, where a partially-informed expert has a signal q which is at most ϵ away from p^*. For such a screening to be possible, we need to make sure that ϵ is sufficiently smaller than the distance between $\bar{p}$ and p^*.

What if $\bar{p} \in \mathcal{P}$? We can use the same contract that was used when $\bar{p} \notin \mathcal{P}$ and observe that this contract is meaningful even when $\bar{p} \in \mathcal{P}$. In such a case we still have that an uninformed expert rejects the contract. We also have that it is the best response of any perfectly-informed expert which accepts the contract to be truthful. Finally, every "sufficiently useful" expert, that is one who knows p^* which is sufficiently separated from $\bar{p}$, will accept the contract (with some care, we can make sure that if p^* is too close to $\bar{p}$ an informed expert will reject). In other words, Alice manages to incentivize useful reports and deter uninformed or non-useful reports. This is different than the original goal of screening experts, but is arguably even more natural.

We next move to present few examples that demonstrates the usefulness of our characterization.

1.1.1 Examples

As argued above, given a prior P on a closed set $\mathcal{P}$ such that $\bar{p} \notin \mathcal{P}$, we can show that screening experts is possible. This can only happen when $\mathcal{P}$ is non-convex. Actually, there is always a prior over non-convex sets for which the average distribution $\bar{p}$ is not in $\mathcal{P}$; Screening is thus possible with such prior, for such non-convex sets of distributions. We will now mention several natural examples of such scenarios.

- **Binary outcomes.** Consider a binary event X, say an outcome of a coin, and let p be the probability that $\omega = 1$. There are 3 cases regarding $\mathcal{P}$ that we discuss:

[4] For this statement, we assume that $\mathcal{P}$ is closed or alternatively that not only $\bar{p}$ is missing from $\mathcal{P}$ but also some neighborhood around $\bar{p}$. This is still very intuitive, as if the distribution p^* may be arbitrarily close to $\bar{p}$ then the usefulness of getting the signal from an informed expert may be arbitrarily small.

- When p is known to be one of *three* probabilities (e.g., $p \in \{1/4, 1/2, 3/4\}$), there exists a prior under which screening becomes impossible (when a mixture of the two extremes has expected bias equal to the middle one) as an uninformed expert will accept the contract and report the expectation. This is the case that $\bar{p} \in \mathcal{P}$. Clearly there also exists another prior for which $\bar{p} \notin \mathcal{P}$, in that case screening is possible.
- If p can be any value in $[0, 1]$ the set $\mathcal{P}$ is closed and convex. In this case for *any* prior P screening is *impossible* as the expectation $\bar{p}$ always belong to $\mathcal{P}$.
- If p is known to be one of *two* probabilities. For example, a coin may be either fair ($p^* = 1/2$) or biased ($p^* = 2/3$). In this case, for *any* non-trivial prior over these two possible probabilities we have that $\bar{p} \notin \mathcal{P}$, which makes screening possible for *every* non-trivial prior.

- **Two samples.** Consider a case where $p \in [0, 1]$ is the distribution of a binary event, but now Alice observes *two* independent samples, each is one with probability p. For example, think about a factory trying to estimate the defect rate in its production chain by inspecting two independent samples. In this case, there are actually four possible realizations for the outcome: $\{0, 0\}, \{0, 1\}, \{1, 0\}, \{1, 1\}$ with respective probabilities of $p^2, p(1 - p), (1 - p)p, (1 - p)^2$. The set of possible distributions is

$$\mathcal{P} = \{p^2, p(1\text{-}p), (1\text{-}p)p, (1\text{-}p)^2 | p \in [0, 1]\}$$

  and this set is not convex, and moreover, every non-trivial convex combination of elements in $\mathcal{P}$ is not in $\mathcal{P}$.[5] This implies that for every non-trivial prior on $\mathcal{P}$ the expected distribution $\bar{p}$ is not in $\mathcal{P}$. Thus, an immediate implication of our results is that with two samples, screening is possible for *every* (non-trivial) prior on $\mathcal{P}$.

- In general, many standard families of distributions are not closed under the mixture operator (that is, averaging with some prior). For example, taking the average of two (discrete analogs of) normal distributions may result in a double-peaked distribution that can be far from any normal distribution. Another example is the set of uniform distributions over subsets of Ω which is clearly not closed under mixing as well.

1.2 Screening with no Prior

We now consider the case where neither an uninformed expert nor Alice knows a prior over $\mathcal{P}$. This is the setting of the negative result of Olszewski and Sandroni [16]. We observe that their result generalizes to the case where p^* belongs to $\mathcal{P}$, as long as $\mathcal{P}$ is compact and convex. This observation should not be very surprising as the main technical tool of [16] is a Min-Max Theorem. We complement this result by a positive result - if $\mathcal{P}$ is missing one of its convex combination (and some neighborhood around it) then there exists a contract that screens informed experts from uninformed experts. This is the case, for example, for a binary random variable with only two possible probabilities.

Aren't all experts valuable when there is no prior? When we assumed that a prior existed, it was easy to argue that an expert telling Alice that p^* in fact equals $\bar{p}$ is not useful. After all, Alice did not gain better prediction power for the value of ω. But when Alice does not know anything about

[5]To see this, consider the set of points $G = \{\ (p^2, (1 - p)^2)\ \}_{p \in [0,1]}$ in $\mathbb{R}^2$. We will argue that every non-trivial convex combination of elements in G does not belong to G, and since G is composed of the first and last coordinates of vectors in $\mathcal{P}$ it will follow that any (non-trivial) convex combination of elements in $\mathcal{P}$ is not in $\mathcal{P}$. Indeed, the set G can be represented by the equation $y = (1 - \sqrt{x})^2$; Since this is a strictly convex function ($y'' > 0$), any non-trivial convex combination of its points will fall strictly above the function's curve, that is, is not of the form $(q^2, (1 - q)^2)$ (for some $q \in [0, 1]$).

p^*, how can learning p^* not be useful? To address this question we recall that Alice needs information regarding ω in order to select an action in such a way that will maximize her utility. Very loosely, we show that when $\mathcal{P}$ is compact and convex, then regardless of Alice's utility function, there is some value q^* such that learning that $p^* = q^*$ does not help Alice to improve her utility (see Theorem 16 for the formal statement). In this respect an expert who knows $p^* = q^*$ is not useful to Alice.

It is interesting to note that our proof of the above claim (regarding the existence of q^*) heavily relies on the negative result of [16]. Therefore, we can reinterpret their result as telling us that whenever $\mathcal{P}$ is compact and convex, then not all experts are useful (and thus screening is indeed impossible).

