Chapter 34. Prediction Markets

One of the central concepts in finance is the "efficient markets hypothesis," that prices will (somehow) incorporate all relevant information. Given the widespread belief in the informational efficiency of markets, it is natural that economists and others have designed new trading institutions to generate price-based predictions about important events, e.g., sales outcomes and elections. This chapter considers information aggregation in laboratory markets and in parallel field settings that are popularly known as *prediction markets*. The V*e*conlab Political Event Market can be used to set up a web-based trading exchange in the field, with payoffs that depend on the outcome of a subsequent event. Each student can then be given a trading account of hypothetical "classroom dollars" to use in this market, which can be "cleared" before class each day if after-hours trading is allowed. In the lab, the Limit Order program can implement an asset market experiment in which partially informed investors trade on the basis of their private information signals about the unknown "common value" of an asset.

I. The Rationale for Prediction Markets

The idea behind a prediction market is to create a tradable "share" or "contract" for each possible event. In the 2004 U.S. Presidential election, for example, one could buy and sell contracts for Bush, Kerry, and Nader using the Iowa Electronic Markets web site, which is run by experimentalists in economics, accounting, and finance at the University of Iowa. In these types of markets, an initial investment of a dollar provides the trader with a "market portfolio" of one contract for each major candidate, and one for "rest of field," so that all possible outcomes are covered. Each share has a cash payoff determined by how well that candidate performs in the subsequent election. In this manner, the transactions prices provide an indication of what investors expect.

The two main types of political event contracts are *winner-take-all* and *vote share*. In a vote share market, a contract associated with a particular candidate can be redeemed for an amount equal to the percentage of the vote subsequently obtained by that candidate. In 2004, Bush garnered about 51.5 percent of the vote, so shares of Bush were redeemed for about 51 and a half cents each. (In some markets, all amounts are scaled up into dollars instead of pennies.) In contrast, there can be more risk in winner-take-all markets, since all shares have a zero payoff, except for those of the winning candidate, which are redeemed for $1.00 each. In either type of market, a person who puts a dollar in and then does not trade at all will have earnings of exactly one dollar, since the sum of the vote shares is 100 and there is only one winner.

It is possible to make money on arbitrage if prices in event markets are not aligned. For example, if bid prices sum to more than a dollar, then a trader can

buy market portfolios for a dollar each and sell them unbundled for a profit. Conversely, if the sum of asking prices is less than a dollar, a trader could purchase a market portfolio that offers a sure gain. Another opportunity for arbitrage would be if prices are lower in one exchange than in another, which provides an opportunities to buy in one and sell in another. These arbitrage opportunities are relatively minor. For example, Wolfers and Zitzewitz (2004) collected bid and ask prices every 4 hours for shares that would pay $100 if Arnold Schwartzenegger were to be elected Governor of California. There was considerable variation in these prices over time, but the TradeSports and World Sports Exchange prices shadowed each other closely so as to eliminate arbitrage opportunities quickly.

An individual might decide to trade in a vote share market on the basis of a hunch that the prices will go up (buy now) or down (sell now). If most people think that a candidate's shares are over-priced after some unexpected negative publicity, then that price will fall as traders try to sell, and conversely for price decreases. These buy and sell pressures push price to a point where supply equals demand, which would constitute a consensus about the vote share for that candidate. Thus the trading prices in vote share markets convey information about the share of the vote that traders expect the candidate to obtain on election day. In contrast, prices in winner-take-all markets are more closely related to probabilities. To see this, note that if a candidate has no chance of winning, then the winner-take-all price would be zero since there will be no payout. Of course, a strong partisan might want to hold shares with no chance of a payout, but marginal traders should put intense downward pressures on the price if it were to rise much. In the 1988 U.S. Presidential election, for example, some traders wanted to "send a message for Jesse" by buying Jackson shares, but the resulting price increase was quickly reversed. Conversely, arbitrage would drive share prices for a sure winner up to $1.00. If a risk neutral person thinks that a candidate has a 50 percent chance of winning, then this person would be willing to pay any amount below 50 cents and would want to sell shares for any amount above 50 cents. These somewhat loose arguments suggest that transactions prices are correlated with consensus win probabilities in these markets. There have been documented attempts to manipulate prices in political event markets, but these have not succeeded for very long (Wolfers and Zitzewitz, 2005b). To summarize, there is a strong connection between trading prices and consensus win probabilities in winner-take-all markets, and between consensus vote shares and trading prices in vote share markets. The intuition behind these connections is clear, although formal derivations of these connections are only available under special assumptions (Manski, 2005; Gjerstad, 2005; Wolfers and Zitzewitz, 2005b).

