Abstract

Businesses are exploring the use of prediction markets to assist with forecasting. In this study, we focus on the ability of prediction markets to reflect the consensus of trader forecasts as well as the dispersion of these forecasts. Using a real-money, computerized, anonymous double-auction market mechanism, we examine a series of markets forecasting a real-life outcome, i.e., movie box office performance. This continuous outcome was segmented into a small number of mutually exclusive ranges, each associated with a winner-take-all contract. This payoff structure allowed us to determine whether the contract prices reflect the dispersion of the traders’ individual forecasts, which were submitted before trading began. We find that these markets do an excellent job revealing the consensus forecast. The market prices for all contracts are consistent with the entire distribution of trader forecasts for most of the markets. In addition, we find that markets with a wider pool of traders tend to result in superior forecasts.

Keywords: electronic markets, forecasting, information aggregation, double auctions

Authors

Thomas S. Gruca (thomas-gruca@uiowa.edu) is Lloyd J. and Thelma W. Palmer Research Fellow in Marketing and Associate Professor of Marketing in the Tippie College of Business, University of Iowa. His research interests include defensive marketing strategy, health care, pricing and electronic markets.

Joyce E. Berg (joyce-berg@uiowa.edu) is Associate Professor of Accounting in the Tippie College of Business, University of Iowa. Her research on electronic markets has received funding from the NSF, FIPSE and the EPA. She was the co-winner of the 2000 UI President’s Award for Technology Innovation for her work with the Iowa Electronic Market.

Michael Cipriano (cipriano@moore.sc.edu) is a doctoral student in the accounting programme of the University of South Carolina. He received his MBA from the Tippie School of Management, University of Iowa. His research interests include archival and experimental research on capital markets.

INTRODUCTION

Companies ranging from Microsoft and Hewlett-Packard to drug giant Eli Lilly are considering using prediction markets to improve business decision-making (Kiviat 2004, Pethokoukis 2004). Prediction markets are intended to gather and summarize information that may be widely dispersed across a large number of public and private sources through the mechanism of trading. In prediction markets, traders buy and sell securities whose outcome is tied to some future event such as the successful development of a new drug or the quarterly sales of a certain model of computer printer. According to the theory of rational expectations, the price observed in a prediction market is a sufficient statistic for all information available to traders (Lucas 1972; Grossman 1981). By harnessing the ability of markets to aggregate information from say, the far-flung employees of a large organization, managers should become better informed about the likelihood of future events.

In addition, much less is known about their functioning in detail. For example, how do prediction markets capture the private information held by traders? Do the prices in the market merely represent the consensus of the traders’ forecasts or do they also reflect the dispersion of forecasts across traders? This is an important issue to decision makers interested in how opinions may differ with respect to future events (Norris 2003).

A second obvious question surrounds the composition of the pool of traders in the market. In their book Making Markets, Kambil and Van Heck (2002: 155) suggest that one key to successful applications of electronic prediction markets is to ensure that ‘traders represent the peer group that is knowledgeable about the issues the market seeks to address’. Contrasting this view are
those who believe that the larger the pool of traders, the more likely the market is to make use of revealing private information that may be in the hands of few participants.

In this paper, we present the results from a series of prediction markets that address these important research questions. First, we examine how well a computerized, double auction prediction market can aggregate the consensus of traders’ forecasts concerning a future event in the real world. Next, we assess how well prices in the market capture the dispersion of the traders’ forecasts. We then examine how the composition of the traders in the market impacts the accuracy of the market-based predictions. We expect our results to better our understanding of the conditions under which electronic prediction markets might be used to forecast future events for many types of organizations.

BRIEF RESEARCH OVERVIEW

Information aggregation

In an electronic prediction market, traders buy and sell assets whose ultimate value is tied to some future outcome. The motivation for prediction markets comes from rational expectations theory whose central tenet is: ‘price summarizes and reveals all the relevant information in possession of all traders’ (Sunder 1995). Consequently, through the mechanism of price, all traders (and all observers of a market) theoretically share all available information (Plott 2000).