The existence of q^* that is not useful, as discussed above, allows Alice to define a contract that differentiates between uninformed experts and informed experts that are also valuable (in the sense that they allow Alice to significantly increase her utility). Unlike the analogues results discussed in Section 1.1 for the case with a prior, the contract here will have to depend on q^* which may depend on Alice's utility and strategy. A similar (though possibly even stronger), dependance occurs in a different approach which can be referred to as "selling the project". In this approach, Alice offers Bob a contract where his reward equals the additional expected utility Alice got due to his report. This gain can be calculated after the value of ω is observed. With this approach, the expected reward Bob gets out of the contract is identical to his value as an expert to Alice. Thus this is a different way one may argue that experts can be valued as long as they are valuable. This observation is mostly orthogonal to our work, as we show tight characterization for when an expert is guaranteed to be valuable. In addition, as already mentioned, this approach has the disadvantage that the contract has to depend on Alice's utility and strategy, which may be at times rather undesirable. For example, when the utility function and strategy of Alice are proprietary information that should be kept secret even from Bob, and therefore she may be prevented from revealing those to Bob (by specifying to Bob a contract Π which depends on them).

1.3 Related Work

The concern that experts may misreport their forecasts when facing a test can be traced back at least to Brier [4]. Various *scoring rules* have been designed to incentivize the experts to report truthfully (see Brier [4], Good [13] and Savage [20]), without worrying about the possibility of uninformed experts.

A series of work studied the related problem of *forecast testing* or *forecasting.* In forecasting, Alice has an interest in an unbounded sequence of probabilistic events. An expert Bob provides a forecast for each event before it occurs. In this setting, negative results were shown on the ability to *test* if Bob is indeed informed (we omit the formal model from this discussion). It was shown that for any reasonable test, if an informed expert is guaranteed to pass the test then so does a completely uninformed expert who provides a fixed sequence of forecasts (independent of the true probabilities); See, e.g., [10], [11], [15], [19]. Several works searched for ways around these impossibilities. Fortnow and Vohra [9] assumed limitations on the computational power of the uninformed agent. Al-Najjar and Weinstein [2] and Feinberg and Stewart [8] assumed the presence of multiple forecasters, where one of the experts must be informed.

Some recent work by Olszewski and Sandroni [18] and Al-Najjar et al. [1] is the most related to the current paper, in the sense that they assume a structure on the set of possible distributions in forecasting settings. They follow two earlier papers ([6, 17]) that show a non-manipulable forecasting test. In [18] in was shown that in these tests, an uniformed Bob can delay its rejection for an arbitrary number of periods, and, more generally, that such phenomena must exist for every test, as long as the class of possible theories is convex. They construct a counter-example of a non-convex set of theories for which their negative result does not hold. We consider one-shot events and therefore the delayed rejection is not an issue, and go beyond presenting a counter-example by showing (for our single period setting) that our convexity restrictions on the class of distributions are essentially necessary and sufficient.

Al-Najjar et al. [1] were also concerned by the delayed decision of existing tests and thus they also restricted the class of possible probabilistic processes. They presented a convex set $\mathcal{P}^*$ of theories referred to as *learnable and predictable*; they showed that when the possible theories are drawn from this set then experts can be tested in a timely manner. Convexity also plays a major role in their work; While the $\mathcal{P}^*$ class is convex, their test asks the expert to report in advance a period T in which he will send his predictions for future events (after a learning period), but the class of theories that are learnable by time T is not convex. This goes along with the premise of the current paper that non-convexity of the set of distributions may make testing (or screening) accomplishable. It is interesting to note that [1] raised the question of what is an expert. They argue that it is natural to expect an expert to provide a report that satisfies their conditions. In our case we argue that experts that cannot be screened are not just missing valuable expertise but rather their reports are not useful to Alice.

Subsequent Work - Forecasting In a subsequent work [3], the authors extend the approach of the current paper to the problem of forecasting. In [3], we again consider the case when the distribution of each event comes from a subset $\mathcal{P}$. We characterize the conditions on $\mathcal{P}$ that allow testing Bob's experts. We consider two types of tests. In the weaker type of test, if Bob is completely uninformed then there exists a sequence of probability distributions (i.e. a choice of nature), which will fail him (this is a complement to the strong impossibility mentioned above). In the strong type of tests, Bob passes the test if and only if his stream of forecasts essentially merge to the stream of probability distributions.

While the setting of forecasting is technically more involved than the "one shot" setting considered in the current paper, the classifications on $\mathcal{P}$ we offer in both papers is in close correspondence (and in both cases relates to convexity conditions on $\mathcal{P}$). Similarly, the interesting special cases are shared by the two papers. For example, [3] gives a test of the strong type whenever at each round the observed event is composed of two independent samples of the same distribution (unknown to Alice but known to an informed expert).

Organization In Section 2 we present the model and background. In Section 3 we consider the case that a prior over the true distributions exists, while in Section 4 we drop this assumption and assume that no prior is given.

2 Model and Preliminaries

Suppose that a principal, Alice, does not know the distribution of a future random outcome. Let Ω be the set of the possible outcomes, and assume Ω is finite. Let $\Delta(\Omega)$ denote the set of all possible distributions over Ω. Let $\mathcal{P} \subseteq \Delta(\Omega)$ be a set of distributions over Ω, where we allow $\mathcal{P}$ to be a strict subset of $\Delta(\Omega)$. For a distribution $p \in \mathcal{P}$ let $p(\omega)$ be the probability that ω is obtained. A natural assumption on $\mathcal{P}$ would be that it is closed and thus it is compact.

Assume that the true distribution is $p^* \in \mathcal{P}$. A self-proclaimed expert, named Bob, claims to know almost perfectly the distribution p^*. Bob may be either informed or uninformed:

- Fix $\epsilon \geq 0$. Bob is *ϵ-informed* if he has a signal q which is a distribution of Euclidian distance at most ϵ from p^*. Equivalently, with the Euclidian metric it holds that $q \in \mathcal{P} \cap B_\epsilon(p^*)$, where $B_\epsilon(p^*)$ is the closed ball of radius ϵ around p^*. We say that ϵ is the *accuracy* of Bob's signal.

- We say that Bob is *perfectly informed* if he is ϵ-informed for $\epsilon = 0$. In other words Bob knows p^* exactly (i.e., $q = p^*$).

- Bob is *uninformed* if he gets no signal.

Both Alice and Bob know that the distribution is an element of $\mathcal{P}$. Bob knows if he is informed or not (and to what accuracy), but Alice does not know if Bob is informed or not.

