2008

Group Prediction in Information Markets With and Without Trading Information and Price Manipulation Incentives

Dorina Tila

David Porter
Chapman University, dporter@chapman.edu

Group Prediction in Information Markets With and Without Trading Information and Price Manipulation Incentives

Comments
Working Paper 08-04

Group Prediction in Information Markets With and Without Trading Information and Price Manipulation Incentives

Dorina Tila

George Mason University
Interdisciplinary Center for Economic Science
4400 University Drive, MS 1B2
Fairfax, VA 22030
Email: dtila@gmu.edu

David Porter

Chapman University
One University Drive, BK 308D
Orange, CA 92866
Email: dporter@chapman.edu

Abstract: The ability of individuals and groups to forecast a future event, with incomplete information, by using the trading history of an asset market is analyzed in the laboratory. The results show: (1) when forecasters observe the summary of market-transacted prices, they do not perform as well as when they are provided with a complete real-time sequence of bids, asks and contract prices; (2) groups do not outperform individuals in forecasting, and when the market does not have price manipulation incentives, individual prediction is better than the group prediction; (3) in markets with manipulators, where only a summary of contract prices is provided, both groups and individuals are unable to predict better than flipping a coin. This inability to aggregate information is remedied when forecasters see the complete evolution of market bids, asks and contracts.

Keywords Information Aggregation, Prediction Markets, Futures Markets

1.0 Introduction

According to the efficient market hypothesis, market prices are an accurate indicator of the true value of traded assets since all publicly available past and current information is absorbed by the prices through the market mechanism. Fama's 1970 article provides strong empirical support for the efficient market hypothesis.[1] Information markets have captured the interest of a large number of scholars who have tested its characteristics through experimental, theoretical, and empirical models. Several theoretical studies have shown that market prices reflect the collective information of the system as the efficient market hypothesis claims. However, there is a long list of reasons that might lead prices to imperfectly aggregate information, such as costly information (Grossman and Stiglitz, 1980), dependence on traders' beliefs, budgets, and risk preferences (Manski, 2005).

The claim that information aggregation is reflected in market prices has been tested in the laboratory as well as in the field, with mixed results. For example, Plott and Sunder (1988) find a convergence to rational expectation (RE) equilibrium in contingent claims or single-security markets with the same preferences across traders, but failure of convergence in the single-security markets when traders have diverse preferences. Plott and Sunder (1982) also find full convergence to RE prices when insiders are fully informed and failure when insiders have uncertain information about the state of nature. Many features of the market can potentially play a role in hindering information aggregation. Such limitations are known as “information traps” (Noeth, Camerer, Plott, and Webber, 1999) or “information cascades” (Holt and Anderson, 1997; Plott and Hung,

[1] The article claims that security markets are extremely efficient in reflecting the complete and accurate information about the fundamental asset value.

2001). In addition, the presence of manipulators seems to have successfully influenced prices in the 1999 Berlin election market on the Iowa Electronic Markets (IEM) (Hansen, Schmidt, and Strobel, 2004). However, Oprea et al. (2006), found that manipulation did not damage the information content of prices in a laboratory environment with manipulation incentives. Manipulation seems to depend on the conditions in the environment (Chakraborty and Yilmaz, 2004).[2]

In addition to experiments in a laboratory setting, field experiments have been conducted. Camerer (1998) showed that efforts to manipulate odds in paramutual betting at racetracks failed. Results from the IEM have shown that these markets outperform polls (Forsythe, Nelson, Neumann, and Wright, 1992; Rhode, Koleman, and Strumpf, 2003). Plott and Chen (2002) have shown that internal prediction markets at Hewlett-Packard have outperformed the company's standard estimation to forecast its printer sales.

These studies of information markets have tested the ability of prices to represent the collective information of the crowd. This literature has expanded to include how decision-makers would interpret the information they observe from market prices to predict future events. This question has been analyzed in the laboratory by Oprea et al. (2006) who found that forecasters use market information in their forecasts and these predictions are extremely accurate, even when some traders have incentives to manipulate the market price. This paper examines the prediction quality of forecasters under a variety of different treatments. In particular, we examine decision-making when only the history of contract prices is provided versus a treatment in which the complete

[2] The article shows that given a long enough horizon, manipulators may trade against their information and undertake short-term losses.

time sequence of bids, asks, and contracts is provided to forecasters. This treatment is based on the observation that in typical field prediction markets, only price and volume history is routinely provided to individuals and not the full information of offers to buy and sell for each asset unit.

In addition, we examine how forecast quality is impacted when predictions are made by either individuals or groups. In practice, many decisions in government, business firms, and family are made by a group rather than an individual. The experimental literature has found that individuals and groups behave differently in strategic games, where groups are considered more "rational" than individuals as their decision is more aligned with the game theoretic solution. This hypothesis is shown in several strategic games such as the centipede game[3] (Bornstein et al., 2004), one-shot ultimatum game[4] (Bornstein and Yaniv, 1998), trust game[5] (Cox, 2002), one-shot gift-exchange game[6] (Kocher and Sutter, 2002), beauty-contest games, where groups exhibit faster learning than individuals[7] (Kocher and Sutter, 2005).

In contrast with these findings, in dictator games, groups have a higher level of sharing than individuals, which departs from the theoretical solution (Cason and Mui, 1997). In a strategic market game, such as common value auctions, groups are found less rational than individuals, and their performance deteriorates when there are more signals

[3] The game theoretic solution, through backwards induction, is for player 1 to end the game at the first node. While both individuals and groups failed to end at node one, groups on average exit the games earlier than individuals.
[4] When the decision maker is a three-person group, player 1 sends a lower amount, and player 2 has lower rejection rate.
[5] While no significant difference is found in the sender's behavior, the group responder's behavior is closer to the game theoretic solution (send nothing back).
[6] Player 1 decides the 'gift' or the wage level, and player 2 decides the effort level. Groups choose smaller wage levels and lower effort levels, which is closer to the game theoretic solution.
[7] There is no difference on average between the choices of inexperienced decision maker types: group and individual. However, in repeated games, groups were faster learners of the dynamics of the game and outperformed individuals.

available (Cox and Hayne, 2006). Other findings have shown no significant difference between group and individual behavior (Prather and Middleton, 2001[8]).

The results of our experiments are clear. Individuals make better forecasts than groups, and access to the real time sequence of bids, asks and contracts as opposed to just a history of contracts increases forecast accuracy.

2.0 Experimental Design

This section will provide the design of the two market information treatments. The difference in the two treatments is the amount of market information provided to the neutral forecasters in the market. These forecasters are neutral because they do not posses any private information. In one treatment, they observe market information through the real time sequence of bids, asks, and contracts; in the static treatments they observe only the contract prices. In addition, with the limited market information treatment, we examine the forecasting accuracy of groups versus individuals. In both treatments we add the possibility of price manipulation to determine its effect on prediction accuracies.

2.1 Design of Baseline Market Information Treatment

A prediction market was created with eight traders. Traders were endowed with a fixed amount of cash and a fixed number of tickets. Tickets had a life of one round, and at the end of the round they would generate a dividend of either 0 or 100 with a prior of equal occurrence. Throughout the round, subjects did not know the actual dividend of the

[8] The empirical results from this paper are unable to show a performance difference between group-managed and individually managed funds from September 1981-1994.

ticket, but they did receive a clue about its true value. First, with equal probability, we randomly and privately selected one of the two values (states) $v=\{0,100\}$. Conditional upon the state, each trading participant received a clue (Black or White[9]) randomly selected (with replacement) from a distribution where two out of three times the clue would comply with the realized value of the ticket. These conditional probabilities are provided in equations (1) and (2).

$$\Pr(\mathit{Clue} = \text{Black} \mid \text{v} = 100) = \Pr(\mathit{Clue} = \text{White} \mid \text{v} = 0) = \frac{2}{3} \qquad (1)$$

$$\Pr(\mathit{Clue} = \text{Black} \mid \text{v} = 0) = \Pr(\mathit{Clue} = \text{White} \mid \text{v} = 100) = \frac{1}{3} \qquad (2)$$

In addition to the traders, five uniformed forecasters where able to view the market activity, which included all the transaction prices and offers to buy and sell submitted by the traders as they occurred. These sets of experiments will be referred to as real time markets and constitute our baseline treatment. The five forecasters had no private information, i.e. they were given no clues. At the end of each round, forecasters made a private prediction about the value of the ticket only observing the real-time market transactions and knowing the general clue structure described in (1) and (2).

Forecasters were paid based on the accuracy of their prediction and traders were paid based on the value of the tickets they held and their remaining cash as shown in equations (3) and (4) respectively.

