

PREDICTION MARKETS AS INSTITUTIONAL FORECASTING SUPPORT SYSTEMS

Preprint of

van Bruggen, Gerrit / Spann, Martin / Lilien, Gary / Skiera, Bernd (2010): "Prediction Markets as Institutional Forecasting Support Systems", *Decision Support Systems*, 49(4), 404-416.

Gerrit H. Van Bruggen
RSM Erasmus University
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
Phone: +31 10 4082258
Fax: +31 10 4089011
Email: gbruggen@rsm.nl

Martin Spann
Ludwig-Maximilians-University (LMU)
Geschwister-Scholl-Platz 1
80539 Munich, Germany
Phone +49 89 2180-72050
Fax: +49 89 2180-72052
Email: spann@spann.de

Gary L. Lilien
Penn State
484 Business Building
University Park, PA 16803 USA
Phone +1 82 863 2782
Fax: +1 814 863 0413
Email: GLilien@psu.edu

Bernd Skiera
Goethe-University Frankfurt
Grüneburgplatz 1, 60323 Frankfurt am Main, Germany,
Phone +49 69 798-34649
Fax: +49 69 798-35001
Email: skiera@skiera.de

PREDICTION MARKETS AS INSTITUTIONAL FORECASTING SUPPORT SYSTEMS

ABSTRACT

An attractive feature of Prediction Markets (PMs, also called information markets or virtual stock markets) is that they provide economic incentives for informants to share unique information, i.e., information that no other informants possess. This feature and the broad availability of the Internet have led to applications of prediction markets in a number of fields. Most reported applications to date contain large numbers of informants, so it is unclear if PMs are appropriate for applications with few knowledgeable informants (i.e. small PMs) as is the case for most institutional forecasting tasks. Hence, we compare the performance of small PMs with more traditional judgment-based approaches which we call the Combined Judgmental Forecasts (CJF) approach, and the Key Informant (KI) approach. Our results show that forecasts from small PMs outperform more traditional approaches in settings of high information heterogeneity (i.e. where the amount of unique information possessed by informants is relatively high) and does no worse in settings of low information heterogeneity.

Keywords: Prediction Markets, Institutional Forecasting

PREDICTION MARKETS AS GROUP FORECASTING SUPPORT SYSTEMS

INTRODUCTION

One of the most challenging tasks firms face is to make the most effective use of their extant internal information. That challenge applies especially to forecasting, which the forecasting literature (see Evans (2002), for example) describes as both “art and science”, with “art” referring to forms of managerial judgment and the “science” referring to statistical or econometric analysis based on historical data. But forecasts in organizations are often required in situations where there are no data-based approaches available. Furthermore, cost or secrecy concerns often preclude customer surveys as data sources for such forecasts. Hence, many organizations rely on informants (sales representatives, research analysts, business managers and project team members) for the development of (judgment-based) forecasts (Fildes and Hastings 1994), what we will refer to here as “institutional forecasting.”

Forecast accuracy is often improved when informant-based forecasts rely on *multiple* rather than a single informant because (a) informants seldom have access to all relevant information and (b) using multiple informants lowers the error component of the group’s forecast (Armstrong 2001). Group discussion enables informants to share information so that groups can access a larger pool of information than can any single informant acting alone (Shaw 1981; Dennis 1996). However, research has also shown that groups are sometimes ineffective at exchanging information (Hightower and Sayeed 1996) and that much unique information known to a single or only some individuals is never shared with the group (Stasser 1992).

Stasser and Titus (1985) claim that groups can benefit from pooling members’ information, particularly when members individually have partial and biased information but collectively can compose a less biased characterization of the decision alternatives. However, in an empirical application of their approach they find that group members often fail to effectively pool their information because discussion is dominated by commonly held information and information that supports members’ existent viewpoints. This result implies that groups that share information interactively through group discussion will normally reach consensus but will neither appropriately correct for nor effectively pool members’ complementary expertise and knowledge.

The increasing ease of interconnectivity and the proliferation of web conferencing tools like LiveMeeting

and GotoMeeting are continually making it simpler for groups of people to work together through computer networks. A Group Forecasting Support System within such an environment is a communication and coordination process that structures the process of communication and information sharing. Dennis (1996) studied the benefits of Group Support Systems (GSSs) and found that groups using a GSS exchanged 50% more information than verbally interacting groups, permitting them to include the best alternative amongst those considered for selection. However, he also found that very few groups ended up selecting the optimal decision, indicating that the GSS was not able to help the groups to process the information that they had optimally. Sia, Tan, and Wei (2002) report that anonymous or dispersed computer-mediated communication settings are required for group discussion to lead to the strong polarization that may be necessary to identify a non-consensus choice.

Prediction Markets (PMs), also called information markets or virtual stock markets, represent an information technology based forecasting platform that may have the potential to address some of the challenges that traditional GSS's face in the institutional forecasting context. PMs enable informants to exchange information and should be considered as a type of organizational Group Forecasting Support System (GFSS). However, unlike traditional Group Support Systems (DeSanctis and Gallupe 1987; Dennis 1996), PMs place an incentive on exchanging “unique” information with other informants in the market because doing so will lead to higher pay-offs for the informant possessing such unique information.

PMs have been successfully applied to predict election outcomes (e.g., the Iowa Electronic Markets (Berg, Nelson, and Rietz 2008)), the success of movies and impact of stars (e.g., the Hollywood Stock Exchange (Pennock, Lawrence, Giles, and Nielsen 2001; Elberse 2007; Foutz and Jank 2008)), sports results (Spann and Skiera 2009), product concepts and new product ideas (Dahan, Lo, Poggio, Chan, and Kim 2007; Soukhoroukova, Spann, and Skiera 2010); and future economic outcomes (e.g., economicderivatives.com). In an inter-organizational setting Guo, Fang, and Whinston (2006) propose a prediction market for information sharing within supply chains. LaComb, Barnett, and Pan (2007) report that several GE businesses experimented with PMs to support idea generation and group-decision making. Ostrover (2005) cites several PM applications in organizations: e.g. Hewlett Packard (HP) uses PMs to forecast sales, financial, and

accounting results while Eli Lilly uses a PM to identify those drugs in the early stages of development most likely to win US Federal Drug Administration approval. The results of these applications demonstrate the potential value of the PM approach for institutional forecasting in settings with a *large* number of participants. However, little work has been reported on PMs in settings with few knowledgeable participants; indeed, scholars have stressed that such situations can lead to markets with low liquidity where small changes in supply and/or demand can have a large impact on market prices (Fung, Mok, and Wong 2004). Research on traditional GSSs (Gallupe, Dennis, Cooper, Valacich, Bastianutti, and Nunamaker 1992) suggests that these systems are also much more effective for larger groups. Yet, many practical institutional forecasting situations actually involve relatively few (knowledgeable) informants, which underlines the need for approaches that are effective for small groups.

Spann and Skiera (2003) report a study of a (single) small PM with twelve employees at a large German mobile phone operator to forecast the usage of five different mobile phone services in a specific month. The PM showed better forecasting accuracy than several competing model-based forecasting approaches. The authors suggest that as the heterogeneity in market informants' knowledge increases, the PM's forecasting error declines relative to other approaches. If this suggestion is correct, it implies that small PMs can be effective in general and may do especially well relative to alternatives if the informants possess relatively large amounts of "unique" or "partially shared" information (Dennis 1996).

Our objective in this paper is to analyze whether the forecasting accuracy observed for PMs in other settings applies to forecasting settings with *few* knowledgeable informants. We consider *whether*, *why* and *when* small PMs are likely to outperform more traditional judgment-based approaches towards institutional forecasting (Kumar, Stern, and Anderson 1993), which we call the Combined Judgmental Forecasts (CJF) approach, or the Key Informant (KI) approach. We compare the accuracy of the PM approach with these alternatives for several forecasting tasks (i.e., forecasting the future value of two financial indices and forecasting the point spread of two college football games).

Our results show that the PM approach performs at least as well as these alternatives in markets with few participants. We find that for tasks we characterize as high in information-heterogeneity (i.e., predicting

football point spreads), forecasts made through the PM approach are more accurate than those developed by the CJF approach or by the KI approach, while for low information-heterogeneity tasks (i.e., predicting financial indices), we find that the PM approach performs no worse than the other approaches.

The paper proceeds as follows: we first review alternative approaches to support institutional forecasting, including the opportunities afforded by the PM approach. Then we review literature on how PMs operate and how their operation may be affected by the number of market informants. We contrast the differences between the PM approach and the more traditional approaches (CJF or KI), which leads to our hypothesis about the expected differences in forecasting accuracy between those three approaches. We next describe the design of our empirical study and present our results. We conclude with a discussion of those results and their implications.

REVIEW OF INSTITUTIONAL FORECASTING APPROACHES

We review the most commonly used approaches to (institutional) forecasting--the (single) key informant and multiple informants approaches. We then compare them with the PM approach.

Key Informant (KI) Approach. Perhaps the most widely used approach in practice because of its simplicity is the key informant approach, where a *single* informant is most often selected because of knowledge and willingness to communicate that knowledge. This approach suffers from significant drawbacks (Kumar et al. 1993), including bias, random error, and the inability to aggregate information spread across multiple informants. Armstrong (2001) summarizes much research that shows that single informant reports are systematically outperformed by those of multiple informants when their reports are appropriately combined.

Multiple Informant Approaches combine knowledge of multiple informants into an overall forecast. The strategies to combine the information of multiple informants differ depending upon whether or not the informants interact and are able to exchange information.