In a series of lab experiments using very simple information sets (e.g., event A does or does not occur), Plott and Sunder (1982, 1988) show that markets have the ability to disseminate information from informed traders to uninformed traders. In addition, markets can aggregate information so that the market price accurately reflects information made available to traders. These seminal papers show that markets are capable of the information processing tasks required by rational expectations theory to ensure that prices reflect all information available to traders.

In a lab setting, the experimenter controls the information that traders see. Traders then use that information to buy and sell assets whose value is tied to an outcome that will be revealed to the traders in the future. The outcome of the market (‘the true state of nature’) is also determined by the experimenter. Plott and Sunder (1982, 1988) demonstrate that the characteristics of the information provided to traders can influence whether the market prices accurately reflect the eventual outcome being predicted by the market.

When the information is perfect (without uncertainty), information is effectively communicated from traders who know what will happen in the future to those who lack such knowledge (Plott and Sunder 1982). In situations where the information provided to traders is complete (the sum of information provided to traders as a whole reveals the true state of nature), then the market prices provide accurate predictions of the unknown event. Thus, markets do a great job in aggregating perfect, complete information. However, if we introduce uncertainty into the information provided to traders or the set of information provided to traders as a whole is incomplete, then the prices can deviate from their expected values (Sunder 1995). In other words, the market prices may lose their power to predict events accurately.

Private information

Kambil and Van Heck (2002: 155) identify a number of factors that potentially influence the accuracy of electronic prediction markets including the number of traders, incentives, etc. However, what makes electronic prediction markets valuable to organizations is their ability to aggregate the private information held by traders (Plott 2000). When traders bring private information to a market, this information will be reflected in the market’s prices and may be observed by others, including those not participating in the market. In laboratory studies, the experimenter controls the traders’ private information. In prediction markets associated with real events, traders bring their own private information to the market. Unfortunately, there is very little research about how well markets aggregate private trader information in these more realistic circumstances.

Traders in the IEM election markets rely on some combination of public information (e.g., political polls) as well as their own private information for their trading. While prior studies have examined the effect of public information on the forecasts made by IEM traders (Forsythe et al. 1992), there have been no studies of the private information that these traders bring to the market. Consequently, we do not know how well these markets are aggregating the private information held by traders.

For the electronic prediction markets described in this study, we collected forecasts from traders before the markets opened for trading. The timing of this data collection is critical since traders learn from each other through the prices that are set and transactions that take place. Thus, we have a measure of the private information that traders bring to the market. Understanding how well markets aggregate traders’ private information is important especially if electronic prediction markets are to replace current approaches to gathering and processing information within organizations.

Information being aggregated

The information set for the prediction markets in this study is much more complex than in the traditional
economics experiment. Traders in a lab setting are usually trying to determine whether one of a small set (e.g., hypothetical event A, B or C) of mutually exclusive events has occurred (as chosen by the experimenter). The outcome being predicted in our markets is continuous and has a very wide range. The experimenters do not control the information used by traders to determine their forecasts. These sources of information are of varying usefulness and verifiability.

The prediction task for traders is much more difficult than the IEM election markets where popular vote shares are strictly bounded between zero and 100 and, in practice, vary only a few percentage points above and below 50% for major candidates.

We split the entire outcome space into a small set of mutually exclusive ranges. Each of these ranges is associated with a winner-take-all contract. If the future outcome falls within the range of a given contract, that contract pays off at the end of the market and all of the others expire worthless. This makes the prediction task for the traders much more difficult than the IEM election markets that predict popular vote share since small changes in the final outcome can have a major impact on which contract pays off.

There is a compelling reason for splitting a continuous outcome space into discrete ranges. The private information held by traders has two very important characteristics. The first is the consensus of the forecasts. This may be represented by some measure of the average of the forecasts of the individual traders, e.g. mean or median. The second important characteristic is the dispersion of trader forecasts around the overall consensus. This measure represents differences in opinions across traders in the market.

Composition of trader pool

Surprisingly, there is very little research on the impact of the scope of the trading population on the accuracy of electronic prediction markets. Traders in the first IEM political market (in 1988) were restricted to students, faculty and staff at the University of Iowa (Forsythe et al. 1992). Four years later, the pool of potential traders was expanded to include traders from around the country (Oliven and Reitz 2004). Since there are very few prediction markets in which the traders were drawn from a restricted pool of participants, it is difficult to ascertain the impact on such a limitation on the accuracy of the markets’ predictions.

