Useful Forecasting: Belief Elicitation for Decision-Making

Lars Wittrock*

September 15, 2023

Click here for most recent version

Abstract

Having information about an uncertain event is crucial for informed decision-making. This paper introduces a simple framework in which 1) a principal uses the reported beliefs of multiple agents to make a decision and 2) the agents reporting their belief are affected by the decision. Naturally, the question arises how the principal can incentivize the agents to report their belief truthfully. I show that in this setting a direct reporting mechanism using a scoring rule to incentivize belief reports and a fixed decision rule lead to truthful reporting by all agents as the unique Nash equilibrium under precisely two conditions, preference diversity and no pivotality. Moreover, if the principal can only consult a single agent the only mechanism that can guarantee truth-telling requires perfect knowledge of the agent's preferences.

Keywords: principal-agent problem, multiple agents, private information, scoring rules, information transmission.

JEL Classification: D71, D82, C72

*Department of Micro and Public Economics, Maastricht University, The Netherlands.

Contact: l.wittrock@maastrichtuniversity.nl.

For helpful comments I thank: Stefan Terstiege, Elias Tsakas, Andrew Mackenzie, Peter Wakker, Caspar Oesterheld, Frauke Stehr, David Albrecht, Ferdinand Pieroth, Martin Strobel and the seminar audience at FUR Ghent.

1. Introduction

Having accurate information about an uncertain event is crucial for informed decision-making. Sometimes a principal in charge of making a decision is poorly informed themselves but is able to consult a group of experts for advice. In this case the principal may be interested to obtain predictions about the event from the experts. However, if the interests of the uninformed principal and the experts are misaligned, the experts may have an incentive to misreport their information in order to influence the decision of the principal in their interest. Such situations can occur, for example, in organizations where a manager is faced with a decision between two projects. Some workers from sales or technical departments may be better informed than the manager about the potential success of the different projects. However, these workers may not report their beliefs about the project's success truthfully as they have additional interests such as job security. Further examples are, the forecasting of economic variables by EU member states resulting in a decision by the ECB, a politician implementing Covid-19 guidelines based on predictions from a group of experts, or forecasting of factors for national security as further explained by one example below.

As the leading example for this paper, consider the decision problem for former US president Obama during the hunt for Bin Laden in 2011. Obama is faced with a decision to attempt capture or wait. If Bin Laden is at his suspected home it is best to attack, otherwise it would be better to wait. Obama himself does not know if Bin Laden is at home, however, his team of experts each have some belief about the probability that Bin Laden is at the suspected location. As documented by Friedman and Zeckhauser (2014), Obama asked all of his advisers to provide him with a precise probability estimate. It is documented that some advisers actively misreported their belief in order to influence Obama's decision. One of the main reasons for misreporting may be that some experts have different preferences over the actions than Obama. Moreover, Obama did not provide explicit monetary incentives for accurate judgements. Naturally, the question arises if the principal can incentivize the experts to report their beliefs truthfully.

I consider a setting in which the state of the world is revealed after the principal chooses an action. In the example at hand this implies that Obama learns if Bin Laden is at the suspected location either immediately or at some later point after he makes a decision. Furthermore, I assume that state-contingent (monetary) transfers between the principal and the agent(s) are possible. Referring again to the leading example, Obama could provide monetary bonus payments for each expert. Bonus payments based on performance are common in many firms and conditioning this bonus payment partly on specific predictions appears to be easily implementable in practice. With the possibility of state-contingent transfers, scoring rule mechanisms are widely regarded as the main tool to elicit private beliefs. Scoring rules are designed to incentivize a person to state their belief truthfully, even if the person's private belief is never known. However, scoring rule mechanisms typically assume that the (monetary) reward is the only factor a person considers when reporting a belief. A frequently omitted aspect of belief elicitation is how the resultant belief is used. It is likely that an agent may not only care about the monetary reward from the scoring rule but also care about how his belief report is used in a subsequent decision problem.

This paper introduces a simple framework in which agents have private information about the state of the world and preferences over the two actions available to a principal. The principal has no knowledge about the state of the world and also does not know the action preferences of the experts. The principal is only interested in choosing her preferred action given the state realisation. She implements a deterministic, threshold-based decision rule and rewards accuracy of agents through a scoring rule mechanism. I study whether this mechanism can lead to truthful reporting by the agent(s).

I show that if the principal only consults a single agent, this agent is likely to misreport his true belief to affect the action being selected by the principal. Misreporting is less likely to occur if the agent's true belief is further away from the decision threshold of the principal, the scoring rule incentives paid by the principal are larger or the agent has weaker preferences preferences over the two actions. The only scoring rule mechanism that makes truth-telling a (strictly) dominant strategy requires perfect knowledge of the expert's preferences over the two actions.

In a setting where the principal can consult multiple agents, a scoring rule mechanism combined with a threshold-based decision rule based on the average reported belief leads to truthful reporting by all agents as the unique Nash equilibrium under precisely two conditions, preference diversity and no pivotality. The condition on preference diversity requires that at least one agent must prefer each action. No pivotality requires that the true mean aggregate belief must not be close to the decision threshold of the principal. Or in other words, no agent must unilaterally be able to change the action being chosen by the principal if all other agents report their beliefs truthfully. Note that this condition is more likely to be satisfied the more experts the principal consults. Moreover, the only time it does not hold is when the principal is (nearly) indifferent between the two actions.

I discuss several alternative mechanisms to the one described above. First, I consider a general method of aggregating belief reports by the agents, instead of using the average. One such alternative may be a principal who has some prior knowledge and applies Bayes' rule to incorporate the additional information provided by the agents. As long as the aggregation method is monotonic, which is the case for Bayesian updating, the results continue to hold, albeit with a slight adaptation of the condition for 'no pivotality'. Second, I discuss two popular methods that are frequently used in practice to elicit beliefs from a (large) group of agents, the Delphi method and prediction markets. I find that both methods are likely to result in a biased aggregate report when this report is used to make a decision affecting the agents.

Taken together, the results imply clear recommendations for a principal faced with a decision under uncertainty and expert advisers that may be impacted by the principal's decision. The principal should provide monetary incentives for accurate forecasts to the experts. Such incentives (even if arbitrarily small in theory) guarantee truthful forecasts by the experts under two simple conditions which the principal can influence. The principal should ensure that she consults experts with differing preferences. Even in large and homogeneous groups, consulting a single expert with different action preferences can lead to truth-telling by the whole group. Moreover, the expert should consult as many experts as possible to ensure no pivotality of any individual expert.

The rest of the paper is organized as follows. Section 2 discusses related work on belief elicitation and decentralized decision-making. Section 3 introduces the model and describes the background on scoring rule mechanisms. Section 4 analyzes a simplified setting with only one expert. Section 5 considers a setting with finitely many experts. Section 6 provides a discussion of some alternative mechanisms. Finally, section 7 provides some concluding remarks.

2. Related Literature

This paper draws on the large literature around scoring rules (e.g. Brier, 1950; Savage, 1971; Winkler et al., 1996, ...) and connects it to the problem of decentralized decision-making (as introduced by Holmström, 1977 and 1984) and strategic information transmission (Crawford and Sobel, 1982).

The vast majority of papers on belief elicitation and scoring rules do not consider subsequent decision-making problems. Gneiting and Raftery (2007) provide an overview of existing work and develop a unifying framework on scoring rules to elicit beliefs from a single agent. Two main types of mechanisms exist that are used to elicit beliefs from a group of agents: prediction markets and prediction polls. Conitzer (2009) studies different mechanisms based on prediction markets. He provides a framework, linking predictions markets and mechanism design, as well as characterizing mechanisms that are incentive compatible. Contrary to prediction markets a prediction poll is a more direct method of belief elicitation. Atanasov et al. (2017) show that prediction polls can be a good alternative to prediction markets in terms of accuracy.