Mechanistic Aggregation—Combined Judgmental Forecasts (CJF) Approaches. CJF algorithms such as forming simple or weighted averages are used when informants *do not interact* and *do not exchange* information (Garthwaite, Kadane, and O'Hagan 2005). Rather, all informants produce forecasts that are then

mathematically combined into a single group forecast. Van Bruggen, Lilien, and Kacker (2002) propose methods for improving weighted averages, showing that the use of confidence and competence scores as weights improves forecasting accuracy. Yet, the mechanistic, mathematical approach can lead to a form of double counting of expertise if the knowledge of various experts overlaps substantially. In addition, and by definition, the approach prevents informants from sharing information to learn from one another (Garthwaite et al. 2005).

Behavioral or Interactive Aggregation permits informants to interact and share information (Garthwaite et al. 2005) and most existing GSSs are designed to facilitate such processes. Approaches differ and depend on whether the sharing is synchronous or asynchronous, if the respondents are anonymous or not and if the informants interact verbally or are supported through Group Support Systems (DeSanctis and Gallupe 1987; Dennis 1996; Hightower and Sayeed 1996; Sia et al. 2002; Dennis and Garfield 2003). All forms of behavioral aggregation can lead to difficulties in reaching consensus and even if they reach consensus, this consensus may be driven by power and personality rather than by knowledge, perhaps explaining the GSS decision biases cited earlier. Garthwaite et al. (2005) and Kumar et al. (1993) discuss the influence of censorship, Groupthink, and the difficulty of organizing and facilitating the needed interactions on forecasting accuracy.

To sharpen our focus and to align our work with practice, we study the (online) survey-based CJF approach and the key informant (KI) approach, a special case of the CJF approach with 100% weight on the key informant. In our survey-based CJF approach, informants do *not* interact and their individual forecasts are combined mathematically.

The Prediction Markets (PM) Approach. The PM approach creates information markets, bringing groups of informants together over electronic communication networks such as the Internet and allowing them to trade shares of virtual stocks that represent a bet on the outcome of a future market situation. They collect, disseminate, and aggregate information that may be widely dispersed across a number of public and private sources, through the mechanism of trading (Sunder 1995; Plott 2000). Public information is information that informants have in common or shared while private information is unique to informants and unshared (Stasser and Titus 1985; Stasser and Titus 1987; Dennis 1996). The ability of markets to reflect public as well

as private information is referred to as their informational efficiency (Fama 1991).

Effective PMs satisfy three criteria (Wolfers and Zitzewitz 2004): they provide *i*) incentives to seek information; *ii*) incentives for truthful information revelation; and *iii*) an algorithm for aggregating diverse opinions of their informants. Ostrover (2005) cites five reasons why the PM approach should be expected to work: *i*) participation: everyone with access to relevant information can contribute; *ii*) motivation: a properly designed reward mechanism incents informants to acquire relevant information and reveal their expectations (Forsythe, Rietz, and Ross 1999); *iii*) anonymity: PMs eliminate fear of reprisal for revealing unpopular expectations; *iv*) coordination: markets provide a natural mechanism for active group interplay; and *v*) computation: markets provide a natural aggregation mechanism.

The question we address here is how well the Prediction Market approach supports institutional forecasting when the number of (trading) informants is relatively small. In a study on a GSS that supported electronic brainstorming Gallupe et al. (1992) found these systems to be less effective for smaller groups. For the PM approach a similar effect may be present. When there are few (active) traders and, hence, few trades, bid and ask prices often differ widely (Fung et al. 2004). Several researchers have addressed the question of whether and how market size affects market efficiency. In small markets, problems like information traps, manipulation and lack of equilibrium have been found to be exacerbated (Chen, Fine, and Huberman 2004), information cascades may occur (Anderson and Holt 1997) and the market may not be information-efficient (Sunder 1992; Spann and Skiera 2003). While a few researchers have demonstrated that small markets often do quite well (Fung et al. 2004), in general larger markets are reported to perform better. When markets attract broad participation, prices encode all the disparate information of these informants (Pennock, Lawrence, Giles, and Nielsen 2000) and there is most likely more information encoded in the market price than with fewer informants. In an absolute sense PMs with a larger number of traders will be better informed and thereby produce more accurate forecasts than those with fewer traders. Furthermore, a low trading volume makes it difficult for markets to react to new information that may not be easily accessible to the few active traders (Barnes 1986). Yet, with the possible exception of Gruca, Berg, and Cipriano (2003), which differs significantly in design from ours, there are no studies that empirically compare the forecasts of the PM

approach with those of the CJF approaches in settings like those we focus on here.

THE PREDICTION MARKETS (PM) APPROACH VERSUS THE COMBINED JUDGMENTAL FORECASTS (CJF) APPROACH: HYPOTHESIS DEVELOPMENT

To guide our hypothesis development, we compare the PM approach and the CJF approach along three key information dimensions: *i*) information collection; *ii*) information dissemination (i.e. broadcasting); and *iii*) information aggregation. This taxonomy (see Table 1) is based on the three functions of markets distinguished by Plott (2000) and describes the activities performed when informant-based forecasts are being developed.

--- INSERT TABLE 1 ABOUT HERE ---

Information Collection is the process of eliciting individuals' forecasts. In the CJF approach, informants provide direct forecasts for the variables of interest and, sometimes, information about their confidence and competence. This approach is straightforward and (relatively) easy for informants. A weakness is that the judgments in these settings may be based on a limited number of mental operations, which potentially lead to biased assessments (Garthwaite et al. 2005). This weakness may also apply for confidence and competence assessments. Furthermore, reasons may exist for informants not to disclose information honestly (Kumar et al. 1993).

With the PM approach, informants' information is collected indirectly through their trading behavior. Furthermore, informants in the PM approach use the actual market price compared to their expectation and trade accordingly. This action (i.e., buy shares if price is considered too low or sell when too high) may be easier than generating a full personal estimate as is needed in the CJF and the KI approaches. The PM approach provides anonymity and incentives for informants to reveal their true expectations (Spann and Skiera 2003; Ostrover 2005). However, the market mechanism may also provide incentives for speculation, and trading in an information market may not be easily understandable for all informants (Stracca 2004). Hence, both approaches have advantages and disadvantages on the dimension of information collection.

Information Dissemination refers to the ability of a method to broadcast unique information held by better

informed informants to less well informed informants (Sunder 1995; Plott 2000). Informants base their responses in the CJF approach and their trading behavior in PMs on the public and private information they possess, where the latter information is unique to or held by a specific informant (Chen, Fine, and Huberman 2003). In the CJF approach, which collects information from each informant independently, informants provide their forecasts based on public and private information, but informants can *not* re-evaluate their opinion in light of the opinion of others.

In PMs the process of trading is dynamic and market feedback may make informants re-assess and update their initial forecasts, leading to changes in the collective forecast (Plott 2000). In PMs, groups of informants trade shares of virtual stocks that represent bets on the outcome of future (market) situations, where their value depends on the realization of these market situations. Once the outcome of a market situation is known, each share receives a payoff according to that outcome. The price of a share of a virtual stock corresponds to the PM's current, aggregate expectation of the event's outcome and, therefore, to the expected payoff of a share of the stock. If an individual's assessment of the outcome of a future event is different from the market's assessment it will be profitable to either buy or sell. By these trading activities informants will reveal the private/unique information they possess because the market provides incentives to reveal unique information. In equilibrium, price summarizes all the relevant information the traders have (Hayek 1945; Grossman 1976). The core service the market provides is thus to facilitate the exchange of items or, in this case, information between individuals (Pennock et al. 2001). The theory of rational expectations posits that prices reflect the total of all information available to all market informants. Even when some agents have exclusive access to inside (unique) information, prices equilibrate as if everyone had access to all information (Grossman 1981). Through the market, previously unshared information will thus be shared and become common information.

Markets can thus aggregate and disseminate information, but that capacity may not be perfect (Plott 2000): the wisdom of crowds is sometimes the ignorance of crowds (Surowiecki 2005). For example, Anderson and Holt (1997) describe the risk of information cascades, where individuals overweight private information of other traders that has become public during the trading process and underweight their own, possibly higher

quality, private information—“I guess they know something I don't.” Or the information that is exchanged may be biased. However, Malkiel (2003) concludes that such market imperfections are exceptions rather than the rule and the market can generally be trusted.

To summarize, in PMs, the price system makes information publicly available and thereby transfers it from informed to uninformed traders (Grossman and Stiglitz 1980); when the informants as a group are wise, that information exchange process is an advantage the PM approach holds over the CJF approach where information cannot be exchanged between informants.

Information Aggregation is the procedure that combines the informants' forecasts to arrive at a single forecast. With the CJF approach, the researcher does the aggregation, employing either a simple unweighted average or some more advanced form of weighted average (Van Bruggen et al. 2002; Chen et al. 2004). While these weighting approaches make the aggregation rule transparent, their weaknesses include the possible double-counting of informants whose knowledge is highly correlated as well as the logical problem of averaging two (or more) widely divergent views, based on totally different assumptions.

A PM simultaneously performs information aggregation, dissemination, and conflict resolution (Plott and Sunder 1988): the market mechanism performs the weighting procedure. This approach is efficient and, according to the “crystal ball” hypothesis (Plott and Sunder 1982), the market equilibrium may reflect even more information than the sum of what is available to individual traders. A possible weakness of the PM approach is that trading will be based on the strengths of expectations of traders and these expectations are not necessarily (fully) in line with reality (Stracca 2004). However according to Berg and Rietz (2003) markets aggregate diverse information in ways that prove more efficient than alternative methods.

Prediction markets can thus use the repeated interactions between informants to produce common forecasts that combine available information, avoiding the problem of weighting different expectations. No knowledge of who is more expert on what topics is required: market traders self-select to focus on the topics where they believe they are most expert in and those who are mistaken are punished by the market.

Table 1, summarizing the discussion above, shows that information dissemination is the most discriminating feature between these approaches, suggesting that the PM approach has the potential to

produce more accurate forecasts than the CJF approach at least under some conditions.

