Collective Revelation: A Mechanism for Self-Verified, Weighted, and Truthful Predictions

Sharad Goel　　Daniel M. Reeves　　David M. Pennock

Yahoo! Research
111 W 40th Street, New York, NY 10018
{goel,dreeves,pennockd}@yahoo-inc.com

ABSTRACT

Decision makers can benefit from the subjective judgment of experts. For example, estimates of disease prevalence are quite valuable, yet can be difficult to measure objectively. Useful features of mechanisms for aggregating expert opinions include the ability to: (1) incentivize participants to be truthful; (2) adjust for the fact that some experts are better informed than others; and (3) circumvent the need for objective, "ground truth" observations. Subsets of these properties are attainable by previous elicitation methods, including proper scoring rules, prediction markets, and the Bayesian truth serum. Our mechanism of *collective revelation*, however, is the first to simultaneously achieve all three. Furthermore, we introduce a general technique for constructing *budget-balanced* mechanisms—where no net payments are made to participants—that applies both to collective revelation and to past peer-prediction methods.

Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Algorithms, Design, Economics

Keywords

Forecasting, prediction markets, polls, mechanism design

1. INTRODUCTION

Many predictions, for example regarding the effects of economic policy, are fundamentally subjective in nature. In such cases, a decision maker may seek to elicit and aggregate the opinions of multiple experts. Ideally, the decision maker would like a mechanism that is: (1) *incentive compatible*, or rewards participants to be truthful; (2) *information weighted*, or adjusts for the fact that some experts are better informed than others; (3) *self-verifying*, or works without the need for objective, "ground truth" observations; and (4) *budget balanced*, or makes no net transfers to agents. A variety of techniques—including proper scoring rules, prediction markets, and peer-prediction methods—have been developed to address each of these four features. Previous mechanisms, however, have not simultaneously satisfied even the first three properties. In this paper we introduce *collective revelation*, an elicitation and aggregation mechanism that is incentive compatible, information weighted, self-verifying, and budget balanced.

By information (or confidence) weighted we informally mean that the predictions of agents with more private evidence have greater influence on the final, aggregate prediction. Collective revelation is constructed to elicit both an agent's prediction and its confidence, as quantified by its willingness to update its beliefs in light of hypothetical new evidence. To circumvent the need for ground truth observations, the agents' own private evidence is used as a proxy for public information. Ultimately, the mechanism constructs an aggregate prediction equivalent to what would have resulted had all agents collectively revealed the entirety of their private information.

In the remainder of the Introduction we describe existing prediction mechanisms and their theoretical guarantees. In Section 2 we formally specify the problem and outline our key assumptions. The basic mechanism is developed in Section 3.1 and shown to satisfy properties (1), (2) and (3) above. Section 3.2 describes a general technique for constructing budget-balanced mechanisms, and applies it to collective revelation. We conclude in Section 4 by discussing the interpretations, implications, and limitations of this work. Some results and technical details are relegated to the Appendix.

Background and Related Work

Table 1 summarizes existing prediction techniques and their properties.

Proper scoring rules [1, 7, 15, 22] assess an agent's forecast of an uncertain outcome against the actual observed outcome, with rewards designed so that agents are incentivized to be truthful. For example, for any random variable X and constant $C \in \mathbb{R}$, the Brier scoring rule rewards an agent's prediction r of $\mathbb{E}X$ according to

$$u(X, r) = C - (r - X)^2.$$

If F is the agent's subjective distribution, then

$$\arg\max_r \mathbb{E}_F[u(X, r)] = \mathbb{E}_F X.$$

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EC'09, July 6–10, 2009, Stanford, California, USA.
Copyright 2009 ACM 978-1-60558-458-4/09/07 ...$5.00.

	Incentive Compatible	Information Weighted	Self-Verifying
Proper Scoring Rules	•		
Prediction Markets	◦	•	
Peer Prediction	◦		•
Delphi Method		◦	•
Competitive Forecasting	◦	•	
Polls			•
Collective Revelation	◦	•	•

Table 1: A comparison of properties for common prediction mechanisms. Solid bullets indicate properties that are rigorously satisfied, and circles indicate those that are approximately satisfied. Collective revelation is Bayes-Nash—as opposed to dominant strategy—incentive compatible; hence the open circle.

Hence, a risk-neutral agent—one who seeks to maximize its expected reward—would rationally report its subjective expectation. Scoring rules have been extensively studied for nearly sixty years, both theoretically [27] and experimentally [19, 26], and have been applied to forecasting problems in several disciplines, including meteorology [18], economics [20], and medicine [23].

While scoring rules induce rational agents to act truthfully, they rely on observing the outcome X (i.e., they rely on an objective, ground truth). Recent work [10, 13, 17, 21] has shown that the need for external validation can be obviated by meta predictions. For example, the *Bayesian truth serum* [21] asks agents to report both their own prediction and their prediction of other agents' predictions. With the appropriate reward structure, this framework leads to truthful equilibria, even for obviously subjective assessments, such as, *Do you prefer red or white wine?*. To derive this type of result, it is assumed that agents' personal beliefs are determined by independent random events (e.g., *'heads' I prefer red wine, 'tails' I prefer white wine*) and that the distribution of these random events is common knowledge among the agents, though it need not be known by the mechanism. Peer-prediction mechanisms are incentive compatible and self-verifying.[1] Prior work along these lines, however, has implicitly assumed each individual has the same amount of private information (e.g., one coin flip), and hence has neglected an important practical aspect of aggregating predictions.

Prediction markets encourage truthful behavior [2] and automatically aggregate predictions from agents with diverse information. Consider, for example, prediction markets for U.S. presidential elections. In this case, agents buy and sell assets tied to an eventual Democratic or Republican win. Each share pays $1 if the corresponding event occurs, and is worthless if the event does not occur. Accordingly, the spot price for shares can be interpreted as the market-aggregated estimate of the likelihood of that event, and participants are encouraged to trade when they believe the market does not reflect their subjective probability estimates. Although prediction markets are information weighted, and in some cases incentive compatible [2], they rely on objective outcomes to determine the ultimate value of assets.[2] A graphical prediction mechanism called *competitive forecasting*[3] elicits confidence intervals on predictions, thereby facilitating information weighting. Like prediction markets, competitive forecasting rewards accuracy, though is not rigorously incentive compatible and relies on benchmarking against objective measurements.

Above we have focused on mechanisms that directly reward accuracy, incentivizing truthfulness. There are, however, several effective elicitation and prediction techniques that do not have this property. Simple polls, for example, do surprisingly well at aggregating opinions [24], and the *Delphi method* generates consensus predictions from experts essentially through a process of structured discussion [3].

2. THE SETTING

To have a concrete problem in mind, suppose we are trying to estimate the proportion p of students at the University of Chicago who are male—the original elicitation example used by Winkler in 1967 [25]. Specifically, if X indicates the gender of a randomly selected student, we seek an aggregate estimate of agents' subjective expectation of X. More generally, we are interested in aggregate expectations for arbitrary parametric random variables.