We will consider two settings. In the setting *with a prior* both Alice and Bob have a common prior P over $\mathcal{P} \subseteq \Delta(\Omega)$. Given the prior P over distributions, we let $\bar{p}$ denote the distribution that is the expected distribution according to the prior P, that is $\bar{p} = \int_p p \mathrm{d}\mathsf{P}$ (which is always well defined). In the setting *without a prior* Alice and Bob have no additional information on the selection of p from $\mathcal{P}$.

Contracts and Payoffs In the setting of screening, Alice wishes to offer Bob a contract such that, if informed, given his signal Bob will be willing to accept the contract and report his signal, while, if uninformed, Bob will prefer to reject to contract.

Formally, a contract is represented as a function $\Pi : \mathcal{P} \times \Omega \mapsto \mathbb{R}$ that specifies the payment $\Pi(q, \omega)$ that Bob is awarded when he forecasts q while the outcome ω is observed. Note that the contracts we consider are binding contracts, which can give a negative payment to Bob in some circumstances (in that case, Bob sends money to Alice).

Bob's action is taken from the set $\{\perp\} \cup \mathcal{P}$. Choosing action $\perp$ means rejecting the contract, while choosing $q \in \mathcal{P}$ means accepting the contract and reporting distribution $q \in \mathcal{P}$. By convention, let $\Pi(\perp, \cdot) = 0$. If Bob is uninformed, a (possibly mixed) strategy for Bob is a randomized action $\xi \in \Delta(\{\perp\} \cup \mathcal{P})$. If informed, a (mixed) strategy for Bob is a mapping $\xi : \mathcal{P} \mapsto \Delta(\{\perp\} \cup \mathcal{P})$ from possible signals to randomized actions.

We denote Bob's expected payment from the contract Π when reporting q under the true distribution p^* by $\Pi(q, p^*) = \sum_\omega p^*(\omega)\Pi(q, \omega)$. Bob seeks to maximize his *perceived* payment as defined below. Essentially, given his information (including his signal q when he is informed) he will take expectation when possible and assume worst case when he has no prior (as in [16]):

- **(*Informed experts.*)** If Bob is ϵ-informed with signal q, his perceived payment when choosing a randomized action $\xi(q) \in \Delta(\{\perp\} \cup \mathcal{P})$, for both the cases with or without prior, is defined to be
$$\inf_{p^* \in \mathcal{P} \cap B_\epsilon(q)} \int_{z \in \mathcal{P}} \Pi(z, p^*) \mathrm{d}\xi(q) \ . \tag{1}$$
That is, for each true distribution he is interested in the expected reward for that distribution, but he assumes that the true distribution is the worst distribution that is consistent with his signal.

- **(*Uninformed experts.*)** In the case that Bob is uninformed, Bob's evaluation of his reward depends on whether we are in the case with or without a prior.

  Uninformed experts with a Prior: When there is a prior P Bob's perceived payment would be equal to his expected reward when choosing a randomized action $\xi \in \Delta(\{\perp\} \cup \mathcal{P})$, which is
$$\int_{z \in \mathcal{P}} \int_{p \in \mathcal{P}} \Pi(z, p) \mathrm{d}\mathsf{P} \mathrm{d}\xi = \int_{z \in \mathcal{P}} \Pi(z, \bar{p}) \mathrm{d}\xi \ .$$
  with $\bar{p} = \int_p p \mathrm{d}\mathsf{P}$, the equality arising from the fact that $\Pi(q, p)$ is affine in p.

  Uninformed experts with no Prior: In the case without prior we assume that Bob evaluates his payment according to the *worst*-case distribution. In this case his perceived payment when choosing a randomized action $\xi \in \Delta(\{\perp\} \cup \mathcal{P})$ is defined to be
$$\inf_{p \in \mathcal{P}} \int_{z \in \mathcal{P}} \Pi(z, p) \mathrm{d}\xi \ .$$

Our definitions follow the standard max-min expected utility modeling (see, e.g., [12]), where the payoff from an event is determined as the minimal expected payoff over a class of possible distributions. We note that in [16] the perceived payment for an uninformed experts without a prior is essentially defined to be $\min_{\omega\in\Omega}\int_z \Pi(z,\omega)\mathrm{d}\xi$. This definition coincides with ours in the case that for any outcome the distribution with probability 1 on that outcome belongs to $\mathcal{P}$ (which is the case in [16]). Our definition only requires that for each possible $p^* \in \mathcal{P}$, the uniformed expert will be satisfied with the contract Π in expectation over the outcomes sampled according to p^* (and not for every ω), this only strengthens our results for the cases where we provide a contract that screens uninformed experts.

Given the above definitions for the perceived payments of Bob, we now define the contracts that Bob accepts (or rejects). An informed expert accepts the contract if for every signal that he may receive he has a randomized action that will guarantee a sufficiently large perceived payment.

Definition 1. *Fix any $\delta \in \mathbb{R}$ and $\epsilon \geq 0$. We say that an ϵ-informed expert Bob* δ-accepts *the contract Π if for any $p^* \in \mathcal{P}$ and any $q \in \mathcal{P} \cap B_\epsilon(p^*)$ there exists a randomized action $\xi(q)$ such that Bob's perceived payment (given the contract Π, signal q, randomized action $\xi(q)$ and ϵ) is greater than δ.*

We say that an uninformed expert Bob (with or without prior) δ-accepts *the contract Π if there exists a strategy ξ for Bob for which his perceived payment (given the contract Π) is greater than δ. Otherwise (for any strategy ξ for Bob, his perceived payment is at most δ), we say that he* δ-rejects *the contract.*

Observe that if an ϵ-informed expert Bob δ-accepts the contract Π then so does an ϵ_1-informed expert, for any non-negative $\epsilon_1 < \epsilon$.

Notation: Throughout the paper when δ is clear from the context we might just say that the expert *accepts* the contract instead of saying that he δ-accepts the contract, and *rejects* the contract instead of saying that he δ-rejects the contract.

Scoring rules We would like the contract we design not only to screen the uninformed experts but also motivate any informed expert to reveal the relevant information he possesses. In the terminology of scoring rules we would like to design a "strictly proper" contract. To define that we first recall the terminology for scoring rules.

A scoring rule $s : \mathcal{P} \times \Omega \mapsto \mathbb{R}$ specifies a payment $s(q,\omega)$ when the forecast is q and outcome ω is observed. If the distribution is p the expected payment is $s(q,p)$. The scoring rule is *proper* if for all p, q, $s(p,p) \geq s(q,p)$, and it is *strictly proper* if for all p, q such that $q \neq p$, $s(p,p) > s(q,p)$, that is, p is the unique maximizer of s.

