Institutional Forecasting:
The Performance of Thin Virtual Stock Markets

Gerrit H. Van Bruggen, Martin Spann, Gary L. Lilien
and Bernd Skiera

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2006-028-MKT
Publication	June 2006
Number of pages	32
Persistent paper URL	
Email address corresponding author	gbruggen@rsm.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

ABSTRACT AND KEYWORDS	
Abstract	We study the performance of Virtual Stock Markets (VSMs) in an institutional forecasting environment. We compare VSMs to the Combined Judgmental Forecast (CJF) and the Key Informant (KI) approach. We find that VSMs can be effectively applied in an environment with a small number of knowledgeable informants, i.e., in thin markets. Our results show that none of the three approaches differ in forecasting accuracy in a low knowledge-heterogeneity environment. However, where there is high knowledge-heterogeneity, the VSM approach outperforms the CJF approach, which in turn outperforms the KI approach. Hence, our results provide useful insight into when each of the three approaches might be most effectively applied.
Free Keywords	Virtual Stock Markets, Forecasting, Information Markets, Electronic Markets
Availability	
Classifications	

Institutional Forecasting: The Performance of Thin Virtual Stock Markets

Gerrit H. Van Bruggen, RSM Erasmus University
Martin Spann, University of Passau
Gary L. Lilien, Penn State University
Bernd Skiera, Johann Wolfgang Goethe-University Frankfurt am Main

May 15, 2006

Abstract

We study the performance of Virtual Stock Markets (VSMs) in an institutional forecasting environment. We compare VSMs to the Combined Judgmental Forecast (CJF) and the Key Informant (KI) approach. We find that VSMs can be effectively applied in an environment with a small number of knowledgeable informants, i.e., in thin markets. Our results show that none of the three approaches differ in forecasting accuracy in a low knowledge-heterogeneity environment. However, where there is high knowledge-heterogeneity, the VSM approach outperforms the CJF approach, which in turn outperforms the KI approach. Hence, our results provide useful insight into when each of the three approaches might be most effectively applied.

1. Introduction

Firms and institutions of all sorts rely on forecasts of their business activities for planning, scheduling and investment decisions. The domain of forecast use ranges from territorial sales forecasts, input factor price forecasts, forecasts of competitive response to a new market entry, forecasts of revenues in a new geography, currency forecasts and many more. The forecasting literature (see (Evans 2002), for example) goes to great length to describe the "art and science" of forecasting, where the "art" refers to the use of some form of managerial judgment and the "science" is based on historical data and some form of statistical or econometric analysis.

Institutional forecasts often take place in conditions where there is little or no appropriate historical data or where formal methods or models for developing the forecasts are not available. Furthermore, either cost or secrecy concerns often preclude customer surveys as data sources for such forecasts. Examples of such situations include developing a sales forecast for a *new* market, a *new* sales territory or a *new* product, forecasting the price response of a competitor after the firm introduces a *new* product, predicting the sales response to a *major change* in an advertising program and others. The institutional informants in such situations include sales representatives, marketing analysts, business managers and others. Such institutional forecasts are beneficial particularly in situations where cost and time constraints limit large customer surveys and information about the forecasting topic should remain confidential as is typically the case with new initiatives. The institutional forecasting problems we deal with here share the following characteristics: *i*) there is little or no directly relevant data available; *ii*) significant volatility or uncertainty is involved (typical in new or rapidly changing markets); *iii*) no single informant can know the "true" value in advance. *iv*) there are more than one but still a *small number of informants* whose knowledge can be tapped who may (partially, at least) disagree and differ in expertise; and *v*) the forecasts are updated over a period of time.

Armstrong (2001 p. 376) provides a taxonomy of forecasting methods, with the most appropriate method depending largely on whether or not sufficient objective data are available, whether or not good knowledge of relationships between variables exists, and whether or not large changes in the environment or conditions are expected. The various forecasting approaches differ with respect to their accuracy and ease of implementation (Chen, Fine, and Huberman 2003). Armstrong (2001) identifies formal combining rules as well as approaches as the Delphi method to elicit and combine forecasts from the multiple informants in such situations. Yet, Armstrong (2001) does not cite the so-called Virtual Stock Market (VSM) approach as an alternative. VSMs have been (successfully) applied to predict election outcomes (e.g., the Iowa Electronic Markets), the success of movies (e.g., the Hollywood Stock Exchange), sports results (e.g., Tradesports), and future economic data releases (e.g., economicderivatives.com). Ostrover (2005) cites several examples of VSMs used in institutional forecasting applications: Hewlett Packard (HP) uses VSMs to forecast sales, financial, and accounting results while Eli Lilly uses a VSM to identify those drugs in the early stages of development that are most likely to win approval. The results of these applications suggest that VSMs may be an effective approach to institutional forecasting in settings with a relatively large number of participants. However, we could find no reports of VSMs that were applied and compared to other forecasting techniques in the subset of the institutional forecasting domain where there are only few (knowledgeable) informants.

Our goal is to study the feasibility and accuracy of VSMs in the type of institutional forecasting domain outlined above. We seek insight into the problems of *whether, why,* and *when* a VSM is likely to outperform more traditional approaches, which we call Combining Judgmental Forecasts (CJFs). We provide a reference point for our analysis by also considering the forecasting accuracy of a single, key informant (KI), where the key informant is a most knowledgeable agent of the institution.

We study the feasibility and accuracy of VSMs for two types of forecasting settings. The first involves a situation where the informants are in similar circumstances and thus have access to common information. Regular, regional sales forecasts based on the judgments of a group of sales representatives is a typical example of such a setting. We describe such a setting as one with *low knowledge heterogeneity*. The second, less routine forecasting task is one where there is little common information (perhaps with more private information) and hence where informants may differ significantly in the type and quality of the information they have available. A forecast of the market performance of a new product by the various people (e.g., marketing, R&D) involved in developing and introducing it is an example of this second setting. We characterize such a setting as one with *high knowledge heterogeneity*.

We find that in high knowledge-heterogeneity environments, forecasts made through VSMs are more accurate than those developed by CJFs, which are in turn more accurate than those from KIs. In low knowledge-heterogeneity environments, we find that none of the approaches dominate.

The paper proceeds as follows: we first review alternative approaches to institutional forecasting, including the opportunities afforded by VSMs. We then contrast the differences between VSMs and the more traditional approaches, leading us to several hypotheses about the expected differences in forecasting accuracy between VSMs, CFSs and a KI. We next describe the design of our empirical study and present our results. We conclude with a discussion of those results and their implications.

2. Institutional Forecasting: A Comparison of Approaches

Widely used approaches for institutional forecasting include both using a single (key) informant and using multiple informants, including the Delphi Method (see for example (Armstrong 2001)). We review these below and then contrast them with the VSM approach.

2.1 A Single, Key Informant (KI). Perhaps the most widely used approach in practice, because of its simplicity, is the single or key-informant approach, where the informant is most often selected because of knowledge and willingness to communicate that knowledge. The approach suffers from significant drawbacks (Kumar, Stern, and Anderson 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 those multiple informant reports are appropriately combined.

2.2 Multiple Informants—Combining Judgmental Forecasts (CJF). CJF takes advantage of the knowledge of multiple informants and combines the information from those multiple informants into an overall statement. The strategies to combine varying information differ markedly depending upon whether or not the informants interact. Mechanistic aggregation algorithms such as forming simple or weighted averages are used when informants do not interact (Garthwaite, Kadane, and O'Hagan 2005). Van Bruggen, Lilien, and Kacker (2002) propose ideas for improving weighted averages. In particular, they show that the use of confidence and competence scores as weights improve forecasting accuracy. Yet, the mechanistic 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 aggregation encourages and facilitates aggregation, permitting information sharing (Garthwaite et al. 2005). However, the approach may lead to difficulties in reaching consensus and even if consensus is reached, that consensus may be influenced by power and personality rather than by knowledge. Garthwaite et al. (2005), Kumar et al. (1993) and Sunstein (2005) discuss the influence of censorship, Groupthink, and the difficulty of organizing and facilitating the needed interactions on forecasting accuracy. They also mention the Delphi Method, developed by the RAND Corporation in the 1950's, as an alternative (Schmidt 1997): informants share their forecasts and their reasoning anonymously and the process cycles until a consensus emerges. The approach mitigates the impact of personalities and empirical evidence supports the accuracy of Delphi forecasts (Rowe and Wright 2001). However, the Delphi approach also suffers from logistical problems and eliminates any positive aspects of informants knowing the identity of other informants: one might wish to place greater weight on the opinion of a very experienced manager rather that that of a novice, independent of the quality of the argument of either. The approach is also not conducive to frequent updating as the relevant environment changes.