---

[9] A Black clue has a $\frac{2}{3}$ chance of being associated with the state dividend of 100, and a White clue has a $\frac{2}{3}$ chance of being associated with the state dividend of 0 (zero) as shown in (1) and (2).

$$\textit{Forecaster Payoff} = \begin{cases} \textit{250, if forecaster is correct [prediction = actual state v]} \\ \textit{0, if forecaster is incorrect [prediction} \neq \textit{actual state v]} \end{cases} \tag{3}$$

$$\textit{Trader}_i \textit{ Payoff} = C_i - \sum_{j=1}^{J_i} B_{ij} + \sum_{k=1}^{K_i} S_{ik} + v(N_i + J_i - K_i) \tag{4}$$

Where:

C_i = Endowed Cash for Trader i (=200)

N_i = Endowed Tickets for Trader i (=2)

J_i = Number of Tickets Trader i *buys* in the Market

K_i = Number of Tickets Trader i *sells* in the Market

B_{ij} = Price of Contract j Purchased by Trader i

S_{ik} = Price of Contract k Sold by Trader i

The real time market provides the baseline treatments, and the static time market with individual and group forecasters provide the other treatments. Each is explored in the presence and absence of manipulation incentives. As shown in Table 1, the difference between the non-manipulation and manipulation markets is that in the latter, half of the traders are given an additional incentive to affect the forecaster predictions. The additional financial incentive for manipulators, which was added to (4), is given by equation (5) where T (0,100) is the prediction target given to a manipulator and F_j is the prediction of forecaster *j*:

$$Additional \quad Payoff = 200 - 2\,|T - \frac{1}{5}\sum_{j=1}^{5} F_j\,| \qquad (5)$$

Thus, if the forecasters' predictions match the manipulator's target, the manipulator obtains an additional payoff of 200. The closer the forecasters' predictions are to the target, the more a manipulator obtains. Hence, the manipulators have an incentive to affect market prices in order to lead the forecasters to provide a prediction closer to their target. Table 1 provides the experimental treatments of the baseline experiments.

Table 1: Market Types in Baseline Treatments

Market Types	Real-Time Market (BASELINE)
Non-manipulation	*8 Traders* *5 Forecasters*
Manipulation	*8 Traders →* *4 Traders* *4 Manipulators* *5 Forecasters*

The only difference between non-manipulation and manipulation treatments is the switch of half of the trader roles to manipulators. Manipulators have an additional financial incentive to affect the forecasters' predictions as shown in equation (5).

The real-time market experiment findings are as follows[10]: (1) manipulators attempt to manipulate prices; (2) manipulators succeed in increasing average contract prices by 7 points over the non-manipulation treatment when the target is *100*; (3) prices are correlated with the information in the system despite the efforts of manipulators; and (4) forecasters' predictions are a better estimate of the true state than market prices. The RE model provided a reasonably accurate summary of the market behavior, although prices did not fully converge to the theoretical Bayesian posterior probability. Even

[10] These results are taken from Oprea et al. (2006).

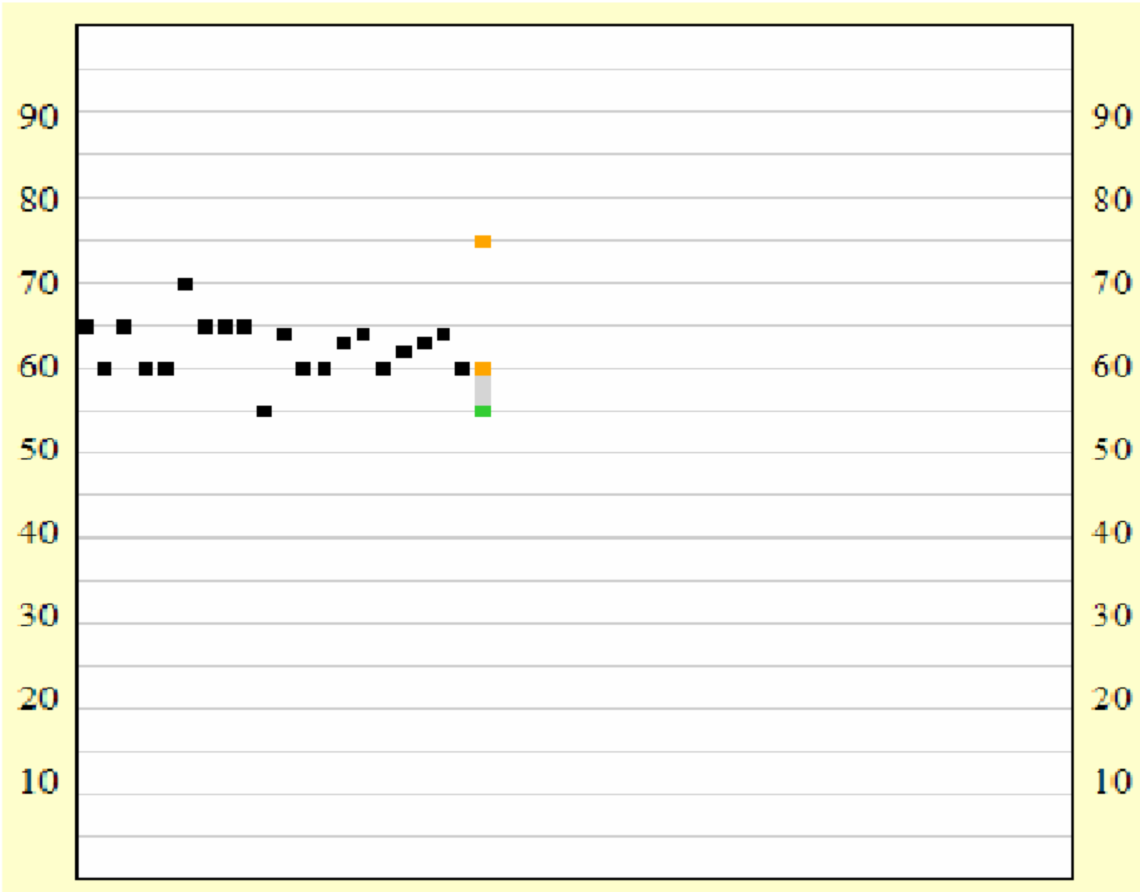
though market prices were not closely correlated with the true state, forecasters made use of them to improve their prediction quality.

2.2 Design of Static Market Treatments

This paper will extend the previous experiments by limiting the information that forecasters have in observing the market. In the new experiments, denoted as <u>static market information</u>, forecasters observe only the history of the market prices as opposed to the real-time markets, where forecasters were provided with complete information of how these prices are reached, through offers to buy and sell in real time. Specifically, for each session that was conducted in real-time trading, the contract price history was retrieved and displayed to the individual and group forecasters. Figure 1 provides an example of a contract series shown to subjects. In the static experiments, the two types of markets, with and without manipulation, are replicated with individual and with three-member group forecasters. The payoff function of forecasters is the same as in baseline treatments as shown in equation (3).[11] Table 2 provides the experimental treatments of our investigation.

[11] When the three-member group prediction matched the realized state, each member of that group received a compensation of 250 as shown in equation (3).

Figure 1: Screenshot Provided to Forecasters in the Static Market Treatment



The black dots show the contracts, orange dots are the last asks, and the green dots are the last bids before the market closed.

Table 2: Matrix of Treatments

Market Type	**Real Time**	**Static Time**	**Static Time**
Decision Maker Type	**Individual Prediction**	**Individual Prediction**	**Group Prediction**
Non-Manipulation	IN_R	IN_S	GN_S
Manipulation	IM_R	IM_S	GM_S

The baseline treatment is represented by individual forecasters without manipulation who have access to real time market information (IN_R). An additional treatment in the baseline is when manipulators are added to the environment (IM_R). Our treatments are represented by the cases where individual forecasters observe only the price history (static market information) from previous non-manipulated (IN_S) and manipulated markets (IM_S). There are also treatments with group predictions where group forecasters have access to static market information from previous non-manipulated (GN_S) and manipulated markets (GM_S).

In the real-time experiments, three sessions were run for each market. Each session had 16 separate prediction market rounds. Since the history of the market prices produced by these traders was used and shown to the forecasters in the static treatments, three sessions were run for each treatment in the static-time experiments with 16 separate prediction rounds. In addition, three sessions each with the same 16 rounds were used for both individual and group predictions.

The parameters of the information structure in our experiments are shown in Table 3. Each manipulator was given the same target in each round; half of the time it was the same as the actual state, and the other half it was the opposite of the actual state.

Table 3: Parameter Table

Round	Positive Signals	Bayesian Decision	Signal Strength	Target	Actual State
1	6	100	2	0	100
2	1	0	3	0	0
3	2	0	2	0	0
4	3	0	1	100	0
5	5	100	1	0	100
6	0	0	4	100	0
7	3	0	1	0	0
8	7	100	3	0	100
9	3	0	1	100	100
10	1	0	3	100	0
11	4	-	0	0	100
12	2	0	2	100	0
13	6	100	2	100	100
14	4	-	0	100	100
15	5	100	1	100	100
16	2	0	2	0	0

Positive signals correspond to the number of black clues (n) that are assigned to the 8 traders. Signal strength is defined as $s = |n - 4|$. The Bayesian decision is the binary (0,100) prediction a forecaster would make if he had all the clue information available to him. The target is the number given to manipulators that determine their bonus for moving forecaster decisions closer to the target.

The Bayesian decision calculated in equation (6) is defined as the choice a forecaster would make if he could see all the clues distributed for the particular round. We define a positive signal as a Black clue which has a $\frac{2}{3}$ chance of being associated with the state dividend of 100. The total number of traders in each session was eight; hence the maximum number of positive signals *(n)* in a round was 8. We define signal strength (s) as $s = |n\text{-}4|$. Thus, signal strength varies from 0 to 4. When $n = 4$, meaning there are 4 positive clues out of 8, the signal strength is 0, and the Bayesian expected dividend (V) of the ticket would be 50.