A few papers (mainly in computer science) have considered settings where beliefs are elicited using a scoring rule and the resultant belief(s) are used by a principal to make a decision (Berg and Rietz (2003), Oesterheld and Conitzer (2019), Othman and Sandholm (2010), Chen and Kash (2011), Chen et al. (2011) and Dimitrov and Sami (2010)). The main difference to the model in this paper is that all of the above papers assume agents to be decision-agnostic. As far as I am aware, Boutilier (2012) is the only one who analyzes a setting where an agent has preferences over the decision made by the principal. The focus lies on a setting with one agent. He suggests the use of a 'compensation function' which assumes that the principal has (im)precise knowledge of the action preferences of the agent. Combining the knowledge of the action preferences with a proper scoring rule allows the creation of a mechanism that induces truth-telling.

Wolinsky (2002) studies a similar model to the one considered in this paper. A decision-maker wants to elicit information from a group of agents to make a decision. The decision affects the payoff of the decision-maker and the agents. However, the agents' preferences differ significantly from the preference of the decision-maker. Contrary to the model of this paper, Wolinsky assumes that no transfers between the decision-maker and the agents are possible, the state of the world is not revealed ex-post and all agents are identical in their preferences. Instead, he focuses on the role of communication between the agents and the role of commitment to a mechanism by the principal. Similarly to the results of this paper, pivotality of the agents plays a crucial role.

Gimpel and Teschner (2014), Choo et al. (2022) and Albrecht (2023) investigate a similar question experimentally. Gimpel and Teschner (2014) and Choo et al. (2022) consider a framework with multiple agents that receive utility depending on the decision made by a principal. A prediction market is used to elicit beliefs. Both papers find that agents misreport their belief strategically. Albrecht (2023) directly provides some agents of the group with manipulation incentives and compares the accuracy of group judgements in face-to-face interactions with the Delphi method. He finds that, while face-to-face is better than the Delphi method, neither method generates fully accurate group judgements.

This paper also relates to the general class of delegation problems as introduced by Holmström (1977, 1984). He poses the problem of an uninformed principal that has to make a decision under uncertainty. The principal may consult a group of informed but biased agents. He suggests that the optimal mechanism for the principal is to delegate the decision to the agent, letting him choose from some (constrained) set of alternatives. Alonso and Matouschek (2008) extend the framework and provide a general characterization of the solution to the delegation problem. Contrary to the model in this paper, they assume that monetary transfers between the principal and the agent are not possible. Krishna and Morgan (2004) study a very similar problem and allow for monetary transfers. They also analyze the role of commitment from the principal and provide a characterization of optimal contracts. A crucial assumption in their model is that the bias of the agent is common knowledge. Other differences to this paper are that the contracts cannot depend on the realized state and the utility of the principal and agent are based on quadratic loss.

The model of this paper is also related to the broader mechanism design problem with correlated information (Cremer and McLean, 1985, 1988; McAfee and Reny, 1992). McAfee and Reny (1992) study a setting where a principal makes a decision that matters to everyone and agents have correlated private information. They find that while having private information typically allows for large rents this is not necessarily the case if the information is correlated. In related work, Riordan and Sappington (1988) find that ex-post information can be used to eliminate rents ex-ante for a privately informed party. While none of these papers discuss the use of scoring rules to elicit beliefs, the general ideas are closely related to the findings of this paper.

3. Model

3.1. General setup

Consider a (female) principal who faces a decision problem $A = \{a_1, a_2\}$. There are two states of the world $\Omega = \{\omega_1, \omega_2\}$. Nature draws the state ω_2 with probability $\bar{\mu} \in [0, 1]$ and vice versa ω_1 is drawn with probability $1 - \bar{\mu}$. The principal has expected utility preferences that depend on the chosen action and her belief about the state of the world,

$$EU^P = \begin{cases} 0 & \text{if } a_1 \text{ is chosen} \\ \mu^P - \alpha & \text{if } a_2 \text{ is chosen} \end{cases}$$

where $\mu^P \in [0, 1]$ denotes the principal's (posterior) belief about the state being ω_2 and $\alpha \in [0, 1]$ indicates a certainty threshold for the choice of a_2 over a_1. The principal is assumed to have no information about the state of the world ex-ante. She adopts the aggregate belief of all agents as her posterior belief.[1] Her preferences imply that for a (posterior) belief $\mu^P < \alpha$ the principal prefers to choose a_1 and for a (posterior) belief $\mu^P > \alpha$ she prefers to choose a_2. In case of indifference, $\mu^P = \alpha$, I assume the principal chooses a_2.

There are n different Bayes-rational and risk-neutral (male) agents. Each agent i receives an independent signal $\mu_i \in \{\mu^1, ..., \mu^K\}$, with $0 < \mu^1 < ... < \mu^K < 1$ and $\mu^k - \mu^{k-1} = \epsilon$, which is correlated with $\bar{\mu}$.[2] This signal denotes the agents belief.[3] Each individual belief may be highly inaccurate but in aggregate, thanks to the law of large numbers, errors tend to cancel, such that with $n \to \infty$, $\frac{1}{n} \sum_{i \in N} \mu_i = \bar{\mu}$, where $N = \{1, ..., n\}$ denotes the set of all agents.[4] For a finite number of agents, $\tilde{\mu}$ denotes the average true belief of all agents. Similar

[1] I leave this deliberately abstract to avoid modeling a formal Bayesian updating problem of the principal. A further extension could be considered in the future. Section 6.1 provides a brief discussion.

[2] The beliefs of all agents are assumed to be discrete with steps of ϵ such that the optimal strategy for each agent is well-defined. This also seems plausible in practice, as agent likely cannot report all real numbers to the principal.

[3] An alternative interpretation might be that all agents receive the same signal but interpret it differently.

[4] This idea is similar to the standard statistical model of collective wisdom. For a further micro-foundation see Hong and Page (2012)

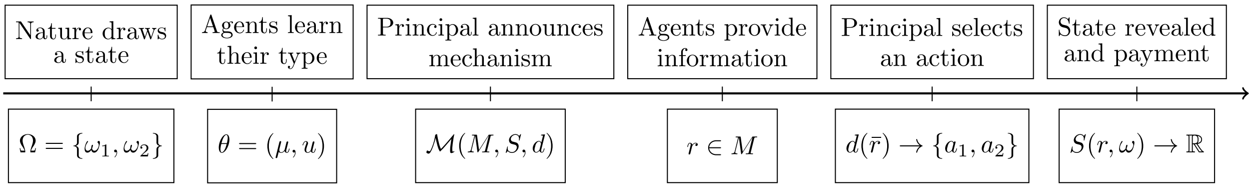

Figure 1: Summary timeline

to the principal, each agent i has preferences over the potential actions being selected by the principal,

$$U_i = \begin{cases} 0 & \text{if } a_1 \text{ is chosen} \\ u_i & \text{if } a_2 \text{ is chosen} \end{cases}$$

with $u_i \in \mathbb{R}$. This implies that, in the absence of other incentives, for $u_i < 0$ agent i prefers the choice of a_1 and vice versa. Agent i's belief and action preference can be summarized as the type, $\theta_i = (\mu_i, u_i)$. The set of all possible types is denoted by $\Theta = [0, 1] \times \mathbb{R}$.

The principal's objective is to choose the best action, maximizing her own expected utility. To do so she relies on information from the agents about the state of the world. The principal adopts a posterior belief μ^P equal to the average of all reported beliefs from the agents, $\bar{r} = \frac{1}{n}\sum_{i \in N} r_i$, where r_i is the reported belief of agent i.[5] Ideally, she would choose a_1 if $\bar{\mu} < \alpha$ and a_2 if $\bar{\mu} \geq \alpha$. Knowing the agents' action preferences is only indirectly relevant to the principal. The principal implements a specific direct mechanism $\mathcal{M}(M, S, d)$ in which the message space M for agents is restricted to belief reports $r \in \{\mu^1, ..., \mu^K\}$, transfers from principal to each agent are based on a scoring rule, $S : \{\mu^1, ..., \mu^K\} \times \Omega \to \mathbb{R}$, and the principal decides on an action based on an intuitive decision rule $d(\bar{r}) \to \{a_1, a_2\}$ given by:

$$d(\bar{r}) = \begin{cases} a_1 & \text{if } \bar{r} < \alpha \\ a_2 & \text{if } \bar{r} \geq \alpha \end{cases}$$

The scoring rule function determines the payoffs for all agents based on their reported belief and the realization of the state. The following section explains further details about scoring rules. Figure 1 provides a summary timeline of the events.