Let $d(q,p)$ be the Euclidian distance between q and p. We say that a scoring rule s is *distance rewarding* if for any q_1, q_2 and p such that $d(q_1,p) > d(q_2,p)$ it holds that $s(q_1,p) < s(q_1,p)$ (the reward decreases as the distance increases). Note that any distance rewarding scoring rule is strictly proper.

An example of a continuous in q, distance-rewarding, strictly-proper scoring rule is the negation of Brier score. Brier score $Br(q,\omega)$ is defined to be $Br(q,\omega) = \sum_{\omega'\in\Omega}(\mathbf{1}_{\omega'=\omega} - q(\omega'))^2$, where $\mathbf{1}_{\omega'=\omega}$ is an indicator, it is 1 if $\omega' = \omega$, and 0 otherwise. Clearly, $s(q,\omega) = -Br(q,\omega)$ is continuous in q, and it can be shown that s is also distance rewarding (as we prove for completeness in Appendix A).

ϵ-proper contracts Any contract Π can be viewed as a scoring rule $s(q,\omega) = \Pi(q,\omega)$, and thus we define a contract to be *distance rewarding* if when viewed as a scoring rule it is distance rewarding. Similarly, we define a contract to be *proper* and *strictly proper* in exactly the same way. These definitions have the properties that a distance-rewarding contract is also strictly proper and a strictly-proper contract is also proper. More interestingly, for a proper contract, a perfectly-informed expert that reports his signal q maximizes his perceived payment. For a strictly-proper contract, this is the unique strategy that maximizes the perceived payment of the perfectly-informed experts.

It would have been desirable to define strictly-proper contracts such that the unique strategy that maximizes the perceived payment of an *ϵ-informed expert* would also be to reveal his signal. Nevertheless, in the general case such contracts do not exist. For example, if the diameter of $\mathcal{P}$ is at most ϵ then being ϵ-informed is meaningless, as for every signal $q \in \mathcal{P}$ we have that every $p^* \in \mathcal{P}$ could be the true distribution. Therefore, if for an ϵ-informed expert that receives the signal q it is the unique strategy to report q then reporting q is also the unique maximizing strategy for an ϵ-informed expert that receives any other signal q'. For specific contracts, and even distance-rewarding contracts (such as the one based on the Brier scoring rule), it is possible to come up with examples where to maximize his perceived payment, an ϵ-informed expert with signal q may choose to report some q' that its distance from q is larger than ϵ.

Exploring to what extent and under which conditions ϵ-informed experts can be incentivized to report their signal, is an interesting question. We note though that this question is to a large extent orthogonal to the main goal of our paper. The reason is that the question of strictly-proper contracts is interesting even if Alice does not try to screen uninformed experts (namely, even if Alice is assured by some other means that Bob is informed), and all she wants to do is to elicit his signal. Still, Lemma 2 (below) makes some progress in this direction.

As we discussed, even distance-rewarding contracts may not elicit the exact signal of an ϵ-informed expert and may in fact elicit reports that are more than ϵ away from q. Nevertheless, it is not hard to show that for such contracts the report that maximizes the perceived payment of an ϵ-informed expert is contained in $B_{2\epsilon}(q)$, thus cannot be very far from the signal q. Interestingly, a similar property does not follow from the weaker notion of accuracy rewarding contract previously studied in the literature [14]. More formally, we say that a contract Π is ϵ'-proper with respect to ϵ-informed experts if for an ϵ-informed expert Bob, for every signal q and for every randomized action $\xi(q)$ which maximizes Bob's perceived payment for q, it holds that with probability 1, the reports made by $\xi(q)$ are contained in $B_{\epsilon'}(q)$.

Lemma 2. *Any distance-rewarding contract Π is 2ϵ-proper with respect to ϵ-informed experts.*

Proof. For any possible signal q and any possible report $q' \notin B_{2\epsilon}(q)$ and every possible true distribution $p^* \in B_\epsilon(q)$ the Euclidean distance of the signal q to p^* is strictly smaller than the distance of q' to p^*. Therefore $\Pi(q, p^*) < \Pi(q', p^*)$. The lemma easily follows. □

Remark 3. *Note that even if Bob truthfully reveals his signal, Alice learns the true distribution only up to his accuracy ϵ. With a 2ϵ-proper contract, Alice will learn the true distribution up to error 3ϵ. As we think of ϵ as being very small, the two scenarios are arguably not that distinct.*

3 Characterization with a Prior

In this section, we show that the ability to design contracts that screen experts depends on the shape of the class of possible distributions $\mathcal{P}$. If $\mathcal{P}$ does not contain the average distribution (which is the distribution that Alice has with no experts around), then it is possible to tell whether Bob is an imposter or not. This can be interpreted as follows: Alice can screen the experts with contracts whenever informed experts always can disclose information that Alice doesn't already have, that is, when experts are always valuable. We first observe that screening is impossible when $\bar{p} \in \mathcal{P}$:

Observation 4. *Assume that a prior P over $\mathcal{P}$ is given and suppose that the prior P is such that $\bar{p} \in \mathcal{P}$.*

Fix any $\delta_A \in \mathbb{R}$ and $\epsilon \geq 0$. For any contract Π, if an ϵ-informed expert δ_A-accepts Π, then an uninformed expert also δ_A-accepts Π.

Proof. We have observed that if an ϵ-informed expert δ_A-accepts the contract Π then so does a perfectly informed expert. Now consider the case that a perfectly informed expert Bob δ_A-accepts the contract Π. Assume that the true distribution is $\bar{p}$ (as $\bar{p} \in \mathcal{P}$ it is possible). In this case the signal q of Bob is $\bar{p}$, as he is perfectly informed. Since Bob δ_A-accepts Π, there must exist a report $q'(\bar{p})$ such that $\Pi(q'(\bar{p}), \bar{p}) > \delta_A$ when he gets the signal $\bar{p}$.

By always announcing $q'(\bar{p})$, an uninformed expert has a perceived payment of $\Pi(q'(\bar{p}), \bar{p}) > \delta_A$ and so he always δ_A-accepts the contract as well. □

Observe that when the true distribution is $\bar{p} \in \mathcal{P}$ the distribution of outcomes is exactly the same as the distribution of outcomes given the prior P. As the distribution of outcomes is the only thing that Alice cares about, in that case the knowledge of the expert is completely unhelpful. As such an unhelpful expert was supposed to accept the contract, an uninformed expert (which in that case knows exactly the same as the informed expert about the distribution of the outcomes) also accepts the contract.

