Eliciting Predictions and Recommendations for Decision Making

YILING CHEN, Harvard University
IAN A. KASH, Microsoft Research
MICHAEL RUBERRY and VICTOR SHNAYDER, Harvard University

When making a decision, a decision maker selects one of several possible actions and hopes to achieve a desirable outcome. To make a better decision, the decision maker often asks experts for advice. In this article, we consider two methods of acquiring advice for decision making. We begin with a method where one or more experts predict the effect of each action and the decision maker then selects an action based on the predictions. We characterize *strictly proper* decision making, where experts have an incentive to accurately reveal their beliefs about the outcome of each action. However, strictly proper decision making requires the decision maker use a completely mixed strategy to choose an action. To address this limitation, we consider a second method where the decision maker asks a single expert to recommend an action. We show that it is possible to elicit the decision maker's most preferred action for a broad class of preferences of the decision maker, including when the decision maker is an expected value maximizer.

Categories and Subject Descriptors: F.0 [**Theory of Computation**]: General; J.4 [**Social and Behavioral Sciences**]

General Terms: Economics, Theory

Additional Key Words and Phrases: Market design, decision markets, prediction markets, decision making, information elicitation, scoring rules

ACM Reference Format:
Chen, Y., Kash, I. A., Ruberry, M., and Shnayder, V. 2014. Eliciting predictions and recommendations for decision making. *ACM Trans. Econ. Comp.* 2, 2, Article 6 (June 2014), 27 pages.
DOI:http://dx.doi.org/10.1145/2556271

1. INTRODUCTION

Consider a company attempting to decide whether to invest in solar or wind energy. To improve its chances of making the right decision it would like to acquire some expert advice. The company needs some method of incentivizing experts to be accurate. That is, it needs some means of paying experts so that they honestly reveal their private beliefs or information.

In this article we characterize two methods of accurately soliciting expert advice for decision making. The first elicits predictions from one or more experts about the likely effects of each available action. If these predictions are accurate then the company can use them to make an informed decision. The second simply asks a single expert to recommend an action. If the expert's incentives are aligned with the decision maker's

Parts of this research initially appeared in Chen and Kash [2011] and Chen et al. [2011].
This work is supported by the National Science Foundation under grant CCF-0953516. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors alone.
Author's addresses: Y. Chen, M. Ruberry, and V. Shnayder, School of Engineering and Applied Sciences, Harvard University; email: {yiling, mruberry, shnayder}@seas.harvard.edu; I. A. Kash, Microsoft Research, Cambridge, UK; email: iankash@microsoft.com.

DOI:http://dx.doi.org/10.1145/2556271

then this action will profit them both. So with the first method our challenge is incentivizing accuracy, and in the second method it is aligning the expert's most profitable recommendation with the decision maker's most preferred action.

The first part of this article (Sections 2–4) focuses on eliciting predictions for decision making. One popular mechanism for eliciting accurate forecasts of the future is a *prediction market*. In a prediction market, traders or experts produce a series of forecasts about future outcomes of interest. For instance, traders may be asked to predict whether it will rain or not on Friday next week, or which nominated film will win the Oscar for best picture. These predictions are probability distributions over the outcomes and are made publicly, allowing experts to review each other's forecasts and update their own predictions accordingly. Eventually, the market closes and the future becomes the present and is observed.

A fundamental property of well-designed prediction markets is that they pay or score predictions for accuracy. If providing an accurate forecast maximizes an expert's score for that prediction, we say the market is *proper*, and if an accurate forecast uniquely maximizes the score, the market is *strictly proper*. Strictly proper prediction markets are theoretically effective at aggregating expert information and providing an accurate forecast of the future under some general conditions [Chen et al. 2012; Iyer et al. 2010; Ostrovsky 2012].

Since strictly proper prediction markets are so useful for forecasting the future, we would like to provide the same incentives to experts when eliciting predictions to make a decision. Extending these incentives is not straightforward: the predictions necessary for decision making are different than those made in a classical prediction market, because making a decision changes the observed future. Put another way, in a prediction market an expert predicts *the* future, but a decision maker is interested in the *many possible futures* that can result from its choice. This implies the same techniques that make a prediction market strictly proper do not apply for the elicitation of predictions for decision making. The differences between the classical prediction and decision making settings are detailed in Section 2.

In Section 3, we introduce a model of eliciting predictions for decision making, and in Section 4, we use this model to characterize strictly proper decision making, extending the incentives of strictly proper prediction markets to decision making with both a single expert and many experts in a market (a *decision market*). Unfortunately, creating this incentive requires the decision maker use a completely mixed strategy to choose an action, *i.e.* all actions are chosen with nonzero probability. Essentially, the decision maker must implement an unbiased estimator of the future, and this requires stochastically sampling from it. This limitation suggests that eliciting predictions for decision making lets a decision maker understand its choice, but requires the decision maker to not always act on this understanding. If the decision maker is initially open to taking any action, then eliciting predictions may increase the likelihood it makes the best available choice, but does not allow the decision maker to always make that choice.

In Section 5, we discuss an alternative method of acquiring expert advice where the decision maker simply asks a single expert to recommend an action. In this setting we explicitly model the decision maker's preferences, and our goal is for the expert to accurately reveal the decision maker's most preferred action. We show that we can incentivize an expert to accurately reveal this action if and only if the decision maker's preferences admit a convex weak utility representation, and that this method no longer requires the decision maker choose an action stochastically.

Related Work. Decision markets were first proposed by Hanson [1999] without an analysis of their incentives. Othman and Sandholm [2010b] showed these proposed decision markets did not provide the same incentive for accuracy as a strictly proper prediction market, and we elaborate on this insight in Section 2. They also described a special case of expert recommendation that we detail and generalize in Section 5.

Other work related to eliciting predictions for decision making has considered external incentives in addition to the market's intrinsic incentives. Shi et al. [2009] considered a prediction market where experts can affect the future by taking some actions and defined *principal-aligned* scoring rules that incentivized them to only take "helpful" actions. These rules are similar in spirit to the methods we develop in Section 5, but in our setting, experts cannot take actions to affect the future except by influencing the decision maker's action through their predictions or recommendations. More recently, Boutilier [2012] discussed decision making with an expert who has its own preferences over the decision maker's actions. In his model, the expert predicts the distribution of a discrete random variable and the decision maker selects an optimal action based on the expert's prediction; however, the random variable being predicted is independent of the decision maker's action. Because the expert has preferences over the decision maker's actions, it has incentives to mislead the decision maker to take an action that the expert prefers. Boutilier introduced *compensation rules* that redress the expert's loss of utility for letting its less preferred actions occur to make the expert indifferent again. Different from Boutilier [2012], experts in our setting do not have preferences over actions, but our model of decision making is more general and allows the decision maker's action to affect the likelihood of outcomes being predicted.

Some prior work considers settings where experts can incur cost to improve their beliefs and studies how to induce an appropriate degree of learning as well as accurate predictions [Osband 1989]. In this article, we do not consider cost of obtaining additional information and assume that experts are endowed with their beliefs.

2. PREDICTION AND DECISION MARKETS

In this section we formally compare classical prediction elicitation and eliciting predictions for decision making. This comparison illuminates the new incentive challenges that come with making a decision. We begin by describing the classical setting.

There are many methods of eliciting predictions about the future. One popular method uses a *scoring rule* [Brier 1950; Gneiting and Raftery 2007] to evaluate a forecast, and similar rules will be the focus of our work. Formally, we let Ω be a finite, mutually exclusive and exhaustive set of *outcomes* and $\Delta(\Omega)$ the probability simplex over Ω. A *forecast* or *prediction* is a probability distribution over Ω (an element of $\Delta(\Omega)$), and a *scoring rule* is any function[1]

$$S : \Delta(\Omega) \times \Omega \to \mathbb{R} \qquad \text{(scoring rule)}$$

that maps a forecast and observed outcome to the reals. Intuitively, a scoring rule compares the forecast to the observed outcome and assigns a real-valued score. For example, we might be curious if it will be sunny or cloudy tomorrow.[2] In this case $\Omega = \{\text{Sunny}, \text{Cloudy}\}$ and a forecast is a binary probability distribution like $(\frac{1}{3} : \text{Sunny}, \frac{2}{3} : \text{Cloudy})$. If tomorrow is sunny then this forecast's score would be $S((\frac{1}{3} : \text{Sunny}, \frac{2}{3} : \text{Cloudy}), \text{Sunny})$.

Predictions from a single expert and many experts in a market are scored differently. We describe working with a single expert first. A single expert produces one prediction

[1]Some authors allow scoring rules to be defined on convex subsets of the simplex and/or take values in the extended reals. We avoid these complexities to focus on decision making but note it is easy to extend our results to handle them. Our definition excludes the logarithmic scoring rule, which can take value $-\infty$.

[2]We assume these outcomes are mutually exclusive (no sunny cloudy days or cloudy sunny days) and exhaustive (it is either sunny or cloudy).

$p \in \Delta(\Omega)$, after which we observe the outcome $\omega \in \Omega$ and score the expert $S(p, \omega)$. If the expert believes a forecast q is the true forecast, then its expected score for a prediction p is

$$S(p, q) = \sum_{\omega \in \Omega} q(\omega) S(p, \omega), \qquad \text{(expected score)}$$

where $q(\omega)$ is the probability the belief q assigns to outcome ω. Not every scoring rule is useful. A desirable property is that a risk-neutral expert is incentivized to accurately reveal its belief. A scoring rule that provides such incentive is *proper* and satisfies

$$\arg\max_{p \in \Delta(\Omega)} S(p, q) \supseteq \{q\},\ \forall q \in \Delta(\Omega). \qquad \text{(properness)}$$

That is, we treat the arg max function as returning a set of maximizing arguments to the expression, and a scoring rule is proper when the expected score is maximized by accurately reporting the belief q. Even a proper scoring rule may not be useful. Always paying or scoring an expert \$5 is proper, but it provides no real incentive to be accurate. Instead, we are interested in *strictly proper* scoring rules, where

$$\arg\max_{p \in \Delta(\Omega)} S(p, q) = \{q\},\ \forall q \in \Delta(\Omega). \qquad \text{(strict properness)}$$

The expected score of a strictly proper scoring rule is uniquely maximized by accurate reporting. (Strictly) Proper scoring rules have been characterized previously in Gneiting and Raftery [2007], McCarthy [1956], and [Savage 1971] with convex functions. We will use the following results in later sections of this article.