The first step in setting up an event market is to define the events, which cover all possible outcomes. For example, a market designed to elicit information about sales forecasts could break down possible sales outcomes into categories, each being an event: sales below 100, between 100 and 110, etc. These categories are analogous to candidates, so it is important that all possible outcomes are covered. Then a winner-take-all market would yield consensus probabilities associated with each of the events. Arbitrage pressures will force this collection of probabilities to sum to one. Therefore, the inferred event probabilities constitute a probability distribution over sales outcomes, which could be represented concisely as a bar graph for subsequent management use.

In addition, contracts can be defined for combinations of events. Such "compound event contracts" will pay off only if all of the listed events occur. During the 2004 U.S. Presidential election campaign, for example, one of the available contracts on the TradeSports site would pay off only if Bush won *and* if Osama Bin Laden was "neutralized" prior to election day. (This would have been a bad investment.) These compound event markets can be used to make inferences about more complicated events. For example, suppose that the desired information is how much Bush's chances of winning the election (event B) would have gone up if Osama Bin Laden had been neutralized (event O). So what we want to know is $Pr(B|O)$. A compound event market ("Bush wins *and* Osama loses") is denoted $P(B,O)$, and a separate single-market with contracts that pay if Bin Laden is neutralized would yield a forecast that is denoted by $Pr(O)$. It follows from the rules of conditional probability that $Pr(B,O) = Pr(B|O)Pr(O)$, since the only way to observe both events (left side) is for Bin Laden to be neutralized first, $P(O)$, and then for Bush to win afterwards, $P(B|O)$, as shown on the right side of the conditional probability formula. Thus markets for the single event and the compound event can be used to "back out" the conditional probability of a Bush win after Bin Laden is neutralized. In a similar manner, Davis and Holt (1994c) used simultaneous price data from contracts in nomination and general election event markets to infer that the front runner with intense partisan support in the Republican nomination market, Oliver North, would not have been as strong a candidate in the general election as the more centrist candidate, James Miller.

There is no requirement that the events be political; these markets are useful whenever traders have different sources of information to be aggregated. For example, health care workers only see people getting sick in their own areas, and simple counts of cases may not adequately convey the intensity of a disease or the speed with which it is spreading among family members and co-workers. A prediction market for flu cases was established in Iowa for the 2004 and 2005 flu seasons (Nelson, Polgreen, and Neumann (2006), a development that the *Economist* (October 13, 2005) termed "Trading in Flu-tures." Another example

is the Hollywood Stock Exchange, a market that has generated forecasts of box office success and Oscar winners; these forecasts have compared favorably with those of industry experts (Pennock, Lawrence, Neilsen, and Giles, 2001). The most infamous example of market-based forecasting was the Defense Advance Research Products Agency (DARPA) plan to allow trading in "terrorist futures," a plan that was quickly abandoned after negative press accounts.

Business applications of prediction markets have attracted a lot of attention recently, as major Internet companies are rumored to be developing them. For example, suppose that the traders are sales managers in differing geographic regions for a particular company, and each one sees local trends, worrisome signs when key clients hold back, etc. Requests for sales forecasts may yield biased information, for example, if sales managers send in conservative predictions to make their subsequent performance look better. A market that collects and processes information from local sales managers could be quite valuable at the corporate headquarters where production decisions are made. One early business forecasting application involved Hewett-Packard printer sales. Chen and Plott (2002) report that the market-generated price forecasts were more accurate than those of internal experts. The most widely studied prediction markets, however, are based on political events, which is the next topic.