Hence, using multiple informants to develop forecasts has two types of benefits. First combining the inputs from multiple forecasters can reduce the random error component in the forecast, improving forecasting accuracy (Batchelor and Dua 1995; Armstrong 2001). Both the PM approach and the CJF approach benefit from using multiple informants. Second, and probably more important, more people will know more and access more heterogeneous information. Spann and Skiera (2003) suggest that such information heterogeneity may have a positive impact on the accuracy of PM forecasts. According to Sunder (1995), asymmetry of information among traders is an essential ingredient for prices to play an informational role and, thus, for markets to perform. Hence, the information dissemination feature of a PM is especially beneficial if significant information-asymmetry exists between informants, permitting individual (private) knowledge to become public and allowing less knowledgeable informants to update and improve their knowledge. The resulting higher average knowledge level of the informants can then be expected to increase the accuracy of the aggregated forecast.

Thus, we hypothesize that since the market allows for the dissemination of information through the pricing mechanism, the accuracy in forecasting for the PM approach should be equal to or exceed the accuracy of the most knowledgeable informant (I), i.e.,

$$Accuracy\ of\ Group\ (PM) \geq Max\ (Accuracy\ I_1, Accuracy\ I_2, \dots, Accuracy\ I_n),$$

while for the CJF approach, by definition

$$Accuracy\ of\ Group\ (CJF) \sim (Weighted)\ Mean\ (Accuracy\ I_1, Accuracy\ I_2, \dots, Accuracy\ I_n)$$

Following our argument, the PM approach should strictly outperform the CJF approach, since the PM approach permit sharing of information while the CJF approach does not. According to the information aggregation hypothesis of rational expectations, the prices should reflect available information (Plott 2000). Hence the PM approach should do no worse than the CJF approach in all cases and should significantly outperform the CJF approach in situations of high information-heterogeneity under any weighting scheme. We thus posit that information sharing and exchange should lead to improved forecasts (i.e., the PM approach should outperform the CJF approach) if there is a difference in knowledge between informants,

that is, if different informants' information is not perfectly correlated (Batchelor and Dua 1995). The fact that a PM is small as such will not make it perform worse than the CJF approach as market size affects the CJF approach as well.

As the Key Informant (KI) approach is a special case of the CJF approach, with all weight given to the specific key informant, the discussion above about the CJF approach applies for the KI approach as well. However, as the KI approach discards information from the non-key informants (who will still possess some "part of the puzzle"), we expect that the KI approach will perform more poorly than both the PM approach and than the CJF approach. Formally:

H₁: The PM approach will outperform the CJF approach in terms of forecasting accuracy while the CJF approach will outperform the KI approach in terms of forecasting accuracy.

There will likely be conditions that moderate the effect hypothesized in *H₁*. We suggest that the amount of information heterogeneity is such a moderator. Information-heterogeneity refers to the variations in knowledge, know-how, information and expertise which a group of forecasters can tap (Rodan and Galunic 2004). A *low information-heterogeneity* situation occurs when informants have access to similar (common) information. Regular, regional sales forecasts based on the judgments of a group of sales representatives represent an example of such an institutional setting. A *high information-heterogeneity* situation occurs when there is little common information and informants differ significantly in the type and quality of information and knowledge they possess. Informants in organizations frequently operate at different hierarchical levels, in different departments, and in different geographical areas. Hence, information and knowledge will often be dispersed across the organization and many informants will possess a substantial amount of unique information. A forecast of the market performance of a new product by the various people (e.g., marketing, R&D, sales) involved in developing it provides an example.

Information-heterogeneity thus refers to the variety of knowledge, know-how, information and expertise to which a group of forecasters has access (Rodan and Galunic 2004) and in general, greater information-heterogeneity can be expected to have advantageous effects (Batchelor and Dua 1995; Rodan and Galunic 2004) on the effectiveness of the PM approach. The market mechanism in the PM approach will transfer

information from knowledgeable and informed informants to less knowledgeable and uninformed informants. This transfer will increase the overall, average knowledge level among informants. Since in the CJF approach there is no exchange of information between informants this approach will benefit less from information heterogeneity. Hence:

H_2 *The PM approach's ability to outperform the CJF approach and the KI approach will increase as information heterogeneity increases*

METHOD

We next describe the task, participants, treatments, procedures and experimental measures in our study.

Task

To test our hypothesis, we sought forecasting tasks where we could compare the forecasting accuracy of the PM approach and the CJF approaches in situations with a small number of participants and that varied in information heterogeneity. The characteristics of such settings are: *i*) the judgment of the informant is an important input to the forecast; *ii*) no single informant can know the “true” value in advance, and *iii*) there are multiple but a *small number of informants* whose knowledge can be tapped, who may (partially, at least) disagree, and who may differ in expertise and background. In addition, we conjecture that at least the following three characteristics distinguish domains of high and low information-heterogeneity: (1) *Presence and Strength of an Anchor Point for the Forecasted Variable*: a strong anchor point affects all informants and leads to more common information and homogeneity; (2) *Amount of Public vs. Private Knowledge*: more public relative to private knowledge will logically lead to more homogeneity; (3) *Inherent Predictability*: high inherent predictability of the variable under study will logically lead to more homogeneity.

Using these criteria, we selected two quite different domains for our experiment: predicting point spreads for two specific college football games and predicting the values of two financial indices. There is a weak anchor point for college football point spreads where changes in lineups, field conditions, the changing "buzz" about specific games and other factors lead to diverse and rapidly changing assessments about the outcome of the event. In contrast, the general stability of many financial market indices makes current prices strong anchors. Also, the “real” market that runs openly for the latter task serves as another strong anchor.

Similarly, for public versus private information: most students saw the financial indices they were trading or forecasting (as well as speculations about them) daily, while only the more knowledgeable and dedicated football fans accessed news group and additional information sources on the Internet about football odds. A similar argument follows for predictability (with college football scores inherently much more unpredictable than financial indices).

The specific football task we chose was to predict the point spread for the score for two college football games of national championship significance to be played on 20 November 2004: Michigan vs. Ohio State (labeled here OSU) and Florida vs. Florida State (labeled here FSU), while for financial indices the task was to forecast the Dow Jones Index and the Crude Oil Spot Market Price (Texas Intermediate) on 20 November 2004.¹

Informants

We sought a subject population representative of those likely to be involved in institutional market forecasts. Hence, business school students (upper level undergraduate and MBAs) represent an appropriate population pool. All had taken at least one course in finance and were familiar with financial indices and instruments. Also, all participants were aware of college football but varied greatly in their knowledge and interest in the game. Overall, 126 business school students participated in our study.

Treatments

We sought a between-subjects design to control for dependencies between tasks and we needed to allow for an updating process to assimilate feedback and market information in a natural setting. Hence, we developed an intertemporal study design. Our institutional estimation framework required relatively small groups of participants to create a “market” (in the PM approach environment) and to aggregate multiple independent forecasts in a formal manner for the CJF approach. Previous research from experimental economics (Sunder 1995; Plott 2000) reported that markets with as few as six participants could efficiently disseminate and aggregate information, the number we chose for our design. Hence, we formed 21 experimental groups consisting of six randomly assigned individuals, each in one of two conditions:

¹ There is a considerable time gap between data collection and the publication of this paper. That gap is due to a analysis and publication delays but does not compromise the relevance of the findings

- Condition 1: consisting of 11 groups predicting the football point spreads through the CJF approach and participating in PMs for financial indices.
- Condition 2: consisting of 10 groups predicting the financial indices through the CJF approach and participating in PMs for football point spreads.²

Our 126 subjects were selected after a pre-experimental assessment of their football and financial knowledge. To qualify, participants had to get 7 or more football questions correct and 8 or more finance questions correct out of 10. Sample questions for the pre-experimental assessment of football knowledge were: “a team must go 10 yards or more to get a first down” (True/False) or “a team gets 7 points for a field goal” (True/False). Sample questions for the assessment of financial knowledge were: “it is better to buy high and sell low than buy low and sell high” (True/False) or “Dow Jones is a well-respected radio financial commentator” (True/False).

Procedures

The PMs were open for 22 days prior to 20 November 2004. As all participants were involved in a CJF approach condition (as well as in a PM), they were required to provide forecasts and confidence scores about their forecasts in the CJF approach task four times during those 22 days via an electronic survey. We assessed participant competence via a separate knowledge questionnaire given when the study began. Sample questions to measure football knowledge were: “the number of time outs in a half is (the Same/Different) between college and NFL football” and “the field size is (the Same/Different) between college and NFL football.” Sample questions to measure knowledge of financial markets were: “shares of common stocks always pay dividends” (True/False) and “preferred stock ownership usually ensures voting rights in a company” (True/False).

Incentives: Participants were told they would receive both a (fixed) participation payment and additional compensation based on their performance. Performance compensation in the PM approach was linearly related to the value of the participant’s portfolio at the end of the study. Compensation in the CJF approach

² To leverage the use of our participants, we had them all participate in both a PM and in a direct prediction (CJF) approach. However the two domains were completely different (sports scores and financial markets) and, hence, the PM approach and the CJF approach conditions were independent.

was based on the participant's mean overall financial index price or point spread prediction accuracy. These compensation schemes were designed to provide significant and roughly equivalent incentives for all participants to apply effort and attention to these tasks. While we cannot exclude the possibility that the two tasks (participating in a PM and filling out questionnaires for the CJF and KI approach) will induce different levels of (intrinsic) motivation, we designed the study to ensure that none of the approaches was favored financially. Similarly, although we cannot rule out the possibility that participants in various groups and conditions exchanged information offline during the time our study was running, we see no evidence that such possible exchanges did or, indeed, could systematically affect our results. We set incentives for individual performance, providing a disincentive to share information. Individual participants were randomly and anonymously assigned randomly to study conditions, making information sharing nearly impossible. Even had they been able to do so, we see no reason for that (unlikely) sharing to favor one condition over another..