We make the following structural assumptions, to be specified in more detail in Section 3.

1. *Common prior.* Agents have a common prior subjective belief on the parameters describing X, and this common prior is also known to the administrator of the mechanism. For example, X may be Bernoulli(p), and agents may have a Beta(α, β) prior on p, with α and β common knowledge.

2. *Independent private evidence.* Each agent i privately observes $n_i \geq 1$ independent realizations of the random variable X: $(x_{i,1}, \ldots, x_{i,n_i})$. Observations are in-

[1]By our results in Section 3.2, these mechanisms can be made budget balanced as well.

[2]The need for external validation is partially alleviated for *conditional* markets, in which transactions are voided if certain stated pre-conditions are not met. For example, a conditional market can elicit the probability of recession conditional on interest rates being lowered within the next three months. If interest rates are not ultimately lowered, then the market has generated a probability estimate for a purely hypothetical—and not externally verifiable—event. Still, these markets effectively function because of the threat of external verification (i.e., they would not function if it were common knowledge that the pre-conditions would never be satisfied).

[3]`http://us.newsfutures.com/home/competitiveForecasting.html`

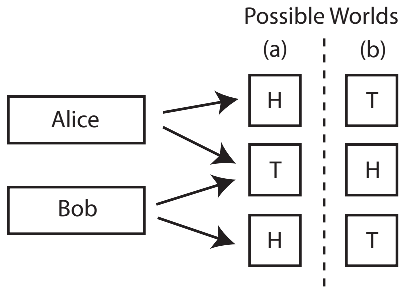

Figure 1: *The difficulties of dependent information.* Even if Alice and Bob each truthfully report their observations, it is still impossible to determine the true state of the world.

dependent across agents as well. Both the outcome vector and the number of trials n_i are private.

3. *Rationality.* Agents update their beliefs via Bayes' rule. In particular, suppose $\theta \in \Theta$ parameterizes the distribution of X, $f(\theta)$ is the density of the common prior, and $p(x \mid \theta)$ is the density of X given parameter θ. Then the posterior density of agent i after observing independent evidence $x_{i,1}, \ldots, x_{i,n_i}$ is given by

$$f(\theta \mid x_{i,1}, \ldots, x_{i,n_i}) = \frac{f(\theta) \prod_{j=1}^{n_i} p(x_{i,j} \mid \theta)}{\int f(\xi) \prod_{j=1}^{n_i} p(x_{i,j} \mid \xi) d\xi}$$

and the agent's corresponding posterior expectation of X is given by

$$\mathbb{E}_{f(\theta \mid x_{i,1}, \ldots, x_{i,n_i})}[X] = \int_{\Theta} \int_{\mathbb{R}} x\, p(x \mid \theta)\, f(\theta \mid x_{i,1}, \ldots, x_{i,n_i})\, dx\, d\theta.$$

4. *Risk neutrality.* Agents act to maximize their expected payoff.

The assumptions of rationality and risk-neutrality are standard in economic and game-theoretic work [14]. The requirement of a common prior, while stronger, is still relatively commonplace. Moreover, when agents' beliefs are largely based on their private evidence, it is often reasonable to assume a *weakly informative prior* [5]. By far the most stringent of our assumptions is the requirement of independent evidence. Without this assumption, however, it seems difficult to model how players would reason about other agents' private information, a key ingredient in peer-prediction approaches, and one that facilitates self-verification. Perhaps for this reason, previous results along these lines have also incorporated the independence assumption [10, 21].

Apart from the question of elicitation, dependent evidence also complicates the problem of aggregation. Consider the following simple example, depicted graphically in Figure 1. Three coins are flipped, and Alice and Bob each observe two of them. Even if Alice and Bob each truthfully report observing one 'heads' and one 'tails', it is still impossible to determine if the true state of the world is two 'heads' and one 'tails', or two 'tails' and one 'heads'.

Our four structural assumptions are idealizations. Nonetheless, they are common idealizations that are difficult to circumvent and that capture, at the very least, an interesting special case on which to build intuition.

3. A MECHANISM FOR COLLECTIVE REVELATION

The best estimate for $\mathbb{E}X$ would result if all agents simply revealed their private evidence; collective revelation generates precisely this ideal estimate. Moreover, the mechanism relies only on the privately observed values of X, not on any publicly verifiable observations.

To gain intuition for the mechanism, suppose that the administrator were able to generate verifiable trials of X. Using standard techniques (i.e., simple scoring rules), one could then elicit each agent's posterior subjective distribution, updated in accordance with its private information. Now, given an agent's updated distribution (and a common knowledge prior), only certain private information is possible. In particular, as we show later, this feasible private information is sufficiently well determined to allow one to weight estimates by their relative information content, in turn producing the final, aggregate estimate.

The above sketch, however, required additional, externally verifiable trials of X, violating the criterion of self-verification. To circumvent this issue, we instead score agents against one another's private information. If a single agent can be induced to truthfully disclose its private information, other agents will then be incentivized to follow suit. Consequently, there is an equilibrium in which agents are induced to collectively reveal their private evidence. A technical hurdle is that we never have direct access to agents' private evidence, even in equilibrium; nonetheless, we can infer their private information well enough to induce truthful equilibrium behavior.

The final ingredient in this mechanism is providing a means by which agents can conveniently describe their subjective distributions. The incentives are such that they should be truthful, but we still need a language for agents to convey their beliefs. Here we incorporate the *hypothetical future sample (HFS)* method [25]: Agents state their subjective expectation of X, and then revise their estimate in light of new, hypothetical evidence. The precise language used to convey distributions is not crucial; for concreteness, however, we state our results in terms of HFS.[4] See Garthwaite et al. [4] for alternative elicitation languages.

Figure 2 outlines the structure of the mechanism. First, agents report their estimate of $\mathbb{E}X$, and their updated expectation given hypothetical evidence. From these reports, the subjective posterior of the agent is reconstructed. From that posterior, the mechanism infers the agent's private information, and also the agent's prediction of other agents' private information. Finally, this prediction of other agents' private information is scored against the inferred private information of other agents, closing the loop and creating an equilibrium with truthful initial reports.

While Figure 2 captures the spirit of our argument, the technical details are sensitive to the precise distribution of the random variable X. We discuss the case of binary events (i.e., Bernoulli random variables) below, and the case of

[4]Agents could convey their subjective distributions by revealing their private evidence directly; for reasons discussed in Section 4, however, we believe HFS to be a more natural language.

normally distributed observations in the Appendix. Similar results hold for Poisson and exponentially distributed outcomes, but the details are not included here.

3.1 The Basic Mechanism

In the following, we make use of some notation and results common in probability and statistics textbooks [5].

Suppose we attempt to elicit predictions for a binary outcome (i.e., $X \sim \text{Bernoulli}(p)$). In this case, a single observation of X, even if public, is not sufficient to elicit detailed information regarding an agent's prior on p.

