INFORMATION EXCHANGE IN PREDICTION MARKETS: HOW SOCIAL NETWORKS PROMOTE FORECAST EFFICIENCY

Completed Research Paper

Liangfei Qiu University of Texas at Austin Austin, Texas 78712 liangfeiqiu@utexas.edu

Huaxia Rui University of Rochester Rochester, NY 14627 huaxia.rui@simon.rochester.edu

Andrew Whinston

University of Texas at Austin Austin, Texas 78712 abw@uts.cc.utexas.edu

Abstract

This paper studies the effects of information exchange and social networks on the performance of prediction markets with endogenous information acquisition. We provide a game-theoretic framework to resolve the question: Can social networks and information exchange promote the forecast efficiency in prediction markets? Our study shows that the use of social networks could be detrimental to forecast performance when the cost of information acquisition is high. Although social networks can provide internal communications among participants, it reduces the incentive to acquire information because of free riding. We also study the effects of social networks on information acquisition in prediction markets. In the symmetric Bayes-Nash Equilibrium, all participants use a threshold strategy, and the equilibrium information acquisition is decreasing in the number of participant's friends and increasing in the network density. The above results are robust to two commonly used mechanisms of prediction markets: forecast-report mechanism and security-trading mechanism.

Keywords: social networks, prediction markets, information exchange

"Under the right circumstances, groups are remarkably intelligent, and are often smarter than the smartest people in them."

– Surowiecki (2004)

"This very fact that crowds possess in common ordinary qualities explains why they can never accomplish acts demanding a high degree of intelligence."

— Gustave Le Bon (1896)

Introduction

The wisdom of crowds is based on the assumption that valuable knowledge in social systems frequently exists only as dispersed opinions, and that aggregating dispersed information in the right way can produce accurate predictions. A prediction market illustrates effective use of the wisdom of crowds in the right circumstances. Assets are created whose final value is tied to a particular event — for example, whether the next U.S. president will be a Republican or a Democrat. People "place bets" on events that they think are most likely to happen, thus revealing in a sense the nature of their private information and subsequent posterior beliefs. The market mechanisms provide a method of "putting your money where your mouth is" (Fang, Stinchcombe, and Whinston, 2007). Prediction markets can also immediately incorporate new information and provide a real-time forecast. The hope is that by aggregating the private information of a large population, a prediction market can generate fairly accurate predictions of future events.

A number of empirical studies have shown that prediction markets have been successfully used to predict outcomes in many areas (Berg, Neumann, and Rietz, 2009; Guo, Fang, and Whinston, 2006). But researchers in most of the previous literature have assumed that the participants in the prediction markets are isolated: They receive small bits and pieces of independent information and cannot affect the decisions of other participants. However in reality, people often turn to their friends, colleagues, and family members to assess the probabilities of future events in a world full of uncertainties. People's decisions rely on information that is gathered through personal networks and social relations, which are important conduits of knowledge. That is especially true when it comes to social networking sites, such as Facebook and Twitter, which have dramatically changed people's information exchange and networks of communication. For instance, as was reported in a Financial Times article (Bradshaw and Khan, 2011), a voting application on Facebook has some features of a networked prediction market. Producer Endemol used Facebook's credits system of micropayments to let fans vote for their favorite contestants via a Facebook application. The fans had to put money into their credits accounts; thus, voting for favorite contestants was a typical way of "putting your money where your mouth is." By aggregating the information from voting, the company could obtain a more accurate prediction of the contestants' potentials to be celebrities. It is crucial to notice that the Facebook network plays a key role in online voting. Unlike traditional phone and text-message voting, each vote cast is shared with friends on the site. and fans can exchange views on their Facebook network. Therefore, their decisions are influenced by their social relations.

In addition to Facebook voting, another example of an effective information exchange network was reported by CNBC news: Farmers used their Twitter accounts to post a message, or tweet, about a particularly robust corn crop. Tweeting with fellow farmers has become a way for participants in a far-flung and isolating business to compare notes on everything from weather conditions to new fertilizers.¹ There is a growing network of farmers and traders whose use of tweets is transforming how this multi-trillion dollar industry does business. These tweets are dramatically accelerating the flow of information that gives investors an edge in the commodities market. Are social networks the right circumstances in which groups are remarkably intelligent?

In this paper, we relax the isolated participants assumption and assume that participants in the prediction markets are linked to some others via a social network, and that their predictions are largely influenced by their friends in the social network. Our model also considers endogenous information acquisition. In

¹ The information is from CNBC News, March 8, 2011. The CNBC reporter called the phenomenon "Trading on Twitter." Grisafi, known as @IndianaGrainCo on Twitter, says he tweets with at least 15 farmers on a regular basis to check on crop conditions.

financial markets, small investors often buy financial reports and statements to perform stock fundamental research and investment analysis. Endogenous information acquisition in social networks can play an important role in prediction markets.

The goal of the present paper is to study the effects of information acquisition and exchange in social networks on the performance of prediction markets. The wisdom of crowds works well under certain conditions and less well under others. Is a networked prediction market the right circumstance for the crowd to be wise? We combine social networks and endogenous information acquisition in our prediction market model to address this important question. To the best of our knowledge, our paper is the first to study the effects of information exchange and social networks on the forecast efficiency with endogenous information acquisition. In our model, information acquisition is costly and endogenously determined in equilibrium. In addition, the participants of the prediction markets are embedded in a social network. We conduct an analysis using an incomplete information network game framework to resolve the question: Can social networks and information exchange promote the forecast efficiency in prediction markets? There are a series of related questions: (1) When does the use of networked prediction markets improve the forecast efficiency? (2) Are the results in question (1) robust to different mechanisms of prediction markets? (3) What are the effects of social networks on information acquisition in prediction markets? (4) How does the social network structure, such as network density, affect the forecast performance? Our analytical model provides a game-theoretic framework, and the numerical simulations analyze these effects quantitatively.

In the model, each participant is a node in the network graph and is connected to some other participants who are specified by the social networks. We relax the assumption that people have complete information on the network structure and assume that they do not know the whole network but are aware who their friends are. It captures the idea that people always have good knowledge about their friends, but they may have no idea about their friends' social connections. Each participant can access a private costly information source and observe the information acquired by her friends. In this network game, the equilibrium action of information acquisition is non-increasing in the number of the participant's friends, which is called "degree." There is a symmetric Bayes-Nash equilibrium in which all participants use a threshold strategy: A participant acquires information if her degree is less than the threshold value, otherwise; she does not acquire information. The inefficiency of information acquisition in networked prediction market is caused by free riding. Participants can "free ride" on the actions of their friends.