Operation of the PMs. Each PM was comprised of six (anonymous) individuals in a condition and two different stocks. Depending on the condition, the stocks represented either the value of a football point-spread or the value of a financial index on 20 November 2004. The payoff function for the football point-spread stock types gives a cash dividend of \$1 (virtual) for every point in the point-spread:

$$(1) \quad d_i = Z_i$$

where:

$i=1,2$: 1: OSU, 2: FSU,

d_i : Cash dividend of the stock modeling the outcome of the i -th football game,

Z_i : (Absolute value of) point-spread of the i -th football game.

For the financial indices, we used two different payoff functions to adjust for the different scale levels of the financial indices. The shares of the stock for the price of crude oil paid \$1 (virtual) for every \$1 (real) per barrel of crude oil (see Equation (2)). The shares of the stocks for the Dow paid \$1 (virtual) for every 1,000 points of this index (see Equation (3)).

$$(2) \quad d_{crude} = Z_{crude}$$

$$(3) \quad d_{Dow} = \frac{Z_{Dow}}{1,000}$$

where:

- d_{crude} : Cash dividend of the stock modeling the price of crude oil on November 20th, 2004,
 d_{Dow} : Cash dividend of the stock modeling the value of the Dow on November 20th, 2004,
 Z_{crude} : Price of crude oil on November 20th, 2004,
 Z_{Dow} : Value of the Dow on November 20th, 2004.

Thus, the price of a share of stock for a specific football game or financial index represented a prediction of its value on 20 November 2004 by inverting the payoff function. We set the initial quotes for the football point-spreads based on the performances of the teams up to that date and the initial quotes for the financial indices based on their actual value on 27 October 2004.

The experiment ran from 29 October to 19 November 2004. Participants received an initial endowment of 100 shares of each stock type in their group-specific PM and \$2,500 (virtual) cash. Based on their performance in the PM (measured by the value of their final portfolio), participants received a bonus payment. We used a market maker mechanism so that participants could trade anytime, 24 hours per day, seven days a week. There was no trading fee. (See Appendix A for a discussion of the Market Maker Mechanism and Appendix B for sample screen shots characterizing the operation of the markets.)

The CJF Approach Procedure. To create forecasts using the (online) survey-based forecasts in the CJF approach conditions, we created “virtual” groups consisting of 6 persons each. We randomly assigned subjects to groups and ex-post analysis showed no significant differences between groups, either in financial knowledge ($F=1.492$, $df=20, 125$, $p=.100$) or football knowledge ($F=.882$, $df=20, 125$, $p=.882$). Participants in the CJF approach condition provided their forecasts by filling out an online questionnaire.

We developed aggregated forecasts for each 6-person group based on the questionnaire data for each of the 22 days. The values of the aggregated forecasts varied over the 22 days because all participants received the questionnaire the same day, but were allowed to send it back on one of the following days. Hence, responses varied over time and our results are based on an aggregation of the most recent six forecasts at any point in time. Following the approach developed in Van Bruggen et al. (2002), we computed the unweighted average forecast UF_{fw} for group w at day t using Equation (4).

$$(4) \quad UF_{ftw} = \frac{\sum_{i=1}^6 Forecast_{ftwi}}{6}$$

where $Forecast_{ftwi}$ is the forecast of individual i , in group w , at day t , for index f .

$w = 1, \dots, p$ ($p = 11$ for the football indices and $p = 10$ for the financial indices)

$t = 1, \dots, 22$

$f = 1, \dots, 4$ (1 = OSU point spread; 2 = FSU point spread; 3 = Dow Jones Index; 4 = Oil Price)

Following the approach developed in Van Bruggen et al. (2002) we also computed knowledge-based and confidence-based weighted averages. We computed the knowledge-based forecast KBF_{ftw} for group w at day t using Equation (5)

$$(5) \quad KBF_{ftw} = \frac{\sum_{i=1}^6 Knowledge_{fi} * Forecast_{ftwi}}{\sum_{i=1}^6 Knowledge_{fi}}$$

where $Knowledge_{fi}$ is the knowledge score of individual i about variable f measured using the knowledge test items. We computed the confidence-based forecast CBF_{ftw} for group w at day t using Equation (6)

$$(6) \quad CBF_{ftw} = \frac{\sum_{i=1}^6 Confidence_{fi} * Forecast_{ftwi}}{\sum_{i=1}^6 Confidence_{fi}}$$

where $Confidence_{fi}$ is the confidence of individual i about variable f . Study participants rated their confidence about the correctness of their forecasts on a 5-point scale anchored by “not confident at all” and “very confident”.

We also developed a Key-Informant Forecast KIF_{ftw} for each 6-person group. We selected the key informant as the one with the highest knowledge score within the group according to the knowledge responses to our knowledge questions. We broke ties by selecting one key informant in a group at random.

Measures

We compare the forecasting accuracy of the CJF-based measures and the PM-based forecasts by

computing the Mean Absolute Percentage Error (MAPE) of the deviations of the forecasts for the four variables Z_1 , Z_2 , Z_{crude} and Z_{Dow} from the actual outcome on November 20 (AV_f). The MAPE, computed as in Equation (7), is invariant to scale, is not influenced by outliers.

$$(7) \quad MAPE_{fmt} = \left| \frac{\text{Forecast}_{fmt} - AV_f}{AV_f} \right|$$

where $MAPE_{fmt}$ is the Mean Absolute Percentage Error for variable f for method m , where 1 = unweighted aggregated CJF forecast, 2 = Key Informant, and 3 = PM-based forecast at day t .

The actual outcomes of the four variables forecasted were as follows:

Actual Point Spread Florida vs. Florida State: 7

Actual Point Spread Ohio State vs. Michigan: 16

Actual Crude Oil Price: \$48.90

Actual Dow Jones (divided by 1,000): \$10.46

RESULTS

Table 2 presents the average MAPE values of the key informants, of the various CJF-based forecasts and of the PM-based forecasts for the financial indices and for the football point spreads across 12 days. While our PMs ran for 22 days, we focus on the middle 12 days for analysis here, eliminating the first and the last five days of trading. As is common in markets where the organizer sets the initial price, there is a transient period of volatility before the market settles to set a (new) price (Das 2005). We dropped the last five days of trading to provide a fair comparison between the PM approach and the CJF - and KI approaches.³ During these last five days the PM participants were able to trade and react to informational events and thus update their forecasts, while the vast majority of participants in the CJF- and KI approaches were not able to do so as they had already provided their last forecast. The results are graphically presented in Figure 1 and 2. (We replicated the analysis with the full data set and the results were similar to those reported here).

³ As our task involves forecasting, it is appropriate to establish a time interval between the PM-established valuation and the realization of the actual valuation. In addition, as trading took place after the last direct forecasting task, including these days in the sample would have inappropriately favored the PM approach.

--- INSERT TABLE 2, AND FIGURE 1 AND 2 ABOUT HERE ---

Overall our results show that the values of the financial indices were more accurately predicted than the football point spreads ($F=79.998$, $df=1, 209$, $p=.000$). The average MAPE across the several CJF approaches and the PM approach is .072 for the financial indices while it is .488 for the football point spreads. In contrast with the findings reported in Van Bruggen et al. (2002), we do not find weighted averages to do better than unweighted averages. A possible explanation for this finding (in line with other research results (Armstrong 2001)) might lie in the participant screening procedure, which eliminated low knowledge (and low confidence—a highly correlated variable) individuals. This screening procedure decreased variation in knowledge and confidence levels within groups, leading to a reduced impact of these weights. Although there is a chance that biases like overconfidence affected the assessments of individual study participants, we have no reason to believe that such biases would have affected the results of certain groups more than others.

Since we found no differences between weighted and unweighted CJF approaches, we only analyzed the unweighted CJF results further. These results show an interaction effect between the forecasting approaches (CJF approach vs. PM approach) and the forecasting tasks (Football Point Spreads vs. Financial Indices) ($F=2.350$, $df=6, 114$, $p=.035$) (see also Figure 1 and Figure 2).

We found no significant differences in the forecasting accuracy for financial indices between the CJF approaches, the KI approach and the PM approach. This finding thus does not support H_1 . However, closer inspection of Figure 1 shows that there is a difference between the results for the first part of the experimental period and the second part: in the second part of the forecasting period, the CJF - and KI approaches actually do somewhat better than the PM approach. While the reason for this result is not clear, it may be that the complexity of the PM task overshadowed its informational advantages, especially in the latter part of the forecasting period when most relevant information was likely already exchanged.

In predicting football point spreads, however, the PM approach clearly outperformed the CJF approaches, a result that supports H_1 . Figure 2 also reveals that the KI approach forecasts of the football point spreads are less accurate than the results of either the CJF approach or the results of the PM approach, also providing support for H_1 . The positive result for football point spreads (high information heterogeneity) compared to

the null result for financial markets (low information heterogeneity) provides support for H₂.

To elaborate on this latter point we note that our results show that at the beginning of the forecasting period (22 days before the event or the close of trading) the mean absolute percentage error (MAPE) for the football point spreads was .42 while it was .11 for the financial indices ($F=79.99$, $df=1, 209$, $p=.000$). In addition, both the domain of sports scores and that of financial indices have been intensively used for research on forecasting, demonstrating substantial uncertainty and unpredictability (for a recent summary see Andersson, Edman, and Ekman (2005) and Törngren and Montgomery (2004)). Furthermore, in interviews, our participants identified financial markets as a domain where they have little unique information and low information heterogeneity whereas they viewed the task of predicting (American) football game scores and point spreads as a high information heterogeneity domain with less common information.

Hence, we conclude that our two tasks differ on information heterogeneity and that this difference helps explain why the PM approach is especially effective in predicting football points spreads and less so in predicting financial indices.