LEMMA 3.1. *Suppose $X_1, \ldots, X_n$ are independent, identically distributed* Bernoulli(p) *random variables, and that an agent has a subjective distribution $F(p)$ over possible values of p. Then it is possible to elicit the k^{th} moment of F via a proper scoring rule that pays based on the outcome $(X_1, \ldots, X_n)$ if and only if $k \leq n$.*

PROOF. First assume $k > n$. Let $f : \mathbb{R}^m \times \{0,1\}^n \mapsto \mathbb{R}$ be any function that pays an agent based on a report $q \in \mathbb{R}^m$ and the outcome $(X_1, \ldots, X_n) \in \{0,1\}^n$. Then for any report q, the agent's subjective expected payoff is

$$\begin{aligned}\mathbb{E}f(q, \cdot) &= \sum_{o \in \{0,1\}^n} f(q, o)\mathbb{P}(o) \\ &= \sum_{o \in \{0,1\}^n} f(q, o) \int_0^1 \xi^{|o|}(1-\xi)^{n-|o|}\, dF(\xi)\end{aligned}$$

where $|o|$ is the number of 1's in o. Consequently, denoting the i^{th} moment of F by $M_i = \int_0^1 \xi^i \, dF(\xi)$, there is a function g such that $\mathbb{E}f(q, \cdot) = g(q, M_1, \ldots, M_n)$. In particular, $\arg\max_q \mathbb{E}f(q, \cdot)$ does not in general depend on the k^{th} moment of F for $k > n$, establishing the necessity of $k \leq n$.

Now consider $k \leq n$, and define

$$f_k(q, X_1, \ldots, X_n) = -\left[q - \frac{1}{\binom{n}{k}}\binom{\sum_{i=1}^n X_i}{k}\right]^2.$$

It is clear that

$$\arg\max_q \; \mathbb{E}f_k(q, \cdot) = \frac{1}{\binom{n}{k}}\mathbb{E}\binom{\sum_{i=1}^n X_i}{k}.$$

Furthermore,

$$\begin{aligned}\mathbb{E}\binom{\sum_{i=1}^n X_i}{k} &= \mathbb{E}\left[\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} X_{i_1} \cdots X_{i_k}\right] \\ &= \binom{n}{k}\mathbb{E}(X_1 \cdots X_k) = \binom{n}{k}\int_0^1 \xi^k \, dF(\xi).\end{aligned}$$

Consequently, $\arg\max_q \mathbb{E}f_k(q, \cdot) = M_k$ and the result follows. □

Lemma 3.1 shows, in particular, that with a single Bernoulli observation, a scoring rule can only elicit the mean $\mu(F)$ of an agent's subjective distribution $F(p)$. Since the parameter p is itself the expectation of a Bernoulli(p) random variable, $\mu(F)$ is also the agent's subjective expectation of X. That is,

$$\mathbb{E}_F X = \int_0^1 \xi \, dF(\xi) = \mu(F).$$

In short, scoring against a single Bernoulli observation reveals only an agent's expectation, and not the *uncertainty* in its prediction as quantified by the variance of F. This uncertainty, however, is a necessary ingredient for weighting predictions, and so Lemma 3.1 shows that any incentive-compatible and information-weighted mechanism for aggregating predictions for Bernoulli outcomes must at least implicitly score agents against multiple observations.

Now we begin to formally describe the collective revelation mechanism for binary outcomes. Following the outline in Figure 2, we begin by deducing the agent's subjective distribution based on estimates derived from hypothetical future samples.

LEMMA 3.2. *Let Y be a* Bernoulli(p) *random variable, and suppose $F(p) \sim \text{Beta}(\alpha, \beta)$ is a prior on p. Fix $n, s \in \mathbb{Z}$ such that $n > 0$ and $0 \leq s \leq n$, and let $F(p \mid s, n)$ be the updated posterior distribution on p given s successes are observed in n hypothetical trials. Then $g_{s,n}(\alpha, \beta)$ defined by*

$$g_{s,n}(\alpha, \beta) = \left(\mathbb{E}_{F(p)}Y, \; \mathbb{E}_{F(p|s,n)}Y\right)$$

is a bijection from $\mathbb{R}_+ \times \mathbb{R}_+$ to

$$\Omega = \{(a, b) \mid 0 < a < b < s/n\} \cup \{(a, b) \mid s/n < b < a < 1\}$$

and

$$g_{s,n}^{-1}(a, b) = \left(\frac{(nb - s)a}{a - b}, \; \frac{(nb - s)(1 - a)}{a - b}\right).$$

PROOF. First note that the beta distribution is conjugate to the binomial distribution, and in particular, $F(p \mid s, n) \sim \text{Beta}(\alpha + s, \; \beta + n - s)$. Furthermore, since $\mathbb{E}[Y \mid p \sim \text{Beta}(\alpha, \beta)] = \alpha/(\alpha + \beta)$, we have

$$g_{s,n}(\alpha, \beta) = \left(\frac{\alpha}{\alpha + \beta}, \; \frac{\alpha + s}{\alpha + \beta + n}\right).$$

Consequently, for fixed a, b we need to solve the system of equations

$$a = \frac{\alpha}{\alpha + \beta} \qquad b = \frac{\alpha + s}{\alpha + \beta + n}$$

for α, β. The first equation yields $\alpha = \beta a/(1 - a)$. Substituting into the second, we get

$$\beta = (nb - s)(1 - a)/(a - b).$$

Using this expression to solve for α, we have the equation for the $g_{s,n}^{-1}$.

Now note that $(a, b) \in \Omega$ if and only if $0 < a < 1$ and $(nb - s)/(a - b) > 0$. Consequently, $g_{s,n}^{-1}(\Omega) \subseteq \mathbb{R}_+ \times \mathbb{R}_+$. Furthermore, for $(\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_+$ and $g_{s,n}(\alpha, \beta) = (a, b)$, we clearly have $0 < a < 1$. So, if $(a, b) \notin \Omega$, then $(nb - s)/(a - b) < 0$, in which case $(\alpha, \beta) = g_{s,n}^{-1}(a, b) \notin \mathbb{R}_+ \times \mathbb{R}_+$, leading to a contradiction. Hence, $g_{s,n}$ is a bijection between the stated sets.

□

Next we derive a one-to-one correspondence between an agent's subjective beta distribution and predictions regarding other agents' private information.