Using a variety of simulations in social network-embedded prediction markets, our present study shows, surprisingly, that social networks do not always improve the forecast efficiency. When the cost of information acquisition is low, social network can improve forecast performance of prediction markets. However, a social network is not a panacea in terms of improving forecast efficiency. Information exchange in networks could lead to a poorer prediction when the cost of information acquisition is high. The results imply that the use of social networks could be detrimental to forecast performance. Too much social communication can make the group forecast produced by prediction markets less accurate. Therefore, we have the following implications for business practice of designing prediction markets: When the predicted event is simple, we suggest using a social network-based prediction market. When it involves complicated issues, the traditional non-networked prediction market is a better choice. For example, it is rather difficult for people to know some information about the event "Hugo Chavez to no longer be the President of Venezuela before midnight ET 31 Dec 2012" (Intrade Prediction Market). However, it is relative easy to have some ideas about the *Twilight* movie box office (Iowa Electronic Markets).

From the simulation results, we can find two effects of social networks: (1) Free rider effect. Although social networks can provide internal communications among participants, they reduce the incentive to acquire information because of free riding. (2) Correction effect. Information exchange and communication correct the overreaction of market prices to the common prior. The first effect is detrimental to the forecast efficiency, and the second promotes the efficiency. When the cost of information acquisition is small, the second effect dominates. As the cost increases, the first effect becomes the dominant effect. We also show that the social network structure can affect the information acquisition in prediction markets.

This study is related to the growing literature on prediction markets. However, researchers in previous studies have focused on how to elicit dispersed private information, such as some variations of scoring

rules. Hanson suggests a new mechanism for prediction markets, market scoring rule, which combines the advantages of markets and scoring rules. Our present model also melds key aspects of market and strategic incentives, but we consider broader issues, such as information acquisition and social networks in prediction markets. In our paper, participants interact strategically with their friends, taking into account incentives and the effects of their information acquisition, but trade anonymously in prediction markets, taking prices as given. Fang, Stinchcombe, and Whinston (2010) propose a proper scoring rule that elicits agents' private information as well as the precision of the information. In their work, the information of agents is independent. However, in our present social-network-embedded prediction market, the information that participants have is correlated with that of their friends. Using a laboratory experiment, Jian and Sami (2012) compare two commonly used mechanisms of prediction markets probability-report mechanism and security-trading mechanism — and find no significant performance difference between these two mechanisms. In our paper, we examine whether our results on networked prediction markets are robust to a forecast-report mechanism and a security-trading mechanism, and we find similar results using different mechanisms. Qiu, Rui, and Whinston (2011) develop an information system that combines the power of prediction markets with the popularity of Twitter. It is a parimutuel betting platform for a social-network-embedded prediction market.

The present work is also related to the recent work on network games by Galeotti, Goyal, Jackson, Vega-Rendondo, and Yariv (2010), who provide a framework to analyze strategic interactions in an incomplete information network game. Sundararajan (2008) studies the adoption of a network good when consumers are embedded in an underlying social network. The value of the network good depends on the adoption of the "local neighbors". Golub and Jackson (2010) consider the wisdom of crowds in social networks. They discuss how network structure influences the spread of information and show that all opinions in a large society converge to the truth if and only if the influence of the most influential agent vanishes as the society grows. In our paper, we address how the use of networked markets can affect information acquisition and efficiency of a market. The effects of social networks on forecast efficiency are related to the social value of information exchange and communication.

The rest of the paper is organized as follows. In Section 2, we present a model of a prediction market integrated with a social network. In Section 3, we discuss the network game and equilibrium results. In Section 4, we study the effects of social networks on forecast efficiency using numerical simulation. We conclude and offer some directions for future research in Section 5. All proofs are in the Appendix.

A Model of Social Network-Embedded Prediction Market

Model Setup

In this section, we set up a game-theoretical model of information exchange in a prediction market. Our model provides a framework to analyze two commonly used mechanisms of prediction markets, a forecast-report mechanism and a security-trading mechanism, with an embedded social network and endogenous information acquisition.

The social network $\Gamma = (N, L)$ is given by a finite set of nodes $N = \{1, 2, ..., n\}$ and a set of links $L \subseteq N \times N$.² Each node represents a participant in prediction markets. The social connections between the participants are described by an $n \times n$ dimensional matrix denoted by $g \in \{0,1\}^{n \times n}$ such that:

$$g_{ij} = \begin{cases} 1, & if(i,j) \in L \\ 0, & otherwise \end{cases}$$

Let $N_i(g) = \{j \in N : g_{ij} = 1\}$ represent the set of friends of Participant *i*. The degree of Participant *i* is the number of Participant *i*'s friends: $k_i(g) = \#N_i(g)$.

A principal wants to forecast a random variable *V*, and he resorts to *n* agents to obtain an accurate prediction. In business practice, *V* could be movie box office revenue, future demand for electricity, or future stock prices. The agents are the participants in the prediction market described above, and they are embedded in the social network Γ .

 $^{^{2}}$ We assume that the social network is exogenous. A further direction is to consider a endogenous network formation under homophily.

The participants share a common prior on *V*, given by:

$$V \sim N(V_0, 1/\rho_V),$$

where ρ_V is the precision of the prior. Each participant can access a private independent information source at a cost *c*. m_i is an indicator function indicating whether or not Participant *i* acquires information. Participants exchange information over the social network: They can observe their friends' private signals if their friends acquire costly information. More precisely, if Participant *i* acquires information from her private source ($m_i = 1$), she observes a conditionally independent private signal and passes it to her friends:

$$S_i = V + \varepsilon_i, \ \varepsilon_i \sim N(0, 1/\rho_\varepsilon), \tag{1}$$

where ρ_{ε} is the precision of Participant *i*'s information source for i = 1, 2, ..., n.³ The signals' errors $\varepsilon_1, ..., \varepsilon_n$ are independent across participants and are also independent of *V*. We assume that the precisions of all participants' information sources are equal. It implies that no one is especially well informed, and that the valuable information is not concentrated in a very few hands.