Overall, we should expect that markets drawing from a restricted pool of potential traders should result in less accurate predictions than those open to as wide a range of traders as possible. When traders are risking their own money in an electronic prediction market, self-interest leads to self-selection. In other words, only those traders who believe that they have superior information will join the market. Otherwise, a trader would be better off not trading at all.

In the next section, we analyse the results of a series of prediction markets to address the questions of: 1) the ability of an electronic prediction market to capture the consensus of the traders’ private information; 2) its ability to reflect variations in the traders’ private information; and 3) an assessment of the impact of restrictions on trader participation on prediction accuracy.

STUDY OVERVIEW

Market design

For this study, we utilized the IEM, the same market mechanism previously used effectively for predictions of the outcomes of political events (Forsythe et al. 1992). Other IEM markets involve prediction of stock prices (for Microsoft), stock returns (for a set of computer industry stocks) and Federal Reserve Board decisions (FOMC changes in interest rates). For more information, see www.biz.uiowa.edu/iem. The subject of the markets discussed here is the prediction of the domestic box office receipts of movies in wide release in theaters. The design and accuracy of an Internet-based game (i.e., no real money payoffs) with similar goals called ‘The Hollywood Stock Exchange’ is discussed in Pennock et al. (2001).

The IEM is a small-scale, real money futures market. Traders can access the market 24-hours a day through the Internet. Trader investments in the IEM are limited to $500 and there is no short selling allowed. Since the IEM charges no transaction fees, this is a zero-sum market in which all investments by traders are returned to the traders collectively.

IEM futures contracts, whose value is tied to future events, are exchanged using a computerized, anonymous double-auction mechanism. To buy a contract, a trader can execute a market order and buy at the current best price available (lowest ask from another trader) in the market. Alternatively, the trader can submit a limit order. This would include an offer to buy (bid) at another price (usually higher than currently available in the market) and a time limit on the offer. [An analogous process can be followed to sell or offer to sell contracts.] The limit orders (bids/asks) are prioritized by price and submission times. The best prices in each queue are displayed to traders. All trading of individual contracts and the resulting prices in the market are determined by activity between individual traders.

Traders can also acquire contracts from the market in a bundle consisting of one of each of the contracts in the market. These can be purchased from or sold to the IEM exchange at any time for $1, the guaranteed liquidation value (payoff at the end of the market) of the bundle.
This crucial feature of the market allows contract supply to expand and shrink as traders desire without contaminating the individual contract prices as set by the traders. This feature (along with strict limits on the size of traders’ accounts) reduces the incidence of price manipulations.

Contract design

In the IEM election markets focusing on US contests, traders try to predict the share of the popular vote that a particular candidate will garner on election day (vote-share market) or they try to predict which candidate will end up with the larger share of the popular vote (winner-take-all market). In the IEM movie box office markets, traders are faced with a combination of these tasks.

The outcome we attempted to predict was the US box office performance of a particular movie for the first four weeks after its release in theaters (>650 screens). This figure is continuous and nearly unbounded from above making the trader’s forecasting problem more difficult than that of forecasting popular vote share which is bounded between zero and one. Because we want to determine how well the market is aggregating the distribution of traders’ private information, the range of box office receipts is segmented into a small set (usually four to six) of mutually exclusive and collectively exhaustive ranges using historical data from similar movies.

For example, in the fall of 2000, there were four winner-take-all contracts associated with the movie *The Sixth Day*. They were defined as follows:

- **SIX50L** Pays $1.00 if *The Sixth Day*’s box office receipts for the 17 November 2000–14 December 2000 period are lower than or equal to $50 million; zero otherwise.
- **SIX70L** Pays $1.00 if *The Sixth Day*’s box office receipts for the 17 November 2000–14 December 2000 period are higher than $50 million and lower than or equal to $70 million; zero otherwise.
- **SIX90L** Pays $1.00 if *The Sixth Day*’s box office receipts for the 17 November 2000–14 December 2000 period are higher than $70 million and lower than or equal to $90 million; zero otherwise.
- **SIX90H** Pays $1.00 if *The Sixth Day*’s box office receipts for the 17 November 2000–14 December 2000 period are higher than $90 million; zero otherwise.