We also performed additional analyses to develop a deeper understanding of our results. We investigated how knowledge affected the accuracy of both the CJF approach and of the PM approach; our analyses did not identify a relationship between average knowledge levels, knowledge of the most knowledgeable participant or knowledge dispersion and forecast quality. Again, this result could have been driven by our screening out less knowledgeable participants. We did find a nearly significant relationship ($r=-.392$, $n=60$, $p=.088$) between participants' knowledge and forecasting accuracy for the financial market indices, but could identify no other links between knowledge and forecasting ability.

Given the small number of traders and trades in our markets, we investigated the link between trading activity and accuracy. While we observed more active trading in the market for financial indices than for football point spreads, that difference was not statistically significant. And we found no relationship between level of trading activity and forecasting accuracy for either market type, suggesting that there was sufficient trading activity for market efficiency. We did find, however, that within markets more active traders performed better ($r=.875$, $n=65$, $p=.000$ for PMs for financial indices and $r=.362$, $n=60$, $p=.005$ for PMs for

football point spreads), suggesting that trading activity itself might indicate market knowledge.

DISCUSSION

We have investigated the feasibility and accuracy of Prediction Markets (PMs) for forecasting situations characterized by a small number of knowledgeable participants, typical for institutional forecasting. Our results show that PMs are feasible in environments with varying degrees of information-heterogeneity, and can be conducted effectively with group sizes as small as six traders per market.

We found that PMs outperform the commonly applied approaches of Combined Judgmental Forecasts (CJF) and Key Informants (KI) in more difficult-to-predict environments, characterized by high knowledge and information heterogeneity between participants. The approach worked as well as the more traditional CJF - and KI approaches in more homogeneous environments. We attribute this superior performance of the PM approach in high information-heterogeneity environments to their ability to provide for information *exchange* between participants: trader's private, unique and unshared information becomes shared and common through their trades.

We also found that forecasts of both the PM approach and of the CJF approach are more accurate than those of the KI approach in situations we characterize as high information-heterogeneity. This finding may emerge because improved accuracy in any high information-heterogeneity environment requires a mechanism to consider and weight different expectations; both the PM approach and the CJF approach provide such a mechanism while the KI approach does not. Because our screening procedure produced a limited absolute information-heterogeneity level among our participants, our results should be viewed as a conservative test of the relative advantage of the PM approach in situations with high information-heterogeneity and we would expect a stronger advantage for the PM approach in cases of an even higher level of information-heterogeneity. Similarly, the PM approach's benefits should also be expected to be larger if market participants are more experienced traders than those we used in our study (Berg, Forsythe, and Rietz 1996).

Our results raise questions about why and when the PM approach should be expected to perform well in the field. If information exchange and learning change participant's knowledge, then the PM approach has

two intertemporal advantages over the CJF approach. First, participants improve their knowledge over time through the exchange mechanism, mitigating the effect of low-knowledge participants. Second, if the weights used for aggregation in the CJF approach, typically taken at one point in time, are not updated, those weights may become suboptimal if the environment is unstable or highly unpredictable. A PM, operating continuously, can aggregate such environmental and knowledge-based changes naturally and continuously and may also be effective in deriving participants' weights for future analyses (see Chen et al. (2004)). In more stable, low information-heterogeneity environments, those problems do not exist and, hence, the PM approach exhibits no advantage over the KI approach or the CJF approach. Indeed, our results imply that the KI approach, the simplest, most cost-effective mechanism, should generally be chosen in such situations as that the additional use of information technology provides little value.

If our results hold in further testing, they suggest some useful managerial implications. First, Prediction Markets may remedy some of the limitations of existing Group Support Systems (GSS). They seem to be effective with small numbers of participants, a challenging domain for traditional GSS (Gallupe et al. 1992). Secondly, better (institutional) forecasts may be achieved with the PM approach compared with the CJF - and the KI approach, with that improvement being most significant in environments with high information-heterogeneity. Also, the PM approach does no worse than the CJF - or the KI approach in low information-heterogeneity environments, making it a robust choice when little is actually known about environmental information heterogeneity.

Thirdly, since the PM approach is superior to the CJF approach in high information-heterogeneity environments and the KI approach is no worse than the CJF approach in the low information-heterogeneity environment, it may be that the CJF approach should be a third choice option in most institutional forecasting environments.

Our results suggest several opportunities for future study. First, we compared the PM approach with approaches (CJF and KI) that are commonly applied in the practice of institutional forecasting. However, we did not compare them directly with more "traditional" GSSs which, like PMs do allow information exchange between participants. In future research the accuracy of PMs could thus also be compared directly with GSSs

even though these systems are intended to facilitate especially group decision rather than group forecasting processes. Second, we allowed all participants to search for additional information, but did not provide any help in doing so. As such we treated all participants equally. We suspect that different ways to provide additional information lead to trading behavior and a market outcome that differ in quality. Third, our procedure screened out low and medium knowledge individuals so we could not determine how any of the procedures would have fared were they included. Including low and medium knowledge individuals would further increase the knowledge-heterogeneity among the participants, increasing market liquidity and potentially increasing the relative advantage of the PM approach

In line with past work, we have set the size of the PMs to six individuals in our experiment. We have no data on how the performance of the PM approach varies with group size or when a market is too small. How market efficiency varies with group size (is there a minimize market size, and if so, what?) is an important topic for future research. In addition, given our arguments about knowledge sharing in high-information-heterogeneity environments as the explanation for the PM approach's performance, it would be useful to directly compare other repeated information-sharing mechanisms, perhaps of the repeated-Delphi type. (See Cil, Alpturk, and Yazgan (2005), for example).

While there are clearly many other research opportunities, we reemphasize our main findings: The PM approach is feasible in situations with small numbers of participants and most effective in situations with high information-heterogeneity. These are characteristics shared by many common institutional forecasting settings in organizations. And the PM approach appears to provide forecasting accuracy that is superior to other methods that are in common use in such situations.

REFERENCES

- Anderson, L.R. and C.A. Holt (1997), "Information Cascades in the Laboratory," *American Economic Review*, 87 (5), 847-62.
- Andersson, P., J. Edman, and M. Ekman (2005), "Predicting the World Cup 2002 in Soccer: Performance and Confidence of Experts and Non-Experts," *International Journal of Forecasting*, 21 (3), 565-76.
- Armstrong, J.S. (2001), *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Norwell MA: Kluwer Academic Publishers.
- Barnes, P. (1986), "Thin Trading and Stock Market Efficiency: The Case Of The Kula Lumpur Stock Exchange," *Journal of Business Finance and Accounting*, 13 (4), 609-17.
- Batchelor, R. and P. Dua (1995), "Forecaster Diversity and the Benefits of Combining Forecasts," *Management Science*, 41 (1), 68-75.

- Berg, J., R. Forsythe, and T. Rietz (1996), "What Makes Markets Predict Well? Evidence from the Iowa Electronic Markets," in *Essays in Honor of Reinhard Selten*, Wulf Albert and Werner Güth and Peter Hammerstein and Benny Moldovanu and Eric Van Damme, Eds.: Springer Verlag, 444-63.
- Berg, J.E. and T.A. Rietz (2003), "Prediction Markets as Decision Support Systems," *Information Systems Frontiers*, 5 (1), 79-93.
- Berg, J.E., F.D. Nelson, and T.A. Rietz (2008), "Prediction Market Accuracy in the Long Run," *International Journal of Forecasting*, 24 (2), 285-300.
- Chen, K.-Y., L.R. Fine, and B.A. Huberman (2003), "Predicting the Future," *Information Systems Frontiers*, 5 (1), 47-61.
- Chen, K.-Y., L.R. Fine, and B.A. Huberman (2004), "Eliminating Public Knowledge Biases in Information-Aggregation Mechanisms," *Management Science*, 50 (7), 983-94.
- Cil, I., O. Alpturk, and H.R. Yazgan (2005), "A New Collaborative System Framework Based on a Multiple Perspective Approach: InteliTeam," *Decision Support Systems*, 39 (4), 619-41.
- Dahan, E., A.W. Lo, T. Poggio, N. Chan, and A. Kim (2007), *Securities Trading of Concepts (STOC)*, Working Paper,
- Das, S. (2005), "A Learning Market-Maker in the Glosten-Milgrom Model," *Quantitative Finance*, 5 (2), 169-80.
- Dennis, A.R. (1996), "Information Exchange and Use in Group Decision Making: You Can Lead a Group to Information, but You can't Make It Think," *MIS Quarterly*, 20 (4), 433-57.
- Dennis, A.R. and M.J. Garfield (2003), "The Adoption and Use of GSS in Project Teams: Toward More Participative Processes and Outcomes," *MIS Quarterly*, 27 (2), 289-323.
- DeSanctis, G. and R.B. Gallupe (1987), "A Foundation for the Study of Group Decision Support Systems," *Management Science*, 33 (5), 589-609.
- Elberse, A. (2007), "The Power of Stars: Do Star Actors Drive the Success of Movies?," *Journal of Marketing*, 71 (4), 102-20.
- Evans, M. (2002), *Practical Business Forecasting*, Malden MA: Blackwell Publishing.
- Fama, E.F. (1991), "Efficient Capital Markets II," *Journal of Finance*, 46, 1575-617.
- Fildes, R. and R. Hastings (1994), "The Organization and Improvement of Market Forecasting," *Journal of the Operational Research Society*, 45 (1), 1-16.
- Forsythe, R., T.A. Rietz, and T.W. Ross (1999), "Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets," *Journal of Economic Behavior & Organization*, 39 (1), 83-110.
- Foutz, N.Z. and W. Jank (2008), *The Wisdom of Crowds: Pre-release Forecasting for New Products via Functional Shape Analysis of the Online Virtual Stock Market*, Working Paper,
- Fung, J.K.W., H.M.K. Mok, and K.C.K. Wong (2004), "Pricing Efficiency in a Thin Market with Competitive Market Makers: Box Spread Strategies in the Hang Seng Index Options Market," *Financial Review*, 39, 434-54.
- Gallupe, R.B., A.R. Dennis, W.H. Cooper, J.S. Valacich, L.M. Bastianutti, and J.F. Nunamaker (1992), "Electronic Brainstorming and Group Size," *Academy of Management Journal*, 35 (2), 350-69.
- Garthwaite, P.H., J.B. Kadane, and A. O'Hagan (2005), "Statistical Methods for Eliciting Probability Distributions," *Journal of the American Statistical Association*, 100 (470), 680-700.
- Grossman, S. (1976), "On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information," *Journal of Finance*, 31 (2), 573-85.
- Grossman, S.J. and J.E. Stiglitz (1980), "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70 (3), 393-408.
- Grossman, S.J. (1981), "An Introduction to the Theory of Rational Expectations under Asymmetric Information," *Review of Economic Studies*, 48 (4), 541-59.
- Gruca, T.S., J. Berg, and M. Cipriano (2003), "The Effect of Electronic Markets on Forecasts of New Product Success," *Information Systems Frontiers*, 5 (1), 95-105.
- Guo, Z., F. Fang, and A.B. Whinston (2006), "Supply chain information sharing in a macro prediction market," *Decision Support Systems*, 42 (3), 1944-58.
- Hayek, F.A.v. (1945), "The Use of Knowledge in Society," *American Economic Review*, 35 (4), 519-30.