All of the participants decide whether to acquire information simultaneously. The information exchange network models the idea that people often consult their friends, colleagues, and family members before making a decision. They can learn useful information from communication and social relations. In this social network-embedded prediction market, if signal S_i is acquired, it can be observed by Participant *i* and her friends, Participant $j \in N_i(g)$. V_0 , ρ_V , and ρ_ε can be observed by all participants in prediction markets. The principal is not an expert on forecasting and he knows little about *V*; thus, we assume that he cannot observe V_0 , ρ_V , or ρ_ε .

The principal designs a prediction market mechanism to elicit the private information of participants. The mechanism defines the participants' payoff as a function of their actions and the realization of *V*. A general payoff function for Participant *i* is $w(m_i, x_i, v)$, where m_i is information acquisition described before, x_i is participant's prediction, and *v* is the subsequently realized value of the random variable *V*.

In Section 2.2 and Section 2.3, we model two commonly used mechanisms for prediction market — a forecast-report mechanism and a security-trading mechanism — in the information exchange networks. A forecast-report mechanism is similar to a proper scoring rule which elicits the true beliefs of participants as probabilistic forecasts. A proper scoring rule gives the participants the incentives to report truthfully; then, the principal aggregates the private information of all participants. A security-trading mechanism is similar to a competitive financial market, and people trade securities according to their forecasts. The market clears when the aggregate demand for securities equals the supply, and market clearing determines the prediction market price.

Forecast-Report Mechanism

In the forecast-report mechanism, we consider that the principal designs a quadratic loss function.⁴ Note that the use of a quadratic loss function is common. In the motivating example in Section 1, a quadratic loss scoring rule implies that farmers want to forecast the price of the corn as precisely as possible, and an error above the price causes the same loss as the same magnitude of error below the price.

Suppose the participants are risk neutral under a forecast-report mechanism and their quadratic payoff function is

$$w(m_i, x_i, v) = a - b(x_i - v)^2 - m_i c,$$
(2)

where x_i is the prediction reported by Participant *i*, and $b(x_i - v)^2$ is a quadratic penalty term for mistakes in the forecast. Notice that reporting the point prediction is equivalent to reporting the private beliefs in this setup. When the scoring rule is a quadratic loss function, the optimal report for Participant *i* is $x_i^* = E[V|I_i]$, where I_i is the information set of Participant *i*; thus, observing x_i^* reveals a participant's

³ We assume that all participants have equal precision of private information. We can allow agents to have different presisions: some with more precise signals, and others with less precesie signals. The basic results of the paper remain the same.

⁴ But the payoff function can be very general — it simply needs to exhibit strategic substitutes, which we define later.

private belief.

A participant follows a two-step procedure and has to make two related decisions: information acquisition and forecast decision. In the first stage, all of the participants decide whether to acquire information simultaneously, so there is a simultaneous move network game. In arriving at the forecast decision in the second stage, a participant makes use of her signal, as well as of the signals of her friends. In this section, we mainly focus on the optimization problem in the second stage. The incomplete information network game is discussed in Section 3.

In the second stage, Participant *i*'s best prediction x_i^* depends on whether Participant *i* and her friends acquire information, and thus x_i^* is a function of m_i and $m_{N_i(g)}$, where $m_{N_i(g)} \in \{0,1\}^{k_i}$ is the action profile of Participant *i*'s friends, and it represents whether Participant *i*'s friends acquire information.

If Participant *i* acquires information ($m_i = 1$), she forms her private belief from the private signal S_i , as well as from information she obtains from her neighbors. Given the information acquisition of her friends and the realization of *V*, her payoff is:

$$a - b[x_i^*(m_i = 1, m_{N_i(g)}) - v]^2 - c.$$

If the participant has decided not to acquire information ($m_i = 0$), she forms the belief only from her neighbors' signals, and her payoff is:

$$a - b[x_i^*(m_i = 0, m_{N_i(g)}) - v]^2.$$

Given the action profile of her friends, Participant *i* maximizes

$$u(m_{i}, m_{N_{i}(g)}) = E_{V} \left[a - b \left[x_{i}^{*}(m_{i}, m_{N_{i}(g)}) - V \right]^{2} - m_{i}c \right],$$
(3)

where E_V is the expectation with respect to V. $u(m_i, m_{N_i(g)})$ depends on whether Participant *i* and her neighbors acquire information. We have several remarks here: (1) Two participants who have the same degree have the same payoff function. (2) *u* depends on the vector $m_{N_i(g)}$ in an anonymous way: A permutation of $m_{N_i(g)}$ does not change the payoff.

Definition 2.1 We say that a function u exhibits strategic substitutes if an increase in others' actions lowers the marginal returns from one's own actions: For all k, $m_i^{'} > m_i$ and $m_{N_i(q)}^{'} \ge m_{N_i(g)}$,

$$u\left(m_{i}^{'}, m_{N_{i}(g)}^{'}\right) - u\left(m_{i}, m_{N_{i}(g)}^{'}\right) \leq u\left(m_{i}^{'}, m_{N_{i}(g)}\right) - u\left(m_{i}, m_{N_{i}(g)}\right).$$
(4)

When u exhibits strategic substitutes, a participant's incentive to take a given action decreases as more friends take that action.

Security-Trading Mechanism

In a security-trading prediction market, assets are created whose final value is tied to a particular event — for example, whether the next U.S. president will be a Republican or Democrat. In our model, people trade securities according to the outcome of future random variable *V*. The market mechanisms provide a method of "putting your money where your mouth is." Participants interact strategically with their friends, taking into account incentives and the effects of their information acquisition, but trade anonymously in the prediction markets, taking prices as given. We characterize the equilibrium when traders are allowed to trade assets in a competitive prediction market.

We adopt the widely used payoff function in financial markets: the exponential utility function with a constant coefficient of absolute risk aversion (CARA). Suppose the participants are risk averse under a security-trading mechanism and their payoff function is

$$w(m_i, x_i, v) = -\exp[-\gamma x_i(v - P)] - m_i c,$$
(5)

where *P* is the market price of the security tied to *V*, and x_i is the demand for the security. $\gamma > 0$, and it is the coefficient of absolute risk aversion, which measures the degree of risk aversion. If Participant *i* thinks the realization of *V* is higher than *P*, she will buy the risky asset; otherwise, she will short. The cost of

information acquisition, c, enters the payoff function in a linear way. One feature of this market is that there are no short-sale constraints; thus, the participants are allowed to hold a negative quantity of the asset. A participant in this market also has to make two decisions. Similarly, a participant first decides whether to acquire her private signal and then chooses the optimal demand for the risky security. The information set of Participant *i*, I_i , depends on whether Participant *i* and her friends acquire information.