LEMMA 3.3. *For $n \geq 2$, suppose $X_1, \ldots, X_n$ are independent* Bernoulli(p) *random variables, and $F \sim \text{Beta}(\alpha, \beta)$ is a prior on p. Let $S = \sum_{i=1}^n X_i$. Then $h_n(\alpha, \beta)$ defined by*

$$h_n(\alpha, \beta) = \left(\mathbb{E}_{F(p)}\left[\frac{S}{n}\right], \; \mathbb{E}_{F(p)}\left[\frac{1}{\binom{n}{2}}\binom{S}{2}\right]\right)$$

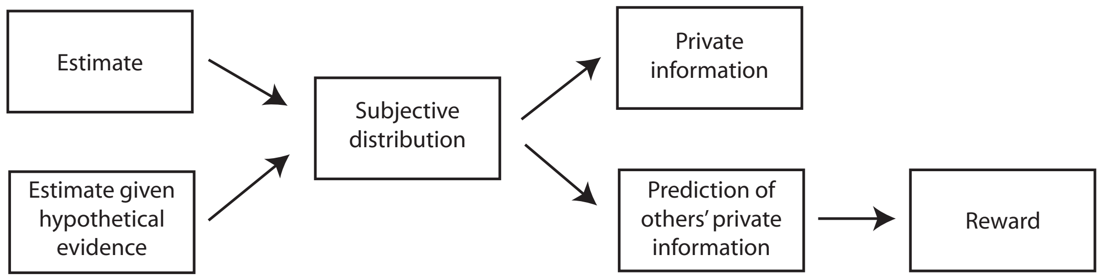

Figure 2: Inducing the collective revelation of private information.

is a bijection from $\mathbb{R}_+ \times \mathbb{R}_+$ *to*

$$\Omega = \{(a, b) : 0 < b < a < \sqrt{b} < 1\}$$

and

$$h_n^{-1}(a, b) = \left(\frac{a(a-b)}{b-a^2},\ \frac{(1-a)(a-b)}{b-a^2} \right).$$

Proof. First, observe that

$$\mathbb{E}_{F(p)}\left[S/n\right] = \mathbb{E}_{F(p)}\left[X_1\right] = \int_0^1 \xi \, dF(\xi) = \frac{\alpha}{\alpha+\beta}$$

and

$$\begin{aligned}\mathbb{E}_{F(p)}\left[\frac{1}{\binom{n}{2}}\binom{S}{2}\right] &= \frac{1}{\binom{n}{2}}\mathbb{E}_{F(p)}\left[\sum_{i<j} X_i X_j\right] \\ &= \mathbb{E}_{F(p)}[X_1 X_2] = \int_0^1 \xi^2 \, dF(\xi) \\ &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}.\end{aligned}$$

Consequently,

$$h_n(\alpha, \beta) = \left(\frac{\alpha}{\alpha+\beta},\ \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \right). \tag{3.1}$$

To compute h_n^{-1}, we need to solve for α, β in

$$a = \frac{\alpha}{\alpha+\beta} \qquad b = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}.$$

The first equation gives $\alpha = a\beta/(1-a)$. Substituting into the second equation, we find $\beta = (x-1)(y-x)/(y-x^2)$. Using this latter expression, we now solve for α, yielding the result.

For $(\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_+$, let $h(\alpha, \beta) = (a, b)$. From (3.1), it is clear that $0 < b < a$ and that $\sqrt{b} < 1$. To show $a^2 < b$, note that for $x, y > 0$, $x/y < (x+1)/(y+1)$ if and only if $x < y$. Consequently, we have that $\text{Image}(h_n) \subseteq \Omega$. Furthermore, for $(a, b) \in \Omega$, it is clear that $h_n^{-1} \subseteq \mathbb{R}_+ \times \mathbb{R}_+$. Hence, h_n is a bijection on the stated sets. □

We are ready to state our main result for predictions about binary outcomes.

Theorem 3.1. *Consider the setting of the Bayesian game in Section 2 with $N \geq 3$ players, $X \sim$ Bernoulli(p), and a* Beta(α_0, β_0) *common prior on the parameter p. Let $F_i(p)$ denote the posterior distribution of agent i after updating according to its private information. Fix $n, s \in \mathbb{Z}$ such that $n > 0$ and $0 \leq s \leq n$. Suppose each agent i plays an action $(a_i, b_i) \in \Omega$ where*

$$\Omega = \{(a, b) \mid 0 < a < b < s/n\} \cup \{(a, b) \mid s/n < b < a < 1\}.$$

Using the notation of Lemmas 3.2 and 3.3, define

$$\begin{aligned} s_{-i} &= \sum_{j \neq i} \left[\mathcal{P}_1 \circ g_{s,n}^{-1}(a_j, b_j) - \alpha_0\right] \\ n_{-i} &= \sum_{j \neq i} \left[\mathcal{P}_1 \circ g_{s,n}^{-1}(a_j, b_j) - \alpha_0\right] + \left[\mathcal{P}_2 \circ g_{s,n}^{-1}(a_j, b_j) - \beta_0\right] \end{aligned} \tag{3.2}$$

where $\mathcal{P}_k$ is projection onto the k^{th} component. For arbitrary constants C_i, set the reward to the i^{th} agent to be

$$\begin{aligned} C_i &- \left[\mathcal{P}_1 \circ h_{n_{-i}} \circ g_{s,n}^{-1}(a_i, b_i) - \frac{s_{-i}}{n_{-i}}\right]^2 \\ &- \left[\mathcal{P}_2 \circ h_{n_{-i}} \circ g_{s,n}^{-1}(a_i, b_i) - \frac{1}{\binom{n_{-i}}{2}}\binom{s_{-i}}{2}\right]^2. \end{aligned} \tag{3.3}$$

Then

$$(a_i, b_i) = \left(\mathbb{E}_{F_i(p)} X,\ \mathbb{E}_{F_i(p|s,n)} X\right) \tag{3.4}$$

is a strict Nash equilibrium.

Proof. Fix attention on agent i, and suppose that for all $j \neq i$, agent j plays the strategy (3.4). Since the beta distribution is conjugate to the binomial distribution, each agent has a beta posterior F_j after updating according to their private information. By Lemma 3.2, $g_{s,n}^{-1}(a_j, b_j) = (\alpha_j, \beta_j)$ gives the parameters of F_j. Suppose s_j, n_j are, respectively, the number of successes and total number of trials privately observed by agent j. Then

$$\alpha_j = \alpha_0 + s_j \qquad \beta_j = \beta_0 + n_j - s_j.$$

Consequently,

$$s_j = \alpha_j - \alpha_0 = \mathcal{P}_1 \circ g_{s,n}^{-1}(a_j, b_j) - \alpha_0$$

and

$$\begin{aligned} n_j &= (\beta_j - \beta_0) + (\alpha_j - \alpha_0) \\ &= \left[\mathcal{P}_1 \circ g_{s,n}^{-1}(a_j, b_j) - \alpha_0\right] + \left[\mathcal{P}_2 \circ g_{s,n}^{-1}(a_j, b_j) - \beta_0\right]. \end{aligned}$$

In particular, $s_{-i} = \sum_{j \neq i} s_j$ and $n_{-i} = \sum_{j \neq i} n_j$. Lemmas 3.2 and 3.3 show that if i plays according to (3.4), then

$$h_{n_{-i}} \circ g_{s,n}^{-1}(a_i, b_i) = \left(\mathbb{E}_{F_i(p)}\left[\frac{S}{n_{-i}}\right],\ \mathbb{E}_{F_i(p)}\left[\frac{1}{\binom{n_{-i}}{2}}\binom{S}{2}\right] \right)$$

where $S = \sum_{k=1}^{n-i} X_k$ for independent Bernoulli(p) random variables X_k. Since we are using the Brier proper scoring rule,[5] this strategy maximizes i's expected reward. Moreover, since $h_{n-i} \circ g_{s,n}^{-1}$ is an injection, this is i's unique best response. Consequently, strategy (3.4) is a strict Nash equilibrium. $\square$

Theorem 3.1 constructs a game in which agents truthfully report their subjective expectations of X, and their updated expectations given hypothetical evidence. To complete the mechanism, Corollary 3.1 describes how to generate an aggregate, information-weighted prediction from these individual reports.

