"PUTTING YOUR MONEY WHERE YOUR MOUTH IS" – INFORMATION ELICITATION MECHANISM DESIGN WHEN AGENTS PRIVATELY KNOW THEIR QUALITY OF INFORMATION

FANG FANG[†], MAXWELL STINCHCOMBE[‡], ANDREW WHINSTON[§]

submitted to **The Workshop in Information Systems and Economics 2004** [1]
(Confidential, For Review Purpose Only!)

1. Introduction

Aggregating dispersed information can produce extremely accurate predictions. When organizational structures are decentralized and/or the information needed to generate good forecasts is highly dispersed, the transaction and agency costs of collecting information increase significantly. The increased speed of modern information technologies opens the possibility of designing more efficient information elicitation and aggregation systems.

People differ in the quality of their information. These differences arise from different experiences or access to information, and the differences are not generally observable. However, people who know they have more accurate information are generally willing to bet more on it, they are willing to put their money where their mouth is. We exploit this regularity to design a mechanism that generates reliable predictions by efficiently quality-weighting each person's information.

The early research on information elicitation mechanism design only focuses on eliciting a single agent's forecasts of a future event (see Savage 1971 and the references cited there). More recent work on eliciting the forecasts from several agents has assumed that the qualities of forecasts are equal (e.g. Chen, Fine and Huberman 2001 and 2004). Weighting observations from different sources according to their reliability has a long history in statistics, *viz.* the heteroskedasticity corrections in any good regression textbook. In the foundational work on information elicitation, Savage (1971) mentions the problem of eliciting and combining several agents' information. He suggests assigning weight to each agent's opinion according to their past experience or, in some other fashion, to "give each the weight you think appropriate." In this research, we systematically weight the opinions as a function of the amount of money the agents are willing to put on what they say.

Market mechanisms provide a method of "putting your money where your mouth is." There is a body of research emphasizing the prediction abilities of futures markets (see e.g. the survey by Wolfers and Zitzewitz, 2003). However, it is not fully understood how reliable such markets are, nor what determines their reliability.[2] This uncertainty limits the use of such market prediction outcomes to support business decision process.

There are two, related theoretical doubts about the reliability of prediction markets. The first is a conceptual problem pointed out by Grossman and Stiglitz (1980). If the market price in a prediction is a good indicator of the future event, as it would be if it aggregated all available information, then no rational person would have an incentive to take part in the market since their private information is almost certainly less accurate than the aggregated information of everyone else. More generally, Milgrom and Stokey (1982) point out that the information that someone else is willing to make a bet against you reveals information making

Date: September 10, 2004
[†] Dept. of MSIS, The Universtiy of Texas at Austin, fang.fang@mail.utexas.edu
[‡] Dept. of Economics, The University of Texas at Austin, maxwell@eco.utexas.edu
[§] Dept. of MSIS, The University of Texas at Austin, abw@uts.cc.utexas.edu.
[1]All proofs in this abstract are available upon request
[2]A recent market failure in the Iowa Electronic Markets is that the market predicts that John Kerry's chance of winning the 2004 Democratic Caucuses is less than 35%.

you less willing to bet against them. Your initial best estimate of the odds is based on your private information. Knowing that someone else, on the basis of their own private information, is willing to bet against you means that you know that their private information contradicts yours. This should decrease your willingness to bet. To resolve these conceptual problems, outside subsidies must be provided to guarantee the agents correct incentives to trade.

We propose a novel betting mechanism in which the principal who needs the forecast will subsidize the bettors. In our mechanism, the prediction is reliable in the sense the principal can estimate the accuracy of the prediction from the bets being put into the market. Besides, each agent's bet also indicates the weight of each individual forecast. We conduct our analysis within a single-principal-multi-agent framework and resolve the following questions.

(1) How do we give agents incentives to truthfully reveal both their private information and the quality of their information.
(2) What do these incentives cost? (How much subsidy the principal should provide to the market place?)
(3) How do we trade off the costs of the incentives and the benefits of the information we gather?