THEOREM 2.1 (GNEITING AND RAFTERY [2007]). *A scoring rule is (strictly) proper if and only if*

$$S(p, \omega) = g(p) - g^\star_p \cdot p + g^\star_p(\omega),$$

where $g : \Delta(\Omega) \to \mathbb{R}$ is a (strictly) convex function and $g^\star_p$ is a subgradient of g at the point p.[3]

COROLLARY 2.2 (GNEITING AND RAFTERY [2007]). *Any proper scoring rule*

$$S(p, \omega) = g(p) - g^\star_p \cdot p + g^\star_p(\omega)$$

satisfies

$$\sum_{\omega \in \Omega} p(\omega) S(p, \omega) = g(p),\ \forall p \in \Delta(\Omega).$$

As mentioned, many experts participating in a market are scored differently from a single expert. A *prediction market* [4] operated using a *market scoring rule* mechanism [Hanson 2003, 2007] opens with an initial forecast p_0 and lets experts make a series of forecasts $p_1, p_2, \ldots$. These forecasts are public, so experts can review prior predictions and update their own accordingly. Eventually the market closes and an outcome $\omega \in \Omega$ is revealed. Instead of being scored for accuracy, however, each forecast

[3]A subgradient of a convex function $g : \mathbb{R}^n \to \mathbb{R}$ at a point $p \in \mathbb{R}^n$ is a vector $g^\star_p$ such that $g(p) - g(q) \leq g^\star_p \cdot (p - q)$ for all $p, q \in \mathbb{R}^n$.

[4]Prediction markets can also be operated using continuous double auctions [Berg and Rietz 2003; Forsythe et al. 1992], automated market makers [Othman and Sandholm 2010a; Othman et al. 2010], and other wagering mechanisms [Mangold et al. 2005; Pennock 2004; Plott et al. 2003]. In this article, we are interested in prediction markets that use scoring rules.

		Outcomes Ω	
		Profit	Loss
Actions $\mathcal{A}$	Solar	$\frac{2}{3}$	$\frac{1}{3}$
	Wind	$\frac{2}{5}$	$\frac{3}{5}$

Fig. 1. An example action-outcome matrix showing an expert's prediction of two possible futures: one resulting from investing in solar energy and the other from investing in wind.

in a market is scored for how much it improves the accuracy of the preceding forecast; the expert who produces forecast p_i is scored or paid $S(p_i, \omega) - S(p_{i-1}, \omega)$ for the forecast.[5] An expert may make multiple forecasts in the market and its total score is the sum of the scores for its forecasts.

This method of scoring is useful since it only rewards experts for improving the accuracy of the prior prediction. Further, we can interpret the last prediction made in the market as a current market or consensus expert belief. After all, if an expert disagrees with the current prediction they have an incentive to change it. If the scoring rule S is proper, then this method of scoring is also proper for multiple experts since

$$\arg\max_{p \in \Delta(\Omega)} S(p,q) - S(p',q) \;=\; \arg\max_{p \in \Delta(\Omega)} S(p,q), \;\forall p,q,p' \in \Delta(\Omega).$$

Intuitively, the score of the previous forecast is fixed and so does not affect the optimization. If the scoring rule S is strictly proper, then this method is strictly proper, too. We describe markets using (strictly) proper scoring rules as (strictly) proper markets.

Note that even in a (strictly) proper market it may be that an expert still expects to profit by misrepresenting its belief. An expert in a proper prediction market maximizes its score for a forecast by being as accurate as possible, and it does not follow that it maximizes its total score by being accurate if it can make more than one prediction in the market. In fact, an expert may find misleading other experts with false predictions to be worthwhile [Chen et al. 2010], since by leading other experts astray the expert can create an opportunity for a large correction. We call an expert *myopic* if when making a prediction it focuses on maximizing the expected score of this prediction and does not consider how this prediction may affect its future profitability in the market. If experts are acting myopically, then we always expect them to accurately report their beliefs in a strictly proper prediction market.

Forecasts for decision making are different from those in the classical prediction setting just detailed. When making a decision we have a set of actions $\mathcal{A}$ and outcomes Ω. We assume both sets are finite, mutually exclusive, and exhaustive. Instead of predicting the unique future, when making a decision, experts are asked to predict the possible futures resulting from a decision maker's choices. This prediction can be represented by an $|\mathcal{A}| \times |\Omega|$ *action-outcome* matrix like the one in Figure 1, with each row representing a probability distribution over possible outcomes if the associated action is taken.

The matrix in Figure 1 contains all the information relevant to making a decision. In this case, if the decision maker believes the prediction is accurate it will prefer investing in solar energy. These forecasts are elicited in a decision market just like in a classical prediction setting, except after elicitation the decision maker selects an action

[5] It is known that market scoring rules can be equivalently implemented as automated market makers [Abernethy et al. 2013; Chen and Pennock 2007; Chen and Vaughan 2010]. We restrict our discussion to the former for technical tractability.

	Prior prediction		**Expert belief**	
	Profit	Loss	Profit	Loss
Solar	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
Wind	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{2}{5}$	$\frac{3}{5}$

Fig. 2. A hypothetical prior prediction (left) and expert belief (right). The expert can improve the market's prediction for what will occur if solar energy is invested in and agrees with the current prediction for wind energy.

based on the final market prediction. Only the outcome of this action is then observed. One intuitive way to think of a decision market is as a collection of predictions markets with one prediction market per action. Instead of observing the outcome of each market, however, we only observe it in one.

We have not discussed how these forecasts are scored. We'd like to design a means of scoring that offers the same incentives for accuracy as strictly proper prediction markets; that is, we want to incentivize experts to accurately reveal their beliefs. Hanson [1999], when introducing the idea of decision markets, suggested that forecasts in a decision market could be treated like forecasts in a set of strictly proper prediction markets, one for each action, and the markets for unchosen actions would simply be voided and unscored. This is a natural proposal, but these markets do not incentivize accuracy, as the following example describes.

Let our decision maker still be deciding whether to invest in solar or wind energy. For simplicity we'll assume the outcome space of interest is simply how likely each is to return a profit, $\Omega = \{\text{Profit}, \text{Loss}\}$. We'll be running a market, and we let the prior prediction and an expert's belief be as in Figure 2. In this example, we further assume that this expert is the last expert in the market and its prediction will be used by the decision maker to select an action.

We can adopt Hanson's proposed scoring scheme using the strictly proper *quadratic scoring rule*

$$S(p, \omega) = 2p(\omega) - \sum_{\omega \in \Omega} p^2(\omega), \qquad \text{(quadratic score)}$$

and assume the decision maker chooses the action most likely to be profitable. Unfortunately, if our expert reports accurately then its expected score is zero: the decision maker will invest in wind energy and the expert did not improve that prediction. Alternatively, the expert can lie and claim wind energy has no chance of becoming profitable. The decision maker will then invest in solar and the expert will expect to score

$$\begin{aligned} &\frac{1}{3}\left(2/3 - (1/3)^2 - (2/3)^2\right) + \frac{2}{3}\left(4/3 - (1/3)^2 - (2/3)^2\right) \\ -&\frac{1}{3}\left(4/3 - (1/3)^2 - (2/3)^2\right) - \frac{2}{3}\left(2/3 - (1/3)^2 - (2/3)^2\right) \\ =&\ 2/9 \ > \ 0. \end{aligned}$$

Thus, misreporting in this decision market is preferred to reporting accurately, and we cannot claim such a market incentivizes accuracy. The intuition of this example was first noted by Othman and Sandholm [2010b] for working with a single expert. When an expert is not the last trader in a market, an additional complication is that the expert's prediction is not the final prediction that is used by the decision maker to select an action, although it may affect future predictions.

Experts' ability to affect which of several possible futures is observed is the salient distinguishing feature of a decision market. Eliciting predictions for decision making has the potential to improve our decisions, but without the right incentives is unlikely to be useful. In the next two sections, we characterize strictly proper decision making that provides the same incentives as strictly proper prediction markets to experts. Section 3 starts by formalizing the decision making and scoring process.

3. ELICITING PREDICTIONS FOR STRICTLY PROPER DECISION MAKING

The key distinction between decision making and the classical prediction setting is that in the latter there is one possible future and in the former experts influence which of multiple possible futures is observed. Adapting the incentives of a strictly proper prediction market to decision making, then, requires accounting both for how the decision maker chooses an action and how accurate an expert's forecast is. In this section, we first formalize our model of eliciting predictions for decision making and selecting an action, then describe how experts are scored and what it means for decision making to be strictly proper. We will see that this definition is different if we are working with a single expert or many experts in a decision market.

3.1. Eliciting Predictions and Decision Making

Until now we have been informal with describing how a decision maker uses the predictions it acquires to make a decision. In this section we formalize this process. We begin by describing what a forecast for decision making is, how these forecasts are acquired in the single expert and market settings, and then conclude with how a decision maker uses these forecasts to select an action.

Let $\mathcal{A}$ be a finite set of possible actions that a decision maker can take and Ω a finite set of mutually exclusive and exhaustive outcomes of interest to the decision maker. In our running energy investment example $\mathcal{A} = \{\text{Solar, Wind}\}$ and $\Omega = \{\text{Profit, Loss}\}$.

Experts are risk-neutral rational agents and have private beliefs representable by $|\mathcal{A}| \times |\Omega|$ *action-outcome* matrices associating actions with distributions over the outcomes. Examples of these matrices appear in Figures 1 and 2. Each row of an expert's *action-outcome* matrix is a probability distribution over outcomes and represents the expert's subjective belief on likely outcomes when the row's action is taken. We denote the set of *action-outcome* matrices as $\mathcal{P}$. Experts are asked to produce forecasts or predictions, which are also *action-outcome* matrices in $\mathcal{P}$, but may not be the same as their beliefs.

We consider eliciting predictions from both a single expert and many experts in a market. When working with a single expert, that expert makes a single prediction $P \in \mathcal{P}$. The decision maker then applies a *decision rule* to this forecast to construct a *decision strategy*—a probability distribution over the available actions.