We next move to prove the converse of the above observation.

Theorem 5. *Assume that a prior P over $\mathcal{P}$ is given and suppose that $\mathcal{P}$ is closed and that $\bar{p} \notin \mathcal{P}$.*

Then there exists a continuous (in q), distance-rewarding contract Π such that for some $\epsilon > 0$ an ϵ-informed expert δ_A-accepts the contract, while an uninformed expert δ_R-rejects the contract.

Furthermore, fix any $\mu > 0$ such that an open ball of radius μ around $\bar{p}$ does not intersect $\mathcal{P}$ (such $\mu > 0$ exists by the closeness of $\mathcal{P}$), for any non-negative $\epsilon \le \mu_0$ an ϵ-informed expert δ_A-accepts Π.

Proof. Without loss of generality we can assume that $\delta_A > \delta_R$, as otherwise we can redefine δ_A to be say $\delta_R + 1$ (observe that if an expert δ-accept the contract then he δ^+-accept the contract for any $\delta^+ > \delta$). Denote by $\mathcal{B}$ the open ball of radius μ around $\bar{p}$. This ball, by the conditions of the theorem, does not intersect $\mathcal{P}$: formally, $\mathcal{B} \cap \mathcal{P} = \emptyset$. Let $s(q, \omega)$ be any continuous (in q), distance-rewarding (thus strictly-proper) scoring rule on the full probability space.

Consider the set $S = \{(p, q) \mid p \in \Delta(\Omega) \backslash \mathcal{B}, q \in B_{\mu_0}(p) \cap \Delta(\Omega)\}$ (recall that $B_{\mu_0}(p)$ is the closed ball of radius μ_0 around p). It is easy to argue that the set S is closed: consider a sequence of elements $(p_i, q_i) \in S$ which converges to (p_∞, q_∞). As $\Delta(\Omega)$ is closed and so is $\Delta(\Omega) \backslash \mathcal{B}$ we have that $p_\infty \in \Delta(\Omega) \backslash \mathcal{B}$ and $q_\infty \in \Delta(\Omega)$ it remains to show that the Euclidian distance $d(p_\infty, q_\infty)$ between p_∞ and q_∞ is bounded by μ_0 which would imply that $(p_\infty, q_\infty) \in S$. But this easily follows from the fact that $d(p_i, q_i) \le \mu_0$ for all i. Define

$$\alpha = \inf_{(p,q) \in S} s(q, p) - s(\bar{p}, p) \ .$$

Recall that s is distance rewarding, and that the distance between $\bar{p}$ and p is at least μ which is larger than μ_0. Therefore, for every $q \in B_{\mu_0}(p) \cap \Delta(\Omega)$ we have that $s(q, p) - s(\bar{p}, p) > 0$. As S is closed and bounded, it is compact. Since the function $s(q, p) - s(\bar{p}, p)$ is continuous in q and its domain is compact the infimum is obtained and it holds that $\alpha > 0$.

Define the contract

$$\Pi(q, \omega) = 2\frac{\delta_A - \delta_R}{\alpha}(s(q, \omega) - s(\bar{p}, \omega)) + \delta_R \ .$$

As the scoring rule s is continuous in q and is distance rewarding, the contract Π is also continuous in q and is distance rewarding.

For all $p \in \mathcal{P}$, we in particular have that $p \in \Delta(\Omega) \backslash \mathcal{B}$. Therefore, as $\epsilon \le \mu_0$, for every $q \in B_\epsilon(p) \cap \Delta(\Omega)$ we have by the definition of α that $s(q, p) - s(\bar{p}, p) \ge \alpha$. Consider now an ϵ-informed expert which receives a signal $q \in \mathcal{P}$ and reports this signal truthfully. This may not be the strategy that maximizes the expert's perceived reward but it certainly gives us a lower bound on the perceived reward. Given the expert's signal q, for every possible value of the true distribution $p^* \in B_\epsilon(p) \cap \mathcal{P}$, we have that $\Pi(q,p^*) \geq \frac{2\delta_A - 2\delta_R}{\alpha}\alpha + \delta_R = \delta_A + (\delta_A - \delta_R) > \delta_A$. Therefore, the expert's perceived reward is at least δ_A and the expert will thus δ_A-accept the contract.

On the other hand, as s is strictly proper, for all $q \in \mathcal{P}$,

$$\Pi(q,\bar{p}) \leq 2\frac{\delta_A - \delta_R}{\alpha}(s(\bar{p},\bar{p}) - s(\bar{p},\bar{p})) + \delta_R = \delta_R \ ,$$

so for any strategy of the uninformed expert, his expected fee is at most δ_R and he always rejects the contract. $\square$

As the contract Π in Theorem 5 is distance-rewarding, then by Lemma 2 it is 2ϵ-proper with respect to ϵ-informed experts. Thus the contract Π not only screen experts, but also motivates any informed expert to report a distribution that is close to the true distribution p^*. In particular, it motivates a perfectly-informed expert to truthfully report his signal.

Remark 6. *The only way that the proof uses the condition that $\mathcal{P}$ is closed is in showing that for some $\mu > 0$ there exists a ball of radius μ around $\bar{p}$ that does not intersect $\mathcal{P}$.*

An alternative interpretation of Theorem 5 We argue that the contract in the proof of Theorem 5 is meaningful even when $\bar{p} \in \mathcal{P}$. Indeed, the assumption that $\bar{p} \notin \mathcal{P}$ was not used for the following claims: (1) The contract is distance rewarding, and (2) The uninformed expert δ_R-rejects the contract. The assumption was used to prove that the ϵ-informed expert δ_A accepts the contract. However, what the proof shows is that as long as p^* is far enough from $\bar{p}$ and ϵ is small enough then the informed expert will accept. To formalize this statement we need the following definition:

Definition 7. *Fix any $\delta \in \mathbb{R}$ and $\epsilon \geq 0$ and $q \in \mathcal{P}$. We say that an ϵ-informed expert Bob (q,δ)-accepts the contract Π if there exist a randomized action $\xi(q)$ such that Bob's perceived payment (given the contract Π, signal q, randomized action $\xi(q)$ and ϵ) is greater than δ.*

We can now state the theorem:

Theorem 8. *Let $\mathcal{P}$ be any subset of $\Delta(\Omega)$ and assume that a prior P over $\mathcal{P}$ is given.*

Fix any two reals δ_A and δ_R. Fix any μ_0 and μ such that $\mu_0 < \mu$. Then there exists a continuous (in q), distance-rewarding contract Π such that for any non-negative $\epsilon \leq \mu_0$ and for every q of distance at least $\mu + \epsilon$ from $\bar{p}$, we have that an ϵ-informed expert (q,δ_A)-accepts the contract, while an uninformed expert δ_R-rejects the contract.