If the movie failed to open in the prescribed time period, the lowest denominated contract would pay $1.

The prices of contacts defined as above can be interpreted as estimates of the probability of the outcome associated with that contract (Plott and Sunder 1988). For example, if the price of the SIX70L contract is $0.80, then traders, as a group, expect that there is an 80% chance that the four-week box office of the movie will be greater than $50 million and less than $70 million.

Description of traders

Traders in these markets included graduate business students as well as other traders with an academic affiliation (student, staff or faculty). As part of a marketing course, students completed a class assignment that included a point estimate of the movie’s box office receipts in the first four weeks of release in theaters. In addition, the students supplied detailed justifications (four to five pages) of their individual forecasts including the data sources they used. These traders did not know the contract definitions until after they had turned in their forecasts. (A sample assignment may be found in Gruca (2000).)

In exchange, student traders were provided with a $5 or $10 trading account that they could redeem for cash after the market liquidated (i.e., $1 is exchanged for each winning contract) provided they executed at least two trades. Students were free to add their own funds to their trading accounts. We have forecasts from the students’ assignments for a total of eleven markets from 1998 to 2002.

The student forecasts in this study are distinctive and valuable data. These forecasts reflect the private information held by traders at the start of a real-money electronic prediction market associated with a real-life event. In addition, the private information for each trader was acquired through the trader’s own search rather than being provided by an experimenter. Thus, traders could draw on a wide variety of information sources and apply any analytical techniques they wished to create their forecasts. This highlights an important difference between these movie markets and the traditional lab markets.

As part of our study design, we examine the influence of restrictions on the pool of traders on the forecast accuracy of the markets. For four of the markets, the trader pool was restricted to those students who had submitted forecasts before the markets opened for trading. Thus, we have a forecast from each and every trader in these markets. We designated these as ‘closed’ markets.

We expanded the trader pool for the other seven markets to include anyone with an academic affiliation. This more inclusive level of trader participation was used in the highly accurate market associated with the 1992 Presidential election (Oliven and Reitz 2004). Following the established practice of the IEM political markets, we did not elicit forecasts from traders who were not part of the student group that turned in forecasts. These markets were designated as ‘open’ markets.
Market operation

The timeline for the IEM movie markets is illustrated in Figure 1.

Since we are interested in how the market aggregates the private information held by traders, we collected the students’ point forecasts before each prediction market opened. These forecasts must precede any trading since traders can acquire information from each other through prices in the market.

Once the forecasts were turned in, the IEM market opened. Trading in the markets began from between four and fourteen days before the opening of the movie in theaters (all of the movies opened on a Friday). Once the movie opened in theaters, trading continued for four weeks.

Nielsen/EDI (entdata.com) tracks movie box office performance on a weekly basis. Daily estimates are also available at other websites, e.g., www.the-numbers.com. After the final four-week receipts are available in print (through Variety), the markets are liquidated. This entailed exchanging $1 for each winning contract held by a trader. Nothing was paid for losing contracts.

Overview of markets

In Table 1, we present some descriptive data about the eleven box-office prediction markets.

With respect to size and duration, these markets fall between the typical small sample, short-duration lab experiment and the large-scale, multi-month IEM political markets.