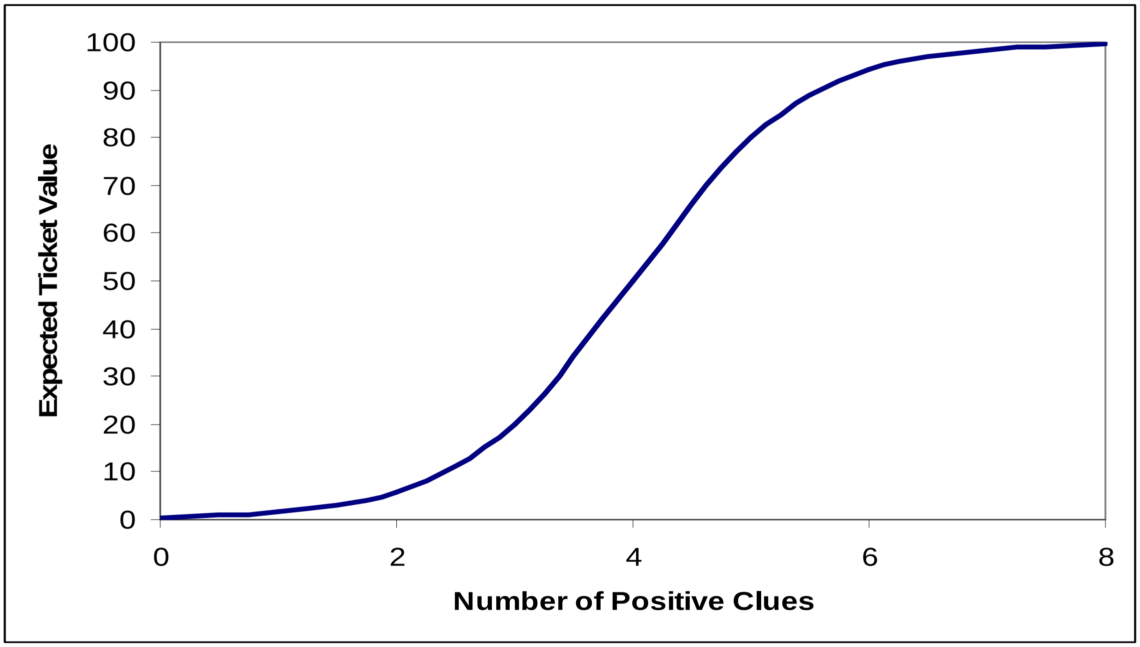
$$\text{Bayesian Decision} = \begin{Bmatrix} 100 \text{ if } \operatorname{prob}(V = 100 \mid n) > 0.50 \\ 0 \text{ if } \operatorname{prob}(V = 100 \mid n) < 0.50 \end{Bmatrix} \tag{6}$$

Equation (7) provides the expected dividend value as a function of the number of positive signals in the market:

$$EV(n) = \frac{(\frac{2}{3})^{n} (\frac{1}{3})^{8-n}}{(\frac{2}{3})^{n} (\frac{1}{3})^{8-n} + (\frac{1}{3})^{n} (\frac{2}{3})^{8-n}} 100 \tag{7}$$

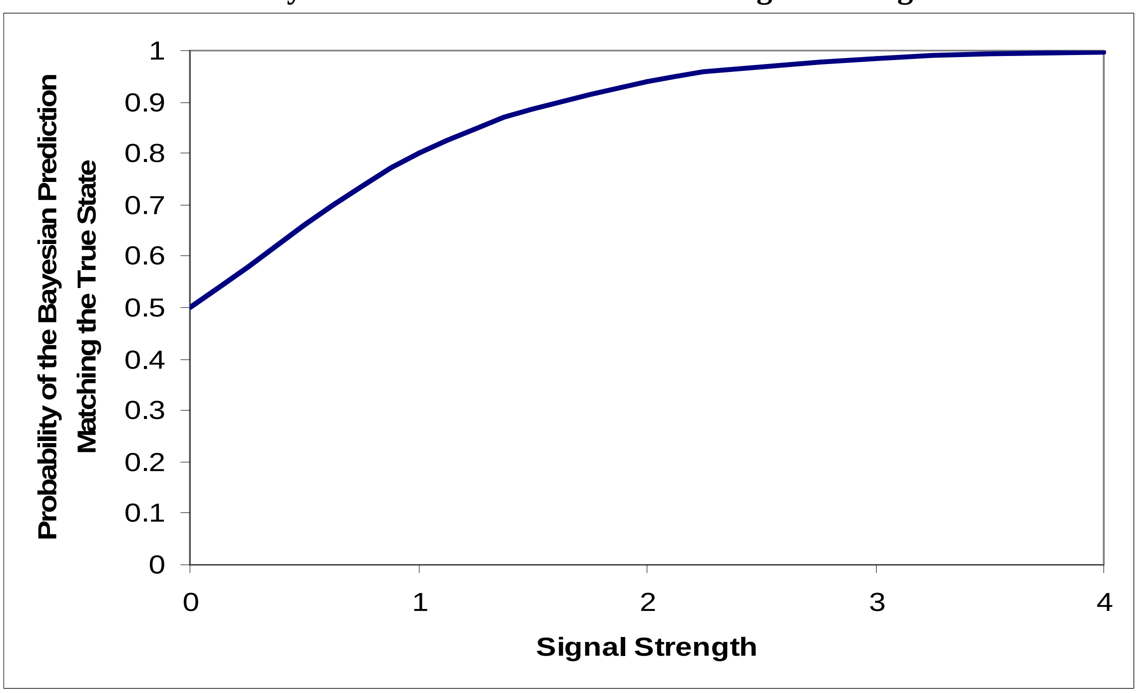
Graph A in Figure 2 charts the Bayesian expected value as a function of positive clues n while Graph B charts the posterior Bayesian probability of predicting the actual value as a function of signal strength s. For example when $n = 8$ (or $s = 4$), the posterior probability of the Bayesian decision being correct (predicting the value to be 100) is 99.6%. When $n=0$ (or $s=4$) the posterior probability of the Bayesian decision being correct (predicting the value to be 0) is 99.6%.

Figure 2 (Graph A): Expected Ticket Value as a Function of the Number of Positive Clues



For each total number of positive clues in the market, the expected dividend value of a ticket is calculated using equation (7). Given that the dividend values are either 0 or 100, if the market were fully aggregating information, under risk neutral assumptions, market price prediction should follow this function.

Figure 2 (Graph B): Probability of Predicting the True Dividend Value Using the Bayesian Decision as a Function of Signal Strength



Each level of signal strength (|n-4|), is charted against the probability that the Bayesian decision listed in equation (6) will accurately predict dividend value of a ticket.

The Bayesian probability of an accurate prediction will be our theoretical benchmark to be compared with the prediction accuracy forecasters. The accuracy of forecasters' predictions in a round is calculated in (8):

$$Prediction \quad Accuracy_{jt} = \frac{1}{5}\sum_{i=1}^{5}[1 - \left|\frac{Bayes \quad Decision_{jt} \ - \ Prediction_{ijt}}{100}\right|] \qquad (8)$$

Where:

Bayes Decision_{jt} = What a Bayesian would predict for the state if he had all the clues of session j in round t

Prediction_{ijt} = The actual dividend prediction of forecaster i of session j in round t

Using the design of the prediction markets, we can compare the average correct prediction, as calculated in equation (8), for individual and group forecasters to the Bayesian probabilities of an accurate prediction as displayed in Graph B in Figure 2. If the accuracy of forecasters on average is positively related to the signal strength, then we can deduce that the forecasters are effectively using the market to predict. In addition, the prediction of the individual forecasters with real-time market information will be compared to the prediction of individual and group forecasters with static market information in order to observe any changes in the prediction quality.

3. 0 Experimental Questions and Procedures

In Oprea et al., 2006, it was found that manipulators affected the contract prices by increasing the average contract by *7* when the target was 100, and not affecting prices when the target was 0. However, the effect of manipulators was stronger in the bids and asks compared to realized prices. Bids were significantly higher when the target was 100, and asks were significantly lower when the target was 0. Thus, manipulators tried to influence price through bids and asks but this did not have an effect on forecaster accuracy. Limiting information to only contract prices would not convey to forecasters this attempt to manipulate through bids and asks. The question we wish to address is whether this lost information will have an impact on prediction quality.

The importance of bids and asks in providing information to participants has been previously discussed by Plott and Sunder (1988), who offer it as one explanation for the better performance of contingent markets relative to single-security markets. If the claim that bids and asks constitute important information to the uninformed forecaster is correct, then we should find diminishing accuracy of forecasters with static market information compared to the real-time market information treatment. In particular, this paper focuses on three main questions:

Question 1*: Does the prediction quality of individual forecasters improve when they observe the real-time evolution of the market trades instead of the price history?*

Question 2*: Are predictions more accurately provided by groups or individuals?*

Question 3: *Is prediction accuracy affected by the presence of manipulators?*

The first question will be explored by comparing the data from real-time information treatment versus data from the static information treatment. The second question will compare the difference between individual and group predictions. The third question will compare predictions of individuals and groups in the non-manipulation treatment to those in the manipulation treatment.

3.2 Experimental Procedures

Subjects were recruited from the undergraduate pool of students at George Mason University.[12] All of the subjects had the role of forecasters and their earnings structure was the same as the previous experiments.[13] The procedures are the same as the ones followed in the first set of experiments. The method of information distribution among the traders, who had generated the contract prices that were given to the forecasters, was explained in detail to the forecasters, paralleling the same process as in the real-time treatments.

Each experiment consisted of written instructions that were read aloud, hands-on demonstration of how clues were generated, two unpaid practice rounds and sixteen paid rounds of decision making.[14] Each session lasted approximately 40 minutes.