[5] For simplicity, I assume that beliefs are aggregated by taking the simple average. The results are robust to other forms of aggregation. A more general method is discussed in section 6.1.

3.2. Scoring Rules

Stimulating agents to report their belief truthfully requires some incentive for accurate belief reports. The realized state is observable before transfers are implemented, hence allowing transfers to depend on the state. Scoring rules are a common method of incentivizing truthful belief reports based on state realizations.

A scoring rule is a function $S : [0,1] \times \Omega \to \mathbb{R}$ chosen by the principal which determines a monetary payoff $S(r, \omega)$ based on the reported belief r and the state of the world ω. Given some scoring rule S and a true belief μ, the agent's expected payoff from reporting r is defined as follows:

$$E_\mu S(r) := \mu S(r, \omega_2) + (1 - \mu) S(r, \omega_1).$$

In the absence of action preferences a scoring rule that is considered proper leads agents to report their belief truthfully in order to maximize their expected payoff. Formally, a scoring rule S is considered proper, whenever

$$E_\mu S(\mu) \geq E_\mu S(r)$$

for every $r \neq \mu$ and every $\mu \in [0,1]$.

Following Gneiting and Raftery (2007), a scoring rule is characterized by a sub-differentiable function $G : [0,1] \to \mathbb{R}$ such that $E_\mu S(\mu) := G(\mu)$ for any $\mu \in [0,1]$. Consider the sub-tangent of $G(\mu)$ at r, $t_r(\mu) := a_r + b_r\mu$. Evaluating the tangent at zero and one shows the payoffs given a reported belief and the state of the world, i.e. $S(r, \omega_1) := t_r(0)$ and $S(r, \omega_2) := t_r(1)$. Figure 2 illustrates the payoffs following some example function G. For some reported belief r, the tangent of G at r shows the payoff $S(r, \omega_1)$ if the state is given by ω_1 and $S(r, \omega_2)$ if the state is given by ω_2. This implies that if an agent reports $r \neq \mu$ his expected score is given by:

$$E_\mu S(r) = G(r) + G'(r) \cdot (\mu - r).$$

Figure 2 also illustrates the expected cost of misreporting a belief. This cost is given by the Bregman divergence $d_G(r, \mu) := E_\mu S(\mu) - E_\mu S(r)$. As shown by Gneiting and Raftery (2007) the underlying function G determines whether a scoring rule is proper. A scoring rule is (strictly) proper if and only if G is (strictly) convex. Proper scoring rules have the desirable characteristic that in the absence of action preferences it is optimal for the agent

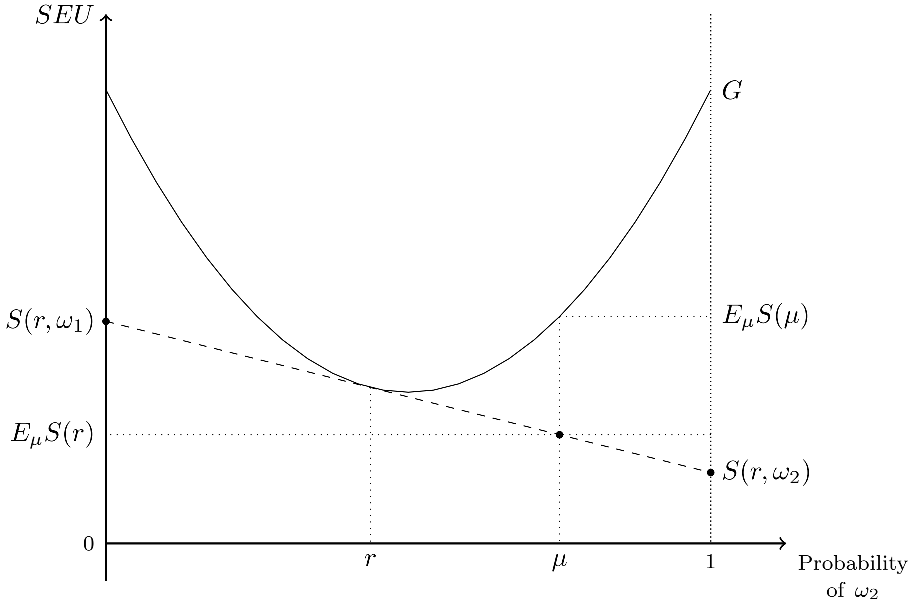

Figure 2: Some function G leads to the scoring rule payoffs as shown by the tangent of G at r.

to report his belief truthfully. From here on I refer to (strictly) proper scoring rules as a scoring rule that makes truth-telling a (strictly) dominant strategy in the absence of action preferences.

4. Single Agent

4.1. Benchmark

Before analyzing the setting with n different agents it is useful to understand the behavior of a single agent. The principal announces a mechanism $\mathcal{M}$ that restricts messages to (discrete) belief reports $r \in \{\mu^1, ..., \mu^K\}$ which are incentivized by some proper scoring rule S.[6] The principal commits to a decision rule d such that:

$$d(r) = \begin{cases} a_1 & \text{if } r < \alpha \\ a_2 & \text{if } r \geq \alpha \end{cases}$$

Note that I assume in case of indifference, $r = \alpha$, the principal chooses a_2. The agent has some belief μ about the probability of the state being ω_2 and some action preferences as given by $u \in \mathbb{R}$. This leads to the following expected utility for the agent,

[6]For convenience, I assume that the decision threshold α is included in the set of feasible belief reports.

$$EU(r) = \begin{cases} E_\mu S(r) & \text{if } r < \alpha \\ E_\mu S(r) + u & \text{if } r \geq \alpha \end{cases}$$

Expected utility depends on two factors, accuracy of the reported belief and the action being chosen by the principal. The two factors are additively separable, hence resulting in a trade-off for the agent. On the one hand, he would like to report his belief truthfully to maximize the expected payoff from the scoring rule. On the other hand, misreporting his belief may influence the action chosen by the principal and hence the outside utility of the agent. This implies that for any proper scoring rule and any outside utility $u \geq 0$ only two reports may be optimal for the agent, reporting truthfully, $r = \mu$, or reporting a belief exactly at the threshold, $r = \alpha$. Conversely, if $u \leq 0$ the only two reports that may be optimal are reporting truthfully, $r = \mu$, or reporting a belief just below the threshold, $r = \alpha - \epsilon$. Taking the scoring rule and outside utility as given it is easy to show that there exists a precise threshold $c_- \in [0, 1]$ such that an agent with a belief $\mu \in (c_-, \alpha)$ finds it optimal to over-report and, conversely, an agent with a belief $\mu \in (\alpha, c_-)$ finds it optimal to under-report. The threshold c_- is such that the utility from optimally misreporting, $E_{c_-}S(\alpha) + u$ or $E_{c_-}S(\alpha - \epsilon) + u$, is exactly equal to the utility of reporting truthfully, $E_{c_-}S(c_-)$. Therefore, c_- is such that $u = E_{c_-}S(c_-) - E_{c_-}S(\alpha)$ if $u > 0$ and respectively $u = E_{c_-}S(c_-) - E_{c_-}S(\alpha - \epsilon)$ if $u < 0$. Figure 3 provides a graphical illustration. Taking together the different observations leads to the optimal report, r^* as given below:

$$r^*(\mu, u) = \begin{cases} \begin{cases} \mu & \text{if } \mu \notin [c_-, \alpha] \\ \alpha & \text{if } \mu \in [c_-, \alpha] \end{cases} & \text{if } u > 0 \\ \begin{cases} \mu & \text{if } \mu \notin [\alpha, c_-] \\ \alpha - \epsilon & \text{if } \mu \in [\alpha, c_-] \end{cases} & \text{if } u < 0 \end{cases}$$

For $\mu = c_-$ the agent is indifferent between reporting truthfully, $r = \mu$, and misreporting to influence the action choice, $r = \alpha$ or $r = \alpha - \epsilon$. For the remainder of the paper I assume the agent reports his belief truthfully in this case.