In our proposed betting mechanism, the principal asks agents to report their forecasts and place money, as a bet, on their reports. After the uncertainty is realized, the agents can get rewards according to how close their reports are to the actual realization of the event and how much money they bet on the reports. With such a mechanism, the agents cannot simply say that they are "experts," they can only demonstrate their "expertise" by betting a lot of money on their report.

We present a family of reward functions with the following two crucial properties: each agent's dominant strategy is to report their true forecast; and agents bet an amount that is monotonically increasing in the precision of their information. We use this family in three market contexts: (1) a simultaneous betting market; (2) a sequential betting market; and (3) a Dutch auction betting market. We compare the three markets and discuss the pros and cons.

In all three markets, the agents' expected gain is positive if they are willing to bet. The gain comes from the principal's incentive cost. The principal can adjust the parameter of the reward function so that only agents who think their forecasts are precise enough will enter the market and bet.

In the simultaneous betting market, the principal can collect the information in one single period. This time advantage makes the simultaneous betting market favorable when the final forecast is needed urgently. Besides, the bet can be kept secret from other participants, which is important if the principal doesn't want to release the prediction for free. However, we found that the incentive cost in such a market is the highest of all three markets. It is because the principal cannot dynamically adjust the rewards as new information is revealed. Moreover, this market is vulnerable to the anonymity issue because an agent can gain by multiple bets using different identities.

A sequential betting market allows the principal to adjust the reward after each round new information has incorporated. We show that the principal can save on incentive costs compared with simultaneous betting if the principal can adopt a public learning strategy in the market. Besides, we found that the optimal sequence which minimizes incentive costs is the one where the agent with the less precise information bets earlier.

The Dutch auction betting market can be regarded as an extension of the sequential betting market, where the principal changes the reward gradually and the agents choose the right time to bet. We show that a symmetric equilibrium exists where all the agents adopt the same timing function. Furthermore, the timing function monotonically decreases with the agent's information precision. We also find that the auction mechanism introduces competition among agents so that the incentive costs are even lower than sequential betting. However, such a mechanism requires the agents to continuously monitor the state of the market and re-act quickly. Such a mechanism is only applicable when agents are willing to commit their time and attention for a reasonably long time period and the communication efficiency between the market and agents are effective.

All the three market designs are anonymous in the sense that no agent needs to have a long-term contract with the principal. Agents do their best in the market and then leave. The anonymity is important to guarantee the agents not to behave strategically. For example, in absence of anonymity, agents will worry about ruining promotion chances if getting the prediction wrong and be hesitant to participate.

The rest of the abstract is organized as follows. In section 2, we introduce the model setup, and analyze it under the three different markets. We conclude in section 3 with discussion of some extensions and interpretations of future research.

2. Model and Analysis

2.1. **The Model.** A risk neutral principal (a firm) wants to forecast the outcome of a future event, which can be represented by a random variable $X \sim N(0,1)$. The firm's payoff from generating a forecast $\widehat{X}$ is $g(\widehat{X}, x) = v - p(\widehat{X} - x)^2$, where x is the realization of X, v is the value to the firm when the forecast is extremely precise, and $p(\widehat{X} - x)^2$ is a quadratic penalty term for mistakes in the forecast.

The firm resorts to N risk-neutral agents, each of whom can access different, independent information sources. Formally, the agents observe signals $s_i = x + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_i^2)$ are independent prediction errors. $\sigma_i^2 \sim U[0,1]$ is private information of agent i.