Table 1. IEM movie market overview

<table>
<thead>
<tr>
<th>Movie</th>
<th>Date market began trading</th>
<th>Date movie opened</th>
<th>Number of forecasts</th>
<th>Type of market</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lost in Space</td>
<td>27 March 1998</td>
<td>3 April 1998</td>
<td>44</td>
<td>Open</td>
</tr>
<tr>
<td>Mercury Rising</td>
<td>27 March 1998</td>
<td>3 April 1998</td>
<td>44</td>
<td>Open</td>
</tr>
<tr>
<td>Enemy of the State</td>
<td>9 November 1998</td>
<td>20 November 1998</td>
<td>88</td>
<td>Closed</td>
</tr>
<tr>
<td>I Still Know What You Did Last Summer</td>
<td>9 November 1998</td>
<td>13 November 1998</td>
<td>88</td>
<td>Closed</td>
</tr>
<tr>
<td>Sleepy Hollow</td>
<td>5 November 1999</td>
<td>19 November 1999</td>
<td>106</td>
<td>Open</td>
</tr>
<tr>
<td>The World is not Enough</td>
<td>5 November 1999</td>
<td>19 November 1999</td>
<td>106</td>
<td>Open</td>
</tr>
<tr>
<td>The Sixth Day</td>
<td>3 November 2000</td>
<td>17 November 2000</td>
<td>91</td>
<td>Closed</td>
</tr>
<tr>
<td>How the Grinch Stole Christmas</td>
<td>3 November 2000</td>
<td>17 November 2000</td>
<td>91</td>
<td>Closed</td>
</tr>
<tr>
<td>Monsters, Inc.</td>
<td>19 October 2001</td>
<td>2 November 2001</td>
<td>34</td>
<td>Open</td>
</tr>
<tr>
<td>Harry Potter and the Sorcerer’s Stone</td>
<td>2 November 2001</td>
<td>16 November 2001</td>
<td>111</td>
<td>Open</td>
</tr>
<tr>
<td>Die Another Day</td>
<td>8 November 2002</td>
<td>22 November 2002</td>
<td>86</td>
<td>Open</td>
</tr>
</tbody>
</table>


**STUDY RESULTS**

The purpose of an electronic prediction market is to aggregate information held by traders. There are two major characteristics of traders’ private information that are important to decision makers. The first is the consensus forecast and the second is the distribution of forecasts across traders. Consequently, we present our results in three major sections. In the first section, we examine how well the contract prices observed in the market reflect the consensus of private information held by student traders, i.e. their forecasts. In the second section, we assess the ability of these electronic prediction markets to reflect the dispersion in traders’ forecasts. Finally, we compare the forecasting accuracy of markets with and without restrictions on trader participation.

Consensus forecast results

For a continuous outcome such as box office receipts, the consensus forecast may be represented by some measure of central tendency. For this study, we used the mean of the students’ individual forecasts. (We obtained similar results using the median as our measure of central tendency. Details are available from the authors.)

The mean of the student forecasts are presented in Table 2 (Column 2).

Means implied by contract prices. To determine how well the market aggregates the students’ private information, we examined the contract prices from the day before the movie opened in theaters. Traders can choose when to enter the market and trade. Therefore, we use a set of contract prices that allow the market to aggregate the information more completely than, for example, the results of the first day of trading.

The contract prices we use are those from Thursday (at midnight). These contract prices are established by traders before anyone actually sees the movie in theaters. The question is: what is the mean forecast that is consistent with these observed prices?

We begin by assuming that the private information reflected by the contract prices is a set of normally distributed point forecasts by individual traders. Under this assumption, we can compute the resulting price of a contract for any mean and standard deviation using the cumulative normal distribution function. The first step is to compute the Z-scores associated with the top and the bottom of the range for a given contract (based on a given mean and standard deviation). The second step is to compute the difference of the cumulative normal distribution evaluated at these Z-scores. This difference is the expected price of the particular contract for a given mean and standard deviation.

To identify the mean and standard deviation that best fit the set of observed contract prices, we reversed this process. Using an optimization algorithm, we conducted an iterative search of a positive two-dimensional space. The first dimension corresponded to the mean of the normal distribution while the second dimension was the standard deviation. At each step, there was a candidate mean and standard deviation. We computed the root mean squared error (RMSE) between the observed contract prices (normalized to sum to one) and contract prices implied by a normal distribution with the candidate mean and standard deviation. Overall, the goal of the optimization was to minimize the RMSE measure. To reduce the chances of obtaining a local minimum solution, we used multiple starting points for each of the markets.

We illustrate the outcome of this process in Table 3 using the contract prices for the fall 2000 IEM box office associated with the movie *The Sixth Day*.

Most of the RMSE’s for these computations were less than 0.02 (and none exceeded 0.05). The resulting means implied by the observed contract prices are presented in Table 2 (Column 3).