The characterization of the optimal report shows that an agent benefits from misreporting his belief whenever his belief is within some range of the decision threshold set by the principal, i.e. (c_-, α) or (α, c_-). Importantly, the principal can influence this range through the choice of the scoring rule function, G. Loosely speaking, a more convex function G leads to a smaller misreporting range for the agent. One way to make G more convex is to offer larger (monetary) incentives for an accurate belief report to the agent.

To summarize, truth-telling can only be achieved if the principal pays substantial incentives (relative to how much the agent prefers one action over the other) or if the agent is indifferent between the two actions.

4.2. Optimal Scoring Rule

Attaching a decision directly to the report made by the agent gives rise to potential misreporting, as shown above. A proper scoring rule that normally makes truth-telling the unique optimal strategy for the agent does not necessarily have the same result if the agent has preferences over the actions. This section characterizes precisely which type of scoring rule mechanism preserves the property that truth-telling is optimal for any belief.

It is useful to define a new function $G^{net} : [0,1] \to \mathbb{R}$ which shows the utility from making a truthful report, including both the scoring rule payoff and the utility that follows from a certain action choice:

$$G^{net}(\mu) = \begin{cases} G(\mu) & \text{if } \mu < \alpha \\ G(\mu) + u & \text{if } \mu \geq \alpha \end{cases}$$

Similarly to before, one can define the sub-tangent of the function G^{net} at $r \in \{\mu^1, ..., \mu^K\}$ as $t_r^{net}(\mu)$. As G^{net} may be non-continuous the tangent $t_r^{net}(\mu)$ shows precisely when a truthful report leads to a lower expected payoff than some other report $r \neq \mu$.

Figure 3 illustrates the different incentives of the agent for some outside utility and some scoring rule incentives. For any belief μ the utility of reporting truthfully is illustrated. The constant labeled u shows the outside utility the agent receives if the principal chooses a_2. The two convex functions show the agents expected utility if he reports his belief truthfully plus the outside utility he receives. Note that the dashed parts of the convex functions are hypothetical for the agent as the principal only chooses a_2 with a report $r \geq \alpha$. The highlighted piecewise convex function therefore shows the expected utility of the agent if he reports his belief truthfully, G^{net}. As explained above, the agent may have an incentive to misreport his belief. The intersection of the tangent, t^{net} for $r = \alpha$ with the function G is given by c_-. This implies that for any true belief $\mu \in (c_-, \alpha)$ the agent's expected utility

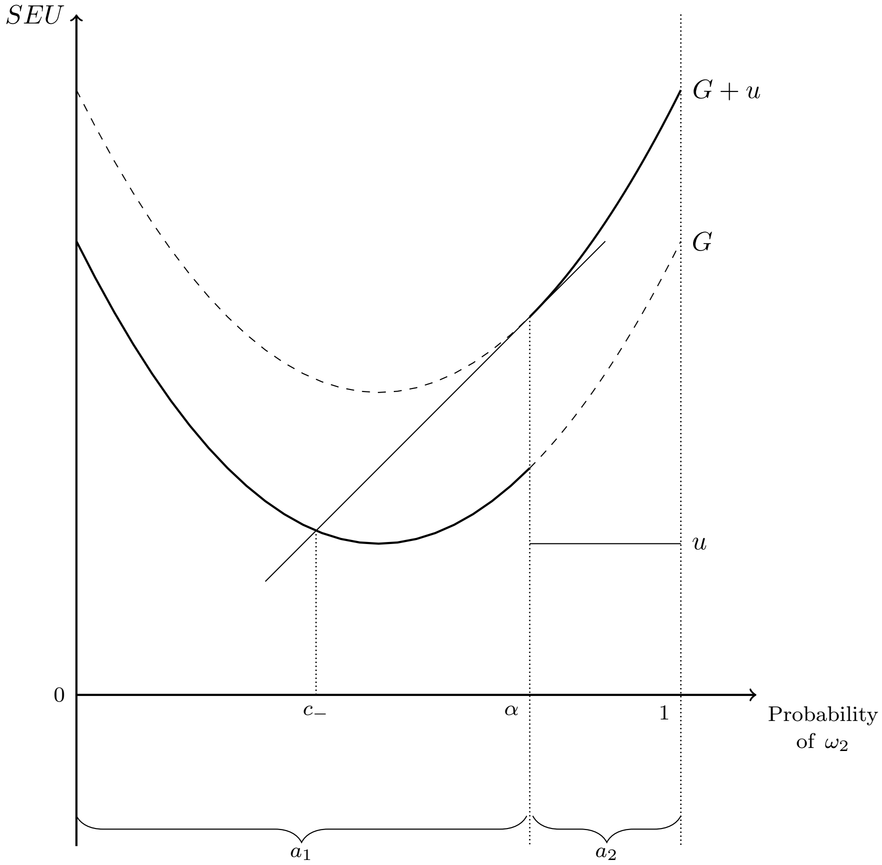

Figure 3: Incentive structure for some example scoring rule and some outside utility. For illustrative purposes a QSR is used. The highlighted convex function illustrates the expected utility of the agent for reporting his belief truthfully. For $\mu \in (c_-, \alpha]$ reporting $r = \alpha$ is optimal.

from reporting $r = \alpha$, such that the principal chooses a_2, is strictly higher than reporting $r = \mu$, i.e. $EU_\mu(\alpha) > EU_\mu(\mu)\ \forall \mu \in (c_-, \alpha)$. Therefore, for any true belief $\mu \in (c_-, \alpha]$, reporting $r = \alpha$ is optimal.

When trying to construct a scoring rule that makes truth-telling a dominant strategy for any belief μ, action preferences of the agent must be taken into account. As illustrated in Figure 3, unless G^{net} is convex, the agent may have some belief that would make it optimal to misreport. The following Theorem shows that indeed only one type of scoring rule mechanism leads to a guaranteed truthful report.

Theorem 1. *For any belief, μ, and some fixed outside preferences, u, truth-telling is a* *dominant strategy if and only if the scoring rule is given by S^* with*

$$S^*(r,\omega) = \begin{cases} S(r,\omega) + u & \textit{if } r < \alpha \\ S(r,\omega) & \textit{if } r \geq \alpha \end{cases}$$

where $S(r,\omega)$ is any proper scoring rule.

Proof. See appendix. □

Theorem 1 shows the unique scoring rule mechanism that guarantees a truthful belief report, $r = \mu$, for any belief μ. However, this scoring rule requires the principal to have precise knowledge of the action preferences of the agent. Action preferences are by definition unknown to the principal. Guaranteeing a truthful report by the agent is therefore impossible for the principal.

5. Multiple Agents

This section focuses on a direct mechanism where agents report their beliefs simultaneously and independently.[7] There are n different agents that all have some knowledge about the state of the world. The principal announces a mechanism $\mathcal{M}$ that restricts messages to belief reports, $r_i \in \{\mu^1, ..., \mu^K\}$ which are again incentivized by a proper scoring rule, S. The principal commits to a decision rule d such that:

$$d(r_1, ..., r_n) = \begin{cases} a_1 & \text{if } \bar{r} < \alpha \\ a_2 & \text{if } \bar{r} \geq \alpha \end{cases}$$

where $\bar{r} = \frac{1}{n}\sum_{i \in N} r_i$

5.1. Agent Behavior

The suggested mechanism implies that the utility for each agent depends not only on his own report but also on the average report of all other agents, $\tilde{r}_{-i} := \frac{1}{n-1}\sum_{j \neq i} r_j$. Formally, the expected utility of agent i is given by:

$$EU_i(r_i) = \begin{cases} E_{\mu_i} S(r_i) & \text{if } \bar{r} < \alpha \\ E_{\mu_i} S(r_i) + u_i & \text{if } \bar{r} \geq \alpha \end{cases}$$

[7] This mechanism is also referred to as a 'prediction poll'.