II. The Success of Political Event Markets

One advantage of prediction markets is that they operate on a "24-7" basis, so the effects of unforeseen events on share prices can be tracked quickly, and inexpensively relative to polls. In the 1994 Virginia Senate race, for example, there was a fairly active IEM market for shares of the main candidates, incumbent Chuck Robb and the Republican challenger, Oliver North, and for two surprisingly strong independent candidates as well. Shares of Robb were selling above those for North throughout the month of August. The final debate on Labor Day weekend was held at Hampton Sidney College (North country), and in an effort to sound tough, Robb said that he would take food "out of the hands of widows and orphans" to control the growth of entitlement programs. The moderator gave him a chance to clarify, but he stood his ground. Most of the traders were college students in the area, and some of the Robb supporters expressed the view that he had won the debate, although North supporters strongly disagreed. Not surprisingly, this difference of opinion generated a high trading volume, and the price of North shares climbed overnight. The polls picked up North's lead in the days that followed, but there had been almost a month between the post-debate poll and the previous poll, so the market provided a quick estimate of the extent of North's surge that was not available otherwise. Several weeks later, there were newspaper accounts of bumper stickers that read "Widows and Orphans for North."

Another aspect of these markets is that, as with other financial markets, prices respond to *new* information, since old information is already incorporated. North's share prices had fallen in the spring campaign season when Ronald Regan questioned the truthfulness of his statements on the Iran-Contra scandal, and soon stores were carrying "Oliver North Pack of Lies" playing cards, with one "lie" per card. But in September when Nancy Regan repeated the charge, it had no effect on prices, despite wide press coverage. The North prices did fall that month, however, when he claimed that (then President) Clinton was not his commander-in-chief. As a result, North's aids launched the "stealth campaign," keeping him out of the spotlight. The press later reported that a motorcyclist caught up with the North motorcade at a stoplight and held up a sign reading "Clinton *is* your Commander in Chief, and Ronald Regan doesn't even like you."

A typical pattern that showed up in this election market was that prices in the winner-take-all market were more volatile than vote share prices, which is explained by the fact that it only takes a relatively small margin to win. For example, North shares in the winner-take-all market doubled in price after the third debate, from about 25 cents to about 50 cents, whereas prices in the vote share market changed by much less. Davis and Holt (1994c) estimated that each 1 percent change in the ratio of the vote share prices for the two leaders resulted in a 3.8 percent change in the ratio of the winner-take-all prices.

A final feature of vote share markets that is worth mentioning is that they are often less volatile than polls, and over the course of a long campaign, the prices are typically closer to the final election outcomes than the major polls (Berg et al., 2003). In the North-Robb campaign, the polls often bracketed the market price movements over the same time period. The intuitive reason for the relative stability of market predictions is that investors in the market are trying to guess what will ultimately happen on election day, i.e. to look past short term events. In contrast, polls reflect individual voter's attitudes about candidates if the election were held that day. Another factor could be that investors are simultaneously considering multiple poll results, with subjective corrections for differing sample sizes and degrees of recency.

At their best, polls have the advantage of being based on representative samples, whereas investors in electronic markets are a self-selected group, with heavy representation of younger and better-educated voters. But pollsters have to make educated guesses about which of their respondents are likely voters. This correction can be quite large, and it usually means scaling down the Democratic percentages, making the results more subjective than is commonly thought. Most people express surprise at the possibility that markets composed of a relatively small group of traders can even come close to major polls in terms of predictive accuracy, but when you consider the variability across polls conducted at the similar times, this does not seem too farfetched. Of course, comparisons are a

little misleading since markets respond to polls. In the North-Robb race, for example, Robb's prices recovered to achieve parity with North's prices in the final days of the campaign, but only after the release of election eve poll results showing Robb to be in the lead.

A comparison of the IEM vote share prices and Gallup poll results in the last five U.S. Presidential elections is tallied in Wolfers and Zitzewitz (2005). For the 13 candidates in these elections, the average absolute error for the final Gallup poll was 1.9 percentage points, as compared with 1.6 percentage points for the IEM markets. The most dramatic failure of the market forecasts came in 1996 when there was a major influx of new cash just prior to the election, but predictions based on the transactions-based average prices for the week prior to this election were quite accurate (Berg et al., 2003). The most dramatic success was for the 1988 U.S. Presidential election, when the polls were all over the place, and the IEM prices were on target (Forsythe et al., 1991, 1992). In this election, George Bush (senior) won by 7.8%, which essentially matched the 8% Bush margin provided by the first vote share market run at Iowa. The election eve Gallup poll gave Bush a 12% edge, and both the Harris and the WSJ/CNN polls had Bush *losing* by about 5%.