Comparison of means. To determine how well the market is able to aggregate the trader’s private information about the movie’s performance, we plotted the mean of the student forecasts (x-axis) versus the mean forecast implied by the market prices on Thursday night (y-axis). The results are presented in Figure 2.

We see that most of the points lie on or near the 45-degree line indicating a close correspondence between the two means. The correlation across markets is 0.99.

This result suggests that the market was able to aggregate the consensus forecast of the student’s private information obtained from their written forecasting assignments. While these results are consistent with those observed in controlled laboratory studies of
information aggregation (Plott and Sunder 1988), important differences make our results very relevant to organizations interested in electronic prediction markets. First, traders in this study generated their own private information (forecasts) rather than their being supplied as part of a controlled experiment. Second, the outcome of the market – the movie’s four-week box office total – is unknown beforehand. Third, and most important, lab experiments usually focus on a binary outcome, i.e. event A did or did not occur. The markets discussed here aggregated the consensus forecast of a continuous outcome, a much more difficult endeavour. Overall, we conclude that electronic prediction markets do a very good job representing the consensus of the traders’ private information.

Dispersion of forecasts

As we discuss above, a key task of prediction markets is to ascertain the degree of variation in the private information held by traders. In this study, the measure of the dispersion of the traders’ private information we report is the standard deviation of the individual point forecasts. These may be found in Table 4 (Column 2).

In the process of identifying the mean of the normal distribution implied by the observed contract prices, we also determined the corresponding standard deviation that best fit the observed contract prices. These figures are reported in Table 4 as well (Column 3).

For every market, the implied standard deviation of the market forecast is smaller than the standard deviation of the students’ point forecasts. This implies a tighter

### Table 3. Implied forecast example: The Sixth Day (Fall 2000 IEM movie market)

<table>
<thead>
<tr>
<th>Contract (Range)</th>
<th>Pre-opening price (Normalized to sum to 1)</th>
<th>Prices if forecasts were distributed (\sim) Normal ((63, 15.9))^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIX50L ((\leq 50\text{MM}))</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>SIX70L ((50+\text{MM}, 70\text{MM}))</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>SIX90L ((70+\text{MM}, 90\text{MM}))</td>
<td>0.31</td>
<td>0.29</td>
</tr>
<tr>
<td>SIX90H ((&gt;90\text{MM}))</td>
<td>0.01</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note: ^1 RMSE=0.02

### Table 4. Standard deviations of forecasts

<table>
<thead>
<tr>
<th>Movie</th>
<th>Standard deviation of student forecasts (in $ millions)</th>
<th>Standard deviation implied by contract prices (in $ millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lost in Space</td>
<td>22.1</td>
<td>6.5</td>
</tr>
<tr>
<td>Mercury Rising</td>
<td>13.1</td>
<td>7.1</td>
</tr>
<tr>
<td>Enemy of the State</td>
<td>52.8</td>
<td>19.1</td>
</tr>
<tr>
<td>I Still Know What You</td>
<td>24.6</td>
<td>7.9</td>
</tr>
<tr>
<td>Did Last Summer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleepy Hollow</td>
<td>22.5</td>
<td>19.0</td>
</tr>
<tr>
<td>The World is not Enough</td>
<td>29.0</td>
<td>27.4</td>
</tr>
<tr>
<td>The Sixth Day</td>
<td>20.4</td>
<td>15.9</td>
</tr>
<tr>
<td>How the Grinch Stole</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christmas</td>
<td>29.1</td>
<td>24.6</td>
</tr>
<tr>
<td>Monsters, Inc.</td>
<td>38.0</td>
<td>31.1</td>
</tr>
<tr>
<td>Harry Potter and the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorcerer’s Stone</td>
<td>53.2</td>
<td>38.9</td>
</tr>
<tr>
<td>Die Another Day</td>
<td>18.8</td>
<td>13.4</td>
</tr>
</tbody>
</table>

Figure 2. Comparison of means
distribution of forecasts around the consensus once the traders began interacting through the prediction market.

One possible reason for this result is that traders with extremely high or extremely low forecasts changed their minds when they saw the consensus forecast of the other traders. If they were not highly confident in their forecasts, they may have sold off those contracts containing their extreme forecasts and purchased contracts nearer the consensus of the other traders. Another possible influence on the standard deviation results may be the number and range of the contracts in the market. Additionally, the assumption of a normal distribution of forecasts generating the observed contract prices may affect the implied standard deviations reported in Table 3.