The main difference to the scenario with just a single agent is that individual agents may not be able to influence the principal's decision. If the principal only consults a single agent this agent knows with certainty that submitting a report $r_i \geq \alpha$ leads to action a_2 being chosen. This is not the case with multiple agents. Only agents that are pivotal can submit a report that may affect the action being chosen. Formally, pivotality is defined as follows.

Definition 1. Agent i is considered to be *pivotal* at $\tilde{r}_{-i}$ if $\frac{n-1}{n}\tilde{r}_{-i} < \alpha \leq \frac{n-1}{n}\tilde{r}_{-i} + \frac{1}{n}$.

Given some $\tilde{r}_{-i}$, $\bar{r} = \frac{n-1}{n}\tilde{r}_{-i}$ is the lowest possible aggregate report that agent i can achieve by reporting $r_i = 0$. Conversely, $\bar{r} = \frac{n-1}{n}\tilde{r}_{-i} + \frac{1}{n}$ is the highest possible aggregate report that agent i can achieve by reporting $r_i = 1$. If the threshold α is such that agent i can influence the action choice of the principal, i.e. reporting $r_i = 0$ leads to $\bar{r} < \alpha$ and reporting $r_i = 1$ leads to $\bar{r} \geq \alpha$, this agent is considered pivotal. Pivotality plays a crucial role for the agent when deciding on a belief to report. If an agent is not pivotal he has no incentive to misreport his true belief. Besides characterizing when an agent is pivotal it is useful to define a pivotal report for each agent, $c_{i,+} \in \mathbb{R}$. If agent i reports $c_{i,+}$ it leads to an aggregate report exactly at the decision threshold $\bar{r} = \alpha$ for any reported beliefs by the other agents, $\tilde{r}_{-i}$.[8] Formally, this pivotal report is given by:

$$c_{i,+} := \alpha + (n-1)(\alpha - \tilde{r}_{-i})$$

Note that an agent may not be pivotal, in which case $c_{i,+} \notin [0,1]$. If the agent is pivotal then for any $r_i \geq c_{i,+}$ the principal chooses action a_2 and, respectively, for any $r_i < c_{i,+}$ the principal chooses action a_1. The optimal report for each agent is defined below.

Proposition 1. *The optimal report for agent i is given by:*

$$r_i^*(\mu, u, \tilde{r}_{-i}) = \begin{cases} \mu_i & \text{if } c_{i,+} \notin [0,1] \\ \begin{cases} \mu_i & \text{if } \mu_i \notin (c_{i,-}, c_{i,+}] \\ c_{i,+} & \text{if } \mu_i \in (c_{i,-}, c_{i,+}] \end{cases} & \text{if } c_{i,+} \in [0,1] \end{cases}$$

if $u_i \geq 0$, and

$$r_i^*(\mu, u, \tilde{r}_{-i}) = \begin{cases} \mu_i & \text{if } c_{i,+} \notin [0,1] \\ \begin{cases} \mu_i & \text{if } \mu_i \notin (c_{i,+}, c_{i,-}] \\ c_{i,+} - \epsilon & \text{if } \mu_i \in (c_{i,+}, c_{i,-}] \end{cases} & \text{if } c_{i,+} \in [0,1] \end{cases}$$

[8] To simplify later notation I omit the dependence on $\tilde{r}_{-i}$ in the notation of $c_{i,+}$.

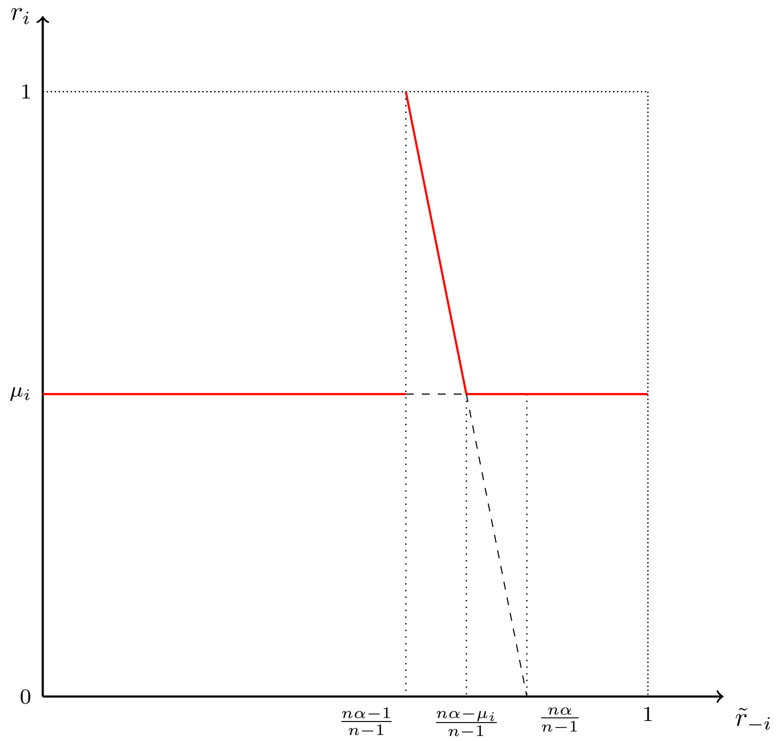

Figure 4: Best response function.

if $u_i < 0$.

As before, $c_{i,-}$ is the report such that agent i is indifferent between misreporting and reporting truthfully, i.e. $u = E_{c\_}S(c_-) - E_{c\_}S(\alpha)$ if $u > 0$ and respectively $u = E_{c\_}S(c_-) - E_{c\_}S(\alpha - \epsilon)$ if $u < 0$. If $\mu_i = c_{i,-}$, I assume the agent reports his belief truthfully.

The optimal report for each agent is similar to the case with only a single agent. The main difference is that while an agent can always influence the decision if he is the only one that is consulted, this is not necessarily the case with multiple agents. As the number of agents increases it is less likely that an individual agent is pivotal to the decision. In this case it is best for the agent to report his belief truthfully. Consider the example as shown in Figure 4. The agent prefers a_2 over a_1. For simplicity assume that the action preferences u_i are sufficiently large relative to the incentives of the scoring rule mechanism such that $c_{i,-}$ plays no role. The optimal report given any value of $\tilde{r}_{-i}$ is then indicated by the red solid line. Every report on or above the diagonal dashed line leads to $\bar{r} \geq \alpha$ and hence a_2 being chosen by the principal. As shown above, if the agent is not pivotal only a truthful report is optimal for him. However, in the interval $[\frac{n\alpha-1}{n-1}, \frac{n\alpha-\mu_i}{n-1})$ it is optimal to over-report.

5.2. Nash Equilibria

For any type profile $(\theta_1, ..., \theta_n)$ only two kinds of (pure strategy) equilibria are feasible, the truth-telling equilibrium and a single misreporting equilibrium. The truth-telling equilibrium is defined as the equilibrium outcome in which every agent reports his belief truthfully, such that the final average is given by $\bar{r} = \tilde{\mu}$, where $\tilde{\mu}$ denotes the average true belief of all agents. An outcome is considered a misreporting equilibrium if at least one agent misreports his belief.

Ideally, the principal would like every agent to submit a truthful report. If the principal can only consult a single expert this cannot be guaranteed. With multiple experts truth-telling is the unique and strict Nash equilibrium under two conditions, diversity and no pivotality. The two conditions are formally defined below.

Theorem 2. *For any number of agents ($n \geq 2$) and any strictly proper scoring rule S, all agents reporting their belief truthfully, $r_i = \mu_i \; \forall \; i$, is the unique and strict Nash equilibrium if,*

1) ***Diversity:*** *the profile of action preferences is such that for at least one agent i $u_i > 0$ and for at least one agent j $u_j < 0$, and*

2) ***No pivotality:*** $\tilde{\mu} \notin [\alpha - \frac{1}{n}, \alpha + \frac{1}{n})$.