- Hightower, R. and L. Sayeed (1996), "Effects of Communication Mode and Prediscussion Information Distribution Characteristics on Information Exchange in Groups," *Information Systems Research*, 7 (4), 451-65.
- Kumar, N., L.W. Stern, and J.C. Anderson (1993), "Conducting Interorganizational Research Using Key Informants," *Academy of Management Journal*, 36 (6), 1633-51.
- LaComb, C.A., J.A. Barnett, and Q. Pan (2007), "The Imagination Market," *Information Systems Frontiers*, 9 (2-3), 245-56.
- Malkiel, B.G. (2003), "The Efficient Market Hypothesis and Its Critics," *Journal of Economic Perspectives*, 17 (1), 59-82.
- Ostrover, S. (2005), "Employing Information Markets to Achieve Truly Collaborative Sales Forecasting," *Journal of Business Forecasting*, 24 (1), 9-12.
- Pennock, D.M., S. Lawrence, C.L. Giles, and F.A. Nielsen (2000), "The Power of Play: Efficiency and Forecast Accuracy in Web Market Games," in NEC Research Institute Technical Report: NEC Research Institute.
- Pennock, D.M., S. Lawrence, C.L. Giles, and F.A. Nielsen (2001), "The Real Power of Artificial Markets," *Science*, 291 (5506), 987-88.
- Plott, C.R. and S. Sunder (1982), "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational Expectation Models," *Journal of Political Economy*, 90 (4), 663-98.
- Plott, C.R. and S. Sunder (1988), "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085-118.
- Plott, C.R. (2000), "Markets as Information Gathering Tools," *Southern Economic Journal*, 67 (1), 2-15.
- Rodan, S. and C. Galunic (2004), "More than Network Structure: How Knowledge Heterogeneity Influences Managerial Performance and Innovativeness," *Strategic Management Journal*, 25 (6), 541-62.
- Shaw, M.E. (1981), *Group Dynamics: The Psychology of Small Group Behavior* (3rd ed.), New York: McGraw-Hill.
- Sia, C.-L., B.C.Y. Tan, and K.-K. Wei (2002), "Group Polarization and Computer-Mediated Communication: Effects of Communication Cues, Social Presence, and Anonymity," *Information Systems Research*, 13 (1), 70-90.
- Soukhoroukova, A., M. Spann, and B. Skiera (2010), "Generating and Evaluating New Product Ideas with Idea Markets," *Journal of Product Innovation Management*, Forthcoming.
- Spann, M. and B. Skiera (2003), "Internet-Based Virtual Stock Markets for Business Forecasting," *Management Science*, 49 (10), 1310-26.
- Spann, M. and B. Skiera (2009), "Sports Forecasting: A Comparison of the Forecast Accuracy of Prediction Markets, Betting Odds and Tipsters," *Journal of Forecasting*, 28 (1), 55-72.
- Stasser, G. and W. Titus (1985), "Pooling of Unshared Information in Group Decision Making: Biased Information Sampling During Discussion," *Journal of Personality and Social Psychology*, 48 (6), 1467-78.
- Stasser, G. and W. Titus (1987), "Effects of Information Load and Percentage of Shared Information on the Dissemination of Unshared Information During Group Discussion," *Journal of Personality and Social Psychology*, 53 (1), 81-93.
- Stasser, G. (1992), "Information Salience and the Discovery of Hidden Profiles by Decision Making Groups: A 'Thought Experiment'," *Organizational Behavior and Human Decision Making Processes*, 52, 156-81.
- Stracca, L. (2004), "Behavioral Finance and Asset Prices: Where do We Stand?," *Journal of Economic Psychology*, 25 (3), 373-406.
- Sunder, S. (1992), "Market for Information: Experimental Evidence," *Econometrica*, 60 (3), 667-95.
- Sunder, S. (1995), "Experimental Asset Markets: A Survey," in *The Handbook of Experimental Economics*, J.H. Kagel and A.E. Roth, Eds.: Princeton University Press, 445-500.
- Surowiecki, J. (2005), *The Wisdom of Crowds: Why the Many Are Smarter Than the Few*, London: Abacus.
- Törngren, G. and H. Montgomery (2004), "Worse than Chance? Performance and Confidence among Professionals and Laypeople in the Stock Market," *Journal of Behavioral Finance*, 5 (3), 246-51.
- Van Bruggen, G.H., G.L. Lilien, and M. Kacker (2002), "Informants in Organizational Marketing Research: Why use Multiple Informants and How to Aggregate Responses," *Journal of Marketing Research*, 39 (4), 469-78.
- Wolfers, J. and E. Zitzewitz (2004), "Prediction Markets," *Journal of Economic Perspectives*, 18 (2), 107-26.

Table 1
Comparison of Main Information Characteristics of the Prediction Market (PM) and the Combined Judgmental Forecasts (CJF) Approach to Institutional Forecasting

| | Combined Judgmental Forecasts | Prediction Markets |
|---|---|---|
| Information Collection <ul style="list-style-type: none"> • About forecasted variable • About knowledge and confidence | <ul style="list-style-type: none"> • Direct and explicit measurement of variables of interest • Information available to each informant • Incentives for providing biased information may be present • Direct (self)assessments of knowledge and/or confidence • Providing responses is straightforward and relatively easy for informants | <ul style="list-style-type: none"> • Indirect measurement of variables of interest through observation of trading behavior • Information provided not directly visible • Incentives for speculation may be present • No direct assessment of knowledge and/or confidence • Market participation can be relatively complex for informants |
| Information Dissemination | <ul style="list-style-type: none"> • Only public information is shared and information of each informant is not broadcasted to other informants | <ul style="list-style-type: none"> • Private information of each informant becomes public through trading and is thereby broadcasted through the market mechanism |
| Information Aggregation | <ul style="list-style-type: none"> • Aggregation explicitly done by the researcher • Various alternative (mostly proportional) weighting schemes can be applied • Knowledge-based weighting possible | <ul style="list-style-type: none"> • Aggregation implicitly done through market mechanism • Disproportional weighting of the input of the strongest believer |