4.0 Experimental Results

The purpose of this study is to analyze the prediction quality of unbiased forecasters when they only observe market price history. The quality of their prediction will be analyzed along two dimensions. First, we ask the question whether the forecasters' prediction quality changes when moving from real to static market information. The second dimension comprises the actual Bayesian decision. While the first dimension distinguishes prediction quality relative to the real-time markets, the second dimension distinguishes prediction quality relative to the prior, which is *50-50*.

[12] Subjects were recruited randomly from a database, excluding students who had participated in the first set of experiments in Fall 2005.

[13] At the end of the experiment, subjects were privately paid their earnings, and for a 40-45 minute experiment, they received $17.25 on average, in addition to a $5.00 show-up fee.

[14] Instructions and procedures can be found at http://ices2.gmu.edu/dorina

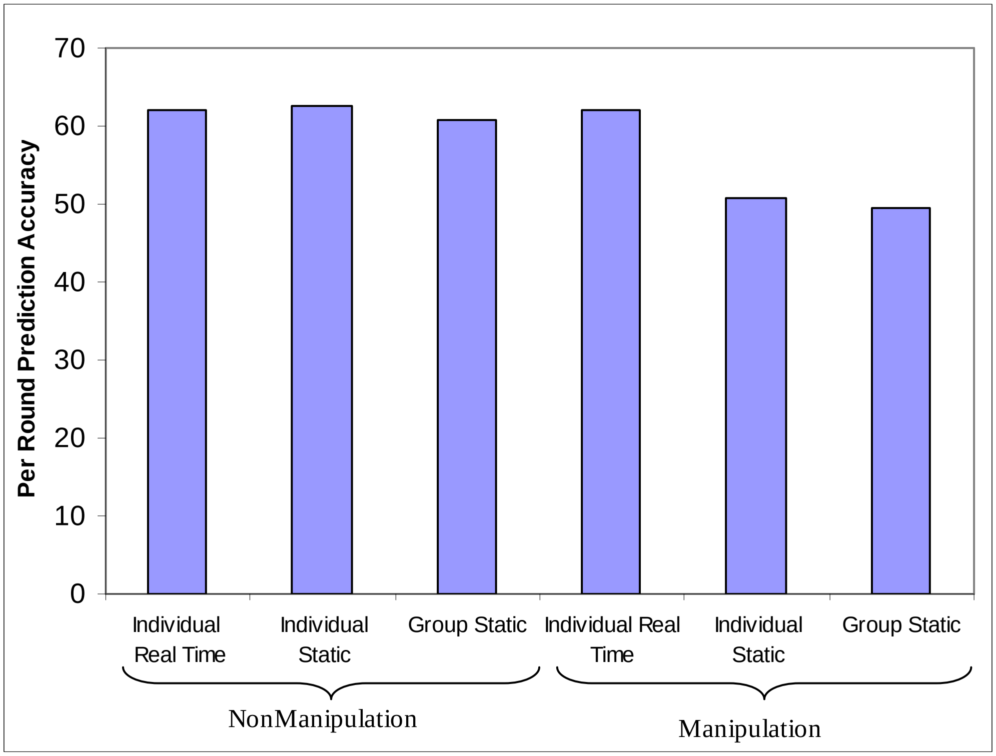
4.1 Relative Prediction Quality of Forecasters

We define prediction accuracy for round t of treatment k through equation (9) where i indexes the forecaster and j denotes the session:

$$Prediction\ Accuracy_{tk} = \frac{1}{15}\sum_{j=1}^{3}\sum_{i=1}^{5}[1-\left|\frac{Bayes\ Decision_{kjt} - Prediction_{kjt}}{100}\right|] \quad (9)$$

Figure 3 charts the per-round prediction accuracy of correct forecasts for each treatment. This figure suggests that the best predictors are the forecasters who observe the real time evolution of the prices in a market without the presence of manipulators. Qualitatively, from Figure 3, individuals predict better than groups and predictions are more accurate with real time information. However, in order to answer our questions quantitatively, we will take a closer look at the data by decomposing these aggregates to the particulars of the market information available in each round and session.

Figure 3: Per Round Prediction Accuracy by Treatment



The treatment prediction accuracy is averaged across all rounds to obtain the percentage of correct predictions per round by the forecasters. Qualitatively, real-time information improves forecast quality and groups do not outperform individuals in predicting the state.

In order to determine whether there is any difference among the treatments, a two-sample Kolmogorov-Smirnov (K-S) test for equality of distribution functions was conducted. The K-S tests in table 4 show that the samples of all treatments come from statistically different distributions.

Table 4: Two-sample Kolmogorov-Smirnov

Treatments	IN_R	IN_S	GN_S	IM_R	IM_S	GM_S
IN_R		0.047	0.000	0.000	0.000	0.000
IN_S			0.000	0.047	0.000	0.000
GN_S				0.000	0.000	0.000
IM_R					0.000	0.000
IM_S						0.047
GM_S						

The table lists the p-values from a two-sample Kolmogorov-Smirnov (K-S) test for equality of distribution functions. In these pairwise comparisons using the K-S test, the null hypothesis of equality of distributions is rejected for all treatments.

4.2 Do Forecasters Aggregate Information in a Static Market?

The odds ratio, as defined in equation (10) below, is the ratio of the probability P of correctly predicting the realized state and the probability (1- P) of incorrectly predicting the realized state. Hence, when the odds-ratio is one, forecaster predictions are correct as many times as they are incorrect, and when the odds-ratio is greater than one, forecasters are correct more often than incorrect. Specifically, P is the amount defined equation (8). Thus, for each session and round we have an observation on P.

$$\text{Odds-Ratio} = \frac{P}{1-P} \qquad (10)$$

The distribution of odds-ratios for each treatment is provided in Figure 4. If the distribution of odds-ratio is skewed to the right (below 1), then it can safely be concluded that forecasters are predicting no better than their prior of 50-50. This seems to be the case for the manipulation treatments with static information for both individual and group forecasters.

In order to determine if forecasters are indeed aggregating information we need to examine prediction behavior as the signal strength changes. The theoretical functional form between the probability of correctly predicting the state and signal strength is shown in equation (11), which is derived from equation (7).[15]

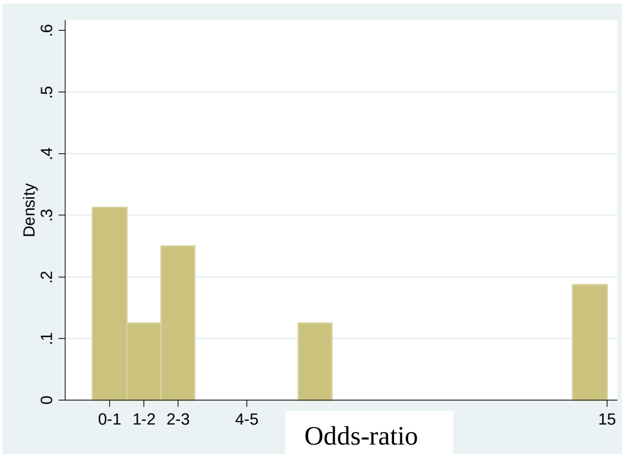
$$\ln\left(\frac{P(s)}{1-P(s)}\right) = \ln(4) * s \text{ , where s is signal strength} \qquad (11)$$

From equation (11)[16], we can derive the values of the odds-ratio depending on the signal strength. For instance, when s = 0, the natural log of odds-ratio is 0 and the odds-ratio is 1. As long as the odds-ratio is greater than one, forecasters are correctly predicting the state at a rate higher than the 50-50 prior. If the odds-ratio increases as the signal strength increases, then it can be safely inferred that forecasters are aggregating this information in their predictions. The further apart from the true functional form, the further apart this prediction is from being efficiently aggregated.

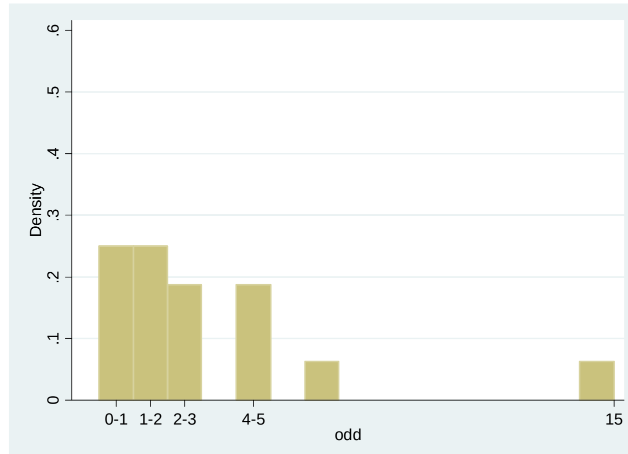
[15] Details of this derivation can be found in the Appendix.
[16] P(s) is defined as *Prediction Accuracy*$_{jt}$ *(P)* now as a function of signal strength s.