Proof. The complete proof can be found in the appendix. The condition on $\tilde{\mu}$ guarantees existence of a truth-telling equilibrium and the condition on agents' preferences ensures uniqueness by eliminating any misreporting equilibrium. □

The 'diversity' condition states that at least one agent must prefer each action. This marks a main difference to the single expert scenario in which the expert can naturally only prefer one of the two actions. This condition on preferences eliminates any potential misreporting equilibria. In practice, it seems plausible that a principal has some degree of control over the group of experts. The principal should thus ensure at least a minimal degree of diversity among the group of experts. While the principal does not know the exact action preferences for any agent, in practice she may know if u is positive or negative for at least some agents. Including some agents from which she knows the direction of u, allows her to make sure that the preference diversity condition is satisfied.

The 'no pivotality' condition requires the true average belief $\tilde{\mu}$ to be such that no agent can individually change the chosen action. Whenever this condition is satisfied, a truth-telling equilibrium is guaranteed to exist. While this condition may not be met in some scenarios it is worth pointing out that the principal is most interested in a truthful aggregate report if the true mean belief is further away from the threshold. In the extreme case of $\tilde{\mu} = \alpha$ the principal is indifferent between the two actions and receiving an aggregate report $\bar{r} \neq \tilde{\mu}$ does not affect the principal. The 'no pivotality' condition largely depends on two factors, the number of agents who reported a belief and the decision threshold chosen by the principal. In practice, the principal should consult as many agents as possible to reduce the chance of individual pivotality as much as possible. Moreover, if the principal had some prior about the true mean belief she may want to choose a decision threshold, α, further away from it.[9]

A second point to note regarding the 'no pivotality' condition is that it requires agents to have reasonably accurate expectations about the average true belief, $\tilde{\mu}$. While agents do not need to know the exact value of $\tilde{\mu}$, if the 'no pivotality' condition is satisfied, agents need to be certain that they are not individually pivotal. Uncertainty about the value of $\tilde{\mu}$ might lead agents to misreport their true belief, hoping to influence the final decision. Agents are likely to have reasonably accurate expectations about $\tilde{\mu}$ if either the interval $[\alpha - \frac{1}{n}, \alpha + \frac{1}{n})$ is sufficiently small or agents know each other. One example may be a committee where agents repeatedly interact.

A final point to mention is that the size of the scoring rule incentives do not matter with this mechanism (as long as the pivotality and diversity condition are satisfied). The principal can choose any strictly proper scoring rule, even with arbitrarily small payoffs. This is a clear difference to the setting where the principal can only consult a single expert. In that case, the agent has a clear trade-off between scoring rule incentives and utility from the action choice. The higher the payoffs from the scoring rule the smaller the range of beliefs the agent would misreport. This is not the case with multiple agents. The precise form of the scoring rule incentives are of lesser importance. The main factor leading agents to report their belief truthfully is that each agent individually must not be able to change the decision of the principal.

[9]The principal would face a trade-off between maximizing her expected utility from choosing the preferred action and incentivizing agents to report their beliefs truthfully.

6. Discussion

The previous sections focused on a direct reporting mechanism in which all agents make one simultaneous report. The principal aggregated all reports by taking the average and committed to an intuitive decision rule. This section provides a discussion of a more general aggregation rule and two popular alternative mechanisms, the Delphi method and prediction markets.

6.1. Robustness

The first alternative is a different, more general method of aggregating reported beliefs for the principal. Throughout the analysis I have assumed that the principal aggregates beliefs by simply taking the mean of all reported beliefs. Of course, other methods are possible. This section discusses a more general form of aggregating beliefs and shows that the main results continue to hold.

I consider a general rule of aggregating reports, $\bar{r} = f(r_1, ..., r_n) \in [0, 1]$, that maintains the property of monotonicity for each agent. Consider the vector of all reports excluding the report of agent i, $\tilde{r}_{-i} = (..., r_{i-1}, r_{i+1}, ...)$. A monotonic aggregation rule maintains the property that if any agent reports a higher belief, the overall aggregate belief of the principal, $\bar{r}$, is higher as well.

Definition 2. The aggregation rule f is strictly monotonic if for every agent i and two arbitrary reports r_i and r_i' with $r_i > r_i'$, it is the case that: $f(r_i, \tilde{r}_{-i}) > f(r_i', \tilde{r}_{-i})$.

As before, it is possible to define a pivotal report $c_{i,+}$, which is the smallest report that leads to a_2 being selected. Previously, the pivotal report was defined as some number $c_{i,+} \in \mathbb{R}$ such that the final aggregate report is exactly equal to the decision threshold, $\bar{r} = \alpha$. For an arbitrary monotonic aggregation rule this may not be possible. Therefore, the pivotal report is only defined if it is a feasible report in the mechanism. Formally,

$$c_{i,+} := \min \left\{ r_i \in \{\mu^1, \ldots, \mu^K\} | f(r_i, \tilde{r}_{-i}) \geq \alpha > f(r_i - \epsilon, \tilde{r}_{-i}) \right\}$$

In words, the pivotal report is the smallest feasible report leading to a decision of a_2 for which also the next smaller feasible report leads to a decision of a_1. Note that, agent i may not be pivotal, in which case no pivotal report $c_{i,+}$ exists that satisfies the condition stated above. If agent i is pivotal, in which case $c_{i,+}$ exists, monotonicity of the aggregation rule ensures that $c_{i,+}$ is unique. It also implies that any report $r_i \geq c_{i,+}$ leads to a_2 being chosen and vice versa, for any report $r_i < c_{i,+}$, a_1 is being chosen.

Other than the adapted definition of $c_{i,+}$, there are no differences to the analysis in Section 5. As before, agents have no incentive to misreport their belief if they are not pivotal and otherwise only misreport if the benefit of misreporting is larger than the foregone payoff from the scoring rule. The optimal report is then defined as in Proposition 1, however, including the adapted definition of $c_{i,+}$. This implies that also Theorem 2 continues to hold albeit with a slightly different condition for 'no pivotality'. Formally, the 'no pivotality' condition in Theorem 2 can be re-written as:

$$\mu_i \notin [\alpha - f(0, \tilde{\mu}_{-i}), \alpha + f(1, \tilde{\mu}_{-i})] \ \forall \ i,$$

where $\tilde{\mu}_{-i}$ denotes the vector of all true beliefs excluding the true belief of agent i. Note that the adapted 'no pivotality' condition depends on the aggregation rule f and it is not possible to define an interval around the true average belief of all agents, $\tilde{\mu}$. Instead, each agents' individual belief, μ_i, must satisfy the condition stated above. All other parts of Theorem 2 remain unchanged and a formal proof using the more general aggregation rule would be nearly identical to the proof of Theorem 2. The different condition for 'no pivotality' has important implications for the practical implementation of the suggested mechanism. Previously, agents needed to have some knowledge about the true average belief of all agents, $\tilde{\mu}$, in order for Theorem 2 to hold. In contrast to that, the adapted definition of 'no pivotality' requires all agents to have precise knowledge of the true beliefs of all other agents. In practice, this assumption may be more difficult to satisfy. While I am not trying to find an optimal aggregation rule for the principal, it seems that the principal may benefit from choosing a simple aggregation mechanism such as the average of all reports.

Throughout this paper I have assumed that the principal directly adopts the average reported belief of all agents as her posterior belief. This may not be what a Bayesian rational principal would do, even if all agents' reports are truthful. A Bayesian rational principal should take into account her own prior belief, which was not formally modelled in this paper. Nonetheless, the results in this paper continue to hold with a Bayesian principal. A Bayesian updating rule fulfills the condition of monotonicity with aggregation presented above. Assuming that agents report their beliefs truthfully, a higher reported belief by any agent would lead to a higher posterior belief of the principal.[10]

6.2. Practical Implementation

This section discusses two popular methods of eliciting beliefs from a group of agents in practice, the Delphi method and prediction markets.