To sharpen our focus and to align our work with practice, we study survey-based CJF approaches here, noting that the key informant (KI) approach is a special case of a CJF, with 100% weight on the key informant.

2.3 Virtual Stock Markets (VSMs). VSMs create information markets, bringing groups of participants 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 (Spann and Skiera 2003). VSMs elicit *and* aggregate information that may be widely dispersed across a large number of *public* and *private* sources, through the mechanism of trading. Such markets play three important roles (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. Gruca, Berg, and Cipriano (2003), Ostrover (2005) and others discuss the benefits of VSMs in terms of cost, speed, sampling, response bias, dealing with outliers, natural updating and the like.

Ostrover (2005) cites five reasons why VSMs should be expected to work in the type of institutional settings that we address here: *i*) participation: everyone with access to relevant information can contribute; *ii*) motivation: a properly designed reward mechanism incents participants to acquire relevant information and reveal their beliefs (Forsythe, Rietz, and Ross 1999); *iii*) anonymity: VSMs eliminate fear of reprisal for revealing unpopular beliefs ; *iv*) coordination: markets provide a natural mechanism for active group interplay; and *v*) computation: markets provide a natural aggregation mechanism.

In a VSM, informants, acting as traders, reveal and share public and private information through their trading behavior. VSMs thus have the ability to disseminate information from informed traders to uninformed traders (Plott and Sunder 1982; Plott and Sunder 1988). If markets are efficient, all available information is reflected in prices (Fama 1970). Thus, the price of a specific stock in an efficient VSM reflects all information (public and private) on the corresponding state (value) of interest and can thus serve as a forecast. Hence, the use of a VSM as an aggregation and forecasting mechanism hinges on its efficiency.

Grossman and Stiglitz (1980) and Oliven and Rietz (2004) argue that perfect information aggregation is impossible. Others (Spann and Skiera 2003; Sunder 1992) note that if the number of insiders is very small, informational efficiency might not be achieved. In markets where different traders have different information signals, the presence of aggregate information uncertainty significantly reduces market efficiency relative to markets with aggregate information certainty (Lundholm 1991). In institutional forecasting settings, complete and certain information will rarely exist. Therefore, it is an open question whether VSMs will do well for institutional forecasting. Furthermore, over-confident traders may affect prices in the market if they stick to their forecasts irrespective of the (perhaps better informed) opinions of other traders. According to Wolfers and Zitzewitz (2004) prediction markets display some of the deviations from rationality that appear in other financial markets as well. For example, speculative bubbles may drive prices away from likely outcomes.

A threat to the informational efficiency of VSMs thus appears when the number of insiders is small (Spann and Skiera 2003; Sunder 1992), of particular relevance to us here. In these so-called *thin markets* problems like information traps, manipulation and lack of equilibrium are exacerbated (Chen et al. 2003; Chen, Fine, and Huberman 2004). A lack of experience in playing in markets may also threaten the accuracy of VSMs relative to those of KI and CJFs. And VSMs may suffer from information cascades (Anderson and Holt 1997), when individuals overweight private information of other traders that has become public during the trading process and underweight their own private information. These latter problems may even out in "thick", but not in "thin" markets.

Of relevance to us here is the extent to which the price reflects the "true" value of the variable of interest at a particular point in time. We conjecture that the answer to this question depends on the characteristics of the information being exchanged. First, when information is without *uncertainty,* it will be effectively communicated from traders who know what will happen in the future to those who lack such knowledge. Second, when information is *complete* (the extent to which the *sum* of information possessed by the various traders reveals the true state of nature), then market prices will provide accurate predictions of unknown events. Forsythe and Lundholm (1990) and Lundholm (1991) show that when uncertainty is introduced into the information provided to traders or if the set of information as a whole is incomplete, then market prices deviate from their true, underlying value.

Summarizing the discussion above, there are several reasons to expect VSMs to do well in the domain of institutional forecasting but there is also research that suggests that VSMs may not work optimally, especially when the number of informants is relatively small. To our knowledge, with the possible exception of the study of Gruca et al. (2003), which differs significantly in design from ours, there are no studies that empirically compare the forecasts of VSMs with those of CJFs in the environments we focus on here.

3. Virtual Stock Markets (VSMs) Versus Combining Judgmental Forecasts (CJF)

To guide our development of hypotheses about expected differences in forecasting accuracy between VSMs and CJFs (and, by implication, the special case of KI), we compare these approaches along three key information dimensions: *i*) information elicitation; *ii*) information exchange; and *iii*) information aggregation.

Information Elicitation is the process of formulating an individual's forecast. It is important to distinguish between the quality of a informant's forecast and the accuracy with which that forecast is obtained. A forecast is elicited well if the value that is derived accurately represents the informant's knowledge, regardless of how good that knowledge is (Garthwaite et al. 2005).

In the CJF approach, informants provide direct forecasts for the variable of interest as well as self-assessments of their confidence and answers to questions measuring their competence. This approach is straightforward and (relatively) easy for informants. A weakness may be that the judgments in these settings may be based on a limited number of mental operations, which potentially may 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. With VSMs, informants provide information in an indirect manner through their trading behavior. This approach provides anonymity and a seemingly clear incentive for participants to reveal their true beliefs, i.e., participants are rewarded based on their market-performance. However, the market mechanism may also provide incentives for speculation and trading in an information market as well as may not be easily understandable for all informants. Hence, both approaches have advantages and disadvantages on the dimension of information elicitation.

Information Exchange: Informants base their responses in the CJF approach and their trading behavior in VSMs on a combination of public and private information, i.e., information unique to or held by one specific trader (Chen et al. 2003). The CJF approach does not allow for information exchange between informants and, hence, they cannot learn from other informants: the mechanism does not lead to consensus and informants with little or highly biased knowledge are not able to learn from others. Since in VSMs private information is exchanged through the market mechanism, the price system makes information publicly available and thereby transfers it from informed to uninformed traders (Grossman and Stiglitz 1980). A potential problem may be that information that is exchanged may be biased. Summarizing, the VSM approach offers clear advantages to the CJF approach on information exchange.

Information Aggregation is the procedure that combines the opinions of individual informant forecasts to arrive at a single forecast. With CJF, the researcher does the aggregation, employing either a simple average or a more sophisticated weighted average. Indeed, some quite complex weighting procedures have been developed, including non-linear aggregation rules (Chen et al. 2004) and Bayesian procedures. Van Bruggen et al. (2002) show that confidence and competence based weights are quite effective and can outperform simple averages. These weighting approaches make the aggregation rule transparent. Weaknesses of these approaches include the possible double-counting of informants (several informants could have highly correlated knowledge) as well as the logical problem of averaging two (or more) widely divergent views which may lead to an inappropriate "compromise". For example, if informants all believe a new product's sales will be "high" if the industry adopts a certain technical standard and "low" otherwise, but differ in their beliefs about the likelihood of that standard being adopted, any non-trivial weighting rule will lead to a forecast of an intermediate level of sales, a result that no individual informant believes will occur. A VSM simultaneously performs information aggregation, dissemination, and conflict resolution (Plott and Sunder 1988): the market performs the weighting procedure. The 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 a VSM is that trading will be based on the strengths of beliefs of traders and these beliefs are not necessarily (fully) in line with reality (Stracca 2004). Summarizing, both approaches have strengths and weaknesses with respect to aggregating information of multiple informants. VSMs also have the advantage that once they have been set up, it is easy for researchers to observe prices (and thus forecasts) on a continuous basis and also see how these forecasts change as a consequence of (important) events.

Table 1 summarizes the main characteristics of the VSM and the CJF approaches. Information exchange is the main difference between VSMs and CJF. This information exchange is beneficial if an information-asymmetry between informants is present: it permits individual (private) knowledge to become public and allows less knowledgeable informants to update and increase their knowledge. This updating process leads to less knowledge-heterogeneity and higher average knowledge.