Figure 4: Distribution of the Odds-Ratio by Treatment

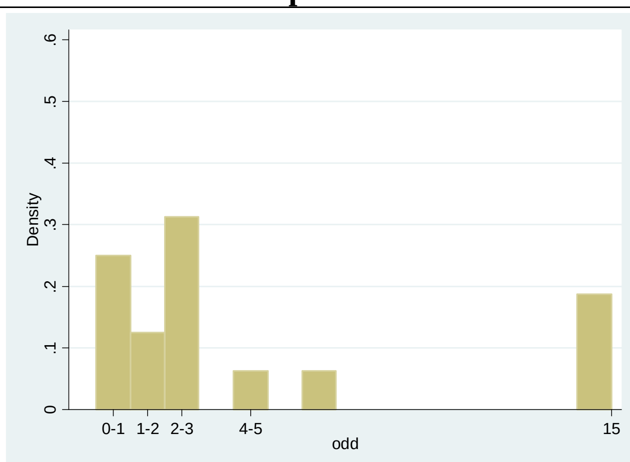
Individual Non-manipulation Real Info



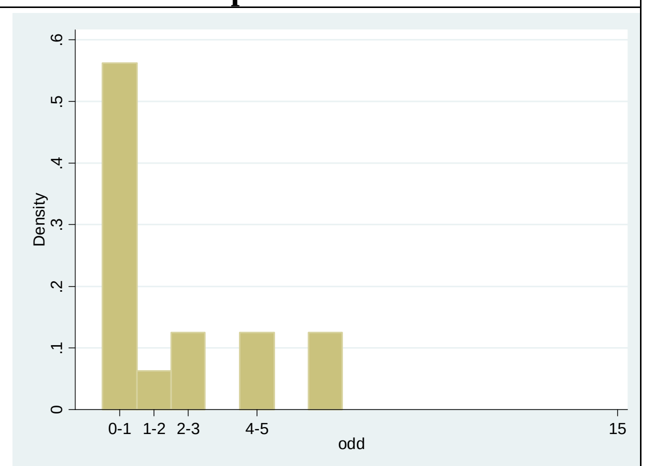
Individual Manipulation Real Info



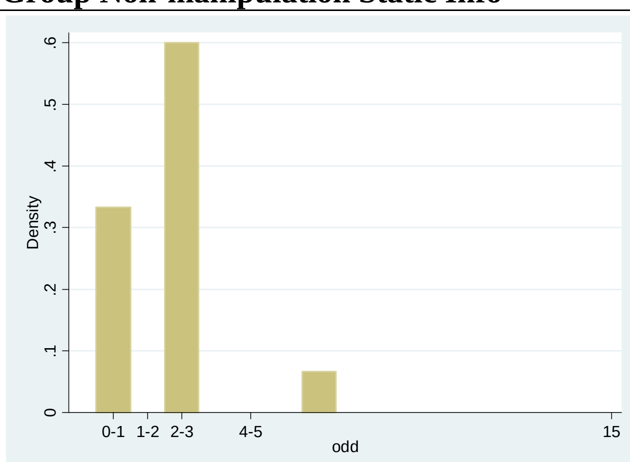
Individual Non-manipulation Static Info



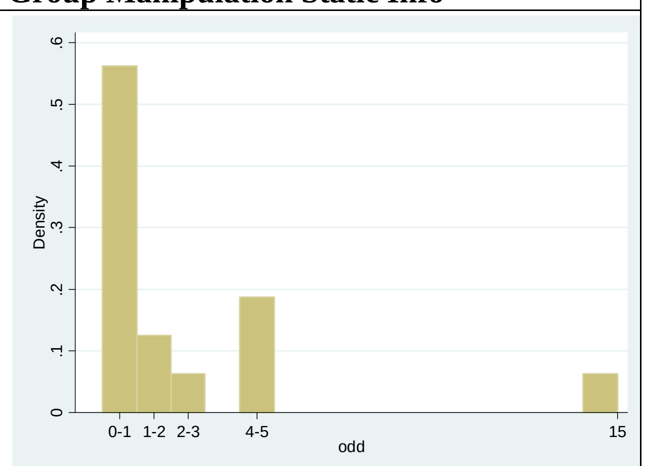
Individual Manipulation Static Info



Group Non-manipulation Static Info



Group Manipulation Static Info



The horizontal axis shows the odds-ratio, while the vertical axis shows the frequency of occurrence. The six graphs in the figure show the frequency of occurrence of the odds-ratio for each treatment. In the treatment with group forecasters and static market information and manipulators, the odds-ratio of less than 1 occurs about 58% of the time.

Using the functional form in equation (11), the following random effects regression is estimated:

$$\ln\left(\frac{PredictionAccuracy_{jt}}{1-PredictionAccuracy_{jt}}\right)=\beta_1 s_{jt}+\beta_2 s_{jt}*m+\beta_3 s_{jt}*i+\beta_4 s_{jt}*m*i+\beta_5 g+\beta_6 g*m+\varepsilon_{jt}+\mu_j \quad (12)$$

In regression (12) t denotes the round and j the session; s_{jt} is the signal strength in round t of session j; m is a dummy variable for whether manipulators were present in the market; i is a dummy variable for our static information treatment; g is the dummy for the group forecaster treatment; * denotes interaction effects; e_{jt} is a random error term assumed to be normally distributed *(N (0; 1))* and u_j is the error term capturing the differences across sessions of the same treatment. Table 5 shows how the dummy variables from regression (12) determine the aggregate coefficients for each treatment. For instance, in real time markets, individual forecast ln(odds-ratio) will increase by β_2 in the presence of manipulators compared to their absence and by a total increase of $\beta_1+\beta_2$ as signal strength increases by an additional unit.

Table 5: Dummy Variables and Coefficients Estimates

Markets \ Decision Makers	Real –Time Individual	Static-Time Individual	Static-Time Group
Non-manipulation	$i = 0$ $g = 0$ $m = 0$	$i = 1$ $g = 0$ $m = 0$	$i = 1$ $g = 1$ $m = 0$
Coefficients	β_1	$\beta_1+\beta_2$	$\beta_1+\beta_3+\beta_5$
Manipulation	$i = 0$ $g = 0$ $m = 1$	$i = 1$ $g = 0$ $m = 1$	$i = 1$ $g = 1$ $m = 1$
Coefficients	$\beta_1+\beta_2$	$\beta_1+\beta_2+\beta_3+\beta_4$	$\beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\beta_6$

The dummy variables are: static information dummy (i) is 1 when forecasters only observe the market price history and 0 for real time market treatments; group dummy (g) is 1 when the decision maker type is a group and 0 otherwise; manipulation dummy (m) is 1 for all treatments where manipulators are present in the market, and 0 otherwise. The values of these dummy variables from regression in (12) will provide the coefficients for each of the six treatments.

The regression estimates can be found in table 6. All of the coefficients are statistically significant. We will use the estimates from table 6 to construct our estimates of the treatment effects on the information aggregation properties of the market in the sections that follow.

Table 6: Regression Estimates

Treatments	Estimated Coefficients	Standard Error	Z	p-value
β_1	0.526875	0.044834	11.75	0.000
β_2	-0.146097	0.053006	-2.76	0.006
β_3	-0.140854	0.053006	-2.66	0.008
β_4	-0.189212	0.074962	-2.52	0.012
β_5	-0.191404	0.054995	-3.48	0.001
β_6	0.150951	0.076382	1.98	0.048

Regression estimates from (12) show that all coefficients are statistically different from 0.

4.2.1 Real versus Static Markets

Result 1.a: *Prediction quality of individual forecasters, who observe only the history of trading prices, is statistically lower than the forecasters with information on the full market evolution, both in the presence and absence of manipulators.*

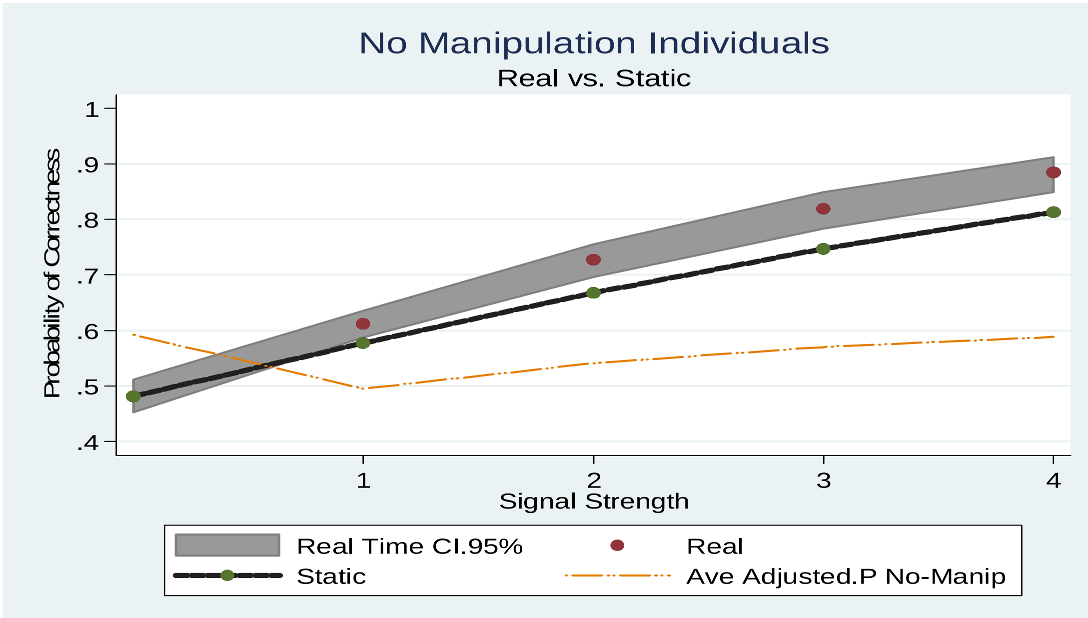
We observe in Table 6 that in the non-manipulation markets, the coefficient for real-time market information treatment is *0.14* (β_3) higher than in the static-time treatment. This is statistically different from *0* at the *1%* level. This coefficient translates into an odds-ratio of *1.69* for signal strength $s = 1$ in real-time information treatment compared to *1.48* in static information treatment. In the manipulation markets, the coefficient in real time is *0.33* ($\beta_3 + \beta_4$) which is statistically higher than the static-information treatment. This coefficient translates into an odds-ratio of *1.46* for signal strength $s = 1$ in real time compared to *1.05* for the static treatment. Figures 5 and 6 provide a visual overview of these findings. In both market types, with and without manipulation, forecasters with real-time information predict statistically better than forecasters with static information. The sample averages in static-information treatment fall out of the *95%* confidence interval of the real-time information treatment. We also supply market price data as a benchmark for the prediction quality of the forecasters. In particular, we examine the average closing price for each treatment based on signal strength. Specifically, for each treatment (k) and particular level of signal strength (s), we calculate the adjusted average price in equation (14) where n is the number of positive clues, m_s indexes the rounds in which the signal strength is s and M_s is the total number of rounds in which the signal strength is s.