Lemma 2.1. *(First Best Aggregation) When all the agents truthfully report the signals s_i and the precision of the information $\tau_i \triangleq \frac{1}{\sigma_i^2}$, $i = 1, ..., N$, the best prediction $\widehat{X}^{FB} = \frac{\sum_{i=1}^N \tau_i s_i}{1+\sum_{i=1}^N \tau_i}$, and $\tau^{FB} = 1 + \sum_{i=1}^N \tau_i$ is the prediction precision.*

However, the agents may not have the incentive to provide the true information. Inspired by the famous saying, "putting your money where your mouth is," we design a betting mechanism to elicit the information. That is, each participant sends a report r_i and places an amount of money B_i as a bet. After x is realized, each participant gets a payoff $f(r_i, B_i, x)$. Given a reward structure $f = f(B_i, r_i, x)$, agent i solves

$$\max_{r_i, B_i} E[\pi_i | s_i, \tau_i] = E[f(B_i, r_i, x) - B_i | s_i, \tau_i]$$

for their optimal strategy $(B_i^*(s_i, \tau_i), r_i^*(s_i, \tau_i))$.

Definition 2.2. *We say that a strategy $(B_i(\tau_i, s_i), r_i(\tau_i, s_i))$ is **fully revealing** if it is invertible for $\forall i \in \Omega = \{1, 2, \cdots, N\}$. If there exists a set $E \subseteq \Omega$, and $(B_i(\tau_i, s_i), r_i(\tau_i, s_i))$ is invertible for $\forall i \in E$, we say that $(B_i(\tau_i, s_i), r_i(\tau_i, s_i))$ is **fully revealing on set** E.*

We consider the class of reward functions

$$\mathcal{F} = \{f_a(B_i, r_i, x) = 2B_i^{\frac{1}{2}}[a - (r_i - x)^2], a \geq 0\}$$

where a is a parameter adjustable by the firm. Proposition 2.1 shows that this family of reward function can reveal a specific set of agents' private information and the precision. [3]

Proposition 2.1. (1) *If $a > \frac{1}{2}$, then all the agents will enter the market and bet. The optimal strategy $(B_i^*(\tau_i, s_i), r_i^*(\tau_i, s_i))$ is fully revealing and agent's expected payoff $(a - \frac{1}{1+\tau_i})^2$.*
(2) *When $a \in (0, \frac{1}{2}]$, only those agents with $\tau_i > \frac{1}{a} - 1$ will enter the market. Thus, the agents' optimal strategy is fully revealing on the set $E_a = \{i : \tau_i > \frac{1}{a} - 1\}$. The agents' expected payoff by playing optimal strategies are $(a - \frac{1}{1+\tau_i})^2$ if $i \in E_a$ and 0 otherwise.*
(3) *When $a \leq 0$, no agent will enter the market to bet. Thus the market shuts down.*

2.2. **Simultaneous betting market.** A simultaneous betting market allows all the agents to submit their reports and bets at the same time. No agent can observe other's behavior in the market. In such a simultaneous betting market, the firm needs to decide the optimal level of a. A large a induces more agents to bet and the market can incorporate more information. However, the firm pays more to each agent in expectation.

Proposition 2.2. (1) *There exists a $\underline{\underline{p}} > 0$ such that for $p \in [0, \underline{\underline{p}})$, the firm shuts down the market by setting $a^* = 0$.*

[3]The families of fully revealing reward functions are not unique. A detailed analysis of the families of reward functions can be found in Fang et al (2004). In this paper, we take the family of reward function as given and look at the impact of different market designs.

(2) *There exists a pair* $(\overline{p}, \underline{p})$ *satisfies* $\overline{p} > \underline{p} \geq \underline{\underline{p}}$ *such that when* $p \in (\underline{p}, \overline{p})$*, an interior solution exists where* $a^* \in (0, \frac{1}{2})$*. Only those agents with precision higher than* $\frac{1}{a^*} - 1$ *will participate in the market.*

(3) *There exists a* $\overline{\overline{p}} \geq \overline{p}$*, so that when* $p > \overline{\overline{p}}$*, the optimal choice is to set* $a = \frac{1}{2}$*. The market is open to all the agents.*

In a simultaneous betting market, it is very critical that the reward function cannot be changed when market is open, although the firm has a strong incentive to do so after observing the bets of first several agents. If the commitment is non-credible, then all the agents will wait till the rewards reach the highest level. Besides, the anonymity of a simultaneous betting market may give agents the opportunity to manipulate the market by placing multiple bets.