Remark 9. *Note that for an ϵ-informed expert to accept, q has to be of distance at least $\mu + \epsilon$ from $\bar{p}$. This is assured as long as p^* is of distance greater than $\mu + 2\epsilon$ from $\bar{p}$. Note that $\mu + 2\epsilon$ can be set to be as close to 3ϵ as we want (by setting $\epsilon = \mu_0$ and μ_0 almost as large as μ). As we think of ϵ as small, if p^* is closer to $\bar{p}$ than 3ϵ then even a perfectly informed expert is not that useful.*

The proof of Theorem 8 is almost identical to that of Theorem 5, and we omit it from this version. As discussed in the introduction, we interpret Theorem 8 as a way to screen "valuable" experts from uninformed experts. A weakness of this interpretation is that when an expert rejects a contract it is not clear if it is because it is uninformed or because it is not valuable. Another possible weakness is that we do not have a tight definition of usefulness such that informed experts accept if and only if they are useful (instead, we know that if an expert is useful he will accept, but otherwise informed experts may either accept or reject).[6] Even with its weaknesses, we argue that Theorem 8 may still be both meaningful and useful.

[6]We note that the theorem can be strengthened such that informed-experts would reject the contract as long as q is too close to $\bar{p}$.

3.1 When screening is possible for every prior

In this section we consider conditions on the the set $\mathcal{P}$ that will ensure that for *every prior* screening will be possible, or will be impossible.

Corollary 10. *Assume that $\mathcal{P}$ is closed and convex. For every prior P and every contract Π, the following holds: For any $\delta_A \in \mathbb{R}$ and $\epsilon \geq 0$, if an ϵ-informed expert δ_A-accepts Π, then an uninformed expert also δ_A-accepts Π.*

Proof. The corollary is a direct result of Lemma 11 and Observation 4. □

Lemma 11. ***[cf., [5]]*** *If $\mathcal{P}$ is closed and convex, then for every prior P over $\mathcal{P}$ it holds that $\bar{p} = \int_p p\mathrm{d}\mathsf{P}$ is in $\mathcal{P}$.*

Corollary 12. *Assume that $\mathcal{P}$ is closed and does not contain any of its non-trivial convex combinations.*

Fix any non-trivial prior P. Fix any $\mu > 0$ such that an open ball of radius μ around $\bar{p}$ does not intersect $\mathcal{P}$ (such $\mu > 0$ exists by Lemma 13). For any two reals δ_A and δ_R and any $\mu_0 < \mu$, there exists a distance-rewarding contract Π such that for any non-negative $\epsilon \leq \mu_0$ an ϵ-informed expert δ_A-accepts the contract, while an uninformed expert δ_R-rejects the contract.

Proof. The corollary is a direct result of Lemma 13 and Theorem 5. □

Lemma 13. ***[cf., [5]]*** *If $\mathcal{P}$ is closed and does not contain any of its non-trivial convex combinations, then, for every non-trivial prior P over $\mathcal{P}$ it holds that $\bar{p} = \int_p p\mathrm{d}\mathsf{P}$ is not in $\mathcal{P}$.*

4 Characterization without Priors

The paper of Olszewski and Sandroni [16] studies the case when there is no prior and shows that when the event is binary and $\mathcal{P} = [0, 1]$ it is impossible to screen the uninformed experts.

Based on a min-max theorem by Fan [7], Olszewski and Sandroni [16] prove the following theorem (adopted to our terminology and the non-binary case with non-specific set $\mathcal{P}$, and emphasizing the convexity of $\mathcal{P}$). We use this result in the proof of theorem 16 below.

Theorem 14 (Essentially due to [16])**.** *Suppose that $\mathcal{P}$ is closed and convex, and no prior is given over $\mathcal{P}$.*

Fix any $\delta_A > \delta \geq 0$ and $\epsilon \geq 0$. For every contract Π, if an ϵ-informed expert δ_A-accepts Π, then an uninformed expert δ-accepts Π.

From our previous results we can prove the converse of the above theorem.

Theorem 15. *Suppose that $\mathcal{P}$ is closed but not convex, and no prior is given over $\mathcal{P}$.*

Fix any two reals δ_A and δ_R. Then, there exists $\mu_0 > 0$ and a continuous (in q), distance-rewarding contract Π such that for any non-negative $\epsilon \leq \mu_0$ an ϵ-informed expert δ_A-accepts the contract, while an uninformed expert δ_R-rejects the contract.

Proof. Consider a point p in the convex hull of $\mathcal{P}$ that does not belong to $\mathcal{P}$. If $\mathcal{P}$ is of Euclidian dimension d there exists a distribution P over $d+1$ points in $\mathcal{P}$ with expectation p. We think of P as a prior over $\mathcal{P}$ and prove the claim using Theorem 5. Let r be the distance of P from $\mathcal{P}$, $r > 0$ as $\mathcal{P}$ is compact and $\mathsf{P} \notin \mathcal{P}$. Fix any μ, μ_0 such that $0 < \mu_0 < \mu < r$. The open ball of radius μ around P does not intersect $\mathcal{P}$. Theorem 5 states that there exists a continuous (in q), distance-rewarding contract Π such that for any non-negative $\epsilon \leq \mu_0$ an ϵ-informed expert accepts the contract, while an uninformed expert rejects the contract.

Note that in Theorem 5 the notions of acceptance and rejection of the contract are with a prior, while we need to prove the claim when there is no prior. The sense in which the informed expert accepts the contract is the same in the case with or without a prior, thus that contract will be accepted by an informed expert also in the case without a prior. The sense in which an uninformed expert rejects the contract is different in the case with no prior than in the case with prior, yet it is easy to see that if an uninformed expert rejects the contract in the sense with prior then he also rejects it in the sense without prior (as the expectation is always at least the infimum). □

When some informed experts are not useful How should we interpret these results? Recall that in the case with prior, the reason that screening was not possible was that we required an informed expert with no new information about the distribution of outcomes to accept the contract. Clearly in the case where there is no prior we can not expect an informed expert to ever be unhelpful in the sense that he brings no information about the distribution of outcomes. Yet we would like to provide a sense in which when $\mathcal{P}$ is closed and convex an informed expert might not be helpful, this will provide a reason why screening is not possible. To do that we next focus on the actions that Alice takes given the information from the expert, and show that when $\mathcal{P}$ is closed and convex for at least one true distribution the information that an informed expert brings is not valuable to Alice as it does not improve Alice's utility.