$$\text{AdjustedAvePrice}_{sk} = \begin{cases} \dfrac{1}{M_{sk}} \displaystyle\sum_{m_{sk}=1}^{M_{sk}} \dfrac{\mathit{Closing\ Price}_{m_{sk}}}{\mathit{100}} & \mathit{if}\ n \geq 4 \\ \dfrac{1}{M_{sk}} \displaystyle\sum_{m_{sk}=1}^{M_{sk}} (1 - \dfrac{\mathit{Closing\ Price}_{m_{sk}}}{\mathit{100}}) & \mathit{if}\ n \geq 4 \end{cases} \tag{14}$$

Result 1.b: *In the absence of manipulators, individual prediction quality with real and static market information is statistically higher than the 50-50 prior and it increases with signal strength. Thus, even though forecasts are more accurate with real-time market information, individuals aggregate information in both cases.*

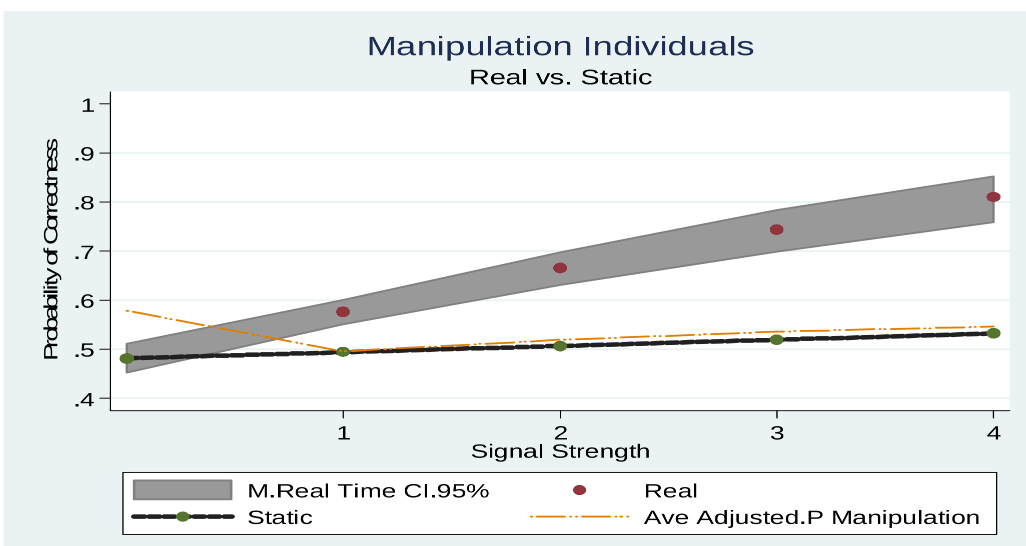
The coefficients are statistically different from *0* for both real-time and static-time, non-manipulated markets. If the coefficient is greater than *0*, the odds-ratio would be greater than *one*. Hence individual forecasters correctly predict the state more often than the prior and the prediction accuracy is positively correlated with the signal strength. From Figure 5 we can also observe that in the non-manipulation markets, individual forecasters outperform the market prices with both real and static market information. The dotted line represents the average of closing prices in non-manipulated markets adjusted with the signal strength as calculated in (14). The average adjusted price can be interpreted as the market posterior probability given signal strength. The market price line is always below the forecasters' probability of predicting the state. However, in Figure 6 we find that when manipulators are present in the market, both individual forecasts and market prices are uninformative if only static market information is provided.

Figure 5: Non-manipulation Individual Forecast Treatments: Real and Static Market Information



The grey area shows *95%* confidence interval (CI.95%) of individual forecast with real time market information as a function of signal strength. The dark line shows the individual mean forecast with static market information and the lower dotted line shows the adjusted price derived from the mean non-manipulation closing market prices.

Figure 6: Manipulation Individual Forecast Treatments: Real and Static Market Information



The grey area shows *95%* confidence interval (CI.95%) of individual forecast with real time market information as function of signal strength. The dark line shows the individual mean forecast with static market information and the lower dotted line shows the adjusted price derived from the mean non-manipulation closing market prices.

4.2.2 Group versus Individual Prediction

Result 2.a: *Prediction quality of individual forecasters is statistically better than the group forecasters in the static market information treatment with no manipulators present.*

From the regression estimates in Table 6, we can calculate the estimated coefficients for each treatment from Table 5. These estimates are reported in Table 7.

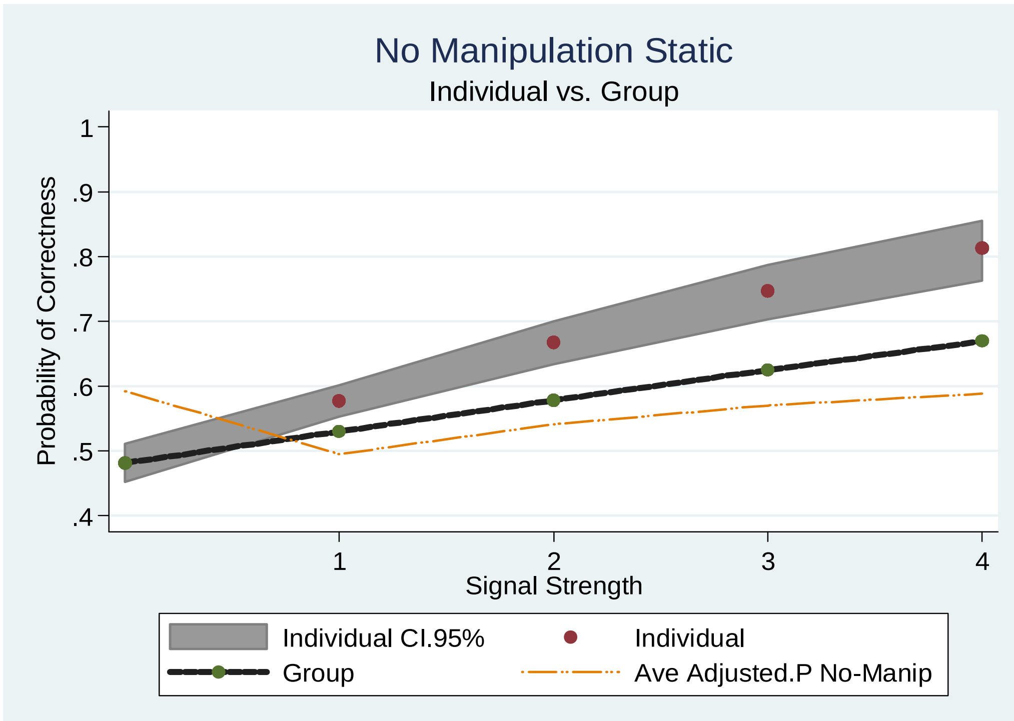
Table 7: Estimated Coefficients from the Random Effects Regression

Treatments	Estimated Coefficients	Standard Error	Z	p-value
IN_R	**0.5268751	0.044834	11.752	0.000
IM_R	**0.3807781	0.069425	5.485	0.000
IN_S	**0.3860209	0.069425	5.560	0.000
IM_S	0.0507122	0.115104	0.441	0.660
GN_S	*0.1946171	0.088568	2.197	0.028
GM_S	0.0102597	0.858385	0.012	0.980

These are the aggregate coefficients for each treatment using the estimated coefficients from regression shown in Table 6 and the aggregate coefficient calculations from Table 5. Coefficients noted (**) are significant at the 1% level, (*) are significant in 5% level.

From the estimates in table 7, we find that prediction quality is not improved when groups forecast. On the contrary, the individuals' odds-ratio is higher than that of groups. Table 6 shows a statistically significant coefficient of -0.19 (β_5 in table 6) for the group dummy. This corresponds to a difference in odds-ratio from *1.48* to *1.22* for a signal strength *one,* which translates to *60%* correct predictions for individuals versus *55%* correct predictions for groups. The odds-ratio increases at an increasing rate as we move to higher signal strengths. Hence, we can conclude that group prediction quality is statistically lower than the individuals. These results are highlighted in Figure 7 where the group prediction lies outside the *95%* confidence interval of the predictions by individuals.

Figure 7: Non-manipulation Static Information Treatments



The grey area shows *95%* confidence interval (CI.95%) of individual forecast with static market information as a function of signal strength. The dark line shows the group mean forecast with static market information and the lower dotted line shows the adjusted-price derived from the mean non-manipulation closing market prices.

Result 2.b*: In the absence of manipulators, both individual and group prediction quality is statistically higher than the 50-50 prior and it increases with signal strength.*

Coefficients corresponding to no manipulation treatments from Table 7 are statistically significantly from zero. This means that the predictions have a higher accuracy rate than the 50-50 prior prediction. Using the estimates from Table 7 we generate Table 8 and Table 9 which show how the odds-ratio and thus prediction accuracies change as the signal strength changes for the non-manipulation individual and group prediction treatments. Specifically, table 9 shows that individual prediction accuracy outperforms group prediction at an increasing rate as signal strength increases.