Definition 3.1 (*Decision Rule*). A decision rule is any function

$$R : \mathcal{P} \to \Delta(\mathcal{A}),$$

mapping predictions in $\mathcal{P}$ to decision strategies in $\Delta(\mathcal{A})$. We let $R(P, a)$ be the probability the decision strategy $R(P)$ assigns to taking action a, and say a decision rule has full support if $R(P, a) > 0$ for all $P \in \mathcal{P}$ and $a \in \mathcal{A}$.

Once the decision maker has its strategy it selects an action according to it and then an outcome $\omega \in \Omega$ is observed. Intuitively, this outcome is the result of the action taken.

Multiple experts in a decision market are treated differently. A decision market opens with an initial prediction P_0 and lets experts make a series of public predictions $P_1, P_2, \ldots$. This is similar to how prediction markets operate, but with matrix forecasts instead of vectors. Eventually the market closes with a final prediction P and the decision maker applies its decision rule to this prediction to construct its decision strategy. We make no assumption regarding market dynamics or how the last expert comes to this final prediction.

We further assume that experts know the decision rule used by the decision maker prior to making their predictions. In Section 4, we will show that this assumption can be relaxed and our results hold as long as experts know that the decision maker will use a decision rule with full support.

3.2. Scoring Predictions

In classical prediction elicitation, forecasts are scored using a scoring rule, a function

$$S : \Delta(\Omega) \times \Omega \to \mathbb{R}. \qquad \text{(scoring rule)}$$

In eliciting predictions for decision making, a decision maker uses the following generalization of a scoring rule instead.

Definition 3.2 (*Decision Scoring Rule*). A *decision scoring rule* is a function

$$S : \Delta(\mathcal{A}) \times \mathcal{A} \times \mathcal{P} \times \Omega \to \mathbb{R} \qquad \text{(decision scoring rule)}$$

mapping a decision strategy, an action taken, a forecast, and an observed outcome to a real number.

A decision scoring rule lets us account for how the decision maker selects its action as well as how accurate the expert's forecast is. In the next section, we'll see that this generalization is essential for strictly proper decision making. Throughout the paper, we assume that experts know the decision scoring rule used prior to making their predictions.

When working with a single expert, the decision maker pays the expert who provides forecast P a score $S(R(P), a, P, \omega)$, when action a, drawn according to the decision strategy $R(P)$, is taken and outcome ω is observed. The expected score of an expert who believes Q and predicts P is

$$\sum_{a \in \mathcal{A}, \omega \in \Omega} R(P, a) Q(a, \omega) S(R(P), a, P, \omega). \qquad \text{(expected score)}$$

Unpacking this, each term is the likelihood an action and an outcome jointly occur, $R(P, a)Q(a, \omega)$, times the value to the expert for that outcome occurring $S(R(P), a, P, \omega)$.

In a decision market, like in a prediction market, experts receive a *net score* that is the difference of their scores and the previous predictions' scores. The net score for prediction P_i is $S(R(P), a, P_i, \omega) - S(R(P), a, P_{i-1}, \omega)$, when the final prediction is P, the decision maker takes action a according to decision strategy $R(P)$, and outcome ω is observed. The expected net score of an expert in a decision market who believes Q and predicts P_i, with final prediction P, is

$$\sum_{a \in \mathcal{A}, \omega \in \Omega} R(P) Q(a, \omega) \big(S(R(P), a, P_i, \omega) - S(R(P), a, P_{i-1}, \omega) \big). \qquad \text{(expected net score)}$$

Note that, unlike in the single expert setting, there is a separation between the prediction P that the decision maker uses to create a decision strategy and an expert's prediction P_i in a decision market.

3.3. Incentives and Strict Properness

In this section, we define strictly proper decision making. Unlike the classical prediction setting, we will use three definitions of strict properness, one for working with a single expert, one for running a market, and one that works for both settings. Also, unlike a prediction market, strict properness is not just a property of the scoring rule or even the decision scoring rule. Instead, an expert's incentives will depend on both the decision rule and the decision scoring rule used. As a result we will describe (R, S) pairs as strictly proper for an expert or for a market, or simply as strictly proper if they work for both. This is a whirlwind of specialized terms, but by going through each one, their necessity should become clear. Also, shortly after introducing pairs that are strictly proper for a market, we will see that we can safely forget about them to focus on the other two. Each version of strict properness, however, brings the same expert incentives as strict properness for the classical prediction setting to the relevant decision making setting.

To begin, we say a decision rule and a decision scoring rule pair is strictly proper for an expert when a single expert uniquely maximizes its expected score for a prediction by revealing its belief. Thus, exactly as with a strictly proper scoring rule, truthful revelation is strictly optimal for a single expert facing such a pair.

Definition 3.3 (*Strictly Proper for an Expert*). A decision rule and decision scoring rule pair (R, S) is *strictly proper for an expert* when

$$\{Q\} = \arg\max_{P \in \mathcal{P}} \sum_{a \in \mathcal{A}, \omega \in \Omega} R(P, a) Q(a, \omega) S(R(P), a, P, \omega), \ \forall Q \in \mathcal{P}.$$

Strict properness *for a market* is defined very differently, and we'll see that these differences are meaningful.

Definition 3.4 (*Strictly Proper for a Market*). A decision rule and decision scoring rule pair (R, S) is *strictly proper for a market* when

$$\begin{aligned} &\sum_{a \in \mathcal{A}, \omega \in \Omega} R(P, a) Q(a, \omega) \big(S(R(P), a, Q, \omega) - S(R(P), a, P_{i-1}, \omega) \big) \\ \geq &\sum_{a \in \mathcal{A}, \omega \in \Omega} R(P', a) Q(a, \omega) \big(S(R(P'), a, P_i, \omega) - S(R(P'), a, P_{i-1}, \omega) \big) \end{aligned}$$

for all $Q, P_{i-1}, P_i, P, P' \in \mathcal{P}$, with the inequality strict if $P_i \neq Q$.

Understanding this definition and how it is different from the prior strict properness for an expert is useful. The expected score of an expert in a market is most notably different because the decision rule may not be applied to the forecast being scored. Instead it is applied to the final forecast made, and for strict properness we require that an expert always expects to strictly maximize its net score by revealing its belief no matter what the decision strategy is.

Intuitively, it is possible that an expert can change the final prediction to affect the resulting decision strategy. Since we make no assumptions regarding the market dynamics and how the final prediction is formed, when an expert changes its prediction from Q to P_i, the final prediction may change from P to P'. What we are ruling out with the preceding definition is that an expert might have an incentive to change the final prediction by predicting against its belief.

While analyzing whether and when a decision market aggregates all private information and produces a consensus prediction with rational participants is arguably the ultimate question of interest, we note that it is not the focus of this article; instead, we aim to understand when a decision market provides incentives for any myopic expert to predict its belief if the expert only cares about its expected payoff of the current prediction, a property that strictly proper prediction markets have but Hanson's decision markets lack. Although strict properness for a market does not allow one to immediately conclude that the final market prediction aggregates all information of rational participants in a decision market, such an incentive is necessary for information aggregation—without it, as shown by the example in Section 2, the last participant of the market may manipulate the market prediction. We hence believe it is fundamental to understand strict properness for a market and such understanding may provide building blocks for characterizing information aggregation in decision markets. In Section 6, we will discuss the implication of strict properness for a market on information aggregation in decision markets with forward-looking rational agents.

Carrying around two definitions of strict properness is cumbersome. We'd like to combine them into one, and we can almost accomplish this with the following.

Definition 3.5 (*Strictly Proper Pair*). A decision rule and decision scoring rule pair (R, S) is strictly proper when a prediction's expected score is independent of the decision strategy

$$\begin{aligned}&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P_i,\omega)\\ =&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P',a)Q(a,\omega)S(R(P'),a,P_i,\omega),\ \forall Q,P_i,P,P'\in\mathcal{P}\end{aligned}\tag{1}$$

and uniquely maximized when an expert predicts its belief

$$\{Q\} = \arg\max_{P_i\in\mathcal{P}} \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P_i,\omega),\ \forall Q,P\in\mathcal{P}.\tag{2}$$

Intuitively, this notion of strict properness makes decision making resemble the classical prediction setting. As in that setting we require that future predictions cannot affect the score of prior predictions, and we demand an expert uniquely maximize its score for a prediction by revealing its belief. Also, this definition nearly combines the previous two, and every strictly proper pair (R, S) is strictly proper for both an expert and a market, as the following proposition formalizes.

PROPOSITION 3.6. *Every strictly proper pair (R, S) is strictly proper for both an expert and a market.*

PROOF. Let (R, S) be a strictly proper pair. For any $P \neq Q$, We have

$$\begin{aligned}&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P,\omega)\\ <&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,Q,\omega)\\ =&\sum_{a\in\mathcal{A},\omega\in\Omega} R(Q,a)Q(a,\omega)S(R(Q),a,Q,\omega).\end{aligned}$$

The inequality follows from Equation (2) and the equality from Equation (1). This implies (R, S) is strictly proper for an expert.

Strict properness *for a market* requires

$$\begin{aligned}&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(S(R(P),a,Q,\omega) - S(R(P),a,P_{i-1},\omega)\big)\\ \geq&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P',a)Q(a,\omega)\big(S(R(P'),a,P_i,\omega) - S(R(P'),a,P_{i-1},\omega),\end{aligned}$$

for all $Q, P, P', P_{i-1}, P_i \in \mathcal{P}$, with the inequality strict if $P_i \neq Q$.

From the definition of strictly proper pairs, we have

$$\begin{aligned}
&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(S(R(P),a,Q,\omega) - S(R(P),a,P_{i-1},\omega)\big)\\
&\qquad - R(P',a)Q(a,\omega)\big(S(R(P'),a,P_i,\omega) - S(R(P'),a,P_{i-1},\omega)\big)\\
&= \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,Q,\omega) - R(P,a)Q(a,\omega)S(R(P),a,P_i,\omega) \;\geq\; 0
\end{aligned}$$

for all $Q,P,P',P_{i-1},P_i \in \mathcal{P}$. The equality follows from Equation (1) and the inequality from Equation (2), and this inequality is strict if $P_i \neq Q$. Thus (R,S) is strictly proper for a market, too. □

In fact, we can go further and say this definition of strict properness defines *most* of strictly proper decision making. For any pair that is strictly proper for a market there is a strictly proper pair that uses the same decision rule and a decision scoring rule that provides experts the same expected net scores as before.