Given I_i , Participant *i* chooses the optimal demand for the risky asset to maximize

$$E_V[-\exp[-\gamma x_i(V-P)] - m_i c |I_i].$$
(6)

From the moment-generating function of normal distribution, it is well known that maximizing function (6) w.r.t. x_i is equivalent to maximizing the following function of x_i :

$$-\exp\left[-\gamma x_i(E[V|I_i] - P) + \frac{1}{2}Var[V|I_i]\gamma^2 x_i^2\right] - m_i c,\tag{7}$$

and Participant i's optimization problem becomes:

$$\max_{x_i} \gamma x_i (E[V|I_i] - P) - \frac{1}{2} Var[V|I_i] \gamma^2 x_i^2.$$
(8)

The first order condition (F.O.C) yields:

$$x_i = \frac{E[V|I_i] - P}{\gamma Var[V|I_i]}.$$
(9)

The price *P* is endogenously given by the market clearing condition: $\sum_{i=1}^{n} x_i = 0$. We can solve *P* by plugging equation (9) into the market clearing condition. Since the information set $I_i(m_i, m_{N_i(g)})$ depends on m_i , and $m_{N_i(g)}$, the optimal demand x_i^* is a function of m_i and $m_{N_i(g)}$.

We make several comments here. A striking difference between prediction markets and financial markets is the number of participants in the market. In a financial market, the number of investors is very large; thus, it is reasonable to assume infinitely many traders. The tractable solutions of rational expectation equilibria (REE) rest on the large economy assumption. However, the number of participants in many prediction markets is relatively small, especially in business practice. Hewlett-Packard uses prediction markets to forecast sales, as well as financial and accounting results. Google and Intel also have similar internal prediction markets. All of these prediction markets involve a small number of participants. There could be no REE in a small network-embedded prediction market. It is questionable to expect the participants in the prediction markets to reach the rational expectation equilibrium.

In finance literature, there is an ongoing debate between REE and differences of opinions (DO) models. The two major approaches for modeling belief heterogeneity are REE and DO.⁵ The difference between the REE and DO models is in the prior beliefs the traders have over their information sets. In an REE, a trader learns both from the private signal and the price. In the DO model, however, the traders are not allowed to condition their decisions on the equilibrium prices. Ottaviani and Sorensen (2009) provide a framework to analyze the parimutuel betting prediction market where participants are not allowed to condition their behavior on the information that is contained in the equilibrium prices. Their Bayes-Nash equilibrium does not converge to an REE as the number of participants increases.

In our security trading prediction market, we relax the REE assumption as the DO models: Individuals cannot make correct inferences about the state of the world from the equilibrium prices. Therefore, If Participant *i* acquires information, the best mean square predictor is:

$$E[V|I_i] = \frac{\rho_V}{(k_a+1)\rho_{\varepsilon}+\rho_V} V_0 + \frac{\rho_{\varepsilon}}{(k_a+1)\rho_{\varepsilon}+\rho_V} \Big(\sum_{i \in A_{N_i(g)}} S_i\Big),\tag{10}$$

and

$$Var[V|I_i] = \frac{1}{(k_a+1)\rho_{\varepsilon}+\rho_V},\tag{11}$$

⁵ Banerjee, Kaniel, and Kremer (2009) discuss the difference between REE and DO models. DO models assume that investors "agree to disagree" or have "differences of opinions." Both approaches share the view that traders have differential information, and prices aggregate the different views during the trading process. They differ in whether agents can agree to disagree. In a DO model, traders disagree even when their views become common knowledge. In an REE model, this type of disagreement is ruled out.

where $A_{N_i(g)} = \{j \in N_i(g): m_j = 1\}$, which is the set of Participant *i*'s friends who acquire information, and $k_a = #A_{N_i(g)}$, which is the number of Participant *i*'s friends who acquire information.

Given the action profile of her friends, Participant *i* chooses m_i to maximize

$$u(m_i, m_{N_i(g)}) = E_V \left[-\exp\left[-\gamma x_i(m_i, m_{N_i(g)})(V - P) \right] - m_i c |I_i| \right].$$
(12)

We will formally describe the information network game and solve the symmetric Bayes-Nash equilibria in Section 3.

Network Game and Equilibrium in Prediction Markets

Following Galeotti, Goyal, Jackson, Vega-Rendondo and Yariv (2010), we assume that participants in prediction markets do not know the structure of the network on which they play. More precisely, Participant *i* deoes not know the degrees of other participants k_j , for $j \neq i$, but are informed only of her own degree, k_i , and a common prior degree distribution in the network game. Here we relax the assumption that people have complete information on the network structure. The incomplete information on social networks captures the idea that people always have good knowledge about the number of their friends, but they may have no idea about their friends' social connections. People don't know the exact network graph, but they have some information about the structure of a network, such as the common prior degree distribution. Dover, Goldenberg, and Shapira (2012) demonstrate that degree distribution significantly affects the diffusion of new products, and propose a method for uncovering the degree distribution from aggregate-level data. The assumption of incomplete information could be also motivated by the privacy issues of social media. For instance, on Google Plus, people put their friends into different "circles." Every time they share information and thoughts, they can specify exactly which circles receive them. People know the number of friends who share information with them, but they don't know the whole network of information sharing. Twitter also has something similar to circle, called "List."

Now let's define the incomplete information network game. Participants start with a common prior degree distribution. Then, each participant observes her own degree k_i , which defines her type, but does not observe the degree or connections of any other participant in the network. She uses this information to update her prior about the network and arrives at a posterior belief:

$$\Pi(\cdot | k_i) \in \Delta\{1, \ldots, k_{\max}\}^{k_i},$$

where k_{\max} is the maximal possible degree, and $\Delta\{1, \dots, k_{\max}\}^{k_i}$ is the set of probability distribution on $\{1, \dots, k_{\max}\}^{k_i}$.

In the incomplete information network game, participants would learn about the network by observing their own degrees. This depends on the assumptions on degree correlations across players. If the degrees of neighboring nodes are independent, knowledge about the participant's own degree reveals no additional information about degrees of her neighbors. By contrast, if there is a positive degree correlation, then a higher degree of one's own indicates a higher expected degree of her neighbors. For simplicity, we make an assumption that neighbors' degrees are all stochastically independent, which means that Participant *i*'s degree is independent from the degree of one of her randomly selected friends. This assumption is true for many random networks, such as the Erdös-Rényi random graph.