To formally show this, we will need to define notations for Alice's utility function and action space. Let Γ be a finite set of actions Alice can take and that $U(\gamma, \omega)$ is the utility of Alice taking the action $\gamma \in \Gamma$ when the outcome is realized to $\omega \in \Omega$. For an action $\gamma \in \Gamma$ we use $U(\gamma, p^*)$ to denote the expectation of $U(\gamma, \omega)$ when ω is sampled from the true distribution p^*, that is $U(\gamma, p^*) = \int_w U(\gamma, \omega) \mathrm{d}p^*$. Alice's actions has no (direct) effect on Bob's utility. For a distribution p^* and a randomized action $\gamma' \in \Delta(\Gamma)$, we use $U(\gamma', p^*)$ to denote the expectation of $U(\gamma, p^*)$ when the action is sampled according to γ', that is $U(\gamma', p^*) = \int_\gamma U(\gamma, p^*) \mathrm{d}\gamma'$.

Alice first interacts with the expert Bob, then she picks an action and then her utility is realized. An *action function* $\mathcal{A}$ for Alice is a mapping from the expert's input to a distribution over Γ, her set of actions. Formally the action function is $\mathcal{A} : \{\mathcal{P} \cup \{\perp\}\} \mapsto \Delta(\Gamma)$, where the input from the expert is either $\perp$ when Bob does not accept the contract, or a distribution in $\mathcal{P}$ if he does, and the output is a distribution over actions, that is, an element of $\Delta(\Gamma)$. The expected utility of Alice when the distribution over outcomes is p^* and she picks a random action $\gamma' \in \Delta(\Gamma)$ is denoted by $U(\gamma', p^*)$.

Assuming that Alice is rational and seeks to maximize her utility, she will pick a *utility-maximizing action function* as defined next. An action function is called a *utility-maximizing action function* if it satisfies the following conditions. (1) for any $p^* \in \mathcal{P}$ the randomized action $\mathcal{A}(p^*)$ satisfies $U(\mathcal{A}(p^*), p^*) \geq U(\gamma', p^*)$ for any randomized action $\gamma' \in \Delta(\Gamma)$. (2) There does not exist a randomized action γ^* such that $U(\gamma^*, p^*) > U(\mathcal{A}(\perp), p^*)$ for all $p^* \in \mathcal{P}$.

The first condition says that if Alice knows that the distribution is p^* she picks a randomized action that maximizes her utility with respect to that distribution. The second says that if she gets no information she never picks a dominated randomized action.

The following theorem shows that even without a prior, if $\mathcal{P}$ is closed and convex then some experts may not be valuable to Alice.

Theorem 16. *Suppose that $\mathcal{P}$ is closed and convex, and no prior is given over $\mathcal{P}$. For any utility function U of Alice, and any utility-maximizing action function $\mathcal{A}$ that Alice uses, there exists a true distribution $p^* \in \mathcal{P}$ such that $U(\mathcal{A}(p^*), p^*) = U(\mathcal{A}(\perp), p^*)$.*

Proof. Fix a utility function U and a utility-maximizing action function $\mathcal{A}$, and assume in contradiction that for every true distribution $p^* \in \mathcal{P}$ it holds that $U(\mathcal{A}(p^*), p^*) \neq U(\mathcal{A}(\perp), p^*)$. Since $\mathcal{A}$ is a utility-maximizing action function, for every true distribution $p^* \in \mathcal{P}$ it holds that $U(\mathcal{A}(p^*), p^*) \geq$

$U(\mathcal{A}(\perp), p^*)$. Therefore, our assumption (for the sake of contradiction) implies that for every true distribution $p^* \in \mathcal{P}$, it holds that $U(\mathcal{A}(p^*), p^*) > U(\mathcal{A}(\perp), p^*)$.

Define the function $t : \mathcal{P} \mapsto \mathbb{R}$ as follows: $t(p^*) = U(\mathcal{A}(p^*), p^*) - U(\mathcal{A}(\perp), p^*)$. For every true distribution $p^* \in \mathcal{P}$ it holds that $t(p^*) > 0$. Lemma 19 shows that t is continuous. As t is a continuous function on a compact domain it holds that its infimum is obtained. Thus there exists $\mu > 0$ such that for every distribution $p^* \in \mathcal{P}$ it holds that $t(p^*) > \mu > 0$. Namely, for every distribution $p^* \in \mathcal{P}$ it holds that $U(\mathcal{A}(p^*), p^*) > U(\mathcal{A}(\perp), p^*) + \mu$.

Now, recall that by Theorem 14 if $\mathcal{P}$ is closed and convex then for any $\delta_A > \delta \geq 0$, for every contract Π, if a perfectly informed expert δ_A-accepts Π, then an uninformed expert δ-accepts Π.

To derive a contradiction we want to show that if for every distribution $p^* \in \mathcal{P}$ it holds that $U(\mathcal{A}(p^*), p^*) > U(\mathcal{A}(\perp), p^*) + \mu$, then there exist a contract Π such that a perfectly informed expert μ-accepts Π while an uninformed expert 0-rejects Π.

Consider the contract $\Pi(q, \omega) = U(\mathcal{A}(q), \omega) - U(\mathcal{A}(\perp), \omega)$. When ω is sampled from the distribution p^*, it holds that when Bob reports q his expected reward is $\Pi(q, p^*) = U(\mathcal{A}(q), p^*) - U(\mathcal{A}(\perp), p^*)$. A perfectly informed expert μ-accepts Π as $U(\mathcal{A}(p^*), p^*) > U(\mathcal{A}(\perp), p^*) + \mu$ for every $p^* \in \mathcal{P}$ implies that $\Pi(p^*, p^*) > \mu$ for every $p^* \in \mathcal{P}$.