--- INSERT TABLE 1 ABOUT HERE ---

We thus posit that information exchange should lead to improved institutional forecasts (i.e., VSMs should outperform CJF) if there is a major difference in knowledge between informants, that is, if different informants possess different types of information or different "pieces of the puzzle." In other words, if the market is efficient, the accuracy in forecasting for a VSM should at least be equal to or exceed the accuracy of the most knowledgeable informant, I, i.e,

$$\textit{Accuracy of Group (VSM)} >= \textit{Max (Accuracy } I_1\textit{, Accuracy } I_{2,\ldots.}\textit{ Accuracy } I_{n\,..}\textit{),}$$

while for CJF,

$$\textit{Accuracy of Group (CJF)} \sim \textit{(Weighted) Mean (Accuracy } I_1\textit{, Accuracy } I_{2,..}\textit{ Accuracy } I_n\textit{)}$$

In fact, following our "different pieces of the puzzle" argument, the VSM should strictly outperform the CJF, since a VSM permits integrating and sharing of information. Hence, if these accuracy suppositions hold, then VSMs should be at least as good as CJF in all cases and should significantly outperform CJF in situations of high information heterogeneity (i.e., where any weighting of the maximal accuracy will be well above the weighted regardless for any weighting scheme).

As the Key Informant (KI) is a special case of CJF, with all weight given to the specific key informant, the discussion above about CJF applies here. However, as the KI approach discards information from the non-key informants (who will still possess some "part of the puzzle"), we expect that KI will perform more poorly than both the VSM and than the CJF method when there is high information heterogeneity. When there is low knowledge-heterogeneity, however, the KI will have internalized that information and, hence, is not likely to do significantly more poorly than either CJF or VSM.

More formally, we hypothesize:

H_1: *If there is low knowledge-heterogeneity between informants, none of the approaches VSM, CJF or KI will differ significantly in terms of forecasting accuracy.*

H_{2a}: *If there is high knowledge-heterogeneity between informants, VSMs will outperform the CJF approach in terms of forecasting accuracy.*

H_{2b}: *If there is high knowledge- heterogeneity, the CJF approach will outperform the KI approach in terms of forecasting accuracy*

4. Method

To test these hypotheses, we sought a forecasting task where we could compare the forecasting accuracy of the VSM and the CJF approach in situations with different knowledge-heterogeneity. We also sought a between-subjects design to control for dependencies between tasks. We wanted to allow for an updating process to assimilate feedback and market information in a natural setting, requiring an intertemporal design. Our institutional estimation framework required groups of a relatively small number of informants to create a "market" (in the VSM environment) and to aggregate in a formal manner for the CJF. In addition, we needed situations where there was the opportunity for both heterogeneous knowledge and homogenous knowledge.

We argue that forecasting environments differ in degree of knowledge across informants and that at least the following four characteristics differentiate between domains of high and low knowledge-heterogeneity: (1) *Presence and Strength of an Anchor Point for the Forecasted Variable:* a strong anchor point affects all informants and leads to more homogeneity; (2) *Amount of Public vs. Private Knowledge:* more public relative to private knowledge should lead to more homogeneity; (3) *Inherent Predictability:* high inherent predictability of the variable under study should lead to more homogeneity. (4) *Environmental Variability:* closely related to (3), low variability in the environment during the forecasting period should lead to more homogeneity. In addition, informants may simply have more inherent knowledge about some items to be forecasted than others.

We had ready access to undergraduate and graduate business students as potential experimental subjects. After screening a number of alternative tasks, we found that a financial market forecasting task for a widely traded commodity fit the low heterogeneity criteria well and forecasting the point spread in a future college football game (absolute difference in score between winning and losing team) matched the high heterogeneity criteria. For example, for the college football point spreads there is a relatively weak anchor point: changes in lineups, field conditions, and the like lead to weak anchors, while the general stability of many financial market indices makes current prices strong anchors. Similarly for public versus private information: most students saw the indices they were trading or forecasting (as well as speculations about them) on a daily basis, while only the more knowledgeable and dedicated football fans used news group and additional information sources on the Internet. A similar argument follows for inherent predictability (with college football scores inherently unpredictable): for example 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 (a difference significant at p=0.000). Arguments for Environmental Uncertainty and Inherent Knowledge (coefficient of variation was 0.302 for football knowledge and 0.198 for financial knowledge) follow similarly.

The specific football point spread task was to predict the score for two games (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.

We formed 21 experimental groups consisting of 6 individual informants, each in one of two conditions:

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

Our 126 subjects were selected after a pre-experimental assessment (Appendix A) 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.

Incentives: Participants were informed in writing that they would receive both a (fixed) participation payment and an additional compensation based on their performance. Performance compensation in the

VSM was linearly related to the value of the participant's portfolio at the end of the study. Their CJF compensation was based on the participant's mean overall financial index price or point spread prediction accuracy. These compensation schemes were designed to provide both significant and similar incentives for all participants to apply effort attention to these tasks.

The VSMs were opened for 22 days prior to 20 November 2004. Participants were also required to provide forecasts in the CJF task four times during those 22 days, via an electronic survey. They also provided confidence scores. We assessed competence via a separate knowledge questionnaire given at the time the study began (Appendix B) where we measured competence as the number of correctly answered questions.

Operation of the VSMs. Each VSM 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 (the cash-out price) for the football point-spread stock types gives $1 (virtual) for every point in the point-spread:

(1) $d_i = Z_i$ i=1,2.

where:

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 use two different payoff functions in order to adjust for the different scale levels of the financial indices. The shares of the stock for the price of crude oil pay $1 (virtual) for every $1 (real) per barrel of crude oil (see equation (2)). The shares of the stocks for the Dow pay $1 (virtual) for every 1,000 points of this index (see equation (3)).

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

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

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 represents 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 VSM and $2,500 (virtual) cash. Based on their performance at the VSM (measured by the value of their final portfolio), participants received a bonus payment. Participants could trade 24 hours per day, seven days a week. There was no trading fee.

Market Maker Trading Mechanism: our VSMs applied a two-sided automated market maker trading mechanism, comparable to the one used at Nasdaq to avoid the thin market problem that can arise due to our experimental setting of only six traders per market (Hanson 2003; Pennock 2004). Our market maker accepted every order from a participant 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 participants 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 (Das 2005). Each trade of a single stock represents a signal to the market maker. Therefore, the quantity 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 VSMs 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. Equation (4) gives our price adjustment function with the parameter values used in our experiment.

$$(4) \qquad p_{j,n} = p_{j,n-1} + \max\left\{ sig_{j,n} \cdot q_{j,n} \cdot \frac{p_{j,\max}^{\;2}}{\gamma^2} \cdot \frac{\left(\sum_{\substack{i=0 \\ sig_{j,n}=sig_{j,i}}}^{I_n} q_{j,n-i} \right)}{I+1}, \alpha \right\}$$

with $sig_{j,n} = \begin{cases} -1 & for\ sale \\ 1 & for\ purchase \end{cases}, I_n = \begin{cases} n & for\ n < I = 10 \\ 10 & for\ n \geq I = 10 \end{cases}$ (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,

$sig_{j,n}$: 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 γ = 500),

α: Minimum tick size (with α = 0.01).

Spann and Skiera (2003) develop prerequisites for a VSM to be efficient; our design, selection, and screening criteria meet those criteria.

In Figure 1a-1d, we provide some illustrative screen shots that depict the environment the participants faced. The market prices in the VSMs during the 22-day trading period represented the forecasts.

--- INSERT FIGURE 1A-1D ABOUT HERE ---

CJF Procedure In order to create forecasts using the survey-based forecast of informants in the CJF 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 with respect to financial knowledge (F=1.207, p=.294) or football knowledge (F=.333, p=.971). 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 informants received the questionnaire at the same day, but were allowed to send it back on one of the following days. Hence, responses varied across the time period 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 use three forecasts: *i*) an unweighted average of the forecasts of the 6 persons in each group; *ii*) a knowledge-based weighted average; and *iii*) a confidence-based weighted average.

We compute the unweighted average forecast for group *w* at day *t* using Equation 5

$$(5) \qquad UnweightedForecast_{ftw} = \frac{\sum_{i=1}^{6} Forecast_{ftwi}}{6}, \text{ where}$$

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

w = 1, .. , p (p = 11 for the football indices and p = 10 for the financial indices)

t = 1, .. , 22

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

We compute the knowledge-based forecast for group *w* at day *t* using Equation 6

$$(6) \qquad KnowledgeBasedForecast_{ftw} = \frac{\sum_{i=1}^{6} Knowledge_{fi} * Forecast_{ftwi}}{\sum_{i=1}^{6} Knowledge_{fi}}$$

where $Knowledge_{fi}$ *is* the knowledge score of individual *i* about variable *f* measured using the test items described in Appendix B.