PROPOSITION 3.7. *For every pair $(R,\bar{S})$ that is strictly proper* for a market*, there exists a strictly proper pair (R,S) such that every prediction has the same expected net score*

$$\begin{aligned}
&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(\bar{S}(R(P),a,P_i,\omega) - \bar{S}(R(P),a,P_{i-1},\omega)\big)\\
&= \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(S(R(P),a,P_i,\omega) - S(R(P),a,P_{i-1},\omega)\big),
\end{aligned}$$

for all $P,Q,P_{i-1},P_i \in \mathcal{P}$.

The proof appears in Appendix A.1.

For all practical purposes, then, we no longer need to consider pairs that are strictly proper for a market. A similar proposition cannot be shown for pairs that are strictly proper for an expert. In the next section we show that strictly proper pairs always have decision rules with full support, but some pairs that are strictly proper for an expert do not. These pairs do, however, create decision strategies with full support for almost all predictions. Hence this distinction is unlikely to be important in practice. We thus say strictly proper pairs describe most of strictly proper decision making.

4. STRICTLY PROPER DECISION MAKING

In this section we characterize strictly proper decision making both with many experts in a decision market and with a single expert. We show that any decision rule with full support is part of a strictly proper pair, and that it is easy to construct such pairs using a strictly proper scoring rule. Unfortunately, a fundamental limitation of this approach to decision making is that it requires the decision maker to use a completely mixed strategy to select an action when running a market, and most of the time when working with a single expert. This suggests that eliciting predictions for decision making can improve the likelihood that a decision maker takes a preferred action, but cannot guarantee it does so. In fact, due to the randomness of the decision strategy, the decision maker may end up taking the least preferred action.

4.1. Strictly Proper Decision Markets

In this section we characterize strictly proper decision markets. Following our discussion in the previous section, instead of working with pairs that are strictly proper for a market, we restrict our attention in this subsection to pairs that are simply strictly proper. We start by showing that a decision rule has full support if and only if it is part of a strictly proper pair, and we provide an easy means of constructing such a pair given a strictly proper scoring rule. We conclude this section with the detailed characterization of these pairs.

We begin by showing that a decision rule must have full support to be part of a strictly proper pair.

THEOREM 4.1 (FULL SUPPORT IS NECESSARY FOR A STRICTLY PROPER PAIR). *If a pair (R, S) is strictly proper then R has full support.*

PROOF. Assume, for a contradiction, that R is a decision rule without full support and S is a decision scoring rule such that (R, S) is strictly proper. Let P^* be a prediction such that $R(P^*, a') = 0$ for some action a', which must exist by our assumption that R does not have full support, and let Q and Q' be two action-outcome matrices differing only on action a'. Then we have

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P^*, a)Q(a,\omega)\big(S(R(P^*), a, P_i, \omega) - S(R(P^*), a, P_{i-1}, \omega)\big)$$
$$= \sum_{a\in\mathcal{A},\omega\in\Omega} R(P^*, a)Q'(a,\omega)\big(S(R(P^*), a, P_i, \omega) - S(R(P^*), a, P_{i-1}, \omega)\big)$$

for all $P_i, P_{i-1} \in \mathcal{P}$. This implies that the same prediction maximizes the expected score of an expert who believes Q and an expert who believes Q'. Yet, since this prediction cannot be both Q and Q', the pair (R, S) violates Equation (2) and so must not be strictly proper, contradicting our assumption. □

Simply put, experts have no incentive to be accurate on actions that are never tested, so a decision rule without full support cannot be strictly proper. This intuition is the same as the one mentioned by Othman and Sandholm [2010b], who showed that any deterministic decision rule cannot be part of a pair that is strictly proper for an expert.

On the other hand, we can constructively demonstrate that any decision rule with full support is part of a strictly proper pair. Given a decision rule R with full support and any strictly proper scoring rule $\bar{S}$, we can create a decision scoring rule

$$S(R(P), a, P_i, \omega) = \frac{1}{R(P, a)}\bar{S}(P_i(a), \omega). \tag{3}$$

The pair (R, S) is strictly proper since the expected score for a prediction P_i given belief Q and decision strategy $R(P)$ is then

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P, a)Q(a,\omega)\Big(\frac{1}{R(P, a)}\bar{S}(P_i(a), \omega)\Big) = \sum_{a\in\mathcal{A},\omega\in\Omega} Q(a,\omega)\bar{S}(P_i(a), \omega),$$

the same expected score as if an expert were participating in $|\mathcal{A}|$ independent and strictly proper prediction markets, one for each action. Intuitively, dividing the scoring rule's score by the likelihood that the decision maker takes an action unbiases how the score is sampled. The following complete characterization shows that all strictly proper pairs are constructed using a similar intuition.

Some additional notation is needed before stating the theorem. We use a colon between two matrices to denote their Frobenius inner product, $A : B = \sum_{i,j} A(i,j)B(i,j)$, and let $G^\star_P$ be a subgradient of the convex function $G : \mathcal{P} \to \mathbb{R}$ at P. This subgradient is also a matrix with the same dimensions as matrices in $\mathcal{P}$ and $G^\star_P(a,\omega)$ is the component of the subgradient with respect to $P(a,\omega)$.

THEOREM 4.2 (STRICTLY PROPER PAIR CHARACTERIZATION). *A pair (R, S) is strictly proper if and only if R has full support and there exists a subdifferentiable, strictly convex function G such that*

$$S(R(P), a, P_i, \omega) = G(P_i) - G^\star_{P_i} : P_i + \frac{G^\star_{P_i}(a, \omega)}{R(P, a)}, \tag{4}$$

where $G^\star_{P_i}$ is a subgradient of G at P_i.

PROOF. We begin by showing that given a decision rule R with full support and a strictly convex G, defining a decision scoring rule S as in Equation (4) makes (R, S) a strictly proper pair.

An expert's expected score for predicting P_i with belief Q and decision policy $R(P)$ is

$$\begin{aligned}
&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P_i,\omega)\\
&= \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\left(G(P_i) - G^\star_{P_i} : P_i + \frac{G^\star_{P_i}(a,\omega)}{R(P,a)}\right)\\
&= \left(\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\left(G(P_i) - G^\star_{P_i} : P_i\right)\right) + Q : G^\star_{P_i}\\
&= G(P_i) - G^\star_{P_i} : P_i + Q : G^\star_{P_i} \qquad \left(\text{since } \textstyle\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega) = 1\right)\\
&= G(P_i) + (Q - P_i) : G^\star_{P_i},
\end{aligned}$$

which is independent of the decision strategy, and the expert's expected score for accurately predicting Q is then

$$G(Q) + (Q - Q) : G^\star_Q = G(Q).$$

Applying the subgradient inequality, we have

$$G(Q) > G(P_i) + (Q - P_i) : G^\star_{P_i},\ \forall P_i \neq Q \in \mathcal{P},$$

implying that (R, S) is a strictly proper pair.

Now we show that given a strictly proper pair (R, S) it is necessary that R has full support and there exists a strictly convex G such that S is as defined in Equation (4). Since Theorem 4.1 proved the necessity of R having full support, we only need to prove the latter condition.

As a shorthand, we define an expected score function

$$V(R(P), Q, P_i) = \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P_i,\omega).$$

Recall from Definition 3.5 that

$$V(R(P), Q, P_i) = V(R(P'), Q, P_i), \forall P, P', P_i, Q \in \mathcal{P},$$

which allows us to write the expected score function simply as $V(Q, P_i)$. Our strictly convex function G will be $G(Q) = V(Q, Q)$. Because (R, S) is a strictly proper pair, $V(Q, Q) \geq V(Q, P_i)$. Hence, $G(Q)$ is the pointwise supremum of a set of convex functions of Q and is also convex. We will shortly verify that it is strictly convex.

We next show that

$$G^\star_{P_i}(a, \omega) = R(P', a)S(R(P'), a, P_i, \omega)$$

for any $P' \in \mathcal{P}$ is a subgradient of G at P_i. This is because the subgradient inequality holds:

$$
\begin{aligned}
&G(P_i) + (Q - P_i) : G^{\star}_{P_i} \\
&= V(P_i, P_i) + \sum_{a \in \mathcal{A}, \omega \in \Omega} (Q(a,\omega) - P_i(a,\omega)) R(P', a) S(R(P'), a, P_i, \omega) \\
&= V(P_i, P_i) + V(Q, P_i) - V(P_i, P_i) \\
&= V(Q, P_i) \\
&< V(Q, Q) \\
&= G(Q)
\end{aligned}
$$

for all $P_i \neq Q \in \mathcal{P}$. The strict inequality follows because (R, S) is a strictly proper pair and this strict inequality implies that G is strictly convex [Hendrickson and Buehler 1971].

We are left to prove that $S(R(P), a, P_i, \omega)$ can be written in the form of Equation (4) for some subgradient of G at P_i. While there may be multiple subgradients at P_i, we use $G^{\star}_{P_i}(a,\omega) = R(P,a)S(R(P), a, P_i, \omega)$ for representing $S(R(P), a, P_i, \omega)$. Thus,

$$
\begin{aligned}
& S(R(P), a, P_i, \omega) \\
= \;& V(P_i, P_i) - V(P_i, P_i) + S(R(P), a, P_i, \omega) \\
= \;& V(P_i, P_i) - \sum_{a \in \mathcal{A}, \omega \in \Omega} R(P,a) P_i(a,\omega) S(R(P), a, P_i, \omega) + \frac{R(P,a) S(R(P), a, P_i, \omega)}{R(P,a)} \\
= \;& G(P_i) - G^{\star}_{P_i} : P_i + \frac{G^{\star}_{P_i}(a,\omega)}{R(P,a)}.
\end{aligned}
$$

This shows that from any strictly proper pair, we can construct a strictly convex G such that S can be expressed in the form of Equation (4) for some subgradient of G. $\square$

Theorem 4.2 shows that while a decision maker can take a preferred action with probability arbitrarily close to one, it must commit to a completely mixed decision strategy. In short, the decision maker must implement an unbiased estimator of the future, and this requires stochastically sampling the actions. Note, however, that it is sufficient for experts to believe they will be scored in a strictly proper fashion, and the decision maker does not have to ex ante design its decision rule. Instead, it can simply review the final prediction, construct any decision strategy with full support, and then score the experts using an appropriate decision scoring rule to create a strictly proper decision market. This insight is analogous in spirit to the observation made by Boutilier [2012] on using compensations rules for prediction elicitation when an expert has preferences over actions. Boutilier [2012] noted that the expert does not need to know the decision rule to be strictly incentivized to predict its belief.