COROLLARY 3.1. *Using the notation of Theorem 3.1, suppose agents play the equilibrium strategy (a_i, b_i) given by (3.4). For*

$$\bar{s} = \sum_{j=1}^{N} \left[\mathcal{P}_1 \circ g_{s,n}^{-1}(a_j, b_j) - \alpha_0\right]$$

$$\bar{n} = \sum_{j=1}^{N} \left[\mathcal{P}_1 \circ g_{s,n}^{-1}(a_j, b_j) - \alpha_0\right] + \left[\mathcal{P}_2 \circ g_{s,n}^{-1}(a_j, b_j) - \beta_0\right]$$

define the aggregate, information-weighted prediction $\tilde{p}$ by

$$\tilde{p} = \frac{\alpha_0 + \bar{s}}{\alpha_0 + \beta_0 + \bar{n}}. \tag{3.5}$$

Let $F(p \mid \{X_{i,j}\})$ be the posterior distribution resulting from the cumulative private evidence of all agents. Then $\tilde{p} = \mathbb{E}_F X$.

PROOF. As shown in the proof of Theorem 3.1, $\bar{s}, \bar{n}$ are, respectively, the total number of successes and the total number of trials among all the agents' private evidence. The posterior distribution $F(p \mid \{X_{i,j}\})$ resulting from this cumulative evidence is thus Beta$(\alpha_0 + \bar{s}, \beta_0 + \bar{n} - \bar{s})$, and the expectation of X under this posterior distribution is precisely $\tilde{p}$. $\square$

Collective revelation essentially induces agents to reveal all their private information. More precisely, the aggregate prediction generated by the mechanism is the same prediction that would have resulted had all agents disclosed their private evidence in its entirety.

3.2 A General Technique for Balancing Budgets

The basic mechanism, as thus far described, is incentive compatible, information weighted, and self-verifying. To achieve budget balance—where no net transfers are made to agents—we develop a general technique that applies both to collective revelation and to past peer-prediction mechanisms (e.g., the Bayesian truth serum [21]).

The idea is to reward agents via a *shared scoring rule* [9, 11, 12]. Suppose there are n agents, and agent i is rewarded by the scoring rule $f_i(r_i, x_i)$ where r_i is the agent's report and x_i the observation the agent is evaluated against. In its simplest form, a shared scoring rule is one in which agents are paid according to their performance relative to one another:

$$\begin{aligned}\tilde{f}_i(r_i, x_i) &= f_i(r_i, x_i) - \frac{1}{n}\sum_{j=1}^{n} f_j(r_j, x_j) \\ &= \frac{n-1}{n} f_i(r_i, x_i) - \frac{1}{n}\sum_{j \neq i} f_j(r_j, x_j). \end{aligned} \tag{3.6}$$

By subtracting the mean reward under the original scoring rule f_j, the new rewards are clearly budget balanced: $\sum_{i=1}^{n} \tilde{f}_i(r_i, x_i) = 0$.

Ordinarily, the observations x_i correspond to an "objective truth" (e.g., the amount of rain tomorrow, or the winner of an election). In these cases, it is reasonable to assume that agent i's report r_i does not affect the original rewards of the other agents, and specifically, that r_i does not affect the sum in Equation (3.6). Consequently,

$$\arg\max_r \tilde{f}_i(r, x_i) = \arg\max_r f_i(r, x_i)$$

and in particular, assuming $\{x_i\}$ are "ground-truth" observations, shared scoring rules preserve incentive compatibility.

In our setting, however, the "observations" x_j are determined precisely by agents' reports, and a simple application of shared scoring rules consequently fails. To circumvent this complication, we effectively decouple *scoring* (computing the original rewards f_j) from *benchmarking* (normalizing rewards to group performance). That is, agent i is rewarded according to its performance relative to a set of agents V_i, but that set of agents V_i is chosen so that agent i cannot affect their scores.

We now formally describe how to balance the budget of a mechanism that naturally restricts to fewer players. For a game G with n players, the reward function is not a priori defined for fewer than the original number of players. A *k-projective family of G* is essentially a family of reward functions that makes sense when applied to $k < n$ players, but otherwise does not alter the structure of the game.

DEFINITION 3.1. *Given a Bayesian game G with n players, a k-projective family of G is a family of games $\{\hat{G}_V\}$ such that for each $V \subseteq \{1, \ldots, n\}$ with $|V| = k$, $\hat{G}_V$ is a Bayesian game restricted to the players V that preserves (for players in V) type spaces, action spaces, and players' beliefs about types. In particular, $\{\hat{G}_V\}$ is determined by a family of reward functions $\{\hat{U}_V\}$ that specifies the expected reward for each player in V resulting from any given strategy profile of those players.*

As a simple example of this definition, consider a first-price sealed-bid auction with n players where types (i.e., private valuations) are independent, identically distributed uniform $[0, 1]$ random variables. Then for any subset of k players $V = \{j_1, \ldots, j_k\}$, define $\hat{G}_V$ to be the usual first-price auction on k players (again with *i.i.d.* $U[0, 1]$ valuations).

THEOREM 3.2. *Consider a Bayesian game G on n players with a (strict) Nash equilibrium $q = (q_1, \ldots, q_n)$. Suppose $\{\hat{G}_V\}$ is a k-projective family of G such that $2 \leq k < n/2+1$. Also suppose that for each $V = \{j_1, \ldots, j_k\}$, $(q_{j_1}, \ldots, q_{j_k})$ is a (strict) Nash equilibrium for $\hat{G}_V$. Then for any constant $C \in \mathbb{R}$, there are player rewards $\tilde{U}_i$ for the n-player game G such that:*

[5] Any proper scoring rule could be used with minimal alteration.