A strategy of Participant *i* is a measurable function $\sigma_i: \{1, \dots, k_{\max}\} \rightarrow \Delta\{0, 1\}$, where $\Delta\{0, 1\}$ is the set of probability distribution on $\{0, 1\}$. This strategy simply says a participant observes her degree k_i , and on the basis of this information she decides whether to acquire information. Notice that $\Delta\{0, 1\}$ means that the participant adopts a mixed strategy: She randomizes her actions with some probabilities in $m_i = 1$ and in $m_i = 0$. $\sigma_{N_i(g)}$ denotes the strategy profile of Participant *i*.

We focus on symmetric Bayes-Nash equilibria, where all participants follow the same strategy σ . A Bayes-Nash equilibrium is a strategy profile such that each participant with degree k_i chooses a best response to the strategy profile of her friends. Let $\phi(m_{N_i(g)}, \sigma, k_i)$ be the probability distribution over $m_{N_i(g)}$ induced by $\Pi(\cdot |k_i)$. The expected payoff of Participant *i* with degree k_i and action m_i is equal to

$$U(m_{i},\sigma_{N_{i}(g)};k_{i}) = E_{m_{N_{i}(g)}}u(m_{i},m_{N_{i}(g)}) = \sum_{m_{N_{i}(g)}}\phi(m_{N_{i}(g)},\sigma,k_{i})u(m_{i},m_{N_{i}(g)}),$$
(13)

where $E_{m_{N_i(g)}}$ is the expectation with respect to $m_{N_i(g)}$. $u(m_i, m_{N_i(g)})$ is the function defined in Section 2. It is important to note that in this section, $u(m_i, m_{N_i(g)})$ is a general function, and it could be given by equation (3) or (12). In other words, the results in this section hold for both a forecast-report mechanism and a security-trading mechanism.

Definition 3.1 Participant *i*'s strategy σ_i is non-increasing if $\sigma_i(k_i)$ first-order stochastic dominates (FOSDs) $\sigma(k_i)$ for each $k_i > k_i$.

This definition implies that if the strategy σ_i is non-increasing, then high degree participants will randomize their actions with less probability in $m_i = 1$ and with greater probability in $m_i = 0$. Therefore, if the strategy σ_i is non-increasing, the higher degree participants are less likely to acquire information.

Definition 3.2 Let $Br_i: \{\Delta\{0,1\}\}^{k_i} \to \Delta\{0,1\}$ denote the best response of Participant *i*, and we define

$$Br_i(\sigma_{N_i(g)}) = \left\{ \sigma_i(k_i) \in \arg\max_{m_i \in \{0,1\}} U(m_i, \sigma_{N_i(g)}; k_i), \forall k_i \in \{1, \dots, k_{\max}\} \right\}.$$

The best response is Participant *i*'s optimal strategy when her friends' decisions are given. Here $\Delta\{0,1\}$ is the set of strategies for Participant *i*, and $\{\Delta\{0,1\}\}^{k_i}$ is the profiles of strategies for the friends of Participant *i*. Then, a Bayes-Nash equilibrium is a strategy profile σ such that $\sigma_i \in Br_i(\sigma_{N_i(g)})$ for all $i \in N$. We define the greatest element in the set $Br_i(\sigma_{N_i(g)})$ as $\overline{Br_i}(\sigma_{N_i(g)})$. The next lemma characterizes the property of the greatest best response in the network game.

Lemma 3.1 The greatest best response $\overline{Br_i}(\sigma_{N_i(q)})$ is an decreasing function of $\sigma_{N_i(q)}$.

The lemma tells us that if more of Participant *i*'s friends acquire information, Participant *i*'s incentive to acquire information decreases. The intuition for this result is the free riding effect: Participant *i* becomes less willing to contribute when more of her friends acquire information. Proposition 3.1 gives us the basic Bayes-Nash equilibrium results.

Proposition 3.1 If Participant i's degree is independent from the degree of one of her randomly selected friends, there exists a symmetric Bayes-Nash equilibrium that is non-increasing in degree in the network game: There exists some threshold $k^* \in \{0, 1, 2, ...\}$, such that the probability $\sigma(m_i = 1 | \cdot)$ of choosing to acquire information satisfies

$$\sigma(m_i = 1|k_i) = \begin{cases} 1, \text{ for } k_i < k^* \\ 0, \text{ for } k_i > k^* \\ (0,1], \text{ for } k_i = k^* \end{cases}$$

Furthermore, the expected payoffs are non-decreasing in degree.

Proposition 3.1 has very clear implications. The participant's equilibrium action is weakly decreasing in her degree. In other words, the more friends she has, the less willing she is to acquire information. Participants can "free ride" on the actions of their friends. If Participant i has more friends, she is more likely to benefit from the signals passed around by her friends. Because the marginal effects of signals in forecasting are decreasing, participants with more friends are less willing to acquire costly information.

A few more remarks should be made here. Since all participants adopt a threshold strategy, Participant *i* believes that the probability of acquiring information for a randomly chosen neighbor is $\theta = \Pr(k_j \le k^*)$, $j \in N_i(g)$. Participant *i*'s belief about the number of informed neighbors follows a binomial distribution given by:

$$f(k_a; k_i, \theta) = \binom{k_i}{k_a} \theta^{k_a} (1 - \theta)^{k_i - k_a},$$

where k_a is the number of participants who acquire information, and $f(k_a; k_i, \theta)$ is the density function of the binomial distribution. Knowing the belief of Participant *i*, we can obtain the expected payoff $U(m_i, \sigma_{N_i(g)}; k_i)$. Since k^* is a threshold, it is determined by the following inequalities:

$$U(m_i = 1, \sigma_{N_i(g)}; k^*) \le U(m_i = 0, \sigma_{N_i(g)}; k^*),$$
$$U(m_i = 1, \sigma_{N_i(g)}; k^* + 1) > U(m_i = 0, \sigma_{N_i(g)}; k^* + 1).$$

The above inequalities simply mean that the k^* degree participant chooses to acquire information, and the $k^* + 1$ degree participant does not have incentive to acquire information.