2.3. **Sequential betting market.** In a sequential betting market, the agents are arranged in a queue with an exogenously determined order. The market allows only one agent to bet at each time. In round t, the firm announces the reward function $f_t(B_t, r_t, x) = 2B_t^{\frac{1}{2}}[a_t - (r_t - x)^2]$ and then the agent reports r_t and bets B_t. After the agent leaves the market, the firm publicize (r_t, B_t). Round $t+1$ begins with a new reward function $f_{t+1}(B_{t+1}, r_{t+1}, x) = 2B_{t+1}^{\frac{1}{2}}[a_{t+1} - (r_{t+1} - x)^2]$.

It takes multiple rounds for the firm to get the prediction in a sequential betting market. So it can only be used for producing predictions which is not needed urgently. In the following analysis, we assume that there is no value discount across time for the firm. Then after round t, the firm's current forecast of the information is:

$$X|\,[(s_1, \tau_1), (s_2, \tau_2), \cdots, (s_t, \tau_t)] \sim N\left(\frac{\sum_{i=1}^t \tau_i s_i}{\sum_{i=0}^t \tau_i}, \sum_{i=0}^t \tau_i\right)$$

where $\tau_0 \triangleq 1$ is the prior precision. The agent i $(i \leq t)$ has the same belief because her private information has already been revealed. For those agents indexed j $(j > t)$, they also update their beliefs according to the information revealed in the market and their beliefs are:

$$X|\,[(s_1, \tau_1), (s_2, \tau_2), \cdots, (s_t, \tau_t), (s_j, \tau_j)] \sim N\left(\frac{\sum_{i=1}^t \tau_i s_i + \tau_j s_j}{\sum_{i=0}^t \tau_i + \tau_j}, \sum_{i=0}^t \tau_i + \tau_j\right)$$

They still have some information advantage. However, the firm only needs to set $a_{t+1} = \frac{1}{\tau^t + 1}$ to induce agent $t+1$'s information, where $\tau^t \triangleq \sum_{i=0}^t \tau_i$. Thus, agent $t+1$'s expected payoff cannot be higher than $(\frac{1}{\tau^t+1} - \frac{1}{\tau^t + \tau_j})^2$, which is much lower than what she can get from simultaneous betting system. In other words, the incentive cost in sequential betting is lower than in simultaneous betting.

Proposition 2.3. $E[\Pi_{seq}] > E[\Pi_{simult}]$ *if there is no time value discount in the prediction.*

In sequential betting market, the order of agents affects the firm's cost of elicitation although the actual prediction accuracy remains the same. Thus, it is interesting to discuss the optimal order of agents which minimize the firm's cost.

Proposition 2.4. *: The optimal queue which minimizes the cost of eliciting information in a sequential betting market satisfies* $\tau_i \geq \tau_j$ *if* $i > j$.

However, it is impossible for the firm to arrange the optimal sequence because τ_i is unobservable.

2.4. **Sequential betting in a Dutch Auction mechanism.** In the Dutch Auction mechanism, the firm increases the reward parameter a gradually until either some agent bets or a reaches $\frac{1}{\tau^t+1}$, where τ^t is the precision of the current market belief. Although agents want to wait to bet until a gets large, they are bearing the risk that a less patient agent will bet earlier and the auction restarts with a less profit potential. For simplicity, we illustrate this idea using the a two-agent example.

Assume that there are only two agents with information precision τ_1, τ_2 independently drawn from $f(\tau)$. W.l.o.g., we assume $a(t) = t$. In this case, each agent knows that no other competitor exists if the other one bets first. So the second bettor can always wait until the reward parameter increases to the lowest level, i.e. $a = \frac{1}{2+\tau_1}$. Thus, the second betters' payoff is always: $\left(\frac{1}{2+\tau_1} - \frac{1}{1+\tau_1+\tau_2}\right)^2$.