We compute the confidence-based forecast for group *w* at day *t* using Equation 7

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

where $Confidence_{fi}$ is the knowledge of individual i about variable f.

We also developed a $Key\text{-}InformantForecast_{ftw}$ for each 6-person group. We selected the key informant as the informant with the highest knowledge score within the group according to the knowledge responses to the questions in Appendix B. We broke ties by selecting one key informant in a group at random.

5. Results

We compare the forecasting accuracy of the CJF-based measures and the VSM-based forecasts by computing the Mean Absolute Percentage Error (MAPE) of the forecasts for the four variables Z_1, Z_2, Z_{crude} and Z_{Dow}. The MAPE is invariant to scale and not influenced by outliers and is computed as in Equation 8.

(8) $$MAPE_{fmt} = \left| \frac{Forecastfm_{fmt} - Actual\,Value\,at\,20November_f}{Actual\,Value\,at\,20November_f} \right|$$

Where:

$MAPE_{fmt}$ is the Mean Absolute Percentage Error for variable f for method m, where 1 = unweighted aggregated CJF forecast, 2 = knowledge-based aggregated CJF forecast, 3 = confidence-based aggregated CJF forecast, 4 = key informant, and 5= VSM-based forecast) at day t.

The actual (true) values 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

Table 2 presents the average MAPE values of the key informants, of the various CJF-based forecasts and of the VSM-based forecasts for the financial indices and for the football point spreads across 12 days.

Our VSMs 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 such markets with a market maker setting the initial price or reacting to a jump in the stock's true value, there is usually a transient period of volatility before the market settles to set a (new) price (Das 2005). And as our task involves forecasting, it is appropriate to establish a time interval between the VSM-established valuation and the realization of the actual valuation. Furthermore, during the last five days before the games were being played, bookmakers quotes were widely available and, per our study design, we wanted to avoid having our football VSMs anchored on these quotes. The results are graphically presented in Figure 2 and 3. (We replicated the analysis with the full data set (see Table 2A) and the results were similar to those reported here.)

--- INSERT TABLE 2, TABLE 2A AND FIGURE 2 AND 3 ABOUT HERE ---

Overall our results show that the values of the financial indices were more accurately predicted than the football point spreads ($F=79.998$, $p=.000$). The average MAPE across the various CJF approaches and VSMs is .072 for the financial indices while it is .488 for the football point spreads. Furthermore, we did not find significant differences between various CJF approaches. In contrast with the findings reported in Van Bruggen et al. (2002), we do not find weighted averages to do any better than unweighted averages. A possible explanation for this finding might lie in the screening procedure we applied before allowing participants to our study. Because of this screening, the variation in knowledge and confidence (these two variables were correlated) was reduced, which also reduced the effectiveness of the knowledge-based and confidence-based weights. This finding of a lack of effectiveness of the weighting approaches is in line with other research results (Armstrong 2001).

Since we did not find any differences between the three CJF approaches we only analyze the unweighted CJF results further. These results show an interaction effect between the forecasting approaches (CJF vs. VSM) and the forecasting tasks (Football Point Spreads vs. Financial Indices) ($F=2.350$, $p=.035$). Furthermore, differences between the CJF approaches and the VSM change over time ($F=1.312$, $p=.051$), a result especially strong for the financial markets (see Figure 2).

When predicting financial indices (low heterogeneity) there are no significant differences in the forecasting accuracy of CJF approaches, the KI approach and the VSMs. This finding supports H1. However, closer inspection of Figure 2 shows a difference between the results for the first part of the experimental period and the second part. In the first part there is no substantial difference in forecasting accuracy between CJF, KI and VSM, while the CJF and KI approaches actually do somewhat better than the VSMs in the second part. The reason for the poor performance of the VSMs is not immediately clear. Participants in VSMs may have been hampered by the complexity of the VSMs relative to that of the task in the CJF conditions. The information exchange characteristics of VSMs may not sufficiently compensate for this complexity, possibly because there is not much additional information to be exchanged in the VSMs for the financial indices in the remaining time period.

In contrast with predicting financial indices, when predicting the football point spreads (high heterogeneity between informants) the VSMs clearly outperform the CJF-based approaches, a result that supports H2a. Figure 3 also reveals that the KI forecasts of the football point spreads are clearly less accurate than both the CJF and the VSM results, providing support for H2b.

Given the rather exploratory nature of this research, we performed additional analyses. We investigated how knowledge affected the CJF or VSM results; our analyses did not identify a relationship between either the average knowledge levels, the knowledge of the most knowledgeable informant or knowledge dispersion and the quality of forecast made through either VSMs or through the CJF approach. This lack of a relationship may, at least partly, be due to the fact that the selection and the screening of our participants were designed to reduce variance in background and knowledge.

One finding of interest surrounds the link between informant knowledge (measured directly) and forecasting error. We found that the knowledge of our key informants was significantly related to the error of their forecast of the financial indices ($r=-.392$, $p=.088$), with the more knowledgeable informants being the more accurate in low heterogeneity environments. We did not find any other links between informant knowledge and forecasting accuracy.

Given the thin nature of our markets, we investigated the relationship 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 significant. And in neither market type did we find a relationship between level of trading activity and forecasting accuracy, suggesting that there seems to be sufficient trading activity even in these thin markets to reach a level of efficiency. We did, however, find a relationship between level of trading activity at the individual level and individual performance, suggesting that trading activity itself might be an indicator of market knowledge ($r=.875$, $p=.000$ for VSMs for financial indices and $r=.362$, $p=.005$ for VSMs for football point spreads).

6. Discussion

We have investigated the feasibility and accuracy of Virtual Stock Markets (VSMs) in institutional forecasting situations, characterized by a small number of knowledgeable informants. We have compared that performance with that of the Combined Judgment (CJF) and Key Informant (KI) approach.

Our results show that VSM are feasible in environments with varying degrees of predictability and knowledge heterogeneity, and can be conducted effectively with group sizes as small as six traders per market. We find that VSMs outperform CJF and KI in more difficult-to-predict environments, characterized by high knowledge and information heterogeneity between informants. The approach works as well as the more traditional CJF and KI approaches in more homogeneous environments. We attribute this superior performance of VSMs in high knowledge-heterogeneity environments to their ability to provide for information exchange between participants: trader's private information becomes public through their trades, thus enabling information dissemination and learning.

We also find that forecasts of both VSMs and CJFs are more accurate in situations of high knowledge-heterogeneity than those of KI. A possible explanation for this finding is that improved accuracy in a high knowledge-heterogeneity environment requires an integration or combining mechanism of some sort; such a mechanism is provided by a VSM or by CJF, but not by a KI. In this particular study, in contrast to past findings such as those by Van Bruggen, Lilien and Kacker (2002) we find no significant performance differences between the different weighting schemes we investigated (equal, confidence or knowledge based) for CJFs. This contrast with past results may be due to our pre-screening of participants, limiting knowledge-heterogeneity and eliminating all low-knowledge participants.

The nature of these results raises interesting questions concerning both why and when VSMs should be expected to perform well in the field. If information exchange and learning change participant's knowledge, than VSMs have two inter-temporal advantages over CJF: First, all 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 CJF, typically taken at a point in time, are not updated, those weights may become seriously suboptimal if the environment is unstable or highly unpredictable. A VSM, operating continuously, has the ability to aggregate such environmental and knowledge-based changes naturally and continuously and may also be effective in deriving informant weights for future analyses (see (Chen et al. 2004)). In more stable, low knowledge-heterogeneity environments, those problems do not exist and, hence, VSM exhibits no advantage over KI or CJF.

Our results exhibit some interesting if speculative managerial implications. First, improved institutional forecasts can be achieved with VSM over the other two approaches, with that improvement being significant in environments with high knowledge-heterogeneity. VSMs are especially attractive in institutional settings where the variable to be forecasted is difficult to predict and where the type and quality of information and knowledge vary between the informants. Judgement-based sales forecasts for new products can often be characterized in such a manner. Also, VSMs do no worse than CJF or KI in low knowledge-heterogeneity environments, making it a relatively robust choice for such forecasting tasks. Secondly, since VSMs are superior to CJFs in the high knowledge-heterogeneity environment and KI is no worse than CJF in the low knowledge-heterogeneity environment, it may be that the CJF approach should be a third choice option in most institutional forecasting environments. Our results also provide critical justification for the widely-used key informant approach, but, going beyond the work of Phillips (1981) and others, specifying conditions when that approach should be expected to perform poorly.