For these reasons, we compared the actual contract prices with those we would expect if the entire distribution of students’ point forecasts were used to determine contract prices. These results are presented next.

Distribution of forecasts

Comparison of expected and actual contract prices. In the previous analysis, we computed the mean and standard deviation of normally distributed forecasts that would result in the contract prices we observe. We then compared them to the actual mean and standard deviation of the student point forecasts. In this section, we take advantage of the nature of the contract structure to determine how well these markets aggregated all of the information we have regarding the distribution of traders’ forecasts.

Contracts in this market correspond to a mutually exclusive and collectively exhaustive set of outcomes. The price of each contract can be viewed as the probability of that particular outcome occurring. Using the frequency data from the point forecasts, we can estimate the probability of each contract paying off $1 at the end of the market. For example, if 10% of the point forecasts lie within the range of a given contract, then its expected price should be 0.10. These expected contract prices should correspond to those we observe in the market if the entire distribution of the students’ private information (i.e., point forecasts) is being reflected in the contract prices.

We compared the expected contract prices to the actual contract prices at the same point as the analysis discussed above (Thursday at midnight). The correlations between these two sets of contract prices are presented in Table 5.

We see that the correlations for seven of the eleven markets are statistically significant (p<0.05 level). Overall, the average correlation across the eleven markets is 0.81. The pooled correlation (across all contracts) is 0.83, which is significant at the p<0.01 level.

The low correlation between the actual and expected contract prices in the Monsters, Inc. market is worth examining further. Before trading began, the majority of the students forecasted that the four-week box office would be less than $120 million. By the opening of the movie, the mean shifted from $124 million to $148 million (see Table 2). Comparing the expected and actual contract prices (from Thursday, midnight prices), we see that prices in this market did not accurately reflect an aggregation of the students’ private information (correlation = −0.05).

There are at least two possible explanations besides market failure. First, there may have been additional information released to the traders that convinced them to move prices in a direction away from the mean of the student forecasts. Alternatively, consider that trading in this market was open to all academic traders. There may have been a number of additional traders, beyond those students whose forecasts we obtained, who were determining contract prices in the market.

Examining the trading records, we found that most of the traders were not members of the class that submitted the forecasts. The students submitting forecasts accounted for about 11% of the trades in the time period from the opening of the IEM market to Thursday at midnight before the movie opened in theaters. Clearly, the other traders brought very different information to the market and moved prices in a direction away from the mean of the student forecasts. We found a similar result in the Harry Potter market. In that market, student traders accounted for 27% percent of the trades between the opening of the market and the opening of the movie in theaters.

In the next section, we see how restrictions on trader participation affected the markets’ overall accuracy.
The effect of trader pool on predictive accuracy

In predicting movie box office results, the relative magnitude of a given level of absolute error varies across markets. For example, a forecast that is within $10 million of the actual box office total of $150 is more acceptable than the same absolute error for a movie earning only $15 million over the same time period. Consequently, we use a relative measure of accuracy: absolute percentage error (APE=absolute value [actual – forecast]/actual). The results are presented in Table 6.

Across the 11 movies, the mean absolute percentage error (MAPE) is 0.29. Comparing the forecasts from the ‘open’ and ‘closed’ markets, we find that the MAPE for the seven open markets is 0.17 compared to 0.50 for the four closed markets. The difference is significant at the p<0.02 levels for a 2-tailed t-test.

The effect of having an open market on forecasting accuracy is well illustrated in the Monsters, Inc. market. Recall from Table 5 that the contract prices on Thursday at midnight before the movie opened in theaters were very different from implied by the student forecasts. However, the APE of the mean implied by the Thursday (at midnight) contract prices (0.23) was much smaller than the original mean of the student forecasts (APE=0.37).

Clearly, the open market allowed the entry of traders who chose to put their own funds at risk in the market. These self-selected traders had a positive influence on the forecasting accuracy of the market in five of the seven open markets. It appears as if traders outside of the students would not participate in the market unless these traders believed that they possessed superior information. The largest improvements were in the Monsters, Inc. and Harry Potter markets.