A revealing analogy to the decision maker in a strictly proper decision market is to an overwhelmed teaching assistant grading a midterm. The teaching assistant does not have the time to grade every question and instead must pick one from each test. If some questions are more likely to be graded than others then students will spend more time on those and neglect the rest, biasing their scores. Only by (1) possibly grading any question and (2) weighting that question's score by the inverse likelihood that the question is graded will the teaching assistant create an unbiased estimator, where the student's expected grade is the same as if every question were reviewed. This encourages students to pay equal attention to each question and not "game the system."

4.2. Strictly Proper Decision Making with a Single Expert

Working with a single expert is different than running a decision market since the expert knows that the decision maker will apply the decision rule to its prediction. In a decision market, on the other hand, the decision rule is applied to the final prediction. This distinction allows pairs that are strictly proper for an expert to use decision rules that do not have full support, although we can formally demonstrate that it is rare for these rules to create decision strategies that are not completely mixed.

THEOREM 4.3. *For any pair (R, S) that is strictly proper for an expert, define $\mathcal{P}_0 \subset \mathcal{P}$ to be the set of forecasts that R maps to decision strategies that are not completely mixed. The set $\mathcal{P}_0$ is nowhere dense in $\mathcal{P}$ with its natural Euclidean topology.*[6]

Intuitively, this means that for any forecast that the decision rule maps to a not completely mixed decision strategy, there is another arbitrarily close forecast that does map to a completely mixed strategy. We think it is unlikely this ability to avoid some actions will be useful in practice.

We conclude this section with a complete characterization of strictly proper for an expert pairs. The statement and its proof are similar to those of Theorem 4.2.

THEOREM 4.4 (STRICTLY PROPER FOR AN EXPERT CHARACTERIZATION). *A pair (R, S) is strictly proper for an expert if and only if there exists a subdifferentiable, strictly convex function G and a subgradient $G^\star_P$ such that $G^\star_P(a) = \vec{0}$ whenever $R(P, a) = 0$ and*

$$S(R(P), a, P, \omega) = G(P) - G^\star_P : P + \frac{G^\star_P(a, \omega)}{R(P, a)},\ \forall R(P, a) > 0. \tag{5}$$

This concludes our discussion of strictly proper decision making, where a decision maker solicits a complete mapping from actions to outcomes. In the next section we discuss an alternative where, instead of this mapping, a decision strategy or recommendation is directly solicited. This alternative allows the decision maker to deterministically take a preferred action, instead of doing so with high probability.

5. RECOMMENDATIONS FOR DECISION MAKING

The previous section demonstrated that strictly proper decision making (almost always) requires the decision maker to use a completely mixed strategy to select an action. Put another way, even if the decision maker learns some actions are undesirable it must risk taking them. This is certainly not ideal and possibly noncredible for the decision maker. In this section we describe an alternative method of using expert advice to make a decision. Instead of asking experts to predict the likely outcome of each action, we instead simply ask a single expert to recommend an action. Since taking an action or following a decision strategy results in a distribution over outcomes, we hope to incentivize the expert to recommend an action that if followed produces an outcome distribution that is most preferred by the decision maker if the decision maker had the same information as the expert did. This method allows the decision maker to always take its most preferred action.

When deciding to invest in wind or solar energy in our running example, the decision maker can run a strictly proper decision market and ask experts to predict the likely

[6] A set is nowhere dense in a topological space if the interior of its closure, with respect to the topological space, is empty.

outcome of each investment. This can increase the decision maker's chances of making the right investment, but with some positive chance it must take the wrong or less preferred action simply to test the experts' accuracy. A simple and useful alternative is to offer a single expert a percentage of the realized profit, and ask them to suggest an action. This expert is no longer interested in making an accurate prediction; instead its incentives are perfectly aligned with the decision maker's to produce a good or preferred outcome. This alignment of the expert's and decision maker's incentives will let the decision maker deterministically act on the expert's recommendation. Formalizing this model of decision making is the topic of this section. We stress that if the expert recommends a single action, the decision maker can deterministically take it, in contrast to the previous result. This is a great benefit of asking for a recommendation. In fact, we'll see that an expert can always recommend a single action since a decision maker will have one action that it (weakly) prefers to the others.

This approach, like eliciting predictions for decision making, also has its limitations. It only lets us solicit a recommendation from a single expert, and eliciting an accurate recommendation is possible if and only if the decision maker's preferences admit a subdifferentiable convex weak utility representation. Still, we think it is an especially interesting option since it uses ideas from scoring rules without asking for a prediction. Instead—intuitively—a scoring rule is used to rank the actions so the expert is incentivized to choose the one the decision maker prefers most—aligning the expert's and the decision maker's preferences.

5.1. A Model for Recommendations

When working with expert recommendations, we consider a single expert reporting a decision strategy $\sigma \in \Delta(\mathcal{A})$ and a prediction $p \in \Delta(\Omega)$ of what is likely to occur if that strategy is adopted. The decision maker then draws an action according to the strategy, observes the outcome ω, and scores the expert using a scoring rule $S(p, \omega)$. Our goal is not to elicit any decision strategy, however, but a decision strategy that, if followed, results in the most preferred outcome distribution p for the decision maker, according to the expert's belief.

Formalizing this statement requires specifying the decision maker's preferences. We let the decision maker's preferences be a binary relation $\preceq$ on $\Delta(\Omega)$, the probability distributions or lotteries over Ω. The decision maker weakly prefers p_1 to p_0 if and only if $p_0 \preceq p_1$ and strictly prefers p_1 to p_0 if and only if $p_0 \prec p_1$. These ordinal preferences admit a *weak utility representation* [Alcantud and Rodríguez-Palmero 1999; Peris and Subiza 1995] if there exists a function $u : \Delta(\Omega) \to \mathbb{R}$ such that if $p_0 \prec p_1$, then $u(p_0) < u(p_1)$ for all $p_0, p_1 \in \Delta(\Omega)$. Further, we say these preferences admit a *(strictly) convex weak utility representation* if there exists a (strictly) convex function u that is also a weak utility representation. We note that an expected value-maximizing decision maker always has preferences that admit a convex weak utility representation. We continue to assume the expert is a risk-neutral expected value maximizer. While this assumption may not always hold, it is arguably reasonable for settings where the reward that the expert can receive is relatively small.

The decision maker's goal is to elicit the decision strategy that if followed results in its most preferred distribution, where the distributions that the decision maker evaluates are with respect to the expert's belief. If the expert has belief $\boldsymbol{Q}$, then the decision maker wants to find a decision strategy $\sigma^* \in \Delta(\mathcal{A})$—a column vector—such that

$$\boldsymbol{Q}^T \cdot \sigma^* \succeq \boldsymbol{Q}^T \cdot \sigma, \ \forall \sigma \in \Delta(\mathcal{A}), \qquad \text{(preferred strategies)}$$

where Q^T is the transpose of Q and hence $Q^T \cdot \sigma$ is the lottery over outcomes created by selecting decision strategy σ. We let Σ_Q^* denote the set of such preferred strategies σ^* and

$$\Phi_Q = \arg\max_{\sigma \in \Delta(\mathcal{A})} \sup_{p \in \Delta(\Omega)} \sum_{a \in \mathcal{A}, \omega \in \Omega} \sigma(a) Q(a, \omega) S(p, \omega)$$

denote the set of decision strategies that maximize the expert's expected score. We say a scoring rule is *a recommendation rule for preferences* $\preceq$ if it always incentivizes the expert to reveal a strategy in Σ_Q^*.

Definition 5.1 (*Recommendation Rule*). A scoring rule S is *a recommendation rule for preferences* $\preceq$ over $\Delta(\Omega)$ when $\Phi_Q \subseteq \Sigma_Q^*$ for all $Q \in \mathcal{P}$.

Intuitively, a recommendation rule translates the decision maker's preferences into a payoff function (scoring rule) for the expert that incentivizes it to reveal the decision maker's most preferred strategy.

To recap, in our recommendation setting there is a decision maker and a single expert. The decision maker shows the expert a scoring rule, and the expert reports a decision strategy and makes a prediction about the outcome of this strategy. The decision maker acts according to the strategy, observes the outcome, and pays the expert based on its prediction and the observed outcome using the scoring rule. If the scoring rule is a recommendation rule then the expert has an incentive to reveal the strategy the decision maker would most prefer taking if it had the same information as the expert did. We note that the expert does not need to know the decision maker's preferences.

5.2. Characterizing Recommendation Rules

In this section, we describe the preferences for which we can construct a recommendation rule where an expert maximizes its expected score by reporting the decision maker's most preferred decision strategy. That is, we describe the preferences for which we can strictly incentivize the expert to reveal strategies in Σ_Q^*. It turns out this is precisely the set of preferences admitting a subdifferentiable convex weak utility representation.

We first show that if we know a subdifferentiable convex function[7] that is a weak utility representation of the decision maker's preferences, we can use it to construct a recommendation rule.

PROPOSITION 5.2 (RECOMMENDATION RULE CONSTRUCTION). *If a subdifferentiable, convex function $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ is a weak utility representation of the decision maker's preferences $\preceq$, then the scoring rule*

$$S(p, \omega) = G(p) - G_p^\star \cdot p + G_p^\star(\omega)$$

is a recommendation rule for its preferences.