From a theoretical standpoint, our results provide a possible explanation for when and why a VSM might be expected to work in an institutional forecasting framework. However, as with much research of a relatively early nature, our results also provide opportunities for future study. First, our procedure screened our low and medium knowledge individuals. A key question arises here about the robustness of any of these approaches to inclusion of such individuals. It may be that thin VSMs are subject to the winner's curse, where several low knowledge individuals drive the VSM to inefficient outcomes. And we have (arbitrarily) set the size of the VSMs to six individuals in our experiment. We have no data on how the performance of VSMs varies with group size. In addition, given our arguments about knowledge sharing in high-knowledge-heterogeneity environments as the explanation for VSM performance, it would be useful to directly compare other repeated information-sharing mechanisms, perhaps of the repeated-Delphi type. Such systems are becoming increasingly available and web-based implementations make such collaborations increasingly easy (see (Cil, Alpturk, and Yazgan 2005), for example).

While there are clearly many other research opportunities here, we reemphasize our main findings: VSMs are feasible for the institutional forecasting problems cited here and they appear to provide forecasting accuracy that is superior to that of methods that are in common use. As our conclusions are based on results from a limited number of variables and markets, future research should address the generalizability of these findings. Furthermore, it will be a challenge for both academics and practitioners to further refine both the benefits and the costs of the VSM approach.

Appendix A

Screener Instrument

<u>Screening Questionnaire:</u>

The purpose of this questionnaire is to determine whether or not you are a good match for our study. Please answer the questions below by circling "true" or "false" to indicate whether you feel the statement is true or false.

General Football Questions

1. Penn State plays in the Big East Conference. True False
2. A team must go 10 yards or more to get a first down. True False
3. A team gets 7 points for a field goal. True False
4. The quarterback gets the ball from the center. True False
5. A football game lasts 60 minutes on the official clock. True False
6. Each team can have up to 15 men on the field. True False
7. A "sack" occurs when the quarterback is tackled behind the line of scrimmage when he is trying to pass. True False
8. A team gets 3 points for a safety. True False
9. Suppose you participated in an office football pool, and you received Penn State + 3 points against Iowa. If Iowa wins by a score of 24-22, you lose your bet. True False
10. Suppose in your office pool, you received Penn State + 3 points against Iowa. (Same as above.) If Penn State wins by 21-20, you win your bet. True False

General Trading Questions

1. It is better to buy high and sell low than buy low and sell high. True False
2. When you buy common stock, you own a percentage of a company. True False
3. Dividends are paid out when a stock splits. True False
4. Dow Jones is a well-respected radio financial commentator. True False
5. The Fortune 500 is the 500 best stock picks of the day. True False
6. You should buy a stock when its price is lower than you think it should be. True False
7. Nasdaq is the name of a stock exchange. True False
8. A barrel of crude oil sells for about $1.80. True False
9. Diversifying your portfolio is a way to increase your level of risk. True False
10. If a stock currently sells at $10, and you believe it will decrease to a $5 selling price by the end of the year, you should sell the stock. True False

Appendix B

Instrument to Measure Knowledge of Football and of Financial Markets

To assess your knowledge of college football, which of these rules differ between college and NFL football:

1. Offsides penalty rule ____ Same ____ Different
2. Pass interference penalty rule ____ Same ____ Different
3. Feet in bounds for a reception rule ____ Same ____ Different
4. Two point conversion rule ____ Same ____ Different
5. Number of time outs in a half ____ Same ____ Different
6. Overtime rules ____ Same ____ Different
7. First down clock rule ____ Same ____ Different
8. Position of hash marks ____ Same ____ Different
9. Field size ____ Same ____ Different
10. Time between downs ____ Same ____ Different

To assess your knowledge of financial markets, please answer the following questions:

1. The New York Stock Exchange is a place where exchange trading takes place. True False
2. Shares of common stock always pay dividends. True False
3. Preferred stock ownership usually ensures voting rights in a company. True False
4. If many people want to sell a stock, and few want to buy, then the price of the stock rises. True False
5. A market order means you can buy, for example, 100 shares of a certain company immediately, regardless of price. True False
6. A stock's P/E ratio is the value of the company divided by the annual earnings per share. True False
7. An investor diversifies their portfolio to increase volatility and risk. True False
8. A prospectus is a document that discloses information about financial consulting/advising firms. True False
9. A portfolio is all securities, real estate, or other investment tools held by a stockholder. True False
10. A treasury bond typically has a time to maturity of more than 1 year after the issue date. True False

Table 1
Comparison of Main Information Characteristics of the Virtual Stock Market (VSM) and the Combining Judgmental Forecasts (CJF) Approach to Forecasting

	Combined Judgmental Forecasts	Virtual Stock Markets
Information Elicitation • About forecasted variable	• Direct and explicit measurement of variables of interest • Information available at the level of the individual informant. • Incentives for providing biased information may be present	• Indirect measurement of variables of interest through observation of trading behavior • Information provided not directly visible • Incentives for speculation may be present,
• About knowledge and confidence	• Direct (self)assessments of knowledge and/or confidence • Straightforward and relatively easy for informant	• No direct assessment of knowledge and or confidence • Market participation relatively complex for informants
Information Exchange	• Only public information shared	• Participants exchange private information, which becomes public through trading
Information Aggregation	• Aggregation explicitly done by the researcher • Various alternative (mostly proportional) weighting schemes can be applied • Knowledge-based weighting possible	• 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 Elicitation while Information Exchange and Information Aggregation considerations do not apply

Table 2
Forecasting Accuracy of the Key Informant (KI), Combined Judgmental Forecast (CJF) and Virtual Stock Markets
MAPE (Standard Deviation) averaged across 12 Days

		Key Informant	Combined Judgmental Forecast			Virtual Stock 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)

Table 2A
Forecasting Accuracy of the Key Informant (KI), Combined Judgmental Forecast (CJF) and Virtual Stock Markets
MAPE (Standard Deviation) averaged across full 22 Days

		Key Informant	Combined Judgmental Forecast			Virtual Stock Markets
			Unweighted	Knowledge-Based Weighted	Confidence-Based Weighted	
Football Point Spreads	OSU	.333 (.121)	.316 (.106)	.312 (.105)	.318 (.095)	.232 (.054)
	FSU	.920 (.978)	.602 (.274)	.642 (.301)	.618 (.266)	.319 (.175)
	Mean	.626 (.750)	.459 (.250)	.477 (.278)	.468 (.248)	.275 (.134)
Financial Indices	Oil Price	.095 (.033)	.108 (.076)	.105 (.067)	.092 (.042)	.113 (.013)
	Dow Jones	.075 (.093)	.037 (.019)	.038 (.021)	.030 (.015)	.079 (.036)
	Mean	.085 (.069)	.072 (.065)	.071 (.059)	.061 (.044)	.096 (.032)
Mean		.369 (.604)	.275 (.269)	.284 (.288)	.274 (.273)	.181 (.131)

Figure 1a
Introductory Screen of the Virtual Stock Market (VSM)