1. *For any strategy profile* s, $\sum_{i=1}^{n} \tilde{U}_i(s) = C$.
2. *The original equilibrium* q *is still a (strict) Nash equilibrium for the modified game (i.e., the game* G *with rewards* $\tilde{U}$*).*

In particular, by setting $C = 0$*, one can alter the rewards so that the game* G *is ex-post strongly budget balanced.*

PROOF. The proof is by construction. For $V \ni i$, let U_i and $\hat{U}_{V,i}$ denote, respectively, the reward functions for player i in the game G and in the game $\hat{G}_V$. For a strategy profile s for G, we slightly abuse notation and write $\hat{U}_{V,i}(s)$ for the reward to player i in the restricted game $\hat{G}_V$. (Technically, $\hat{U}_{V,i}$ is only defined for strategy profiles of the set of players V, whereas s is a complete strategy profile for all n players. However, by $\hat{U}_{V,i}(s)$ we mean to first restrict s to players in V, and then to compute the reward for player i.)

Define the player sets $V_i = \{i, i+1, \ldots, i+k-1\}$ where the players wrap around for $i+k-1 > n$. For example, $V_{n-1} = \{n-1, n, 1, \ldots, k-2\}$. Consider the modified rewards for G:

$$\tilde{U}_i(s) = \frac{C}{n} + U_{V_i,i}(s) - \frac{1}{k-1} \sum_{j=1}^{k-1} U_{V_{i+j},i+j}(s).$$

Now, switching the order of summation

$$\sum_{i=1}^{n} \tilde{U}_i(s) = C + \sum_{i=1}^{n} U_{V_i,i}(s) - \frac{1}{k-1} \sum_{j=1}^{k-1} \sum_{i=1}^{n} U_{V_i,i}(s) = C.$$

Furthermore, observe that

$$\bigcup_{j=1}^{k-1} V_{i+j} = \{i, i+1, \ldots, i+2(k-1)\}.$$

Since $k < n/2+1$, we have $2(k-1) < n$, and so $i \notin V_{i+j}$ for $1 \le j \le k-1$. Consequently, if s_i is the strategy of player i and s_{-i} is the strategy profile of all the other players, then

$$\begin{aligned}\tilde{U}_i(s_i; s_{-i}) &= \frac{C}{n} + U_{V_i,i}(s_i; s_{-i}) - \frac{1}{k-1} \sum_{j=1}^{k-1} U_{V_{i+j},i+j}(s_{-i}) \\ &= U_{V_i,i}(s_i; s_{-i}) + C_i(s_{-i})\end{aligned}$$

where $C_i()$ is a function that does not depend on s_i (i.e., $C_i()$ does not depend on player i's strategy). Since (r_i, r_{-i}) is a Nash equilibrium for $\hat{G}_{V_i}$,

$$\arg\max_{s_i} \tilde{U}_i(s_i; r_{-i}) = \arg\max_{s_i} U_{V_i,i}(s_i; r_{-i}) = r_i.$$

The result now follows. □

The assumptions of Theorem 3.2 hold for the basic collective revelation mechanism of Section 3.1: In equilibrium, players truthfully report both their expectation of X, and their updated expectation in light of hypothetical evidence. Since our Bernoulli collective revelation mechanism requires at least 3 players, there is a natural k-projective family for any $k \ge 3$. It then follows from the theorem that there is a budget-balanced extension of Bernoulli collective revelation for $n \ge 5$ players.

A slightly different construction yields a budget-balanced mechanism for $n = 4$ players. Informally, the agents are split into two groups of two, benchmarking each against its own group but scoring each against the other group. For normally distributed observations (see the Appendix) the basic collective revelation mechanism requires only $k \ge 2$ players. Consequently, in that case, Theorem 3.2 shows how to construct the budget-balanced extension for $n \ge 3$ players.

The constant C in Theorem 3.2 can be seen as a subsidy provided by the mechanism administrator to encourage participation. In particular, positive subsidies make it rational for risk-neutral agents with common priors to compete, circumventing so-called *no-trade theorems* [16] that rule out such behavior in zero-sum games.

4. DISCUSSION

Collective revelation elicits both individual predictions and optimal aggregate estimates, even when agents hold a priori unknown amounts of private information. Moreover, the mechanism can be adapted to several of the most common classes of observations, including Bernoulli and normal outcomes. Together, these features extend past prediction methods. One tradeoff for this versatility, however, is that the mechanism requires as input the common prior of agents; the *Bayesian truth serum*, in contrast, assumes only that there is a common prior, not that this common prior is known to the mechanism. While we see this work as an important first step, it would certainly be interesting if one could reduce—or perhaps even eliminate—the need for specific distributional knowledge.

To further explain the name of our mechanism, we note that agents reveal their subjective distributions essentially because they are rewarded based on predictions that the mechanism makes on their behalf. Specifically, the "predictions of other's private information" step in Figure 2 is done by proxy. Agents are hence incentivized to initially report their subjective distributions truthfully so that these proxy predictions accurately reflect their beliefs. In this sense, collective revelation borrows insight from the proof of the classic *revelation principle* [6, 8], in which a similar proxy argument is used to transform any mechanism with a Bayes-Nash equilibrium to a direct mechanism in which agents truthfully report their types.

As described in Section 3, the initial querying of agents' subjective distributions is accomplished via *hypothetical future samples* (HFS). Agents make a prediction, are given additional, hypothetical evidence, and then revise their prediction in light of this new evidence. Although HFS is a particularly intuitive mechanism for eliciting distributions, it is by no means the only one, and Garthwaite et al. [4] describe several alternatives that are easily incorporated into our mechanism.[6] Ultimately, however, we do assume that agents are sufficiently introspective to accurately impart their beliefs. This assumption of self-awareness is standard in the economic literature. Nonetheless, in practice agents may simply not be able to accurately communicate their beliefs, even if the incentives are correct. Consequently, it would be of significant interest to develop aggregation schemes that do not require such fine-grained information. In particular, it is conceivable that predictions of relative performance (e.g.,

[6]They are primarily interested in situations were cooperative participants would like to convey their information to researchers. In these cases, the question is not one of truthfulness, but rather of ease and accuracy.

I expect to outperform p% of other agents.) may have cognitive advantages, and could lead to more widely applicable, yet still rigorous, mechanisms for belief aggregation.

5. REFERENCES

[1] G. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78:1–3, 1950.

[2] Yiling Chen, Stanko Dimitrov, Rahul Sami, Daniel M. Reeves, David M. Pennock, Robin D. Hanson, Lance Fortnow, and Rica Gonen. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica*, to appear.

[3] Norman Dalkey and Olaf Helmer. An experimental application of the Delphi method to the use of experts. *Management Science*, 9, 1963.

[4] Paul H. Garthwaite, Joseph B. Kadane, and Anthony O'Hagan. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):680–700, 2005.

[5] Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman & Hall, 2nd edition, 2004.

[6] Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41:587–601, 1973.

[7] I.J. Good. Rational decisions. *JRSS-B*, 1952.

[8] Jerry R. Green and Jean-Jacques Laffont. *Incentives in Public Decision-Making*. North-Holland Publishing Company, 1979.

[9] D.J. Johnstone. The parimutuel Kelly probability scoring rule. *Decision Analysis*, 4(2):66, 2007.