Considering all types of elections, the accuracy of vote share markets is increased when the election has high visibility (e.g. Presidential elections), when there is a high trading volume, and when the number of possible contracts is smaller (Berg et al., 2003). An overall assessment would be that the market-based forecasts are surprisingly accurate early in the campaigns, and that they do at least as well as polls in the final days. Moreover, political event markets sometimes attract a fair amount of trading activity in remote locations where polls are not available, e.g. in small local elections.

IV. Information Aggregation and "Common Value Trading"

There is a parallel line of research with laboratory experiments in which the unknown event to be predicted is determined randomly, and each trader is provided with a randomly generated private "signal." In this case, it is possible to measure precisely what the relevant aggregate information is, and to see if this information is reflected in prices. The examples in this section will be based on the V*e*conlab setup that provides traders with independent "signals" about an unknown final redemption value. The signals are analogous to bidders' signals in a common value auction; each person knows something about the final value, but nobody knows what it is exactly.

For example, suppose that an asset is traded for 20 periods, after which it will be redeemed for an unknown amount per share. You might think of period 20 as the time at which results of a drug patent approval decision will be made on the basis of ongoing clinical trials. Each trader may have different sources of

information about the potential profitability of the drug relative to that of competing drugs, and trading prices for the asset may be affected by these different sources of information.

Suppose that the only information that one trader has is that the common redemption value is equally likely to be any amount between $25 and $75, and a second trader only knows that the redemption value is between $55 and $105. Since the first trader knows that the final value is no higher than $75 and the second trader knows that it will be at least $55, together they could reduce the range of possible final values to the $55-$75 range. In this example, each person's initial uncertainty was characterized by a $50 range, but their joint information narrowed the range to $20. It is easy to imagine that if additional traders have independent estimates of the unknown common value, their aggregate information may yield a much tighter range of possible values.

Bostian, Goeree, and Holt (2005) implement this type of situation by providing each trader with a signal that is known to be within plus or minus $20 of the true redemption value. For example, if the true redemption value is $50, then with a range of plus or minus $20, individual signals would be between $30 and $70. A person with a relatively low signal of 35 would know that the value is between 15 and 55, and a person with a relatively high signal of 65 would know that the value is between 45 and 85. Traders were not told anything about the possible level of the redemption value, other then their own signals. A person with a signal of S, for example, was told that the final redemption value would be equally likely to be any amount between S - $20 and S + $20. In this case, the lowest of the signals puts an upper bound on the range of redemption values, and the highest of the signals puts a lower bound on this range. Therefore, the range will tend to be narrower in markets with more traders.

The asset did not pay dividends and there was no interest paid on cash, so the only information that might be revealed in the trading process would be something about others' beliefs about the final value. All sessions lasted for 20 periods, and each trader began with an initial portfolio of 6 shares and a signal as described above. In one treatment, traders were given low cash balances of $500, and in another treatment, they were given $1500. The motivation for the high cash treatment was the observation of stronger price bubbles and crashes in the limit order asset markets discussed in chapter 11. This information aggregation can be seen in Figure 34.1, which shows results for a high cash session. There is a slow, steady bubble that reached about $60 in period 13, before falling to the (unknown) true common value of about $50 in the final period, which is shown by the flat horizontal line. The range of possible common values for this session, based on the private information individual signals, was from $47.38 to $50.34. Despite the final period convergence, the results of these experiments offer mixed

support for the accuracy of prediction markets, since price surges that are not based on fundamentals can occur.