PROOF. Because G is convex, by Theorem 2.1 and Corollary 2.2, the scoring rule

$$S(p, \omega) = G(p) - G_p^\star \cdot p + G_p^\star(\omega)$$

[7] A convex function $G : \mathbb{R}^n \to \mathbb{R}$ is subdifferentiable everywhere in its relative interior. We are requiring, for notational simplicity, it also be subdifferentiable at its relative boundary.

is proper (in the classical sense) with expected score function

$$G(p) = \sum_{\omega \in \Omega} p(\omega) S(p, \omega).$$

The set of decision strategies that maximize the expected score of an expert with belief Q is

$$\Phi_Q = \arg\max_{\sigma \in \Delta(\mathcal{A})} \sup_{p \in \Delta(\Omega)} \sum_{\omega \in \Omega} (Q^T \cdot \sigma)(\omega) S(p, \omega).$$

Since S is proper, given any σ

$$\sum_{\omega \in \Omega} (Q^T \cdot \sigma)(\omega) S(Q^T \cdot \sigma, \omega) \ \geq \ \sum_{\omega \in \Omega} (Q^T \cdot \sigma)(\omega) S(p, \omega)$$

for all Q and p. Thus,

$$\Phi_Q = \arg\max_{\sigma \in \Delta(\mathcal{A})} \sum_{\omega \in \Omega} (Q^T \cdot \sigma)(\omega) S(Q^T \cdot \sigma, \omega) = \arg\max_{\sigma \in \Delta(\mathcal{A})} G(Q^T \cdot \sigma).$$

G is a weak utility representation of $\preceq$ means that $p_0 \prec p_1$ implies $G(p_0) < G(p_1)$. Since

$$Q^T \cdot \sigma^* \ \succ \ Q^T \cdot \sigma, \ \forall \sigma^* \in \Sigma_Q^*, \ \forall \sigma \notin \Sigma_Q^*$$

by the definition of preferred strategies, we know that

$$G(Q^T \cdot \sigma^*) \ > \ G(Q^T \cdot \sigma), \ \forall \sigma^* \in \Sigma_Q^*, \ \forall \sigma \notin \Sigma_Q^*.$$

This means that $\Phi_Q \subseteq \Sigma_Q^*$. Hence, S is a recommendation rule. □

Proposition 5.2 indicates that the decision maker's preferences admitting a subdifferentiable convex utility representation is a sufficient condition for the existence of a recommendation rule for the preferences. In fact, it is also a necessary condition. Theorem 5.3 gives a complete characterization.

THEOREM 5.3 (RECOMMENDATION RULE CHARACTERIZATION). *If the decision maker is considering at least two actions, there exists a recommendation rule S for its preferences $\preceq$ if and only if these preferences admit a subdifferentiable convex weak utility representation.*

PROOF. Proposition 5.2 proves that if the decision maker's preferences admit a subdifferentiable convex weak utility representation there exists a recommendation rule for them. Here we only prove the necessity of this condition.

Assume, for a contradiction, that the preferences $\preceq$ do not admit a subdifferentiable convex weak utility representation but there is a recommendation rule S for them. The expected score of an expert whose recommendation results in lottery q is

$$V(q) = \sup_{p \in \Delta(\Omega)} \sum_{\omega \in \Omega} q(\omega) S(p, \omega),$$

which is a subdifferentiable convex function of the lotteries. Since we assumed that $\preceq$ does not admit a subdifferentiable convex weak utility representation, this implies that there exists q_1 and q_2 such that

$$V(q_1) \ \geq \ V(q_2), \ \text{and}$$
$$q_1 \ \prec \ q_2.$$

Let there be an expert with belief Q such that $Q(a) = q_1$ and $Q(a') = q_2$ for all $a' \neq a$. This expert expects a (weakly) higher score by recommending a less preferred action a. Further, the expert expects to score (weakly) higher by recommending action a than any convex combination of actions because

$$V(q_1) \geq \alpha V(q_1) + (1-\alpha)V(q_2) \geq V(\alpha q_1 + (1-\alpha)q_2),\ \forall \alpha \in [0,1],$$

where the second inequality is due to the convexity of V. Thus, the decision strategy of taking action a with probability 1 is an element in Φ_Q but not in Σ_Q^*. This contradicts our assumption that S is a recommendation rule for preferences $\preceq$. □

It is interesting that a scoring rule is used to rank lotteries in a way that matches the decision maker's preferences over lotteries. This lets us incentivize an expert to reveal the decision maker's most preferred decision strategy. Furthermore, because the decision maker's preferences must admit a convex weak utility representation, it is without loss of generality to restrict the expert to reporting a single action instead of a decision strategy. To see this, let u be the convex function representing the decision maker's preferences, and whenever $p_1 \prec p_2$, we have $u(p_1) < u(p_2)$. By convexity of u, we know that for any $\alpha \in [0,1]$, $u(\alpha p_1 + (1-\alpha)p_2) \leq u(p_2)$, which implies $\alpha p_1 + (1-\alpha)p_2 \preceq p_2$. Thus, any mixed decision strategy (which will create a convex combination of lotteries) is always (weakly) less preferred to the best single action (which leads to the most favorable lottery). The expert can simply recommend a single action for the decision maker to deterministically take.

5.3. Quasi-Strict Properness and Strictly Proper Recommendation Rules

Recommendation rules incentivize an expert to reveal the decision maker's best decision strategy, but not necessarily to accurately reveal their prediction on likely outcomes if that decision strategy is followed. Othman and Sandholm [2010b] define *quasi-strictly proper* scoring rules to be those that incentivize an expert to accurately reveal both for a special case of decision making. In their paper, a decision maker has a finite set of actions and only two outcomes, "good" and "bad." The decision maker solicits an action-outcome matrix from a single expert, then applies a deterministic decision rule to select an action (no mixed decision strategy is allowed). The authors focus on the natural special case of their model where the decision rule selects the action most likely to result in the "good" outcome, and show they can create a *quasi-strictly proper* rule with two nice properties: (1) the action that the expert believes will most likely result in the "good" action is always chosen by the decision maker, (2) the expert accurately reports the likely results of this action. These rules are "quasi-strictly" instead of "strictly" proper since the rest of the action-outcome matrix may not be accurate.

In our setting, we no longer request an entire action-outcome matrix when an expert makes a recommendation, and so we can simply describe our recommendation rules as strictly proper when they incentivize the expert to accurately reveal its belief about the recommended strategy's outcome.

Definition 5.4 (*Strictly Proper Recommendation Rule*). A scoring rule S is a *strictly proper recommendation rule* for preferences $\preceq$ if it is a recommendation rule for $\preceq$, that is, $\Phi_Q \subseteq \Sigma_Q^*$ for all $Q \in \mathcal{P}$, and for all $\sigma^* \in \Phi_Q$,

$$\arg\max_{p \in \Delta(\Omega)} \sum_{a \in \mathcal{A}, \omega \in \Omega} \sigma^*(a) Q(a,\omega) S(p,\omega) = \{Q^T \cdot \sigma^*\}$$

for all $Q \in \mathcal{P}$.

In practice, strictly proper recommendation rules may be interesting as they allow the decision maker to understand and plan for the likely effects of its decision. These rules can be partially characterized immediately as a corollary of our recommendation rule characterization.

COROLLARY 5.5. (STRICTLY PROPER RECOMMENDATION RULE CHARACTERIZATION). *If preferences $\preceq$ admit a subdifferentiable and strictly convex weak utility representation, then there exists a strictly proper recommendation rule for $\preceq$.*

The proof is immediate from that of Proposition 5.2, since a strictly convex function implies that the expert uniquely maximizes its expected score when the prediction p is equal to the resultant lottery $Q^T \cdot \sigma^*$. Note, however, this result is not tight, and we leave open the possibility that other types of preferences may have strictly proper recommendation rules.

6. CONCLUSION

We study the elicitation of predictions and recommendations for decision making. For eliciting predictions for decision making, our characterization shows that strict properness generally requires the decision maker to risk taking an action at random. This implies that a decision maker can improve the likelihood of taking a preferred action, but cannot guarantee it does so. On the other hand, when asking a single expert to recommend a decision strategy, the decision maker can use recommendation rules to incentivize the expert to reveal the decision maker's most preferred decision strategy, allowing the decision maker to take a preferred action deterministically. The existence of these recommendation rules depends on the structure of the decision maker's preferences. Moreover, we show that some of the recommendation rules are strictly proper—they incentivize the expert to accurately reveal not only the most preferred decision strategy but also a prediction on the likely outcomes when the decision strategy is followed. In what follows, we will discuss some implications of our results and possible future directions.

Information Aggregation in Decision Markets. Our definition of strict properness for decision making provides the same dominant-strategy incentive compatibility for myopic experts in decision markets as strictly proper scoring rules offer for prediction markets. We make no assumption about market dynamics and allow an expert's prediction to affect the final market prediction in an arbitrary manner. However, a related interesting question is whether experts' information is properly aggregated in a strictly proper decision market if some reasonable assumptions are made on expert behavior. After all, one of the advantages of using a market, rather than polling multiple experts independently, is to obtain an aggregate prediction.

For strictly proper prediction markets, several researchers have considered this information aggregation question with either myopic or forward-looking rational agents in game-theoretic settings. In their work, a prediction market is modeled as a Bayesian extensive-form game: the market concerns a discrete random variable; risk-neutral experts each receive a private signal related to this random variable; the prior joint distribution of the random variable and experts' signals is common knowledge; and experts make predictions about the distribution of the random variable according to a known, prespecified order of participation. Under this model, in any strictly proper prediction market, information is fully aggregated with myopic experts who predict their posterior distribution of the random variable, given both their private signal and all prior predictions [Chen et al. 2012]. With forward-looking experts, the market prediction eventually incorporates all private information as time (the number of prediction rounds) goes to infinity at any Perfect Bayesian Equilibrium (PBE) [Ostrovsky 2012].[8]

We can consider a similar Bayesian extensive-form game for a decision market. The market concerns a random variable for each action and experts have private signals related to each random variable. In our running example, a decision market concerns two binary random variables, profitability of investment in solar energy and profitability of investment in wind energy. If the experts' signals about a random variable are independent of other random variables (e.g. an expert's signal about profitability of investment in solar energy is independent of profitability of wind energy), a strictly proper decision market provides the same incentive for experts as $|\mathcal{A}|$ independent strictly proper prediction markets. The information aggregation results for prediction markets imply that, in such cases, information is fully aggregated with myopic experts and is fully aggregated in the limit at any PBE with forward-looking experts in strictly proper decision markets. However, when the experts' signals about one random variable are dependent on some other random variable, a strictly proper decision market cannot be viewed as a set of independent prediction markets and the question of when information is fully aggregated in these markets deserves its own treatment and is an interesting future direction.

In addition to understanding information aggregation in strictly proper decision markets for more complex information structures, our work suggests a few other interesting future directions.