[10] Radu Jurca and Boi Faltings. Incentives for expressing opinions in online polls. In *ACM Conference on Electronic Commerce*, 2008.

[11] D.M. Kilgour and Y. Gerchak. Elicitation of probabilities using competitive scoring rules. *Decision Analysis*, 1(2):108–113, 2004.

[12] Nicolas Lambert, John Langford, Jennifer Wortman, Yiling Chen, Daniel Reeves, Yoav Shoham, and David Pennock. Self-financed wagering mechanisms for forecasting. In *ACM Conference on Electronic Commerce*, pages 170–179, 2008.

[13] Nicolas Lambert and Yoav Shoham. Truthful surveys. *Lecture Notes in Computer Science*, 5385:154–165, 2008.

[14] Andrew Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomics Theory*. Oxford University Press, New York, NY, 1995.

[15] James E. Matheson and Robert L. Winkler. Scoring rules for continuous probability distributions. *Management Science*, 22(10):1087–1096, 1976.

[16] Paul Milgrom and Nancy L. Stokey. Information, trade and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.

[17] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Management Science*, 51:1359–1373, 2005.

[18] A. Murphy and R. Winkler. Probability forecasting in meteorology. *Journal of the American Statistical Association*, 79(387):489–500, 1984.

[19] R. Nelson and D. Bessler. Subjective probabilities and scoring rules: Experimental evidence. *Journal of Agricultural Economics*, 71(2):363–369, 1989.

[20] F. O'Carroll. Subjective probabilities and short-term economic forecasts: An empirical investigation. *Applied Statistics*, 26(3):269–278, 1977.

[21] Drazen Prelec. A Bayesian truth serum for subjective data. *Science*, 306:462–466, 2004.

[22] Leonard J. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

[23] D. Spiegelhalter. Probabilistic prediction in patient management and clinical trials. *Stat. Med.*, 5:421–433, 1986.

[24] James Surowiecki. *The Wisdom of Crowds*. Doubleday, 2004.

[25] Robert L. Winkler. The assessment of prior distributions in Bayesian analysis. *Journal of the American Statistical Association*, 62(319):776–800, 1967.

[26] Robert L. Winkler. Probabilistic prediction: Some experimental results. *Journal of the American Statistical Association*, 66(336):675–685, 1971.

[27] Robert L. Winkler, J. Muñoz, J.L. Cervera, J.M. Bernardo, G. Blattenberger, J.B. Kadane, D.V. Lindley, A.H. Murphy, R.M. Oliver, and D. Ríos-Insua. Scoring rules and the evaluation of probabilities. *TEST/Springer*, 5(1):1–60, 1996.

APPENDIX

A. THE CASE OF NORMALLY DISTRIBUTED OBSERVATIONS

In the main text we assume that observations are binary events (i.e., Bernoulli random variables). Here we develop the collective revelation mechanism for normally distributed observations. The corresponding mechanisms for other distributions (e.g., Poisson and exponential) can be derived similarly. In all these cases, the structure of our arguments is identical to that for Bernoulli observations (Section 3.1); the details, however, change sufficiently that generalizing to completely arbitrary distributions seems difficult.

Here we state analogs of Lemma 3.2, Lemma 3.3, Theorem 3.1 and Corollary 3.1. These results together establish the basic collective revelation mechanism, which is incentive compatible, information weighted, and self-verifying. Budget balance follows from the results of Section 3.2.

Lemma A.1. *For fixed μ, σ^2 let Y be a $N(\mu, \sigma^2)$ random variable, and suppose $F(\mu) \sim N(\hat{\mu}, \hat{\sigma}^2)$ is a prior on μ. Fix $n \in \mathbb{Z}_+$, $s \in \mathbb{R}$ and let $F(\mu \mid s, n)$ be the updated posterior distribution on μ given that on n hypothetical trials, $\sum_{i=1}^n X_i = s$ where X_i are independent with the same distribution as Y. Then*

$$g_{s,n}(\hat{\mu}, \hat{\sigma}^2) = \left(\mathbb{E}_{F(\mu)}Y,\ \mathbb{E}_{F(\mu|s,n)}Y\right)$$

is a bijection from $\mathbb{R} \times \mathbb{R}_+$ to

$$\Omega = \{(a, b) \mid a < b < s/n\} \cup \{(a, b) \mid s/n < b < a\}$$

and

$$g_{s,n}^{-1}(a, b) = \left(a,\ \frac{\sigma^2(b-a)}{s-nb}\right).$$

PROOF. *The result follows from a straightforward computation. First note that the normal distribution is self-conjugate, and in particular,*

$$F(\mu \mid s, n) \sim N\left(\frac{\hat{\mu}/\hat{\sigma}^2 + s/\sigma^2}{1/\hat{\sigma}^2 + n/\sigma^2},\ \frac{1}{1/\hat{\sigma}^2 + n/\sigma^2}\right).$$

Furthermore, since $\mathbb{E}[Y \mid \mu \sim N(\hat{\mu}, \hat{\sigma}^2)] = \hat{\mu}$, we have

$$g_{s,n}(\hat{\mu}, \hat{\sigma}^2) = \left(\hat{\mu},\ \frac{\hat{\mu}/\hat{\sigma}^2 + s/\sigma^2}{1/\hat{\sigma}^2 + n/\sigma^2}\right).$$

Consequently, for fixed a, b we need to solve the system of equations

$$a = \hat{\mu} \qquad b = \frac{\hat{\mu}/\hat{\sigma}^2 + s/\sigma^2}{1/\hat{\sigma}^2 + n/\sigma^2}$$

for $\hat{\mu}, \hat{\sigma}^2$. An elementary computation yields the expression for the $g_{s,n}^{-1}$. Finally, note that $g_{s,n}^{-1}(a, b) \in \mathbb{R} \times \mathbb{R}_+$ if and only if $(b-a)/(s-nb) > 0$ (i.e., if and only if $(a, b) \in \Omega$). $\square$

LEMMA A.2. *Suppose $X_1, \ldots, X_n$ are independent $N(\mu, \sigma^2)$ random variables, and $F \sim N(\hat{\mu}, \hat{\sigma}^2)$ is a prior on μ. Let $S = \sum_{i=1}^n X_i$. Then $h_{n,\sigma^2}(\hat{\mu}, \hat{\sigma}^2)$ defined by*

$$h_{n,\sigma^2}(\hat{\mu}, \hat{\sigma}^2) = \left(\mathbb{E}_{F(\mu)}\left[\frac{S}{n}\right],\ \mathbb{E}_{F(\mu)}\left[\frac{S^2}{n^2}\right]\right)$$

is a bijection from $\mathbb{R} \times \mathbb{R}_+$ to

$$\Omega = \{(a, b) : b > a^2 + \sigma^2/n\}$$

and

$$h_{n,\sigma^2}^{-1}(a, b) = \left(a,\ b - a^2 - \frac{\sigma^2}{n}\right).$$

PROOF. First, observe that

$$\mathbb{E}_{F(\mu)}[S/n] = \mathbb{E}_{F(\mu)}[X_i] = \int_{\mathbb{R}} \xi \, dF(\xi) = \hat{\mu}.$$