Table 8: Odds-Ratio in Non-manipulation Markets across Signal Strengths

Type of Market	Real time & non-manipulation	Static time & non-manipulation	Static time & non-manipulation
Decision Maker Type	Individual	Individual	Group
Odds-Ratio (s=1)	1.70	1.48	1.22
Odds-Ratio (s=2)	2.89	2.18	1.49
Odds-Ratio (s=3)	4.90	3.22	1.82
Odds-Ratio (s=4)	8.33	4.76	2.23

The odds-ratios are displayed for different signal strengths, from 1 to 4. The treatments observed are for all markets with no manipulators. These odds-ratios are calculated by using the results from the regressions in Table 7.

Table 9: Forecasters' Prediction Accuracy in Non-manipulation Markets across Signal Strengths

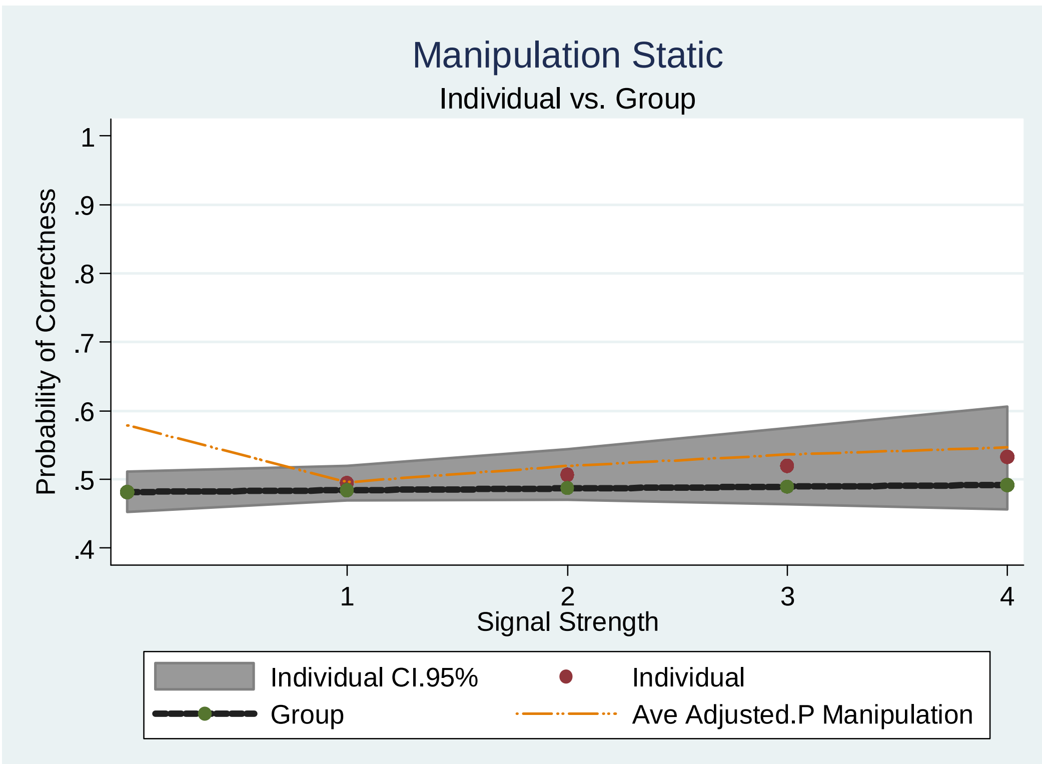
Type of Market	Real time & non-manipulation	Static time & non-manipulation	Static time & non-manipulation
Decision Maker Type	Individual	Individual	Group
P (s=1)	0.63	0.60	0.55
P (s=2)	0.74	0.69	0.60
P (s=3)	0.83	0.76	0.65
P (s=4)	0.89	0.83	0.69

The probability of correctly predicting the state is displayed for different signal strengths, from 1 to 4. The treatments observed are for all markets with no manipulators. These odds-ratios are calculated by using the results from the odds-ratios in Table 8.

Result 2.c: *In the presence of manipulators, individual and group prediction is statistically equivalent.*

In Figure 8, the dark dots show the prediction accuracy of individual forecasters when they observe the history of prices (static information) with manipulators. The grey area shows the *95%* confidence interval of individual forecasts. The black line shows the group forecast accuracy with static information in the presence of manipulators as a function of signal strength, which falls within the 95% confidence interval of the individual forecast accuracy. Both the individual and group predictions are not different than the 50-50 prior with no information. The estimated coefficients from Table 7 are not statistically different from zero, which translates to a prediction accuracy of 50%.[17]

Figure 8: Manipulation Static Information Treatment



The grey area shows *95%* confidence interval (CI.95%) of individual forecast with static market information as function of signal strength when there are manipulators in the market. The dark line shows the group mean forecast with static market information and the lower dotted line shows the adjusted-price derived from the mean manipulation closing market prices.

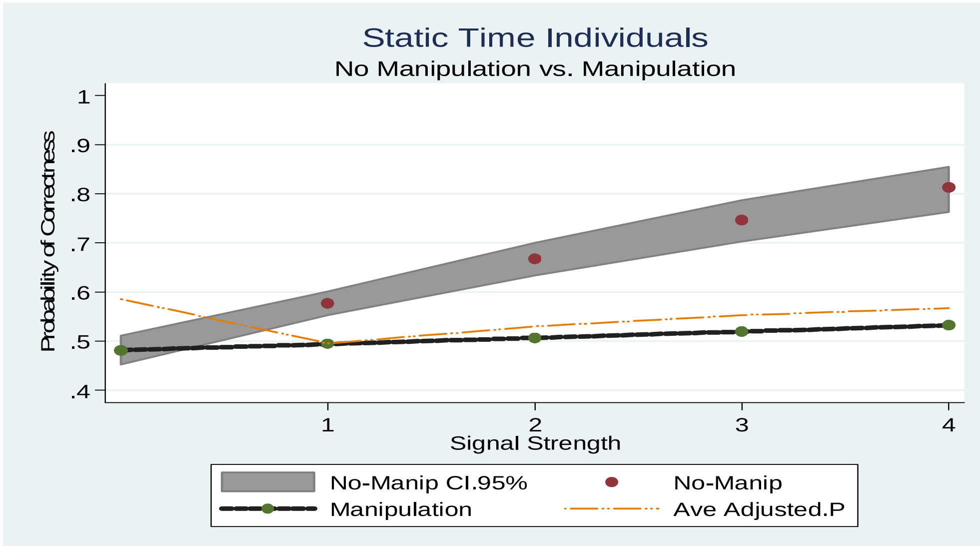
[17] When the coefficient is 0, then ln(odds-ratio)=0 which means that the odds-ratio=1, and p(s)=1-p(s)=50%.

4.2.3 The Effect of Manipulation in the Static Market

Result 3: *With only static market information, prediction accuracy is reduced when there are manipulators in the market.*

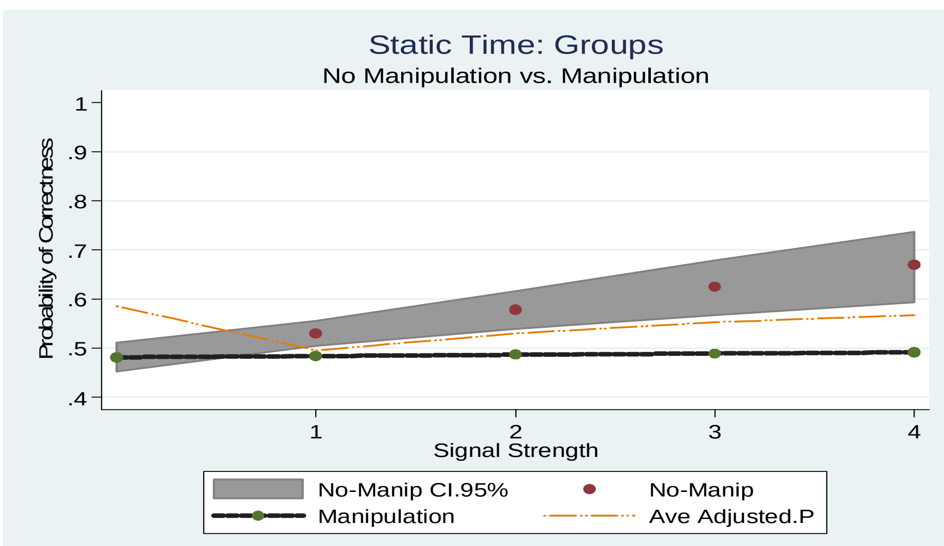
In contrast to the results of Oprea et al. (2006), where the presence of manipulators did not effect prediction accuracy when forecasters had access to real time market information, our results show that with limited market information, manipulators can have a significant effect on forecast accuracy. Specifically, the individual forecast estimated coefficient in the no-manipulation markets is *0.34* higher than that of manipulated markets. This difference is shown by coefficient $\beta_2+\beta_4$ in Table 6 which is statistically significant. This holds true for both individual and group forecasters. These results can be found in Figures 9 and 10. The presence of manipulators has such a dramatic effect when forecasters have limited market information that the predictions are no better than flipping a coin no matter the signal strength.

Figure 9: Individual Prediction Accuracy with Static Market Information



The grey area shows *95%* confidence interval (CI.95%) of individual forecast with static market information as function of signal strength when there are no manipulators in the market. The dark line shows the individual mean forecast with static market information and the lower dotted line shows the adjusted price derived from the closing market prices.