After the decision on information acquisition, each participant submits the report. Under the forecastreport mechanism, each participant reports the best point estimation. Under the security-trading mechanism, each participant chooses the optimal risky asset demand. The hope of prediction markets is that by aggregating the private information of a large population, the principal can generate fairly accurate predictions of future events. How does the principal aggregate these small bits and pieces of relevant information that exist in the opinions and intuitions of diverse individuals? For a security-trading mechanism, it is simple. Prediction market price P is close to the mean belief of traders and thus can be interpreted as an aggregate forecast from prediction markets. Under a security-trading mechanism, the principal's prediction is the market price P. For a forecast-report mechanism, the principal can observe all of the participants' forecasts, but he knows nothing about V_0 , ρ_V , or ρ_{ε} , so the simple averaging rule offers a reasonable starting point. Under a forecast-report mechanism, the principal adopts a simple averaging rule and his prediction is $\frac{1}{n}\sum_{i=1}^{n} x_i^*$. Note that the simple averaging rule is optimal only when all participants' forecasts are independent and equally accurate; however, it is a good operational rule for limited information. Winkler (1971) finds the empirical evidence that supports the use of simple averaging: In examining the forecasts of the outcomes of football games, equal weighting is as accurate as weighting people's forecasts according to the previous accuracy or self-rated expertise. In our networked prediction markets, the principal has limited information: He does not know the social network graph. In this case, the principal cannot propose a weighted averaging rule, and simply follows the operational rule of thumb: "use equal weights unless you have strong evidence to support unequal weighting of forecasts" (Armstrong, 2001, pp. 422).

The Effects of Social Networks on Forecast Efficiency

The unique feature of our model of prediction markets is the information acquisition and exchange in social networks. A natural question arises: Can a social network promote the forecast efficiency in prediction markets? Social networks are credited with creating social contagion and driving product diffusion and peer adoption (e.g. Aral and Walker, 2011). In this section, we conduct numerical simulations to analyze the effects of information acquisition and social networks on the forecast efficiency of two different prediction market mechanisms: forecast-report and security-trading. The most striking finding is that social networks do not always improve the forecast efficiency. When the cost of information acquisition is low, a social network can promote forecast efficiency, but it could decrease the prediction performance if the cost of information acquisition is high.

Information Acquisition in Social Networks

To examine the forecasting performance, we conduct a variety of agent-based simulations in the social network-embedded prediction markets. In every simulation round, a random social network that includes 100 participants is generated, using a 100 × 100 dimensional matrix. Following the literature on random graph, we assume that the link between two participants is formed with independent probability p in our simulation. But in business practice, we can extract network degree distribution from aggregate-level data. Dover, Goldenberg, and Shapira (2012) show how to estimate the parameters of the network's degree distribution. We set the parameter values for the common prior $V \sim N(V_0, 1/\rho_V) = N(50, 1/0.4)$, and the noise of the signal $\varepsilon_i \sim N(0, 1/\rho_{\varepsilon}) = N(0, 10)$. The results are robust for other parameter values. Based on Proposition 3.1, we can compute the fixed point — the threshold degree k^* — and then study some comparative statics on information acquisition. In this subsection, we address some important questions in prediction markets: (1) How do social interactions between participants affect the acquisition and dissemination of information? (2) How does social network structure affect participants' information acquisition? These questions can be answered only in a framework of a network-embedded prediction market.

We study the effects of social networks under forecast-report mechanism first. The following corollary characterizes how cost of information acquisition c, precision of prior information ρ_V , and quadratic penalty b affect information acquisition.

Corollary 4.1 In forecast-report prediction market, the threshold k^* is decreasing in *c* and ρ_v , and increasing in *b* and *p*.

 k^* is the threshold degree of information acquisition. As k^* increases, the probability of acquiring information, $\Pr(k_i \leq k^*)$, increases. The intuition for the corollary is not straightforward as it seems to be. There are two effects of increasing the cost of information acquisition. As the cost increases, a participant has less incentive to acquire information by herself. It is the negative direct effect on the information acquisition. However, there is another positive indirect effect: As the cost increases, the friends of Participant *i* are also less willing to acquire information. Participant *i* cannot "free ride" on the actions of her friends, and it gives her incentive to acquire information. Corollary 2.1 says the negative direct effect dominates, and thus $\Pr(k_i \leq k^*)$, the probability of acquiring information decreases.

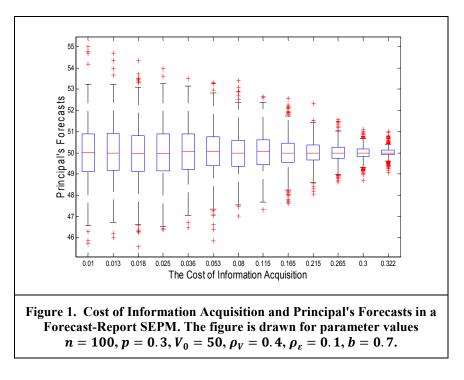
As for quadratic penalty *b* and the precision of prior information ρ_V , the intuition is straightforward: As the quadratic penalty increases, the signals become more valuable, and thus $\Pr(k_i \leq k^*)$ increases. An increase in the prior precision makes the signals less valuable. Thus, participants are less likely to acquire information, and k^* decreases.

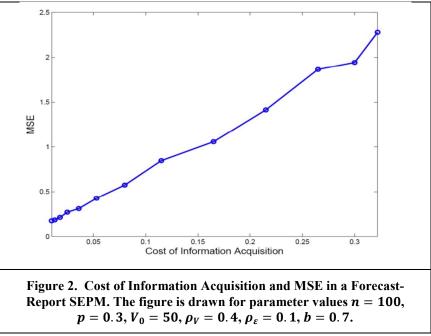
The link between two participants is formed with the independent probability p. If the probability that each pair of participants are connected increases, the networks will become more dense. An increase in network density changes the beliefs of Participant i in the network game. As the linking probability p increases, Participant i with degree k_i thinks that she will have more high degree friends. Because her friends acquire less information, Participant i has more incentive to acquire information, and k^* increases.