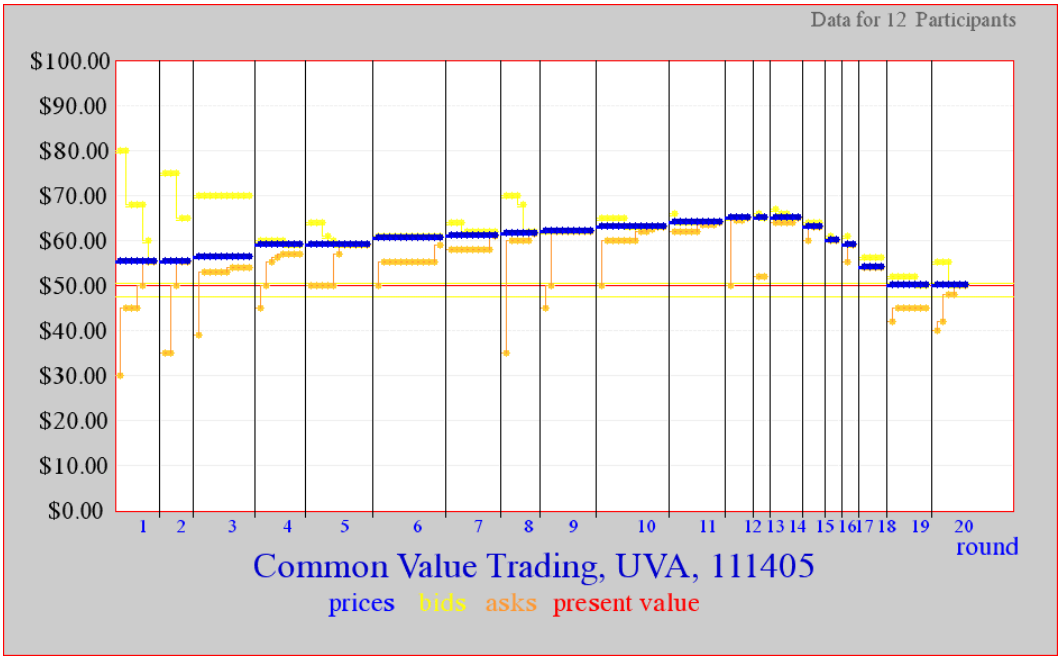

Figure 34.1. A Call Market with an Unknown Common Redemption Value (Dark Horizontal Line). Source: Bostian, Goeree, and Holt (2005)

Price bubbles were observed in all of the high-cash sessions, and with one exception, bubbles were not observed in the low-cash sessions. Regardless of whether or not a bubble formed, however, the final trading price in round 20 tended to be quite close to the narrow range implied by aggregating individuals' disparate information signals (again with one exception). Interestingly, the Iowa Economic Markets are required by federal regulators to limit individual investments to $500.

Another setup option for the V*e*conlab Limit Order Market program is to provide signals to a subset of the traders, i.e. to designated "insiders." In this case, class discussion can focus on whether insiders earn more than uninformed traders, and on whether the trading prices end up revealing the insiders' information. Plott and Sunder (1982) report an experiment with three possible "states" that determine final dividends. Some traders were told the actual state before trading began, although nobody knew identities of other insiders. Prices tended to adjust quickly to the level that would have been expected if all traders had been informed of insider information, i.e. to the price determined by the intersection of the supply and demand curves if all traders knew the true state.

These information-aggregation predictions are known in the literature as "rational expectations" predictions. The adjustment to rational expectations predictions was so quick that the profits of insiders were not noticeably different from the profits of the uninformed traders. Watts (1992) reports a similar experiment in a different informational setting, and in this case, insiders tended to earn higher profits, although price convergence patterns were generally supportive of the information revelation ("rational expectations") predictions.

IV. Extensions

Prediction markets are essentially betting mechanisms, and such mechanisms have arisen naturally in many forms. One example is "pari-mutuel" betting structure used at racetracks. Under pari-mutuel betting, the purse (total amount bet, minus commissions) is divided among those who bet on the winner, in proportion to the amounts that they bet. Suppose, for example, that there are just two horses and the equal dollar amounts are bet on each. Then the payout rates are the same, and the best decision is to bet on the horse that you think will win (which would then raise the price for the horse that you bet on). If instead there is twice as much cash being bet on one of the horses, then its payout rate is only half as high. If you do not think its chances of winning are two-thirds or more, then you should bet on the long shot. Even though the previous argument implicitly assumes risk neutrality, which is a bad assumption for bettors, the point is that the allocation of bets will tend to incorporate the diverse information that betters bring to the racetrack. Plott, Wit, and Yang (1997) ran experiments motivated by pari-mutuel betting. There were six alternative assets, and only one of them had been selected at random to have a positive payout. The payout was divided equally among the investors in that asset. Each subject received a private signal with some information about which asset had been selected, but the order of decision making was not exogenous as in the information cascade experiments discussed in chapter 31. Individuals could observe others' decisions as they were made. In this manner, the information dispersed among the investors could become aggregated and incorporated into the prices of the six assets. In most cases, the asset prices ended up signaling which asset would actually produce a positive payout. In some cases, however, an initial surge of purchases for a particular asset would stimulate others to follow, and the result was that the market price rose for an asset that had no payout in the end, as would be the case with an incorrect cascade.