Figure 10: Group Prediction Accuracy with Static Market Information



The grey area shows *95%* confidence interval (CI.95%) of group forecast with static market information as function of signal strength when there are no manipulators in the market. The dark line shows the group mean forecast with static market information and the lower dotted line shows the adjusted price derived from the closing market prices.

2.5 Concluding Remarks

Using markets in order to aggregate dispersed information about the likelihood of a future event is a powerful tool. Uniformed observers can then use the information conveyed in market transactions by informed traders to improve their forecasts and decision making. This paper examined the quality predictions by uninformed forecasters under a variety of conditions. Our results show that forecasters use the market information to improve their forecasts. However, our findings show that when forecasters observe only a summary of transaction prices, they do not perform as well as when they are provided with real time access to the price discovery process. In addition, we find that the presence of manipulators lowers the prediction quality of the forecasts when provided only with the history of the transacted prices. In fact, the prediction quality drops to a level no different than the uninformative prior. However, when forecasters are provided real time access to bids, asks and contracts, their predictions significantly improve even when manipulators are present in the market.

The literature on comparing group and individual decision-making is growing at a rapid pace, and yet the findings are inconclusive. We have added to this literature to examine the prediction quality of groups relative to individuals in our markets. We find that group prediction does not perform as well as individuals in accurately forecasting the state. This suggests that in a non-strategic setting, individual decision-making is likely to result in superior predictions than if the decision must be arrived at by a group.

Appendix A: Derivation of Functional Form

We shall start the calculations from the Bayesian expected value of a ticket as a function of the number of positive signals (n):

$$EV(n) = \frac{\left(\frac{2}{3}\right)^n \left(\frac{1}{3}\right)^{8-n}}{\left(\frac{2}{3}\right)^n \left(\frac{1}{3}\right)^{8-n} + \left(\frac{1}{3}\right)^n \left(\frac{2}{3}\right)^{8-n}} \bullet (100)$$

$$= \frac{\left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^{8-n} 2^n}{\left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^{8-n} 2^n + \left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^{8-n} 2^{8-n}} \bullet (100)$$

$$= \frac{\left(\frac{1}{3}\right)^8 2^n}{\left(\frac{1}{3}\right)^8 (2^n + 2^{8-n})} \bullet (100)$$

$$= \frac{2^n}{(2^n + 2^{8-n})} \bullet (100)$$

$$= \frac{1}{\frac{(2^n + 2^{8-n})}{2^n}} \bullet (100)$$

$$= \frac{1}{1 + \frac{2^{8-n}}{2^n}} \bullet (100)$$

$$= \frac{1}{1 + 2^{2(4-n)}} \bullet (100)$$

We have defined the signal strength (s) in relation to positive signals (n) as $s = |n-4|$. The expression inside the absolute value will change signs depending on the value of n, but the signal strength will always take a positive value between 0 and 4. Hence, the probability of correctly predicting the state can be derived from the Bayesian prediction as shown in eq (1.10).

$$\Pr(Guess = v \mid s) = P(s) = \begin{cases} \dfrac{1}{1+2^{2(4-n)}} & \textit{if } n > 4 \textit{ or } (s = n - 4) \\ 1-\dfrac{1}{1+2^{2(4-n)}} & \textit{if } n < 4 \textit{ or } (s = 4 - n) \\ .5 & \textit{if } n = 4 \textit{ or } (s = 0) \end{cases}$$

Thus,

$$P(s) = \frac{1}{1+2^{-2s}}$$

so,

$$\frac{P(s)}{1-P(s)} = \frac{\dfrac{1}{1+2^{-2s}}}{1-\dfrac{1}{1+2^{-2s}}} = \frac{\dfrac{1}{1+2^{-2s}}}{\dfrac{2^{-2s}}{1+2^{-2s}}} = 2^{2s}$$

Hence,

$$\ln\left(\frac{P(s)}{1-P(s)}\right) = \ln(4) \bullet s$$

References:

Allen, F., & Gale, D. (1992). Stock-Price manipulation. *The Review of Financial Studies*, 5, 503--529.

Anderson, L., & Holt , C. (1997). Information Cascades in the Laboratory. *American Economic Review*, 87(5), 847-62.

Blander, A. (2000). Are Two Heads Better than One?: An Experimental Analysis of Group vs. Individual Decision Making. *NBER Working Paper* 7909.

Bornstein, G., Kugler, T., & Ziegelmeyer, A. (2004) Individual and group decisions in the centipede game: Are groups more rational players?, *Journal of Experimental Social Psychology*, 40, 599-605.

Bornstein G., & Yaniv, I. (1998). Individual and Group Behavior in the Ultimatum Game: Are Groups more "Rational" Players. *Experimental Economics*, 101-108.

Camerer,C. (1998). Can asset markets be manipulated? A Field Experiment with Race-track Betting. *Journal of Political Economy,* 106, 457-482.

Cason, T. N., & Mui, V. (1997). A laboratory study of group polarization in the team dictator game. *Economic Journal*, 107, 1465–1483.

Chakrabortya, A., & Ylmaz, B. (2004). Manipulation in market order models. *Journal of Financial Markets*, 7, 187-206.

Cox, J. C. (2002). Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. In R. Zwick & A. Rapoport (Eds.), *Advances in Experimental Business Research*. Dordrecht: Kluwer Academic Publishers.

Cox, J., & Hayne, S. (2006). Barking up the right tree: Are small groups rational agents? *Experimental Economics*, 9(3), 209-222.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25(2), 383-417.

Forsythe, R., Nelson, F., Neumann, G., & Wright, J. (1992). Anatomy of an experimental political stock market. *American Economic Review*, 82, 1142-1161.

Forsythe, R., & Lundholm, R. (1990). Information aggregation in an experimental market. *Econometrica*, 58, 309-347.

Grossman, S.J. & Stiglitz, J.E. (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review*, 70(3), 393-408.

Grossman, S.J. & Stiglitz, J.E. (1976). Information and Competitive Price Systems. *American Economic Review,* 66(2), 246-253.

Gjerstad, S. (2005). Risk aversion, beliefs and prediction market equilibrium. *Economic Science Laboratory,* University of Arizona.

Greene, W. (2002). *Econometric Analysis*, 5th ed. New York: Prentice Hall.

Hansen, J., Schmidt, C., & Strobel, M. (2004). Manipulation in political stock markets – preconditions and evidence. *Applied Economics Letters,* 11, 459-463.

Hanson, R., & Oprea, R. (2004). Manipulators increase information market accuracy, *George Mason University Working Paper.*

Hanson, R., Oprea, R. & Porter, D. (2006). Information Aggregation and Manipulation in an Experimental Market. *Journal of Economic Behavior and Organization,* 60, 449-459.

Hanson, R., Oprea, R., Porter, D., Hibbert, C. & Tila, D. (2006). Interpretation and Manipulation in Imperfect Prediction Market, *George Mason University Working Paper*.

Kocher M.G., & Sutter M. (2005). The Decision Maker Matters: Individual versus Group Behavior in experimental Beauty-Contest Games. *Economic Journal*, 115, 200-223.

Kocher, M., & Sutter, M. (2003), Individual versus group behavior and the role of the decision making process in gift-exchange experiments. *Royal Economic Society Annual Conference*,125.

Kumar, P., & Seppi, D. (1992). Futures manipulation with 'cash settlement'. *Journal of Finance,* 47, 1485-1502.

Manski, C. (2004). Interpreting the Predictions of Prediction Markets. *NBER Working Paper 10359*.

Muth, J. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica,* 29, 315-335.

Noeth, M., Camerer, C. F., Plott, C. & Webber, M.R. (1999). Information Aggregation in Experimental Asset Markets: Traps and Misaligned Beliefs. *Social Science Working Paper No. 1060*. Pasadena: California Institute of Technology.

Plott, C., & Sunder,S. (1982). Efficiency of experimental security markets with insider information: an application of rational-expectations models. *The Journal of Political Economy,* 90, 663-698.

Plott, C., & Sunder,S. (1988). Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*, 56, 1085-1118.

Plott, C. & Hung, A. (2001). Information Cascades: Replication and an Extension to Majority Rule a Conformity-Rewarding Institutions. *American Economic Review, 91(5),* 1508-1520.

Prather, L.J., & Middleton, K.L. (2002). Are N + 1 heads better than one? The case of mutual fund managers. *Journal of Economic Behavior & Organization*, 47(1), 103–120.

Strumpf, K., & Rhode, P. (2003). Historical presidential betting markets. *Journal of Economic Perspectives*, 18, 127-141.

Tucker, J., & Meirowitz A. (2004). Learning from Terrorist Markets. *Perspectives on Politics*, 2, 331-336.

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, 18, 107-126.

Wolfers, J., & Zitzewitz, E. (2005). Interpreting prediction market prices as probabilities. *NBER Working Paper 12200.16*

Economic Science Institute Working Papers

2008

08-04 Caginalp, G., Hao, L., Porter, D. and Smith, V. Asset Market Reactions to News: An Experimental Study.

08-03 Thomas, C. and Wilson, B. Horizontal Product Differentiation in Auctions and Multilateral Negotiations.

08-02 Oprea, R., Wilson, B. and Zillante, A. War of Attrition: Evidence from a Laboratory Experiment on Market Exit.

08-01 Oprea, R., Porter, D., Hibbert, C., Hanson, R. and Tila, D. Can Manipulators Mislead Prediction Market Observers?