A Comparison of Forecast Performance: A Network-Embedded Market versus a Non-Networked Market

There are two key features of our prediction market model: Information acquisition is costly and endogenously determined in equilibrium, and the participants are embedded in a social network. The effects of social networks on forecast efficiency in prediction markets are related to the social value of information exchange and communication. Prediction markets work well under certain conditions and less well under others. Is a networked prediction market the right circumstance for the crowd to be wise? In this subsection, we try to find the conditions that are necessary for networked prediction markets to work well. The numerical simulations help us compare the prediction performances and the efficiency of information aggregation. Recall that for a forecast-report mechanism, the principal makes a combined forecast by using the statistical mean of all participants' forecasts. The mean square errors (MSE) of a principal's forecasts are our measure of prediction market performance. For simplicity, SEPM is short for a social-network-embedded prediction market, and NNPM is short for non-networked prediction market.

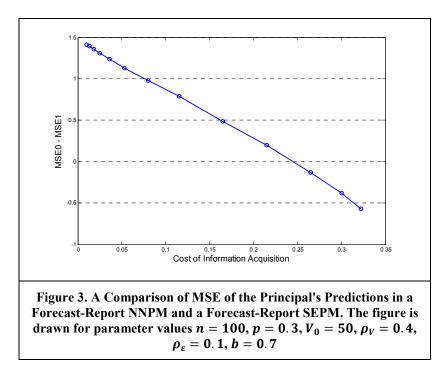
For each cost level of information acquisition, we run 1,000 round simulations in both SEPM and NNPM, and then we compute the estimated MSE of SEPM and NNPM under a forecast-report mechanism. Figure 1 plots the principal's forecasts (1,000 round simulations for each given cost level) against the cost of information acquisition. Figure 2 shows that MSE increases with c in a forecast-report SEPM. These two figures tell us that the variance of principal's forecasts is decreasing in c, but the MSE is increasing in c. Since MSE is equal to the sum of the variance and the squared bias of the principal's forecast, the squared bias must be increasing in c. The underlying reason is simple. When the cost of information acquisition becomes higher, participants tend to acquire information less and place more weights on the common prior. The principal aggregates all participants' forecasts and hence the variance of his forecasts for 1,000 round simulations decreases, and the bias increases as the cost of information acquisition increases.

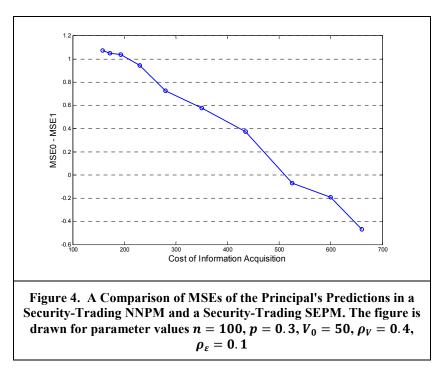




The comparison of MSE under a forecast-report mechanism is shown in Figure 3. *MSE*0 represents the MSE computed in the NNPM, and *MSE*1 represents the MSE in the SEPM. Figure 3 examines the impact of information acquisition cost on prediction market performances of SEPM and NNPM. When *c* is small, MSE0 - MSE1 > 0, which means a SEPM outperforms a NNPM. As *c* increases, MSE0 - MSE1 decreases, and when *c* is large enough, a NNPM performs better than a SEPM.

The pattern for a security-trading mechanism is similar. Figure 4 shows the comparison of MSE under a security-trading mechanism. When *c* is small, a SEPM outperforms a NNPM. As *c* increases, a NNPM performs better than a SEPM.

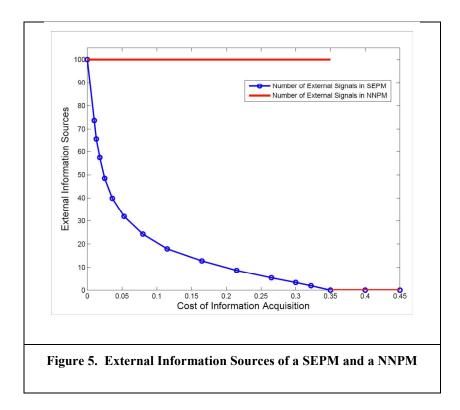




There are two implications in Figure 3 and Figure 4: First, when the information acquisition cost is low, social network and information exchange can promote forecast efficiency in prediction markets. Second, Information exchange could be detrimental to the prediction performance if the information acquisition cost is high. The second result is surprising. Unlike in the previous literature, it means more social communications could result in an inefficient information aggregation. Social communications serve as a focal point for the group and can have a detrimental effect on the performance of prediction markets. "Too much communication, paradoxically, can actually make the group as a whole less intelligent" (Surowiecki, 2004). The key is that in our present model, we study the effects of social networks on the

aggregate forecast instead of an individual forecast, so it is different from social contagion and product diffusion. Social networks could reduce people's incentive to acquire information and could then be detrimental to the forecast efficiency of prediction markets.

These implications are critical to understand how to use social networks to improve the performance of prediction markets. Our present results suggest the following guidance in business practice of prediction markets: When the predicted event is simple, which is interpreted as low information acquisition cost, we suggest using a social network-based prediction market. When it involves complicated issues, which can be interpreted as high information acquisition cost, the traditional non-networked prediction market is a better choice. In a laboratory experiment, Healy, Linardi, Lowery and Ledyard (2010) compares four different prediction mechanisms (the double auction, the market scoring rule, the parimutuel, and the poll). They find that the performance of the mechanisms is significantly affected by the complexity of the environment. Our study also has a similar managerial implication, but in terms of social networks. Whether or not to use social networks in prediction markets depends on the information acquisition cost.



Our simulation study identifies the conditions that are necessary for prediction markets to perform well. A further question is the underlying reason why social networks can promote forecast efficiency in some given cost environments but not in others. Figure 5 depicts the external information input of a SEPM and a NNPM under a forecast-report mechanism. The external information input is the number of independent signals acquired by the participants. For NNPM, each participant cannot observe others' signals; thus, all of them acquire information when the cost information acquisition is less than 0.35, and none of them acquire information when the cost is greater than 0.35. In this case, the external information input decreases with the cost of information acquisition. From Figure 5, it is not difficult to explain why a NNPM outperforms a SEPM when c is large. Although a social network can provide internal communications among participants, it reduces the incentive to acquire information because of free riding. As c increases (as long as $c \le 0.35$), the external information input of a SEPM decreases, but the external information input of a NNPM remains the same. Thus, when c is sufficiently close to 0.35, a NNPM outperforms a SEPM in terms of prediction accuracy.