Automated Market Makers for Decision Markets. It is well known that prediction markets using market scoring rules can be equivalently implemented as automated market maker mechanisms, where a set of Arrow-Debreu securities[9] are offered, one for each outcome, and the market maker sets prices for these securities and accepts trades at the prices [Chen and Pennock 2007; Chen and Vaughan 2010]. In practice, the automated market maker implementation is sometimes preferred, especially for experts who are familiar with security trading. For strictly proper decision markets, an interesting and practically important direction is to develop equivalent automated market maker implementations. The immediate challenge is to design natural securities for a decision market. As an expert's net score for a prediction given an outcome depends on the decision strategy, which is not known ex ante, securities that have fixed payoff given an outcome may not be the best candidates to consider.

Extensions of Recommendation Elicitation. Our model of eliciting recommendations uses scoring rules to not only evaluate the accuracy of a forecast but also rank alternatives, aligning the expert's and decision maker's incentives. Two fascinating directions for future work are expanding this model to work with multiple experts, and identifying other situations when scoring rules can be used to rank alternatives. How multiple experts in this recommendation setting are to aggregate their beliefs and be individually assigned credit for their contributions, like in a decision market, is an open challenge. Finally, we have left two more immediate questions open—characterizing the

[8]Ostrovsky [2012] considers a setting where experts can predict the expected value, rather than the entire probability distribution, of a random variable. He characterizes a *separability* condition about the random variable and the prior joint distribution such that, when the condition is satisfied, information is fully aggregated in the limit at any PBE with forward-looking experts. Predicting the entire probability distribution of a discrete random variable can be viewed as predicting the expected values of a set of binary random variables, each being an indicator variable for a possible value of the original random variable. The results of Ostrovsky [2012] then imply that the separability condition is always satisfied when predicting the distribution of the random variable.

[9]An Arrow-Debreu security for an outcome pays off $1 if the outcome happens and $0 otherwise.

necessary conditions for the existence of strictly proper recommendation rules and understanding when preferences have a strictly convex weak utility representation.

APPENDIXES

This appendix contains the proof of Proposition 3.7, which appeared in Section 3 and the proofs of Theorems 4.3 and 4.4, which appeared in Section 4.2.

A.1. Proof of Proposition 3.7

PROPOSITION 3.7. *For every pair* $(R,\bar{S})$ *that is strictly proper* for a market*, there exists a strictly proper pair* (R,S) *such that every prediction has the same expected net score*

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(\bar{S}(R(P),a,P_i,\omega)-\bar{S}(R(P),a,P_{i-1},\omega)\big)$$
$$=\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(S(R(P),a,P_i,\omega)-S(R(P),a,P_{i-1},\omega)\big)$$

for all $P,Q,P_{i-1},P_i\in\mathcal{P}$.

PROOF. Consider any pair $(R,\bar{S})$ that is strictly proper for a market. Pick an arbitrary $P^*\in\mathcal{P}$ and define a new decision scoring rule as

$$S(R(P),a,P_i,\omega)=\bar{S}(R(P),a,P_i,\omega)-\bar{S}(R(P),a,P^*,\omega).$$

We first prove that (R,S) is a strictly proper pair. From the definition of strict properness for a market we have

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\left(\bar{S}(R(P),a,Q,\omega)-\bar{S}(R(P),a,P_{i-1},\omega)\right)$$
$$\geq\sum_{a\in\mathcal{A},\omega\in\Omega} R(P',a)Q(a,\omega)\left(\bar{S}(R(P'),a,P_i,\omega)-\bar{S}(R(P'),a,P_{i-1},\omega)\right),\ \forall Q,P,P',P_i,P_{i-1}\in\mathcal{P}.$$

This implies

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\left(\bar{S}(R(P),a,Q,\omega)-\bar{S}(R(P),a,P_{i-1},\omega)\right)$$
$$=\sum_{a\in\mathcal{A},\omega\in\Omega} R(P',a)Q(a,\omega)\left(\bar{S}(R(P'),a,Q,\omega)-\bar{S}(R(P'),a,P_{i-1},\omega)\right),$$

$\forall\ Q,P,P',P_{i-1}\in\mathcal{P}$. Rearranging terms, we obtain

$$\sum_{a\in\mathcal{A},\omega\in\Omega} Q(a,\omega)\left(R(P,a)\bar{S}(R(P),a,Q,\omega)-R(P',a)\bar{S}(R(P'),a,Q,\omega)\right)$$
$$=\sum_{a\in\mathcal{A},\omega\in\Omega} Q(a,\omega)\left(R(P,a)\bar{S}(R(P),a,P_{i-1},\omega)-R(P',a)\bar{S}(R(P'),a,P_{i-1},\omega)\right),\quad(6)$$

$\forall\ Q,P,P',P_{i-1}\in\mathcal{P}$. Equation (6) implies

$$\sum_{a\in\mathcal{A},\omega\in\Omega} Q(a,\omega)\left(R(P,a)\bar{S}(R(P),a,P_i,\omega)-R(P',a)\bar{S}(R(P'),a,P_i,\omega)\right)$$
$$=\sum_{a\in\mathcal{A},\omega\in\Omega} Q(a,\omega)\left(R(P,a)\bar{S}(R(P),a,P^*,\omega)-R(P',a)\bar{S}(R(P'),a,P^*,\omega)\right),$$

$\forall\ Q, P, P', P_i \in \mathcal{P}$. Rearranging terms gives

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\left(\bar{S}(R(P),a,P_i,\omega) - \bar{S}(R(P),a,P^*,\omega)\right)$$
$$= \sum_{a\in\mathcal{A},\omega\in\Omega} R(P',a)Q(a,\omega)\left(\bar{S}(R(P'),a,P_i,\omega) - \bar{S}(R(P'),a,P^*,\omega)\right),\ \forall Q, P, P', P_i \in \mathcal{P}.$$

This means

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P_i,\omega) = \sum_{a\in\mathcal{A},\omega\in\Omega} R(P',a)Q(a,\omega)S(R(P'),a,P_i,\omega),$$

$\forall Q, P, P', P_i \in \mathcal{P}$. Thus, a prediction's expected score under (R, S) is independent of the decision strategy. Furthermore, it is easy to see that the expected score is uniquely maximized when an expert predicts its belief since from the definition of strictly properness for a market,

$$\arg\max_{P_i\in\mathcal{P}} \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(\bar{S}(R(P),a,P_i,\omega) - \bar{S}(R(P),a,P^*,\omega)\big) = \{Q\},\ \forall Q, P \in \mathcal{P}$$

and hence

$$\arg\max_{P_i\in\mathcal{P}} \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P_i,\omega) = \{Q\},\ \forall P, Q \in \mathcal{P}.$$

The pair (R, S) is strictly proper.

We next show that (R, S) and $(R, \bar{S})$ provide the same expected net score for every prediction in a decision market. By the definition of S,

$$S(R(P),a,P_i,\omega) - S(R(P),a,P_{i-1},\omega) = \bar{S}(R(P),a,P_i,\omega) - \bar{S}(R(P),a,P_{i-1},\omega),$$

$\forall P, Q, P_i, P_{i-1} \in \mathcal{P}$. It is immediately evident that

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(\bar{S}(R(P),a,P_i,\omega) - \bar{S}(R(P),a,P_{i-1},\omega)\big)$$
$$= \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)\big(S(R(P),a,P_i,\omega) - S(R(P),a,P_{i-1},\omega)\big)$$

$\forall P, Q, P_i, P_{i-1} \in \mathcal{P}$. □

A.2. Proof of Theorem 4.3

THEOREM 4.3. *For any pair (R, S) that is strictly proper for an expert, define $\mathcal{P}_0 \subset \mathcal{P}$ to be the set of forecasts that R maps to decision strategies that are not completely mixed. The set $\mathcal{P}_0$ is nowhere dense in $\mathcal{P}$ with its natural Euclidean topology.*

PROOF. Let (R, S) be any pair that is strictly proper for an expert. Let $\mathcal{P}_a$ denote the set of forecasts that R maps to distributions assigning zero probability to action a. The set of all forecasts that R maps to distributions without full support is a finite union of such sets. So, proving P_a is nowhere dense in $\mathcal{P}$ for all $a \in \mathcal{A}$ proves the theorem.

Assume, for a contradiction, that $\mathcal{P}_{a'}$, is not nowhere dense in $\mathcal{P}$ for some $a' \in \mathcal{A}$. This assumption implies that there exists distinct forecasts $\bar{P}$ and $\hat{P}$ in the interior of the closure of $\mathcal{P}_{a'}$ such that $\bar{P}$ and $\hat{P}$ differ only on the prediction for a' and are the same everywhere else. The expected score of a prediction P given (R, S) and belief Q can be written as a function

$$V(Q, P) = \sum_{a \in \mathcal{A}, \omega \in \Omega} R(P, a) Q(a, \omega) S(R(P, a), a, P, \omega).$$

Considering $V(\bar{P}, \bar{P})$, we have

$$\begin{aligned}
&V(\bar{P}, \bar{P}) \\
&= \sum_{a \in \mathcal{A}, \omega \in \Omega} R(\bar{P}, a)\bar{P}(a, \omega) S(R(\bar{P}), a, \bar{P}, \omega) \\
&= \sum_{a \neq a', a \in \mathcal{A}, \omega \in \Omega} R(\bar{P}, a)\bar{P}(a, \omega) S(R(\bar{P}), a, \bar{P}, \omega) && (\text{by } R(\bar{P}, a') = 0) \\
&= \sum_{a \neq a', a \in \mathcal{A}, \omega \in \Omega} R(\bar{P}, a)\hat{P}(a, \omega) S(R(\bar{P}), a, \bar{P}, \omega) && (\text{by } \bar{P} \text{ and } \hat{P} \text{ differing only on } a') \\
&= \sum_{a \in \mathcal{A}, \omega \in \Omega} R(\bar{P}, a)\hat{P}(a, \omega) S(R(\bar{P}), a, \bar{P}, \omega) && (\text{by } R(\hat{P}, a') = 0) \\
&= V(\hat{P}, \bar{P}).
\end{aligned}$$

Because (R, S) is strictly proper for an expert, $V(\hat{P}, \bar{P}) < V(\hat{P}, \hat{P})$. Thus, we obtain $V(\hat{P}, \hat{P}) > V(\bar{P}, \bar{P})$.