DISCUSSION AND CONCLUSIONS

There are a number of organizations actively studying the use of electronic prediction markets to gather widely dispersed information in order to improve decision-making. A leader in this field is Hewlett-Packard, which has conducted a number of successful prediction markets to assist with sales forecasting (Plott and Chen 2002). While there is a lot of enthusiasm surrounding prediction markets, their ability to function as an aggregator of complex information is not very well understood.

Electronic prediction markets are intended to aggregate privately held information that is brought to the market by the traders themselves. The results presented above confirm that markets are able to fulfill this function well. We find strong evidence that markets are able to aggregate information about consensus of the private information (Table 2, Figure 2) as well as the dispersion of private information (Table 4). These results confirm the ability of markets to accurately aggregate the two most important pieces of information about any distribution of information. For managers, this is important given their interest in both the consensus of forecasts as well as differences in opinions.

Kambil and Van Heck (2002) suggest electronic prediction markets run for the benefit of a particular organization should limit participation to members of the organization with demonstrable expertise. In an important way, the conditions of our ‘closed’ markets mimic those of a market run in-house since the traders were very homogeneous. We found that the markets open to a wider range of participants resulted in better predictions. While the number of markets is too small at present to make a definitive statement, the results described above suggest that opening the markets to interested outsiders who are risking their own funds in the market might actually be a more effective approach. The self-selection process coupled with risk aversion when one’s own funds are at stake seems to result in more accurate market predictions.

That being said, there are other possible explanations beyond trader restrictions that may account for these results. For example, the box office performance of some movies is harder to predict than others. In the 2000 markets, the movie How the Grinch Stole Christmas was clearly a surprise hit while The Sixth Day was clearly a surprise bomb. We did not randomly assign trader restrictions (i.e., open v. closed market status) to a particular movie. Therefore, we cannot completely easily

Table 6. Forecasting accuracy

<table>
<thead>
<tr>
<th>Movie</th>
<th>Mean forecast implied by Thursday midnight prices (in $ millions)</th>
<th>4-week box office total (in $ millions)</th>
<th>Absolute percentage error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lost in Space</td>
<td>43.0</td>
<td>59.1</td>
<td>0.27</td>
</tr>
<tr>
<td>Mercury Rising</td>
<td>32.3</td>
<td>28.5</td>
<td>0.13†</td>
</tr>
<tr>
<td>Enemy of the State</td>
<td>97.1</td>
<td>74.4</td>
<td>0.31†</td>
</tr>
<tr>
<td>I Still Know What You Did Last Summer</td>
<td>48.7</td>
<td>36.4</td>
<td>0.34†</td>
</tr>
<tr>
<td>Sleepy Hollow</td>
<td>54.0</td>
<td>82.9</td>
<td>0.35</td>
</tr>
<tr>
<td>The World is not Enough</td>
<td>111.0</td>
<td>101.3</td>
<td>0.10</td>
</tr>
<tr>
<td>The Sixth Day</td>
<td>63.1</td>
<td>33.4</td>
<td>0.89†</td>
</tr>
<tr>
<td>How the Grinch Stole Christmas</td>
<td>105.1</td>
<td>199.8</td>
<td>0.47†</td>
</tr>
<tr>
<td>Monsters, Inc.</td>
<td>148.4</td>
<td>194.9</td>
<td>0.24</td>
</tr>
<tr>
<td>Harry Potter and the Sorcerer’s Stone</td>
<td>239.2</td>
<td>234.4</td>
<td>0.02</td>
</tr>
<tr>
<td>Die Another Day</td>
<td>120.0</td>
<td>134.4</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Note: † closed markets
separate out these two effects given the number of observations available at this point in time.

In order for electronic prediction markets to become a serious alternative information system to support decision-making, research in this area must move beyond its primary focus on predictive accuracy. There should be less emphasis on documenting successes in predicting future events and more study of how conditions far removed from a lab setting might affect the functioning of prediction markets. By documenting a prediction market’s ability to aggregate both the consensus of private information as well as the entire distribution of that private information, our results make an important contribution to this body of research.

References