Conclusion

In this paper, we develop a game-theoretic framework for the analysis of information exchange in social network-embedded prediction markets with endogenous information acquisition. Our model melds key aspects of strategic incentives in an incomplete information game with market equilibrium: Participants take into account incentives and the effects of their friends' information acquisition, and then they take prices as given and trade anonymously in prediction markets. In the network game, all participants adopt a threshold strategy, and the equilibrium action of information acquisition is non-increasing in participant's degree.

Unlike in much of the previous literature, we find that the use of social networks could be detrimental to forecast performance. Our results provide important managerial implications in business practice of prediction markets: When the predicted event is simple, we suggest using a social-network-embedded prediction market. When it involves complicated issues, the traditional non-networked prediction market is a better choice. We also discuss the effects of network structure on information acquisition in prediction markets. As network density increases, the incentive to acquire information increases. The above results are robust to two commonly used mechanisms of prediction markets: a forecast-report mechanism and a security-trading mechanism.

A future research direction is to examine the incentives for sharing information in a social network. In the present setup, we assumed that participants exchange information according to reciprocity and norms of fairness, rather than focusing on the incentives for sharing information. However, whether participants have incentives to share information, and what those incentives are, remain open questions. A well-developed literature of information sharing has been developed in the setup without networks. For example, Gal-Or (1985) shows that no information sharing is the unique equilibrium in an oligopolistic market. Studying the incentives for sharing information in social network-embedded prediction markets would be an interesting extension of information sharing. In this paper, we assume that social networks are exogenous. It is interesting to consider a prediction market model with endogenous network formation. Another future research direction is the interaction of different social networks. Participants in prediction markets might not be restricted to acquiring information from one social network also depends on the costs occurred in other social networks. The costs for a piece of information in two social networks for one participant could be different due to distinct characteristics of network structure.

Appendix: Proof

Proof of Lemma 3.1

Proof. Reversing the order of the strategy and the type space of Participant *i*, we can show that $u(m_i, m_{N_i(g)})$ is supermodular in (m_i, k_i) and has increasing differences in $(m_i, m_{N_i(g)})$. Using Proposition 8 in Van Zandt and Vives (2007) and reversing the order back, we obtain that $\overline{Br_i}(\sigma_{N_i(g)})$ is an decreasing function of $\sigma_{N_i(g)}$.

Proof of Proposition 3.1

Proof. The results follow from Lemma 3.1.

Proof of Corollary 4.1

Proof. The payoff function of our model satisfies monotone comparative statics (MCS) in games with strategic substitutes (Roy and Sabarwal, 2010). Thus, the results follow from Theorem 3 in Roy and Sabarwal (2010).

References

- Aral, S. and Walker, D. 2011. "Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks," *Management Science, Forthcoming*.
- Armstrong, J. S. 2001. Combining forecasts. Principles of Forecasting: A Handbook for Researchers and Practitioners, Norwell, MA: Kluwer Academic, pp. 417-439.
- Banerjee, S., Kaniel, R., and Kremer, I. 2009. "Price Drift as an Outcome of Differences in Higher Order Beliefs," *Review of Financial Studies*, (22:9), pp. 3707-3734.
- Berg, J. E., Neumann, G. R., and Rietz, T. A., 2009. "Searching for Google's Value: Using Prediction Markets to Forecast Market Capitalization Prior to an Initial Public Offering," *Management Science*, (55:3), pp. 348-361.
- Bradshaw, T. and Khan, M. 2011. "Facebook voting comes to UK's 'Big Brother'," *The Financial Times*, August 27, available at

http://www.ft.com/intl/cms/s/0/dfbc94b8-cffa-11e0-81e2-00144feabdc0.html.

- Dover Y., Goldenberg, J., and Shapira D. 2012. "Network Traces on Penetration: Uncovering Degree Distribution from Penetration Data," *Marketing Science, Forthcoming*.
- Fang, F., Stinchcombe, M., and Whinston, A. B. 2007. "Putting Your Money Where Your Mouth Is --Betting Mechanism Design for Better Prediction," *Review of Network Economics*, 6(3), pp. 214-238.
- Fang, F., Stinchcombe, M., and Whinston, A. B. 2010. "Proper Scoring Rules with Arbitrary Value Functions," *Journal of Mathematical Economics*, (46:6), pp. 1200-1210.
- Galeotti, A., Goyal, S., Jackson, M. O., Vega-Rendondo, F. and Yariv L. 2010, "Network Games," *Review* of Economic Studies, (77:1), pp. 218-244.
- Gal-Or, E. 1985. "Information Sharing in Oligopoly," *Econometrica*, (53:2), pp. 329-343.
- Golub, B., and Jackson M. 2010. "Naive Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, (2:1), pp. 112-149.
- Guo, Z., Fang, F., and Whinston, A. B. 2006. "Supply Chain Information Sharing in a Macro Prediction Market," *Decision Support Systems*, (42:3), pp. 1944-1958.
- Jian, L., and Sami, R. 2012. "Aggregation and Manipulation in Prediction Markets: Effects of Trading Mechanism and Information Distribution," *Management Science* (58:1), pp. 123--140.
- Healy, P. J., Linardi S., Lowery, J. R., and Ledyard J. O. 2010. "Prediction Markets: Alternative Mechanisms for Complex Environments with Few Traders," *Management Science*, (56:11), pp. 1977-1996.
- Le Bon, G. 1896. The Crowd: A Study of the Popular Mind, New York, NY: The Macmillan Co.
- Ottaviani, M., and Sorensen P. N. 2009. "Surprised by the Parimutuel Odds?" American Economic Review, (99:5), pp. 2129-2134.
- Qiu, L., Rui, H., and Whinston, A. B. 2011. "A Twitter-Based Prediction Market," Proceedings of the 32th International Conference on Information Systems (ICIS).
- Roy, S., and Sabarwal T. 2010. "Monotone Comparative Statics for Games with Strategic Substitutes," *Journal of Mathematical Economics*, (46:5), pp. 793-806.
- Sundararajan, A. 2008. "Local Network Effects and Complex Network Structure," B.E. Journal of *Theoretical Economics*, (7:1), Article 46.
- Surowiecki, J. 2004. The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations, New York, NY: Doubleday.
- Van Zandt, T. and Vives X. 2007. "Monotone Equilibria in Bayesian Games of Strategic Complementarities," *Journal of Economic Theory*, (134:1), pp. 339--360.
- Winkler, R. L. 1971. "Probabilistic prediction: Some experimental results," *Journal of the American Statistical Association*, (66:336), pp. 675-685.