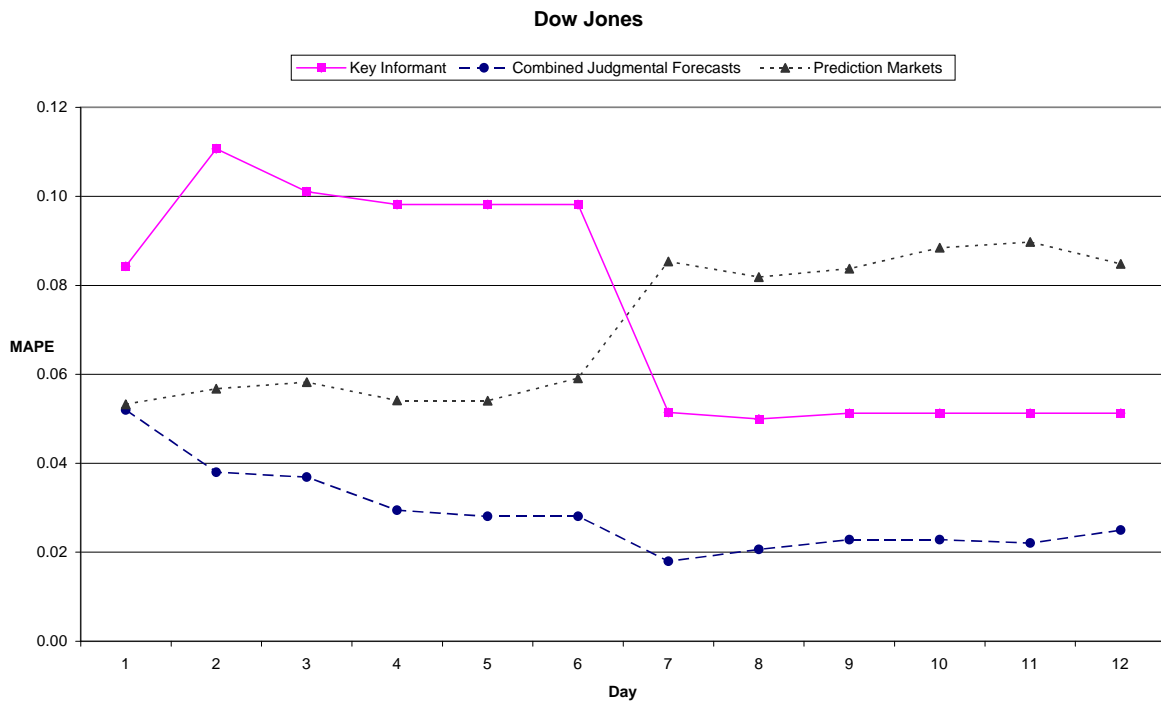
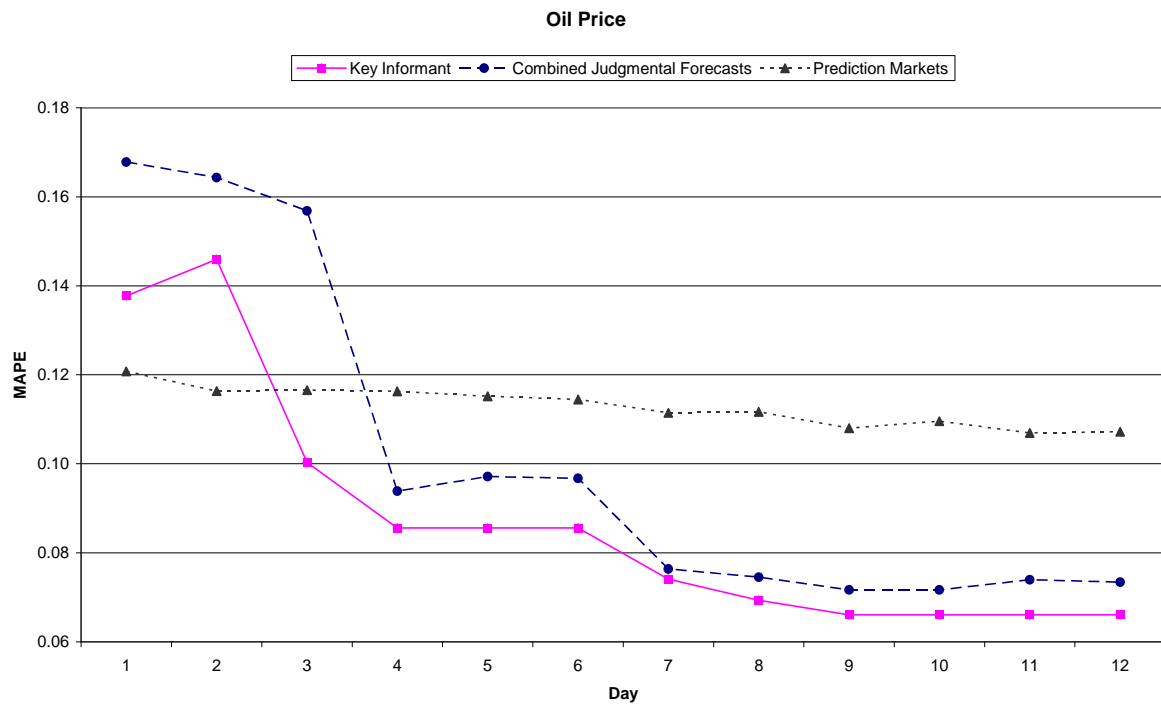
Note: The Key Information approach, a special case of Combined Judgmental Forecasts, shares the Combined Judgmental Forecasts characteristics for Information Collection while Information Broadcasting (i.e. Information Dissemination) and Information Aggregation considerations do not apply

Table 2

Forecasting Accuracy of the Key Informant (KI), Combined Judgmental Forecasts (CJF) and Prediction Markets
 MAPE (Standard Deviation) averaged across 12 Days

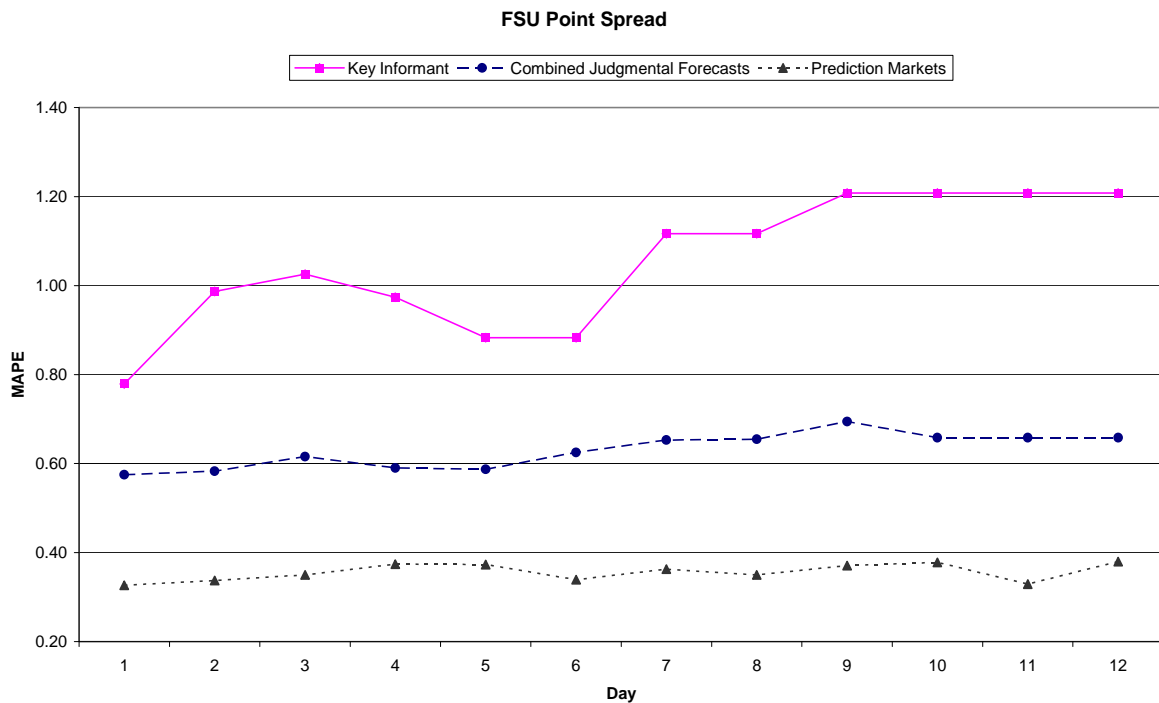
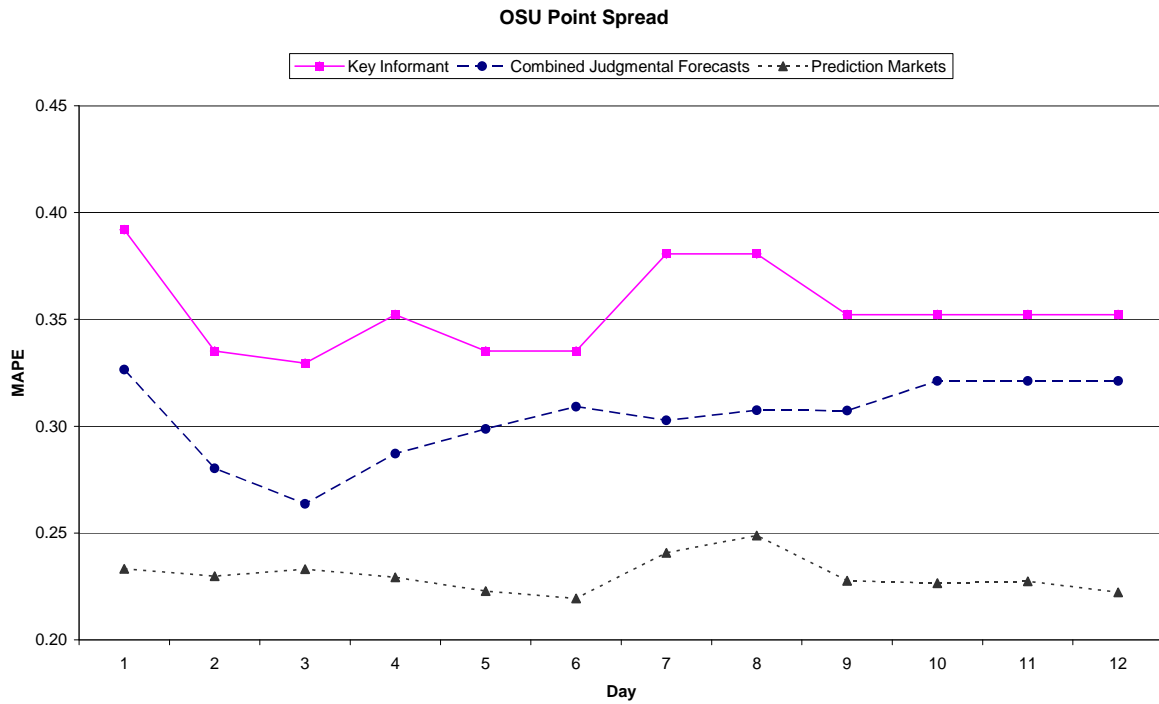
| | | Key Informant | Combined Judgmental Forecasts | | | Prediction Markets |
|------------------------|-----------|---------------|-------------------------------|-----------------------------|-------------------------------|--------------------|
| | | | Unweighted | Knowledge-Based Weighted | Confidence- Based Weighted | |
| Football Point Spreads | OSU | .354 (.216) | .304 (.121) | .301 (.129) | .307 (.111) | .230 (.060) |
| | FSU | 1.050 (1.12) | .630 (.305) | .672 (.337) | .640 (.292) | .356 (.194) |
| | Mean | .702 (.866) | .467 (.281) | .486 (.313) | .473 (.275) | .293 (.154) |
| Financial Indices | Oil Price | .087 (.028) | .102 (.050) | .099 (.045) | .090 (.030) | .113 (.017) |
| | Dow Jones | .075 (.101) | .029 (.014) | .030 (.019) | .023 (.010) | .071 (.044) |
| | Mean | .081 (.072) | .065 (.052) | .065 (.049) | .056 (.041) | .092 (.039) |
| Mean | | .406 (.697) | .276 (.288) | .286 (.311) | .275 (.290) | .188 (.149) |

Figure 1: Mean Absolute Percentage Errors (MAPE) for the Financial Indices Forecasts



Note: Graphs exclude first and last five days of the market; Day 1 refers to 6th day of trading

Figure 2: Mean Absolute Percentage Errors (MAPE) for the Football Point Spread Forecasts



Note: Graphs exclude first and last five days of the market; Day 1 refers to 6th day of trading

Appendix A

MARKET MAKER MECHANISM

Our main criterion for selecting the market mechanism was that it should maintain a sufficient level of liquidity, it should be easy to understand, and it should lead to efficient prices. The most common market mechanism in prediction markets - the double auction, requires a matching bid-and-ask order pair in order to determine the price of a trade. This competitive mechanism operates well in the case of a large number of traders that are willing to sell and buy stocks. However, in small PMs that last several days or weeks like our PMs, a rather low trader-to-stock ratio is likely. The reasons for this are that the number of informants is small and the informants trade asynchronously. In this situation, a double auction might limit trading opportunities, which might decrease the informants' interest in the PM. The use of a trading agent similar to a market maker on NASDAQ or the "Virtual Specialist" at the Hollywood Stock Exchange avoids this shortcoming. The automatic version of the trading agent uses an algorithm to automatically set price quotes that allow the buying and selling of stocks at any time.