Delphi Method

The Delphi method was developed as a group forecasting tool in the 1950s by RAND.[11] A group of experts are asked to give individual estimates which are aggregated by a principal. The principal announces the aggregate report to the agents who can then revise their initial estimate. This process is repeated until consensus is achieved.

In the framework of this paper it could serve as a way to implement the mechanism analyzed above. Theorem 2 shows that truth-telling can be guaranteed in the one-shot reporting mechanism if the 'diversity' and 'no pivotality' condition are satisfied. However, the implicit assumption is that all agents know the true average belief $\tilde{\mu}$. This assumption is reasonable in settings where agents know each other well but may not be fulfilled in some other settings. If agents do not know $\tilde{\mu}$ perfectly, they may hold incorrect expectations about the reports of all other agents, $\tilde{r}_{-i}$. Agents may incorrectly assume that they are pivotal.

One may try to overcome this problem with a mechanism similar to the Delphi method. With this mechanism, agents are to report their beliefs simultaneously to the principal who incentivizes accuracy with a scoring rule mechanism, as before. After all reports are collected, the principal announces the aggregate report. The agents are then asked to report their beliefs again. This process is repeated until the aggregate reported belief no longer changes. The suggested mechanism could correct agents' expectations about $\tilde{r}_{-i}$ and therefore, lead to truth-telling by all agents, given that the two conditions from Theorem 2 are satisfied.

The main advantage of the Delphi method, correcting agents' expectations about others, also introduces several new problems. First, agents may try to learn from the reports of the other agents and correct their own reported belief. This may be undesirable for the

[10]For this to be true, neither the principal nor the agents should have perfect information about the state of the world. I assume this to be the case. The principal is initially uncertain about the state of the world and all agents have imperfect information, i.e. $\forall\ i$, $0 < \mu_i < 1$.

[11]See: https://www.rand.org/topics/delphi-method.html.

principal who may want to elicit agents' true initial beliefs. Truth-telling by all agents would essentially require that agents are convinced of their own true belief, μ_i, and do not update this belief based on the announced aggregate report, $\bar{r}$.[12] A second and more serious problem is that agents may want to misreport their beliefs to influence how other agents behave in later rounds of the game. The sequential nature of the game makes the existence of a truth-telling equilibrium less likely if agents are sophisticated. Truth-telling by all agents would require the assumption that agents are myopic in every round of the game. Agents should not take into account the effect their report may have on other agents in future rounds of the game.

In view of this discussion, the Delphi method is not guaranteed to deliver accurate results in a setting where the final aggregate report is used to make a decision that affects the agents.

Prediction Markets

Prediction markets are frequently used in practice to aggregate beliefs from a group of agents. They are employed to forecast election outcomes, other geo-political and financial events, or the replication of academic articles[13] In standard settings, without a decision attached to it, prediction markets are said to "overwhelmingly outperform conventional forecasting methods" (Choo et al., 2022, p.6716). In prediction markets, agents can buy and sell an asset that ultimately yields a payoff of zero or one depending on the realisation of a state of the world. With each trade of an asset the market price of the asset updates. Prediction markets often run over a time period of hours/days/weeks and agents can buy or sell as many assets as is possible in the given market. Buying and selling assets at a given market price is interpreted as agents reporting their belief about the uncertain state of the world.

Prediction markets differ in two main ways to the mechanism analyzed in detail in this paper. First, agents trade sequentially and sometimes repeatedly. Second, agents can choose themselves to what extent their belief should influence the final aggregate report. Both

[12] I assume that all agents have equally accurate information about the uncertain state of the world. In practice, agents may have differently accurate information and some agents updating their own true beliefs based on the announced aggregate report may be desirable. Parenté and Anderson-Parenté (1987) show that Delphi accuracy improves over rounds when relatively less accurate agents update their beliefs more than relatively more accurate agents. Rowe et al. (2005) provide empirical support for this assumption.

[13] Examples include: The Iowa Electronic Market, PredictIt, the Good Jugdement Project, and see Gordon et al. (2021) for a recent overview of prediction markets in academic studies used to predict replications.

of these differences are problematic if the aggregate report is used for a decision that at least some agents care about (a lot). With sequential and potentially repeated interactions by agents often times no truth-telling equilibrium exists. The reason is that sequential reporting facilitates coordination by the agents to jointly misreport their beliefs. The second difference is that the weight of each agent in the final aggregate is endogenous. This is especially problematic for the existence of a truth-telling equilibrium. With agents being able to choose their own weight in the final aggregate, each agent could become pivotal to the decision taken by the principal. That implies that agents with strong preferences regarding the action choice have a strong incentive to misreport and manipulate the action choice.

In general, prediction markets seem not well suited to elicit and aggregate beliefs from multiple agents when a decision is directly attached to the outcome of the market which affects the agents. Both the sequential nature of eliciting beliefs as well as the endogenous weight each agent has in the final aggregate report are problematic for truthful reporting.

7. Conclusion

This paper considered a setting in which an uninformed principal is faced with a decision. The principal can consult multiple experts for advice, however, the experts are affected by the principal's decision. The principal incentivizes agents to report their beliefs truthfully with a scoring rule mechanism and uses a deterministic, threshold-based decision rule to implement one of two actions based on the aggregate reported belief. I show that with just one agent, truth-telling can only be achieved if the agent's preferences happen to be aligned with the ones of the principal or if the principal pays substantial incentives for truthful reporting (relative to how much the agent prefers one action over the other). With multiple agents, the principal can implement a direct reporting mechanism, similar to a 'prediction poll', commit to an intuitive decision rule and pay minimal incentives in the form of a scoring rule. All agents reporting their belief truthfully is the unique Nash equilibrium if no agent is individually pivotal and not all agents prefer the same action to be chosen by the principal. In practice the principal should therefore focus on consulting as many agents as possible rather than offering high incentives to few agents for their advice. I further show that the results of this paper are robust to any monotonic aggregation of beliefs from multiple agents.

Finally, two common methods of aggregating beliefs from multiple agents, the Delphi method and prediction markets, may not work well in a setting where the final belief is used to make a decision that affects the agents. While truth-telling could be sustained under some strong assumptions for the Delphi method, it seems implausible that complete truth-telling could occur in prediction markets.

References

Albrecht, D. (2023). Mitigating manipulation in committees: Just let them talk! *Working Paper*.

Alonso, R. and Matouschek, N. (2008). Optimal delegation. *The Review of Economic Studies*, 75(1):259–293.

Atanasov, P., Rescober, P., Stone, E., Swift, S. A., Servan-Schreiber, E., Tetlock, P., Ungar, L., and Mellers, B. (2017). Distilling the wisdom of crowds: Prediction markets vs. prediction polls. *Management Science*, 63(3):691–706.

Berg, J. E. and Rietz, T. A. (2003). Prediction markets as decision support systems. *Information Systems Frontiers*, 5(1):79–93.

Boutilier, C. (2012). Eliciting forecasts from self-interested experts: Scoring rules for decision makers. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, AAMAS '12, page 737–744, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3.

Chen, Y. and Kash, I. (2011). Information elicitation for decision making. In *AAMAS '11: The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*.

Chen, Y., Kash, I., Ruberry, M., and Shnayder, V. (2011). Decision markets with good incentives. In *WINE*, pages 72–83.

Choo, L., Kaplan, T. R., and Zultan, R. (2022). Manipulation and (mis) trust in prediction markets. *Management Science*, 68(9):6716–6732.

Conitzer, V. (2009). Prediction markets, mechanism design, and cooperative game theory. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, UAI '09, page 101–108, Arlington, Virginia, USA. AUAI Press.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431.

Cremer, J. and McLean, R. P. (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345.

Cremer, J. and McLean, R. P. (1988). Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247.

Dimitrov, S. and Sami, R. (2010). Composition of markets with conflicting incentives. In *EC '10*.

Friedman, J. and Zeckhauser, R. (2014). Handling and mishandling estimative probability: Likelihood, confidence, and the search for bin laden. *Intelligence and National Security*, 30:77–99.

Gimpel, H. and Teschner, F. (2014). Market-based collective intelligence in enterprise 2.0 decision making.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *American Statistical Association*.