By a symmetric argument, we know that $V(\hat{P}, \hat{P}) = V(\bar{P}, \hat{P})$. Strict properness for an expert means that $V(\bar{P}, \hat{P}) < V(\bar{P}, \bar{P})$. Hence, $V(\bar{P}, \bar{P}) > V(\hat{P}, \hat{P})$. We have a contradiction regarding the values of $V(\bar{P}, \bar{P})$ and $V(\hat{P}, \hat{P})$.

We conclude that $\mathcal{P}_{a'}$ must be nowhere dense in $\mathcal{P}$. Since a' was chosen arbitrarily, the set of all forecasts that R maps to distributions with full support is also nowhere dense in $\mathcal{P}$. □

A.3. Proof of Theorem 4.4

THEOREM 4.4 (STRICTLY PROPER FOR AN EXPERT CHARACTERIZATION). *A pair (R, S) is strictly proper for an expert if and only if there exists a subdifferentiable, strictly convex function G and a subgradient $G^\star_P$ such that $G^\star_P(a) = \vec{0}$ whenever $R(P, a) = 0$ and*

$$S(R(P), a, P, \omega) = G(P) - G^\star_P : P + \frac{G^\star_P(a, \omega)}{R(P, a)}, \ \forall R(P, a) > 0. \tag{5}$$

PROOF. Given a decision rule R, strictly convex function G, and a subgradient $G^\star_P$ of G, with $G^\star_P(a) = \vec{0}$ whenever $R(P, a) = 0$, we prove that the pair (R, S) with S written as in Equation (5) whenever $R(P, a) > 0$ is strictly proper for an expert.

The expected score of forecast P with belief Q is:

$$\begin{aligned}
&\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P,\omega)\\
&= \sum_{a\in\{a'|R(P,a')>0\},\omega\in\Omega} R(P,a)Q(a,\omega)\left(G(P)-G^\star_P : P + \frac{G^\star_P(a,\omega)}{R(P,a)}\right)\\
&= G(P)-G^\star_P : P + \sum_{a\in\{a'|R(P,a')>0\},\omega} G^\star_P(a,\omega)Q(a,\omega)\\
&= G(P)-P : G^\star_P + Q : G^\star_P \qquad\qquad (\text{recalling } R(P,a)=0 \implies G^\star_P(a)=\vec{0})\\
&= G(P)+(Q-P) : G^\star(P).
\end{aligned}$$

Similarly, the expected score of forecast Q with belief Q is:

$$\begin{aligned}
&\sum_{a\in\mathcal{A},\omega\in\Omega} R(Q,a)Q(a,\omega)S(R(Q),a,Q,\omega)\\
&= \sum_{a\in\{a'|R(Q,a')>0\},\omega\in\Omega} R(Q,a)Q(a,\omega)\left(G(Q)-G^\star_Q : Q + \frac{G^\star_Q(a,\omega)}{R(Q,a)}\right)\\
&= G(Q)-Q : G^\star_Q + Q : G^\star_Q\\
&= G(Q).
\end{aligned}$$

Because G is strictly convex, $G(P)+(Q-P) : G^\star(P) < G(Q)$, $\forall\, Q \neq P$. Hence,

$$\sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P,\omega) < \sum_{a\in\mathcal{A},\omega\in\Omega} R(Q,a)Q(a,\omega)S(R(Q),a,Q,\omega),$$

$\forall\, Q \neq P$. The pair (R,S) is strictly proper for an expert.

Now consider the other direction. Given a (R,S) pair that is strictly proper for an expert, we will construct a strictly convex function G and a subgradient of G satisfying the theorem's criteria. As in Theorem 4.2, we consider the expected score function

$$V(Q,P) = \sum_{a\in\mathcal{A},\omega\in\Omega} R(P,a)Q(a,\omega)S(R(P),a,P,\omega)$$

and let $G(Q) = V(Q,Q) = \max_{P\in\mathcal{P}} V(Q,P)$. G is convex because it is the pointwise supremum of a set of convex functions. Define a subgradient of $G(P)$ as

$$G^*_P(a,\omega) = R(P,a)S(R(P),a,P,\omega).$$

We can verify that it satisfies the subgradient inequality:

$$\begin{aligned}
&G(P)+(Q-P) : G^\star_P\\
&= \sum_{a\in\mathcal{A},\omega\in\Omega} (P(a,\omega)+Q(a,\omega)-P(a,\omega))R(P,a)S(R(P),a,P,\omega)\\
&= \sum_{a\in\mathcal{A},\omega\in\Omega} Q(a,\omega)R(P,a)S(R(P),a,P,\omega)\\
&= V(Q,P)\\
&< V(Q,Q)\\
&= G(Q)
\end{aligned}$$

for any $Q \neq P$, where the inequality holds because (R, S) is strictly proper for an expert. This also implies G is strictly convex.

Note also that $R(P, a) = 0 \implies G^\star_P(a) = \vec{0}$ for all P. We show in the following that S can be written as in Equation (5) using G and $G^\star_P$ for any $a \in \{a' | R(P, a') > 0\}$:

$$\begin{aligned} & G(P) - G^\star_P : P + \frac{G^\star_P(a, \omega)}{R(P, a)} \\ = \; & \frac{G^\star_P(a, \omega)}{R(P, a)} \\ = \; & \frac{R(P, a) S(R(P), a, P, \omega)}{R(P, a)} \\ = \; & S(R(P), a, P, \omega). \end{aligned}$$

□

ACKNOWLEDGMENTS

The authors are grateful to Craig Boutilier for helpful discussions about this work and anonymous reviewers and associate editor for feedbacks on earlier drafts.

REFERENCES

Jacob Abernethy, Yiling Chen, and Jennifer Wortman Vaughan. 2013. Efficient market making via convex optimization, and a connection to online learning. *ACM Trans. Econ. Comp. 1*, 2 (May), 1–39.

José C. R. Alcantud and Carlos Rodríguez-Palmero. 1999. Characterization of the existence of semicontinuous weak utilities. *J. Math. Econ. 32*, 4 (Dec.), 503–509.

Joyce E. Berg and Thomas A. Rietz. 2003. Prediction markets as decision support systems. *Inform. Syst. Frontier 5*, 79–93.

Craig Boutilier. 2012. Eliciting forecasts from self-interested experts: Scoring rules for decision markets. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 737–744.

Glenn W. Brier. 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Rev. 78*, 1, 1–3.

Yiling Chen, Stanko Dimitrov, Rahual Sami, Daniel M. Reeves, David M. Pennock, Robin D. Hanson, Lance Fortnow, and Rica Gonen. 2010. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica 58*, 4, 930–969.

Yiling Chen, Ian Kash, Mike Ruberry, and Victor Shnayder. 2011. Decision markets with good incentives. In *Proceedings of the 7th International Conference on Internet and Network Economics (WINE)*. 72–83.

Yiling Chen and Ian A. Kash. 2011. Information elicitation for decision making. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 175–182.

Yiling Chen and David M. Pennock. 2007. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence (UAI)*. 49–56.

Yiling Chen, Mike Ruberry, and Jenn Wortman Vaughan. 2012. Designing informative securities. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence (UAI)*. 185–195.

Yiling Chen and Jennifer Wortman Vaughan. 2010. A new understanding of prediction markets via no-regret learning. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*. 189–198.

Robert Forsythe, Forrest Nelson, George R. Neumann, and Jack Wright. 1992. Anatomy of an experimental political stock market. *Amer. Econ. Rev. 82*, 5, 1142–1161.

Tilmann Gneiting and Adrian E. Raftery. 2007. Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc. 102*, 477, 359–378.

Robin Hanson. 1999. Decision markets. *IEEE Intell. Syst. 14*, 3, 16–19.

Robin D. Hanson. 2003. Combinatorial information market design. *Inform. Syst. Frontiers 5*, 1, 107–119.

Robin D. Hanson. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. *J. Predict. Markets 1*, 1, 1–15.

Arlo D. Hendrickson and Robert J. Buehler. 1971. Proper scores for probability forecasters. *Ann. Math. Statist. 42*, 6, 1916–1921.

Krishnamur Iyer, Ramesh Johari, and Ciamac C. Moallemi. 2010. Information aggregation in smooth markets. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*. 199–205.

Bernard Mangold, Mike Dooley, Rael Dornfest, Gary W. Flake, Havi Hoffman, Tejaswi Kasturi, and David M. Pennock. 2005. The tech buzz game. *IEEE Comput. 38*, 7, 94–97.

John McCarthy. 1956. Measures of the value of information. *Proc. Nat. Acad. Sciences USA 42*, 9, 654–655.

Kent Osband. 1989. Optimal forecasting incentives. *J. Polit. Econ. 97*, 5, 1091–1112.

Michael Ostrovsky. 2012. Information aggregation in dynamic markets with strategic traders. *Econometrica 80*, 6, 2595–2648.

Abraham Othman and Tuomas Sandholm. 2010a. Automated market making in the large: The Gates Hillman prediction market. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*. 367–376.

Abraham Othman and Tuomas Sandholm. 2010b. Decision rules and decision markets. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 625–632.

Abraham Othman, Tuomas Sandholm, David M. Pennock, and Daniel M. Reeves. 2010. A practical liquidity-sensitive automated market maker. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*. 377–386.

David M. Pennock. 2004. A dynamic pari-mutuel market for hedging, wagering, and information aggregation. In *Proceedings of the 5th ACM Conference on Electronic Commerce (EC)*. 170–179.

Josep E. Peris and Begoa Subiza. 1995. A weak utility function for acyclic preferences. *Econ. Lett. 48*, 1, 21–24.

Charles R. Plott, Jorgen Wit, and Winston C. Yang. 2003. Parimutuel betting markets as information aggregation devices: Experimental results. *Econ. Theory 22*, 2, 311–351.

Leonard J. Savage. 1971. Elicitation of personal probabilities and expectations. *J. Amer. Statist. Assoc. 66*, 336, 783–801.

Peng Shi, Vincent Conitzer, and Mingyu Guo. 2009. Prediction mechanisms that do not incentivize undesirable actions. In *Proceedings of the 5th International Workshop on Internet and Network Economics (WINE)*. 89–100.

Received August 2012; revised April 2013; accepted September 2013