In our study the PMs thus applied a two-sided automated market maker trading mechanism, comparable to the one used at NASDAQ to avoid lack of trading opportunities that can arise due to our experimental setting of only six traders per market. Our market maker accepted every order from a informant and executed it at a pre-announced price that is identical for purchase or sale orders. That price was adjusted after every executed order by an automatic price adjustment procedure.

The use of the market maker trading mechanism allowed informants to trade anytime at a pre-announced price. Purchases increased the price p for the next order, sales decreased this price. The goal of our price adjustment mechanism was to set a price p according to an estimate of the stock's true value V based on traders' order flow: informed traders are aware of changes in V and trade accordingly. Each trade of a single stock represents a signal to the market maker. Therefore, the quantities of stocks per transaction as well as the number of transactions in the same direction (i.e. the number of purchases or sales) were indicators of the possible magnitude of the deviation between p and V . We applied this principle in determining the price adjustment based on a moving window of the last I transactions, accounting for volume and direction of each transaction. To increase robustness, we also used information about the maximum

possible value of V to scale the magnitude of price adjustment per share. We tested our mechanism both numerically and empirically before the application in our PMs and set a maximum order quantity of 50 shares to stabilize the markets. The latter characteristic led to more frequent price adjustments in case of large orders. An important goal for the design of the market maker mechanism together with the initial portfolio size and the amount of cash the informants received was to provide market informants with the possibility and freedom to fully express their opinions in their trading behavior. The following equation gives our price adjustment function with the parameter values used in our experiment.

$$p_{j,n} = p_{j,n-1} + \max \left\{ \text{sig}_{j,n} \cdot q_{j,n} \cdot \frac{p_{j,\max}^2}{\gamma^2} \cdot \frac{\left(\sum_{i=0}^{I_n} q_{j,n-i} \right)}{I+1}, \alpha \right\}$$

$$\text{with } \text{sig}_{j,n} = \begin{cases} -1 & \text{for sale} \\ 1 & \text{for purchase} \end{cases}, I_n = \begin{cases} n & \text{for } n < I = 10 \\ 10 & \text{for } n \geq I = 10 \end{cases} \quad (n > 0, j \in J),$$

where:

$p_{j,n}$: Market maker price for j -th stock after the n -th trade,

$q_{j,n}$: Quantity of order of j -th stock at the n -th trade,

$\text{sig}_{j,n}$: Sign of the order of j -th stock at the n -th trade,

$p_{j,\max}$: Maximum price for j -th stock,

I_n : Length of moving average window,


J : Index set of stocks,

γ : Scaling parameter (with $\gamma = 500$),

α : Minimum tick size (with $\alpha = 0.01$).

The market prices in the PMs during the 22-day trading period represented the forecasts.

Appendix B Introductory Screen of the Prediction Market (PM)



All Stocks

Homepage

Log in

Username

Password

Welcome to the virtual stock market.

You have an initial endowment of \$ 2,500, and 100 shares in each of the securities in your portfolio. The stock market is open 24 hours a day.

Your goal is to maximize the value of your portfolio, (your cash plus the liquidation value of your stocks) on November 20. You will receive a performance bonus based on the value of you portfolio relative to the value of those of the others playing in the same stock market. (There are up to six traders in your market). Enjoy trading!


Bonus Payment: Everyone who completes all the tasks will receive a performance bonus. There will be \$80-\$120 in a BONUS POOL for your stock market depending on the number of participants in that market. Here is how we calculate that PERFORMANCE BONUS.

- + You will be participating in a stock market along with a number of anonymous other participants.
- + At the end of the study, you will have some cash in your account and your stocks will be liquidated (sold) at the latest market price at that time for those securities. The total of that cash plus the liquidated value of your stocks is called your PORTFOLIO VALUE. The same will happen to everyone else in your stock market.
For example, if you have \$10,000 in cash and 100 shares of Stock A, priced at \$12 per share at the end of the study, your portfolio value will equal $\$10,000 + 100 \times \$12 = \$11,200$.
- + Your BONUS SHARE = YOUR PORTFOLIO VALUE / (TOTAL PORTFOLIO VALUE OF EVERYONE IN YOUR STOCK MARKET).
- + Your PERFORMANCE BONUS = Your BONUS SHARE X BONUS POOL.
- + So, if your PORTFOLIO VALUE = \$10,000 and total of the portfolio values of everyone in your stock market (including you) = \$20,000, you will receive $\$10,000 / \$20,000$ or 50% of the bonus pool. If your bonus pool were \$100, you would receive a \$50 bonus.
- + If you perform about as well as everyone else in your market, you can expect to receive about a \$20 bonus.

The menu bar at the left will aid you in navigating the site.

- + The **homepage** lists instructions, basic information on the market, and contact information if you have questions.
- + The **stocks** page lists the three stocks that you may trade with. Click on each stock to reveal historical price data of the stock and detailed stock information. From this screen you may also click on *buy* or *sell* to buy and sell shares of your stock. You will arrive at another screen that will prompt you for the quantity of which you wish to buy or sell, and that asks you to confirm your order.
- + **My portfolio** lists the stocks in your portfolio. This page displays the quantity of the stock that you own, the *current price*, and your *current cash position*. You may also click on "history" to view a list of your past

Explanatory Screen of Prediction Market (PM) for Oil Prices



All Stocks

Homepage

Stocks

My Portfolio

Log out

Crude Oil

Crude Oil


The following cases illustrate how the Crude Oil share market works.

- + Case 1: You have 10 shares of Crude Oil that are worth \$42 in the real market on November 20th.
- + Case 2: You have 10 shares of Crude Oil that are worth \$44 in the real market on November 20th.

If you bought the 10 shares of Crude Oil at \$43, when the shares are redeemed on November 20...

- + Case 1 Answer: Your shares will be liquidated on November 20 at $10 \times \$42 = \420 , for a **loss** of $\$420 - (\$43 \times 10 = \$430) = \10 , and you will retain \$420 in your portfolio.
- + Case 2 Answer: Your shares will be liquidated on November 20 at $10 \times \$44 = \440 , for a **profit** of $\$440 - (\$43 \times 10 = \$430) = \10 , and you will retain \$440 in your portfolio.


Stock Information



| | |
|-------------|---------|
| Start price | \$55.17 |
| Last price | \$54.52 |
| Last Volume | 10 |

[Disclaimer](#) | [Conditions of Use](#) | [Privacy](#)

Explanatory Screen of Prediction Market (PM) for Football Point Spread Shares



All Stocks

Florida vs. Florida State

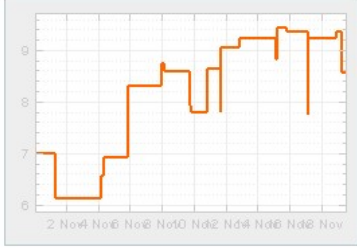
Florida vs. Florida State
The following case illustrates how the point-spread share market works.

- + Case 1: If Florida beats Florida State on November 20 by a score of 21-10, the point spread = 11.
- + Case 2: If Florida State beats Florida by a score of 20-13, the point spread = 7.

If you own 100 Florida-Florida State point spread shares purchased at an average price of 8 (i.e., an outlay of \$800), when the shares are redeemed after the game on November 20th...

- + Case 1 Answer: Your shares will be liquidated on November 20th at $100 \times 11 = \$1,100$, for a **profit** of $\$1,100 - \$800 = \$300$, and you will retain \$1,100 in your portfolio.
- + Case 2 Answer: Your shares will be liquidated on November 20th at $100 \times 7 = \$700$, for a **loss** of $\$700 - \$800 = \$100$, and you will retain \$700 in your portfolio.


Stock Information



| | |
|-------------|--------|
| Start price | \$7.00 |
| Last price | \$8.58 |
| Last Volume | 25 |

[Disclaimer](#) | [Conditions of Use](#) | [Privacy](#)

Portfolio Screen of Prediction Market (PM) for Financial Shares



All Stocks

My Portfolio

| Stock | Quantity | Average price | Current | | Total value | ! |
|---------------------------|----------|---------------|---------|---------|-------------|--|
| Crude Oil | 100 | \$55.17 | \$54.58 | -1.07 % | \$5,458.00 | buy sell |
| Dow Jones | 100 | \$9.76 | \$11.48 | 17.62 % | \$1,148.00 | buy sell |
| | | | | | Cash | \$2,500.00 |

For more information click on the respective stock.

[Disclaimer](#) | [Conditions of Use](#) | [Privacy](#)

Fertig