Define $\theta(\tau)$ as the agent's optimal timing strategy, which can be derived from the following equation:

$$\underset{\hat{\tau}_1}{max} \int_{\hat{\tau}_1}^{+\infty} \left(\frac{1}{2+\tau_2} - \frac{1}{1+\tau_1+\tau_2} \right)^2 f(\tau_2) d\tau_2 + \left(\theta(\hat{\tau}_1) - \frac{1}{1+\tau_1} \right)^2 F(\hat{\tau}_1)$$

Proposition 2.5. *The equilibrium with symmetric timing function $\theta(\tau)$ uniquely exists, and $\theta(\tau)$ strictly decreases with τ.*

Proposition 2.6. *When there are only 2 agents with precision $\tau_1, \tau_2 \sim f(\tau)$, the expected cost of eliciting the information in the auction betting market is always less than the minimal expected cost in a sequential betting market described in Proposition 2.4.*

3. Conclusion and Future Research

In this abstract, we described a betting mechanism to elicit agents' private information and the quality of information simultaneously. The information elicited is useful for the firm to generate reliable forecast. In such a betting mechanism, the firm pays each agent for contributing to the output. Each agent's expected gain is decided based on their own performance (e.g. the actual distance between r_i and x) and their potential to improve the quality (e.g. the precision τ_i). Agents are self-selected to make contributions and those who don't have enough potential will be discouraged to participate.

In our setup, if the agents are risk averse, they will bet less, and/or bet only when their precision is higher. If the agents risk-aversion can be learnt or estimated, the same general principle, **invertibility**, can still be applied. However, if the agents' risk-aversion is also a private information, more complicated mechanisms need to be designed. In Chen et al (2004), a two-round market mechanism is introduced with the first round specifically designed to estimate the agent's risk-aversion. However, such a mechanism is only applicable in small group forecast. When the number of potential agents is large, more complicated mechanism needs to be designed to elicit the agents' all relevant information simultaneously.

We have assumed that the firm's loss from a prediction error is symmetric and quadratic. In certain applications, the loss could be asymmetric. For example, the firm's loss from having a more conservative prediction of the expected sale of a new prototype may be much more less than an optimistic one. Then the task is to design a mechanism which can discourage the agent's unrealistic high forecast.

Sometimes, the agents incur costs to access information sources. Thus, given a specific reward function, some of the agents may not collect the information if they see the expected gain from the market cannot cover the expected cost. In this case, the sequential betting market and the auction betting market are not efficient because the agents needs to decide whether to learn the information after observing the previous agents' bets. This expands the expected time period the firm needs to get a satisfactory forecast.

References

[1] K.Y.Chen, L.R.Fine and B.A.Huberman. 2001. Forecasting uncertain events with small groups. Proc. ACM EC'01 Conf. ACM, Tampa, FL.

[2] K.Y.Chen, L.R.Fine and B.A.Huberman. 2004. Eliminating Public Knowledge Biases in Information-Aggregation Mechanisms. Management Science. 50(7) pp.983-994

[3] F. Fang, M. Stinchcombe and A. B. Whinston. 2004. Optimal Elicitation Mechanism Design when People Differ in the Quality of Information. Working Paper.

[4] Grossman, Sanford J & Stiglitz, Joseph E, 1980. On the Impossibility of Informationally Efficient Markets, American Economic Review, 70(3) pp.393-408.

[5] Paul Milgrom and Nancy Stokey. 1982. Information, Trade, and Common Knowledge. Journal of Economic Theory. 26 pp.17-27.

[6] O'Hara. Market Microstructure Theory. Cambridge, Mass : Blackwell Publishers. 1995.

[7] Savage, Leonard J. 1971. Elicitation of Personal Probabilities and Expectations. Journal of the American Statistical Association. 66(336) pp.783-801.

[8] Spann, M. and B. Skiera. 2003. Internet-Based Virtual Stock Markets for Business Forecasting. Management Science 49(10) pp.1310-1326

[9] Justin Wolfers and Eric Zitzewitz. Prediction Markets. Journal of Economics Perspectives. 18(2) pp.127-141