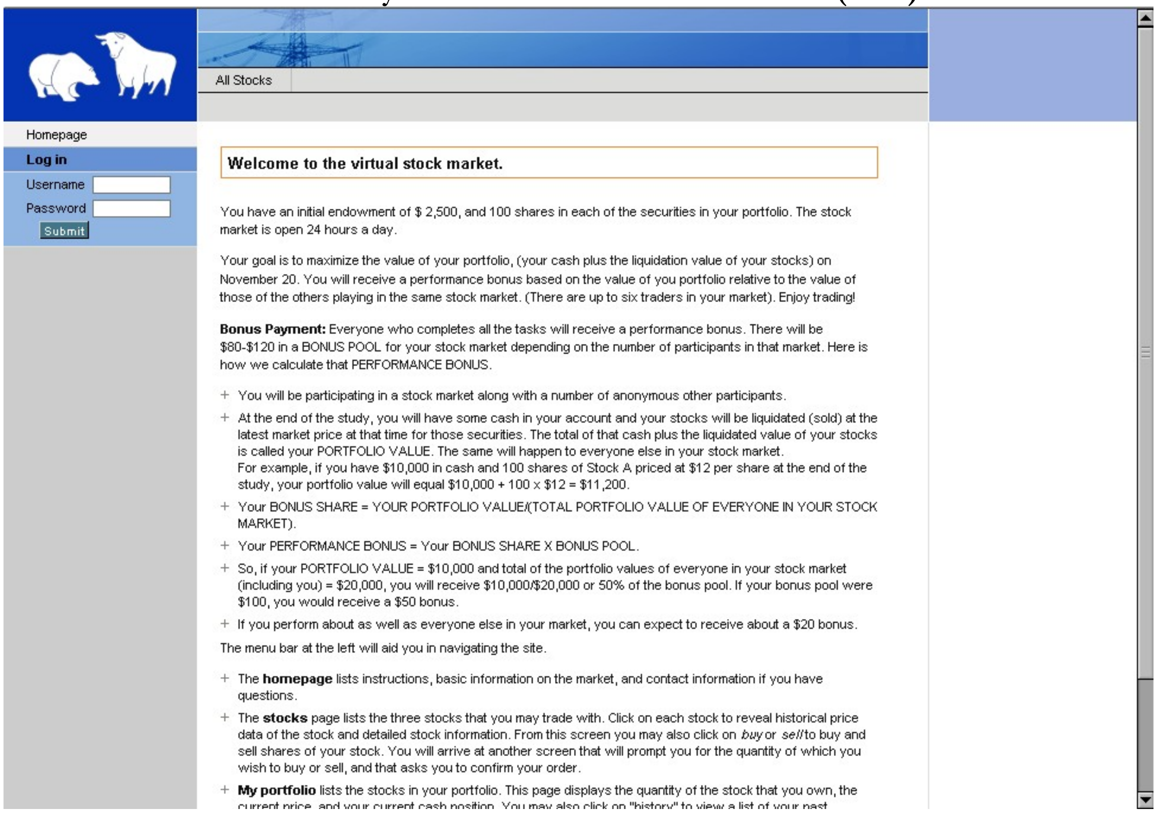


Figure 1b
Explanatory Screen of Virtual Stock Market (VSM) for Oil Prices

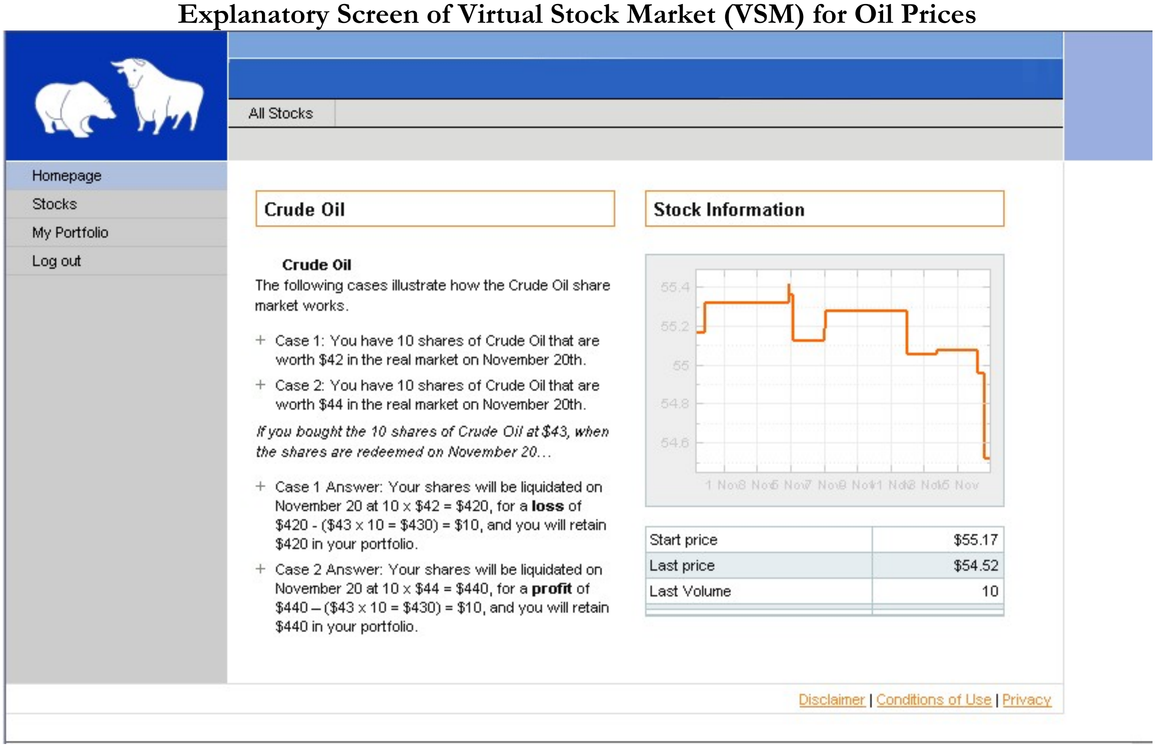


Figure 1c
Explanatory Screen of Virtual Stock Market (VSM) for Football Point Spread Shares