Gordon, M., Viganola, D., Dreber, A., Johannesson, M., and Pfeiffer, T. (2021). Predicting replicability - analysis of survey and prediction market data from large-scale forecasting projects. *PLOS ONE*, 16(4):e0248780.

Holmström, B. (1977). On incentives and control in organizations.

Holmström, B. (1984). *On the Theory of Delegation*, chapter Bayesian Models in Economic Theory, pages 115–141. North Holland.

Hong, L. and Page, S. E. (2012). Some microfoundations of collective wisdom. In *Collective Wisdom*, pages 56–71. Cambridge University Press.

Krishna, V. and Morgan, J. (2004). The art of conversation: eliciting information from experts through multi-stage communication. *Journal of Economic Theory*, 117(2):147–179.

McAfee, R. P. and Reny, P. J. (1992). Correlated information and mecanism design. *Econometrica*, 60(2):395.

Oesterheld, C. and Conitzer, V. (2019). Eliciting information for decision making from individual and multiple experts.

Othman, A. and Sandholm, T. (2010). Decision rules and decision markets. volume 1, pages 625–632.

Parenté, F. J. and Anderson-Parenté, J. K. (1987). Delphi inquiry systems. *Judgmental forecasting*, 2:23.

Riordan, M. H. and Sappington, D. E. (1988). Optimal contracts with public ex post information. *Journal of Economic Theory*, 45(1):189–199.

Rowe, G., Wright, G., and McColl, A. (2005). Judgment change during delphi-like procedures: The role of majority influence, expertise, and confidence. *Technological Forecasting and Social Change*, 72(4):377–399.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*.

Winkler, R. L., Muñoz, J., Cervera, J. L., Bernardo, J. M., Blattenberger, G., Kadane, J. B., Lindley, D. V., Murphy, A. H., Oliver, R. M., and Ríos-Insua, D. (1996). Scoring rules and the evaluation of probabilities. *Test*.

Wolinsky, A. (2002). Eliciting information from multiple experts. *Games and Economic Behavior*, 41(1):141–160.

A. Proofs

A.1. Theorem 1

Suppose $S^*(r, \omega)$ is given as in Lemma 1. Then G^{net} is given by

$$G^{net}(\mu) = \begin{cases} G(\mu) + u & \text{if } \mu \geq \alpha \\ G(\mu) + u & \text{if } \mu < \alpha \end{cases}$$

This implies that G^{net} is strictly convex and hence proper. Therefore, truth-telling is the dominant strategy, as shown in section 3.2.

Conversely, suppose that the agent reports his belief truthfully for every belief, μ, and some fixed outside preferences, u. Hence, $t_r^{net}(\mu) \leq G^{net}(\mu) \quad \forall \mu \in [0, 1]$. Therefore, $G^{net}(\mu)$ must be convex, as shown for example by Gneiting and Raftery (2007). Hence, the scoring rule is given as stated in Theorem 1.

A.2. Proposition 1

The proof consists of two parts. Part 1 shows that given some $\tilde{r}_{-i}$, if agent i is not pivotal, for any (strictly) proper scoring rule S it is (strictly) optimal for the agent to report his belief truthfully, $r_i = \mu_i$. Part 2 shows that given some $\tilde{r}_{-i}$, such that agent i is pivotal, the only report $r_i \neq \mu_i$ that could be optimal is given by $r_i = c_{i,+}$ if $u > 0$ or $r_i = c_{i,+} - \epsilon$ if $u < 0$. Taking together part 1 and part 2 leads to the optimal report as stated in Proposition 1.

Part 1. Agent i not being pivotal implies that $\tilde{r}_{-i}$ is such that $\alpha \notin (\frac{n-1}{n}\tilde{r}_{-i}, \frac{n-1}{n}\tilde{r}_{-i} + \frac{1}{n}]$. Consider two cases. 1) $\tilde{r}_{-i} < \frac{n}{n-1}\alpha - \frac{1}{n-1}$: Then for any r_i it is the case that $\bar{r} = \frac{n-1}{n}\tilde{r}_{-i} + \frac{1}{n}r_i < \alpha$. This implies that $EU_i(r_i) = E_{\mu_i}S(r_i) + u_i$. Hence, $r_i = \mu_i$ is optimal.

2) $\tilde{r}_{-i} \geq \frac{n}{n-1}\alpha$: Then for any $r_i \in$ it is the case that $\bar{r} = \frac{n-1}{n}\tilde{r}_{-i} + \frac{1}{n}r_i \geq \alpha$. This implies that $EU_i(r_i) = E_{\mu_i}S(r_i) + u_i$. Hence, $r_i = \mu_i$ is optimal.

Part 2. Consider agent i with true belief $\mu_i < c_{i,+}$ and $u > 0$. A truthful report would thus lead to $\bar{r} < \alpha$ and a_1 being chosen by the principal. Suppose the agent reports $r_i' > c_{i,+}$ such that $\bar{r}' > \alpha$. Both reports, $c_{i,+}$ and r_i', lead to the same action being chosen by the principal, a_2. Hence, the agent would have a higher expected utility from reporting $r_i = r_i' - \epsilon < r_i'$. Thus r_i' is never optimal. The same holds true vice versa for an agent with $\mu_i > c_{i,+}$ and $u < 0$.

A.3. Theorem 2

Existence: Let $\tilde{\mu}_{-i}$ denote the average truthful report of all agents excluding agent i, i.e. $\tilde{\mu}_{-i} := \frac{1}{n-1}\sum_{j\neq i}\mu_j$. The no pivotality condition implies that for each agent i, $\frac{1}{n}r_i + \frac{n-1}{n}\tilde{\mu}_{-i} \geq \alpha$ or $\frac{1}{n}r_i + \frac{n-1}{n}\tilde{\mu}_{-i} < \alpha \quad \forall\ r_i$. By Proposition 1, for any strictly proper scoring rule S, $EU_i(\mu_i) > EU_i(r_i)\ \forall\ r_i \neq \mu_i$. Hence, each agent i reporting $r_i = \mu_i$ is a Nash Equilibrium.

Uniqueness: Suppose a Nash equilibrium exists in which at least one agent i misreports his belief, i.e. $r_i \neq \mu_i$. Consider 3 cases for the aggregated average belief $\bar{r}$. Case 1) $\bar{r}$ is not near the decision threshold α, i.e. $\bar{r} > \alpha$ or $\bar{r} < \alpha - \epsilon$: Any agent that reported $r_i \neq \mu_i$ would be strictly better off by reporting $r_i' - \epsilon$ if $r_i > \mu_i$ or $r_i' + \epsilon$ if $r_i < \mu_i$. The reason is as follows. $\bar{r} \neq \alpha$ and $\bar{r} \neq \alpha - \epsilon$ implies that reporting r_i' does not change which action is selected by the principal. Hence, $EU_i(r_i') > EU_i(r_i)$. Therefore, $\bar{r} \neq \tilde{\mu}$ with $\bar{r} < \alpha - \epsilon$ or $\bar{r} > \alpha$ is not a Nash equilibrium. Case 2) $\bar{r} = \alpha$: Diversity implies that at least one agent i has action preferences such that $u_i < 0$. In this case, $EU_i(\mu_i - \epsilon) > EU_i(\mu_i)$ as it would lead to $\bar{r} = \alpha - \epsilon'$ and the choice of a_1 by the principal. Hence, $\bar{r} = \alpha$ is not a Nash equilibrium. Case 3) $\bar{r} = \alpha - \epsilon$: Diversity implies that at least one agent i has action preferences such that $u_i > 0$. In this case, $EU_i(\mu_i + \epsilon) > EU_i(\mu_i)$ as it would lead to $\bar{r} = \alpha$ and the choice of a_2 by the principal. Hence, $\bar{r} = \alpha - \epsilon$ is not a Nash equilibrium.

Taken together, this implies that all agents reporting $r_i = \mu_i$ is the unique Nash equilibrium for any profile of beliefs and action preferences that satisfy (a) no pivotality and (b) diversity.