Furthermore, since a linear combination of independent normal random variables is also normal, $S/n \sim N(\mu, \sigma^2/n)$, we have

$$\mathbb{E}_{F(\mu)}\left[\frac{S^2}{n^2}\right] = \int_{\mathbb{R}} \left(\frac{\sigma^2}{n} + \xi^2\right) dF(\xi) = \frac{\sigma^2}{n} + \hat{\sigma}^2 + \hat{\mu}^2.$$

Consequently,

$$h_{n,\sigma^2}(\hat{\mu}, \hat{\sigma}^2) = \left(\hat{\mu},\ \frac{\sigma^2}{n} + \hat{\sigma}^2 + \hat{\mu}^2\right). \tag{A.1}$$

To compute h_{n,σ^2}^{-1}, we need to solve for $\hat{\mu}, \hat{\sigma}^2$ in

$$a = \hat{\mu} \qquad b = \frac{\sigma^2}{n} + \hat{\sigma}^2 + \hat{\mu}^2.$$

An elementary computation gives the expression for h_{n,σ^2}.

Finally, for $(a, b) \in \mathbb{R} \times \mathbb{R}$, it is clear that $h_{n,\sigma^2}^{-1}(a, b) \in \mathbb{R} \times \mathbb{R}_+$ if and only if $(a, b) \in \Omega$. $\square$

THEOREM A.1. *Consider the setting of the Bayesian game in Section 2 with $N \geq 2$ players, $X \sim N(\mu, \sigma^2)$, and a $N(\hat{\mu}_0, \hat{\sigma}_0^2)$ common prior on μ (σ^2 is common knowledge). Let $F_i(\mu)$ denote the posterior distribution of agent i after updating according to its private information. Fix $n \in \mathbb{Z}_+$, $s \in \mathbb{R}$. Suppose each agent i plays an action $(a_i, b_i) \in \Omega$ where*

$$\Omega = \{(a, b) \mid a < b < s/n\} \cup \{(a, b) \mid s/n < b < a\}.$$

Using the notation of Lemmas A.1 and A.2, define

$$s_{-i} = \sigma^2 \sum_{j \neq i} \left[\frac{\mathcal{P}_1 \circ g_{s,n}^{-1}(a_j, b_j)}{\mathcal{P}_2 \circ g_{s,n}^{-1}(a_j, b_j)} - \frac{\hat{\mu}_0}{\hat{\sigma}_0^2}\right]$$

$$n_{-i} = \frac{1}{\sigma^2} \sum_{j \neq i} \left[\frac{1}{\mathcal{P}_2 \circ g_{s,n}^{-1}(a_j, b_j)} - \frac{1}{\hat{\sigma}_0^2}\right]$$

where $\mathcal{P}_k$ is projection onto the k^{th} component. For arbitrary constants C_i, set the reward to the i^{th} agent to be

$$C_i - \left[\mathcal{P}_1 \circ h_{n_{-i},\sigma^2} \circ g_{s,n}^{-1}(a_i, b_i) - \frac{s_{-i}}{n_{-i}}\right]^2 - \left[\mathcal{P}_2 \circ h_{n_{-i},\sigma^2} \circ g_{s,n}^{-1}(a_i, b_i) - \left(\frac{s_{-i}}{n_{-i}}\right)^2\right]^2.$$

Then

$$(a_i, b_i) = \left(\mathbb{E}_{F_i(\mu)} X,\ \mathbb{E}_{F_i(\mu \mid s,n)} X\right) \tag{A.2}$$

is a strict Nash equilibrium.

PROOF. The proof is identical in structure to that of Theorem 3.1. What remains to be shown is that

$$s_{-i} = \sum_{j \neq i} \sum_{k=1}^{n_j} X_{j,k}$$

and that $n_{-i} = \sum_{j \neq i} n_j$. To see this, suppose that the posterior distribution of agent i on μ is $N(\hat{\mu}_i, \hat{\sigma}_i^2)$, and that $s_i = \sum_{j=1}^{n_i} X_j$. From the posterior hyperparameters $\hat{\mu}_i$ and $\hat{\sigma}_i$, we would like to infer n_i and s_i. Now, by the Bayesian update rule, we have

$$\hat{\mu}_i = \frac{\hat{\mu}_0/\hat{\sigma}_0^2 + s_i/\sigma^2}{1/\hat{\sigma}_0^2 + n_i/\sigma^2} \qquad \hat{\sigma}_i^2 = \frac{1}{1/\hat{\sigma}_0^2 + n/\sigma^2}.$$

An elementary computation shows that

$$n_i = \frac{1}{\sigma^2 \hat{\sigma}_i^2} - \frac{1}{\hat{\sigma}_0^2 \sigma^2} \qquad s_i = \frac{\sigma^2 \hat{\mu}_i}{\hat{\sigma}_i^2} - \frac{\sigma^2 \hat{\mu}_0}{\hat{\sigma}_0^2}.$$

Finally, in the equilibrium (A.2), $g_{s,n}^{-1}(a_i, b_i) = (\hat{\mu}_i, \hat{\sigma}_i^2)$ by Lemma A.1, and so the result follows. $\square$

COROLLARY A.1. *Using the notation of Theorem A.1, suppose agents play the equilibrium strategy (a_i, b_i) given by* (A.2). *For*

$$\bar{s} = \sigma^2 \sum_{i=1}^{N} \left[\frac{\mathcal{P}_1 \circ g_{s,n}^{-1}(a_i, b_i)}{\mathcal{P}_2 \circ g_{s,n}^{-1}(a_i, b_i)} - \frac{\hat{\mu}_0}{\hat{\sigma}_0^2}\right]$$

$$\bar{n} = \frac{1}{\sigma^2} \sum_{i=1}^{N} \left[\frac{1}{\mathcal{P}_2 \circ g_{s,n}^{-1}(a_i, b_i)} - \frac{1}{\hat{\sigma}_0^2}\right]$$

define the aggregate, confidence-weighted prediction $\tilde{\mu}$ by

$$\tilde{\mu} = \frac{\hat{\mu}_0/\hat{\sigma}_0^2 + \bar{s}/\sigma^2}{1/\hat{\sigma}_0^2 + \bar{n}/\sigma^2}.$$

Let $F(\mu \mid \{X_{i,j}\})$ be the posterior distribution resulting from the cumulative private evidence of all agents. Then $\tilde{\mu} = \mathbb{E}_F X$.

PROOF. As shown in the proof of Theorem A.1, $\bar{n}$ is the total number of private trials among all players, and $\bar{s} = \sum_{i,j} X_{i,j}$. The estimate $\tilde{\mu}$ is then obtained by Bayesian updating. $\square$