To show that an uninformed expert 0-rejects Π we need to show that for some $p^* \in \mathcal{P}$, for any randomized action $\xi \in \Delta(\{\perp\} \cup \mathcal{P})$ it holds that $\Pi(\xi, p^*) = U(\mathcal{A}(\xi), p^*) - U(\mathcal{A}(\perp), p^*) \leq 0$, where $\Pi(\xi, p^*)$ denotes the expected reward when the distribution is p^* and the uninformed expert picks the randomized action ξ. Assume in contradiction that for all $p^* \in \mathcal{P}$ it holds that $\Pi(\xi, p^*) = U(\mathcal{A}(\xi), p^*) - U(\mathcal{A}(\perp), p^*) > 0$. Then $\mathcal{A}$ is not undominated as for $\gamma^* = \mathcal{A}(\xi)$ it holds that $U(\gamma^*, p^*) > U(\mathcal{A}(\perp), p^*)$ for all $p^* \in \mathcal{P}$, a contradiction. □

Remark 17. *The first condition in the definition of a utility-maximizing action function can be dropped at the price of weakening Theorem 16 just a bit. This would leave us with a very moderate assumption on the rationality of Alice (i.e., only Condition (2)). The weaker theorem, under the weaker assumption on $\mathcal{A}$, would imply that for every μ there exists a true distribution $p^*_\mu \in \mathcal{P}$ such that $U(\mathcal{A}(p^*_\mu), p^*_\mu) < U(\mathcal{A}(\perp), p^*_\mu) + \mu$. In other words, the usefulness of informed experts may be arbitrarily small.*

Acknowledgments

We thank Yuval Peres for very helpful conversations.

References

[1] Al-Najjar, N. I., Sandroni, A., Smorodinsky, R., and Weinstein, J. Testing theories with learnable and predictive representations. *Journal of Economic Theory 145*, 6 (2010), 2203 – 2217.

[2] Al-Najjar, N. I., and Weinstein, J. Comparative testing of experts. *Econometrica 76*, 3 (2008), 541–559.

[3] Babaioff, M., Blumrosen, L., Lambert, N. S., and Reingold, O. Testing Expert Forecasts for Parameterized Sets of Distributions. In preparation.

[4] Brier, G. Verification of Forecasts Expressed In Terms of Probability. *Monthly Weather Review 78*, 1 (1950), 1–3.

[5] Choquet, G. *Lectures on Analysis, Vol.I, II, III.* W. A. Benjamin, Inc., New York., 1969.

[6] Dekel, E., and Feinberg, Y. Non-bayesian testing of a stochastic prediction. *Review of Economic Studies 73*, 4 (2006), 893–906.

[7] Fan, K. Minimax theorems. *Proceedings of the National Academy of Sciences 39* (1953), 42–47.

[8] Feinberg, Y., and Stewart, C. Testing Multiple Experts. *Econometrica 76* (2006), 561–582.

[9] Fortnow, L., and Vohra, R. V. The complexity of forecast testing. *Econometrica 77*, 1 (2009), 93–105.

[10] Foster, D., and Vohra, R. Calibrated Learning and Correlated Equilibrium. *Games and Economic Behavior 21*, 1-2 (1997), 40–55.

[11] Fudenberg, D., and Levine, D. An Easier Way to Calibrate. *Games and economic behavior 29*, 1-2 (1999), 131–137.

[12] Gilboa, I., and Schmeidler, D. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*, 2 (1989), 141–153.

[13] Good, I. J. Rational Decisions. *Journal of the Royal Statistical Society. Series B (Methodological) 14*, 1 (1952), 107–114.

[14] Lambert, N. S., Pennock, D. M., and Shoham, Y. Eliciting properties of probability distributions. In *ACM Conference on Electronic Commerce* (2008), pp. 129–138.

[15] Lehrer, E. Any Inspection is Manipulable. *Econometrica 69*, 5 (2001), 1333–1347.

[16] Olszewski, W., and Sandroni, A. Contracts and Uncertainty. *Theoretical Economics 2*, 1 (2007), 1–13.

[17] Olszewski, W., and Sandroni, A. A nonmanipulable test. *Ann. Statist. 37*, 2 (2009), 1013–1039.

[18] Olszewski, W., and Sandroni, A. Strategic manipulation of empirical tests. *Math. Oper. Res 34* (2009), 57–70.

[19] Sandroni, A. The reproducible properties of correct forecasts. *International Journal of Game Theory 32*, 1 (2003), 151–159.

[20] Savage, L. Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association 66*, 336 (1971), 783–801.

[21] Winkler, R., Muñoz, J., Cervera, J., Bernardo, J., Blattenberger, G., Kadane, J., Lindley, D., Murphy, A., Oliver, R., and Ríos-Insua, D. Scoring Rules and the Evaluation of Probabilities. *TEST 5*, 1 (1996), 1–60.

A Additional Lemmas and Proofs

Fact 18. *Define the scoring rule s by $s(q,\omega) = -Br(q,\omega)$, then s is distance rewarding.*

Proof. First we note that by definition, $Br(q,\omega) = 1 - 2q(\omega) + \sum_{\omega'\in\Omega}(q(\omega'))^2$. Now,

$$
\begin{aligned}
& Br(q,p) \\
= \; & \sum_{\omega\in\Omega} p(\omega)\left(1 - 2q(\omega) + \sum_{\omega'\in\Omega}(q(\omega'))^2\right) \\
= \; & 1 - \sum_{\omega\in\Omega} 2p(\omega)q(\omega) + \sum_{\omega'\in\Omega}(q(\omega'))^2 \\
= \; & 1 + \|p-q\|_2^2 - \|p\|_2^2,
\end{aligned}
$$

where $\|\cdot\|_2^2$ denotes the ℓ_2 norm squared. The lemma easily follows. □

Lemma 19. *Fix any $\mathcal{A}(\perp) \in \Delta(\Gamma)$. If the function $t : \mathcal{P} \mapsto \mathbb{R}$ is defined to be $t(p^*) = U(\mathcal{A}(p^*), p^*) - U(\mathcal{A}(\perp), p^*)$, then t is continuous.*

Proof. For any fixed $\mathcal{A}(\perp)$ the function $U(\mathcal{A}(\perp), p^*)$ is continuous in p^* as it is linear in the probabilities of the outcomes.

Define $r : \mathcal{P} \times \mathcal{P} \mapsto \mathbb{R}$ as follows: $r(q,p) = U(\mathcal{A}(q), p)$. We thus need to show that $r(p,p) = U(\mathcal{A}(p), p)$ is continuous in p. For any fixed q the function $r(q,p)$ is linear in p and thus continuous in p. As $\mathcal{A}$ is a utility-maximizing action function, $r(p,p) \geq r(q,p)$ and $r(q,q) \geq r(q,p)$ hold that for every p, q.

Fix $\epsilon > 0$ and consider p and p' of distance at most ϵ. As $r(q,p)$ is continuous in p it holds that there exists $\delta > 0$ such that $r(p',p') \geq r(p,p') \geq r(p,p) - \delta$. Symmetrically, $r(p,p) \geq r(p',p) \geq r(p',p') - \delta$. Thus $|r(p,p) - r(p'p')| < 2\delta$ and the claim follows. □