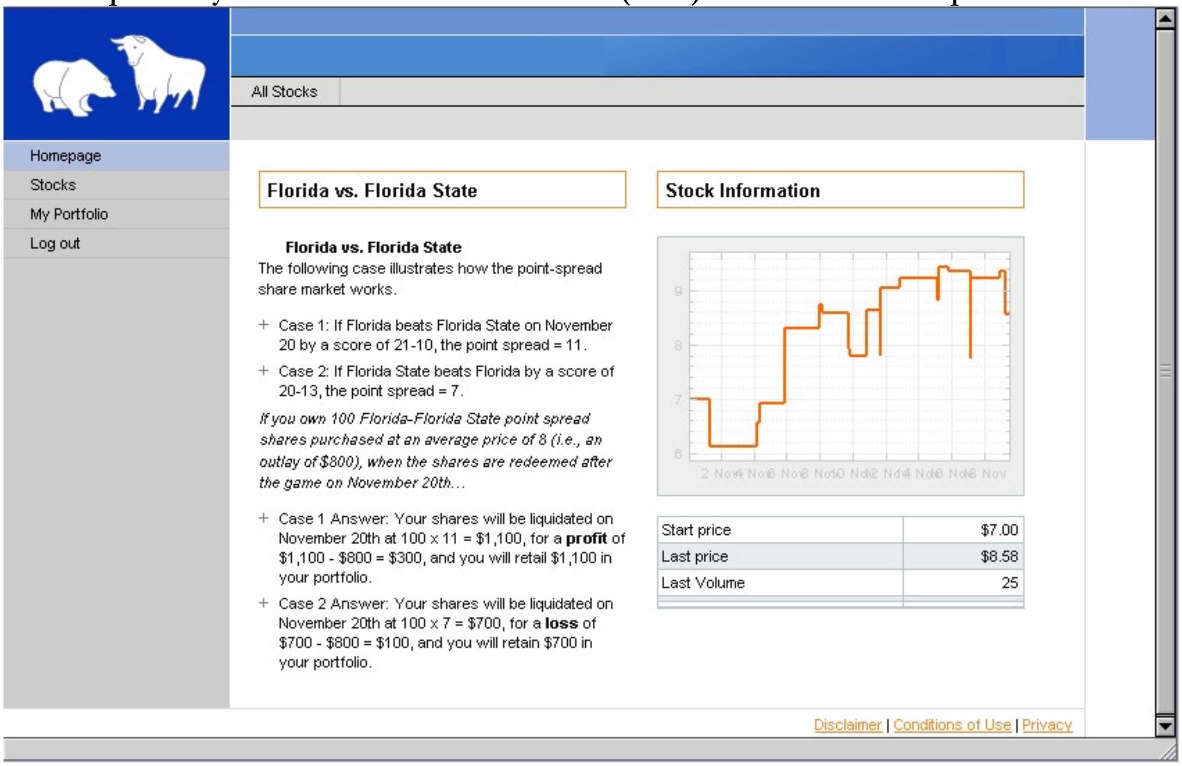


Figure 1d
Portfolio Screen of Virtual Stock Market (VSM) for Financial Shares

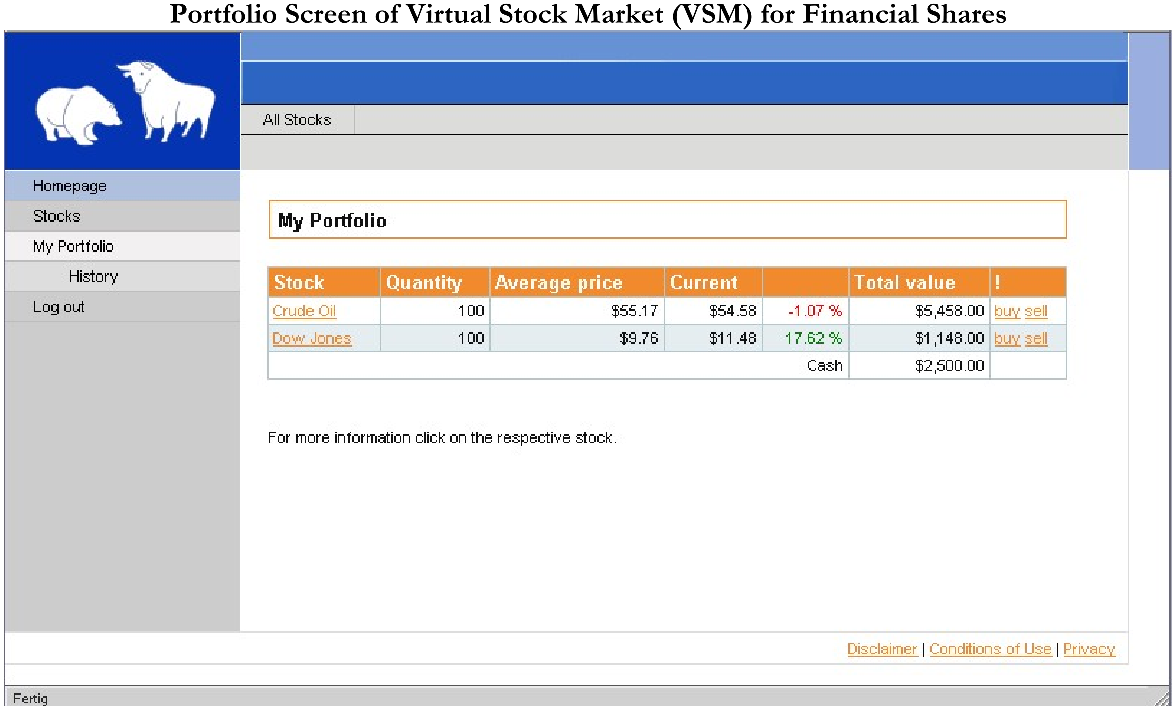
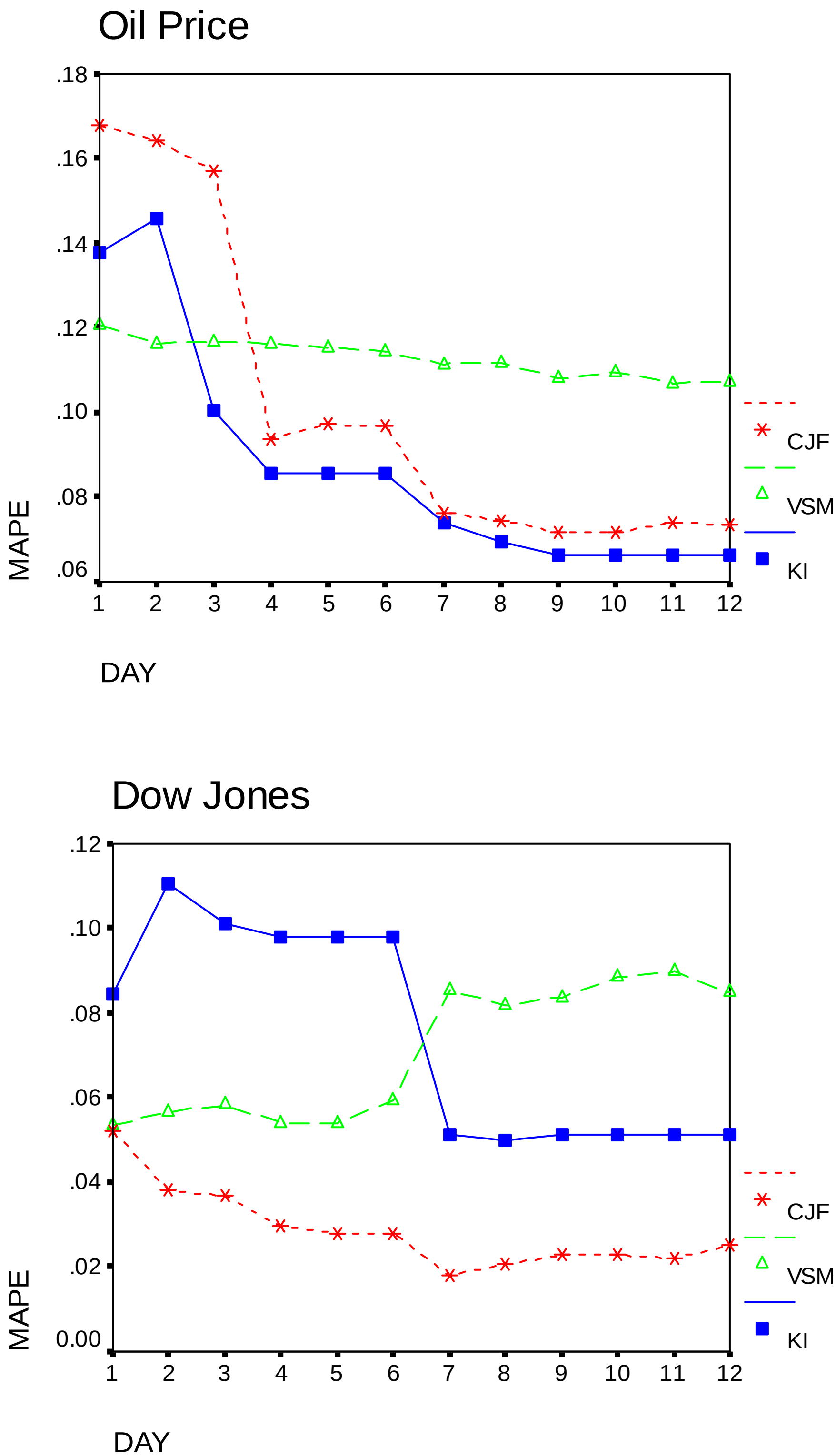
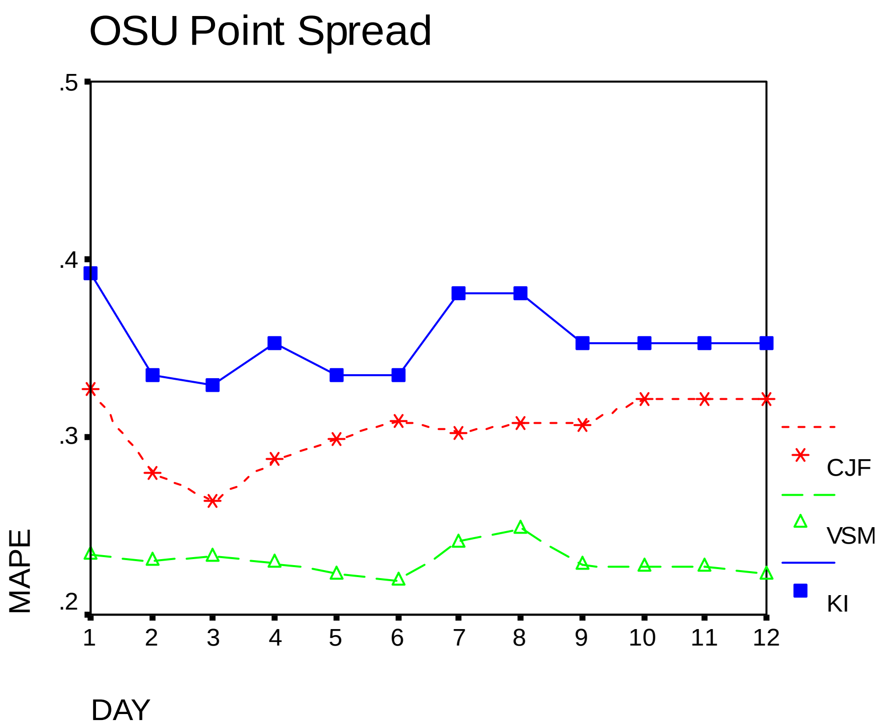


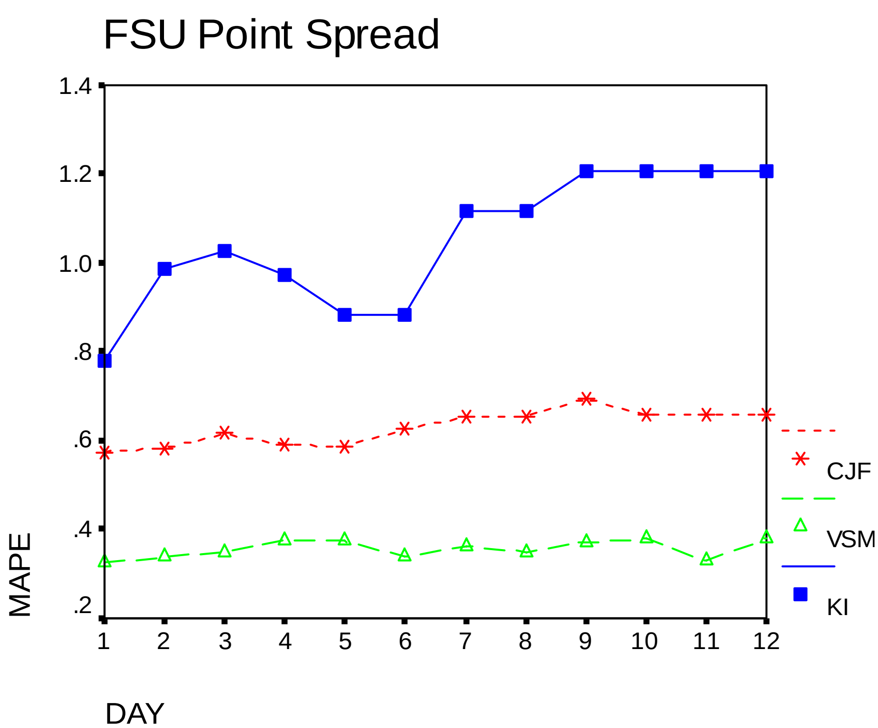
Figure 2: Mean Absolute Percentage Errors for the Financial Indices Forecasts



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

Figure 3: Mean Absolute Percentage Errors 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

References

Anderson, Lisa R and Charles A Holt (1997), "Information Cascades in the Laboratory," *American Economic Review*, 87 (5), 847-862.

Armstrong, J. Scott (2001), *Principles of Forecasting: A Handbook for Researchers and Practioners*. Norwell MA: Kluwer Academic Publishers.

Chen, Kay-Yut, Leslie R Fine, and Bernardo A Huberman (2004), "Eliminating Public Knowledge Biases in Information-Aggregation Mechanisms," *Management Science*, 50 (7), 983-994.

Chen, Kay-Yut, Leslie R. Fine, and Bernardo A. Huberman (2003), "Predicting the Future," *Information Systems Frontiers*, 5 (1), 47-61.

Cil, Ibrahim, Oguzhan Alpturk, and Harun R Yazgan (2005), "A New Collaborative System Framework Based on a Multiple Perspective Approach: InteliTeam," *Decision Support Systems*, 39 (4), 619-641.

Das, Sanmay (2005), "A Learning Market-Maker in the Glosten-Milgrom Model," *Quantitative Finance*, 5 (2), 169-180.

Evans, Michael (2002), *Practical Business Forecasting*. Malden MA: Blackwell Publishing.

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

Forsythe, Robert and Russell Lundholm (1990), "Information Aggregation in an Experimental Market," *Econometrica*, 58 (2), 309-347.

Forsythe, Robert, Thomas A. Rietz, and Thomas W. Ross (1999), "Expectations and Actions: A Survey on Price Formation in Election Stock Markets," *Journal of Economic Behavior & Organization*, 39, 1142-1161.

Garthwaite, Paul H, Joseph B Kadane, and Anthony O'Hagan (2005), "Statistical Methods for Eliciting Probability Distributions," *Journal of the American Statistical Association*, 100 (470), 680-700.

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

Gruca, Thomas S., Joyce Berg, and Michael Cipriano (2003), "The Effect of Electronic Markets on Forecasts of New Product Success," *Information Systems Frontiers*, 5 (1), 95-105.

Hanson, Robin (2003), "Combinatorial Information Market Design," *Information Systems Frontiers*, 5 (1), 107-119.

Kumar, Nirmalya, Louis W Stern, and James C Anderson (1993), "Conducting Interorganizational Research Using Key Informants," *Academy of Management Journal*, 36 (6), 1633-1651.

Lundholm, Russell J. (1991), "What Affects the Efficiency of a Market? Some Answers from the Laboratory," *Accounting Review*, 66 (3), 486-515.

Oliven, Kenneth and Thomas A Rietz (2004), "Suckers Are Born but Markets Are Made: Individual Rationality, Arbitrage, and Market Efficiency on an Electronic Futures Market," *Management Science*, 50 (3), 336-351.

Ostrover, Steven (2005), "Employing Information Markets to Achieve Truly Colloborative Sales Forecasting," *Journal of Business Forecasting*, 24 (1), 9-12.

Pennock, David M. (2004), "A Dynamic Pari-Mutuel Market for Hedging, Wagering, and Information Aggregation." New York, NY, USA: ACM Press.

Phillips, Lynn W. (1981), "Assessing Measurement Error in Key Informant Reports: A Methodological Note on Organizational Analysis in Marketing," *JMR, Journal of Marketing Research*, 18 (4), 395-415.

Plott, Charles R and Shyam Sunder (1982), "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational Expectation Models," *Journal of Political Economy*, 90, 663-698.

---- (1988), "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085-1118.

Rowe, Gene and George Wright (2001), "Expert Opinions in Forecasting: The Role of the Delphi Technique," in Principles of Forecasting: A Handbook for Researchers and Practioners, J. Scott Armstrong, Ed. Norwell, MA: Kluwer Academic Publishers.

Schmidt, Roy C (1997), "Managing Delphi Surveys using Nonparametric Statistical Techniques," *Decision Sciences*, 28 (3), 763-774.

Spann, Martin and Bernd Skiera (2003), "Internet-Based Virtual Stock Markets for Business Forecasting," *Management Science*, 49 (10), 1310-1326.

Stracca, Livio (2004), "Behavioral Finance and Asset Prices: Where do We Stand?," *Journal of Economic Psychology*, 25 (3), 373-406.

Sunder, Shyam (1992), "Market for Information: Experimental Evidence," *Econometrica*, 60 (3), 667-695.

Sunstein, C. R. (2005), "Group Judgments: Statistical Means, Deliberation, and Information Markets," *New York University Law Review*, 80 (3), 962-1049.

Van Bruggen, Gerrit H, Gary L Lilien, and Manish Kacker (2002), "Informants in Organizational Marketing Research: Why use Multiple Informants and How to Aggregate Responses," *Journal of Marketing Research*, 39 (4), 469-478.

Wolfers, J. and E. Zitzewitz (2004), "Prediction Markets," *Journal of Economic Perspectives*, 18 (2), 107-126.

Publications in the Report Series Research* in Management

ERIM Research Program: "Marketing"

2006

The Effect of Feedback and Learning on DSS Evaluations
Ujwal Kayande, Arnaud de Bruyn, Gary Lilien, Arvind Rangaswamy and Gerrit Van Bruggen
ERS-2006-001-MKT
http://hdl.handle.net/1765/7302

Interaction Between Shelf Layout and Marketing Effectiveness and Its Impact On Optimizing Shelf Arrangements
Erjen van Nierop, Dennis Fok and Philip Hans Franses
ERS-2006-013-MKT
http://hdl.handle.net/1765/7632

Institutional Forecasting: The Performance of Thin Virtual Stock Markets
Gerrit H. Van Bruggen, Martin Spann, Gary L. Lilien and Bernd Skiera
ERS-2006-028-MKT

Irritation Due to Direct Mailings from Charities
Merel van Diepen, Bas Donkers and Philip Hans Franses
ERS-2006-029-MKT
http://hdl.handle.net/1765/7832

---

* A complete overview of the ERIM Report Series Research in Management:
https://ep.eur.nl/handle/1765/1

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship