

NBER WORKING PAPER SERIES

PREDICTION MARKETS

Justin Wolfers
Eric Zitzewitz

Working Paper 10504
<http://www.nber.org/papers/w10504>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2004

Prepared for the Journal of Economic Perspectives. The authors would like to thank David Pennock, Emile Servan-Schreiber of NewsFutures, David Dempsey and John Delaney of Tradesports, Alison Fealey and Oliver Frankel of Goldman Sachs, and George Neumann of IEM for help with data. Thanks to Kay-Yut Chen, Robin Hanson, Jim Hines, Andrew Leigh, Betsey Stevenson, Timothy Taylor, Hal Varian, Michael Waldman, Craig Yee and William Ziemba for stimulating discussions, and to seminar audiences at Stanford, Kansas and Wharton. Doug Geysler, Chris Lion, Mike Goelzer, Paul Reist, Eric Snowberg, and Ravi Pillai provided outstanding research assistance. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

©2004 by Justin Wolfers and Eric Zitzewitz. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Prediction Markets
Justin Wolfers and Eric Zitzewitz
NBER Working Paper No. 10504
May 2004
JEL No. D7, D8, E3, E6, G1, H8, Q4

ABSTRACT

We analyze the extent to which simple markets can be used to aggregate dispersed information into efficient forecasts of uncertain future events. Drawing together data from a range of prediction contexts, we show that market-generated forecasts are typically fairly accurate, and that they outperform most moderately sophisticated benchmarks. Carefully designed contracts can yield insight into the market's expectations about probabilities, means and medians, and also uncertainty about these parameters. Moreover, conditional markets can effectively reveal the market's beliefs about regression coefficients, although we still have the usual problem of disentangling correlation from causation. We discuss a number of market design issues and highlight domains in which prediction markets are most likely to be useful.

Justin Wolfers
Graduate School of Business
Stanford University
518 Memorial Way
Stanford, CA 94305
as of July 2004:
The Wharton School
University of Pennsylvania
and NBER
jwolfers@nber.org

Eric Zitzewitz
Graduate School of Business
Stanford University
518 Memorial Way
Stanford, CA 94305
ericz@stanford.edu

1. Introduction

In July 2003, press reports began to surface of a project within the Defense Advanced Research Projects Agency (DARPA), a research think tank within the Department of Defense, to establish a Policy Analysis Market that would allow trading in various forms of geopolitical risk. Proposed contracts were based on indices of economic health, civil stability, military disposition, conflict indicators and potentially even specific events. For example, contracts might have been based on questions like: “How fast will the non-oil output of Egypt grow next year?” or “Will the U.S. military withdraw from country A in two years or less?” Moreover, the exchange would have offered combinations of contracts, perhaps combining an economic event and a political event. The concept was to discover whether trading in such contracts could help to predict future events, and how connections between events were perceived. However, a political uproar followed. Critics savaged DARPA for proposing “terrorism futures,” and rather than spend political capital defending a tiny program, the agency dropped the proposal.¹

Ironically, the aftermath of the DARPA controversy provided a vivid illustration of the power of markets to provide information about probabilities of future events. An offshore betting exchange, Tradesports.com, listed a new security that would pay \$100 if the head of DARPA, Admiral John Poindexter, was ousted by the end of August 2003. Early trading suggested a likelihood of resignation by the end of August of 40 percent, and price fluctuations reflected ongoing news developments. Around lunchtime on July 31, reports started citing credible Pentagon insiders who claimed knowledge of an impending resignation. Within minutes of this news first surfacing (and hours before it became widely known), the price spiked to around 80. These reports left the date of Poindexter’s proposed departure uncertain, which explains the remaining risk. As August dragged on, the price slowly fell back toward 50. On August 12, Poindexter then issued a letter of resignation suggesting that he would resign on August 29. On the 12th, the market rose sharply, closing at a price of 96.

¹ Looney (2003) provides a useful summary of both the relevant proposal and its aftermath. Further, Robin Hanson has maintained a useful archive of related news stories and government documents at: <http://hanson.gmu.edu/policyanalysismarket.html>.

This anecdote describes a new—and emerging—form of financial market, often known as a prediction market, but also going by the name “information market” or “event futures.” Analytically, these are markets where participants trade in contracts whose payoff depends on unknown future events. Much of the enthusiasm for prediction markets derives from the efficient markets hypothesis. In a truly efficient prediction market, the market price will be the best predictor of the event and no combination of available polls or other information can be used to improve on the market-generated forecasts. This statement does not require that all individuals in a market be rational, only that the marginal trade in the market is motivated by rational traders. Of course, it is unlikely that prediction markets are literally efficient, but a number of successes in these markets, both within firms and with regard to public events like presidential elections, have generated substantial interest among both political and financial economists.

Although markets designed specifically for information aggregation and revelation are our focus in this article, the line between these kinds of prediction markets and the full range of contingent commodities -- from stock in your company to betting on the Super Bowl -- can become blurry. However, we will generally lean away from discussing markets where the primary focus is holding or trading risk that may be intrinsically enjoyable, as in sports betting and other gambling markets. We will also lean away from focusing on markets that are substantial enough in size to allow a significant extent of risk-sharing and pooling by matching risky assets with risk-acceptant investors, as the major financial markets do.² However, most contingent commodity markets involve some mix of risk-sharing, fun and information transmission, so these distinctions are not impermeable.

We begin by describing the types of contracts that might be traded in prediction markets, before proceeding to survey several applications. We then draw together a rough and fairly optimistic description of what we have learned from early experiments, raise some market design issues, and conclude with some evidence on the limitations of prediction markets.

² For a vision of how prediction markets, if they develop sufficient liquidity, may also prove useful for those wishing to hedge against specific risks, see the discussions in Athanasoulis, Shiller, and van Wincoop (1999) and Shiller (2003).

2. Types of Prediction Markets

In a prediction market, payoffs are tied to unknown future events. The design of how the payoff is linked to the future event can elicit the market’s expectations of a range of different parameters. We will speak as though the market is itself a representative “person” with a set of expectations. However, the reader should be warned that there are important but subtle differences between, say, the market’s median expectation and the median expectation of market participants.

Table 1 summarizes the three main types of contracts.

Table 1: Contract Types—Estimating Uncertain Quantities or Probabilities

Contract	Example	Details	Reveals market expectation of...
Winner-takes-all	Event y : Al Gore wins the popular vote	Contract costs $\$p$ Pays \$1 if and only if event y occurs Bid according to value of $\$p$	Probability that event y occurs, $p(y)$
Index	Contract pays \$1 for every percentage point of the popular vote won by Al Gore	Contract pays $\$y$.	Mean value of outcome y : $E[y]$
Spread	Contract pays even money if Gore wins more than y^* % of the popular vote.	Contract costs \$1 Pays \$2 if $y > y^*$ Pays \$0 otherwise. Bid according to the value of y^* .	Median value of y .

First, in a “winner-takes-all” contract, the contract costs some amount $\$p$ and pays off, say, \$1 if and only if a specific event occurs, like a particular candidate winning an election. The price on a winner-take-all market represents the market’s expectation of the probability that an event will occur (assuming risk neutrality).³

³ The price of a winner-takes-all security is essentially a state price, which will equal an estimate of the event’s probability under the assumption of risk neutrality. The sums wagered in prediction markets are typically small enough that assuming that investors are not averse to the idiosyncratic risk involved seems reasonable. But if the event in question is correlated with investors’ marginal utility of wealth, then probabilities and state prices can differ. In what follows, we leave this issue aside and use the term probability to refer to risk-neutral probability.

Second, in an “index” contract, the amount that the contract pays varies in a continuous way based on a number that rises or falls, like the percentage of the vote received by a candidate. The price for such a contract represents the mean value that the market assigns to the outcome.

Finally, in “spread” betting traders differentiate themselves by bidding on the cutoff that determines whether an event occurs, like whether a candidate receives more than a certain percentage of the popular vote. (Another example of spread betting is point-spread betting in football, where the bet is either that one team will win by at least a certain number of points, or will not.) In spread betting, the price of the bet is fixed, but the size of the spread can adjust. When spread betting is combined with an even-money bet (that is, winners double their money while losers receive zero), the outcome can yield the market’s expectation of the median outcome because this is only a fair bet if a payoff is as likely to occur as not.

The basic forms of these relevant contracts will reveal the market’s expectation of a specific parameter: a probability, mean, or median, respectively. But in addition, prediction markets can also be used to evaluate uncertainty about these expectations. For instance, consider a family of winner-takes-all contracts that pay off if and only if the candidate earns 48 percent of the vote, 49 percent, 50 percent and so on. This family of winner-takes-all contracts will then reveal almost the entire probability distribution of the market’s expectations. A family of spread betting contracts can yield similar insights. An even-money bet in a spread contract will define the median, as explained above. But for similar reasons, a contract that costs \$4 and pays \$5 if $y > y^*$ will elicit a value of y^* that the market believes to be a four-fifths probability, thus identifying the 80th percentile of the distribution. As a final alternative, non-linear index contracts can also reveal more information about the underlying distribution. For instance, consider a market with two index contracts, one which pays in a standard linear form, and another that pays according to the square of the index, y^2 . Market prices will reveal the market’s expectation of $E[y^2]$ and $E[y]$, which can be used to make an inference about the market’s beliefs regarding the standard deviation of $E[y]$, more commonly known as the standard error. (Recall that the standard deviation can be expressed as $\sqrt{(E[y^2]-E[y]^2)}$, or the square root of the mean of the squares less the square of the means.) By the same logic,

adding even more complicated index contracts can yield insight into higher order moments of the distribution.

Applications and Evidence

Perhaps the best-known prediction market among economists is the Iowa Electronic Market, run by the University of Iowa. The original Iowa experiment, run in 1988, allowed trade in a contract that would pay 2 ½ cents for each percentage point of the popular vote in the presidential election won by Bush, Dukakis, or others. More recently, it has run markets based on the 2003 California gubernatorial election, the 2004 presidential election, the 2004 Democratic presidential nomination, and how the Federal Reserve will alter the federal funds interest rate. Universities in other countries have also started running event markets about their own elections, like the Austrian Electronic Market run by the Vienna University of Technology or the University of British Columbia Election Stock Market that focuses on Canadian elections.

There are a growing number of web-based event markets, often run by companies that provide a range of trading and gambling services. Some prominent examples include Tradesports.com and Betfair.com, as well as pseudo-markets (in which participants trade virtual currency) such as Newsfutures.com and Ideosphere.com. These websites often take the lead on defining a contract (as in the example of Poindexter's departure from DARPA described earlier), but then allow individuals to post their offers and to accept the offers of others.

More recently, Goldman Sachs and Deutsche Bank have launched markets on the likely outcome of future readings of economic statistics, including employment, retail sales, industrial production, and business confidence. The Chicago Mercantile Exchange is planning to open a market in inflation futures. Some event markets also forecast private-sector returns. The Hollywood Stock Exchange allows people to use virtual currency to speculate on movie-related questions like opening-weekend performance, total box office returns, and who will win Oscars. In several cases, private firms have found innovative ways to use prediction markets as a business forecasting tool.

Table 2 lists some of these prediction markets. Drawing on these and other experiences with event markets, it is possible to start suggesting some generalizations

about how prediction markets work, both in terms of their accuracy and whether arbitrage or market manipulation is possible.

Table 2: Prediction Markets

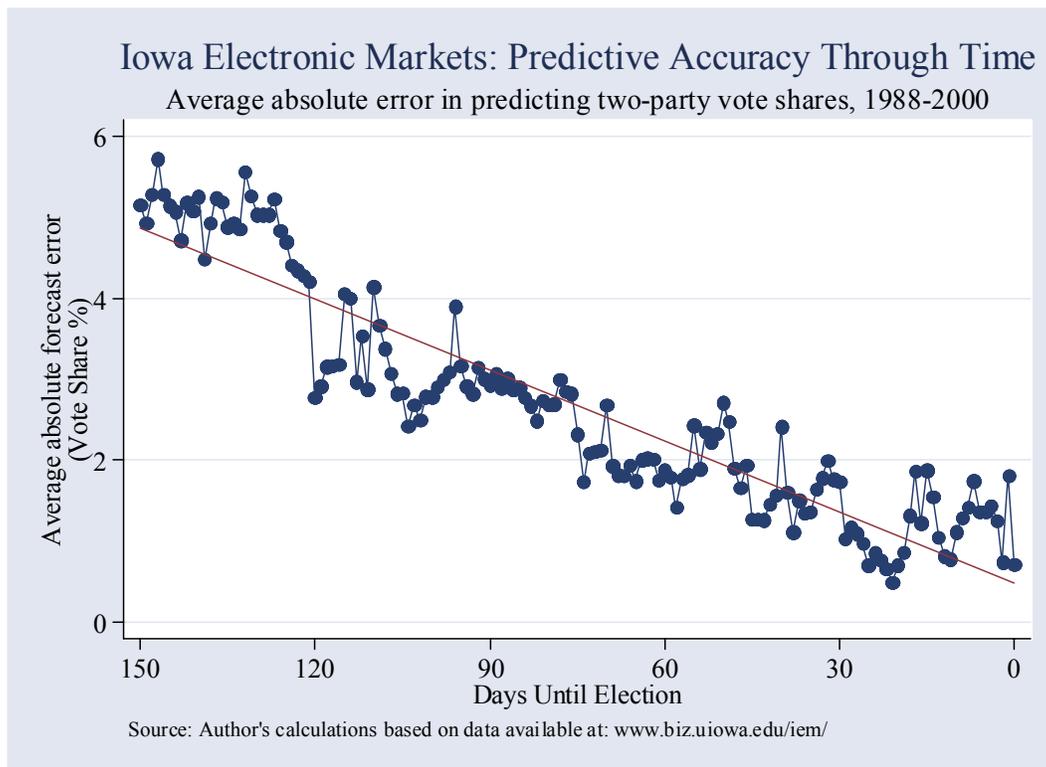
Market	Focus	Typical turnover on an event (\$US)
Iowa Electronic Markets www.biz.iowa.edu/iem <i>Run by University of Iowa</i>	Small-scale election markets. Similar markets are run by: UBC (Canada) www.esm.buc.ca and TUW (Austria) http://ebweb.tuwien.ac.at/apsm/	Tens of thousands of dollars (Traders limited to \$500 positions)
TradeSports www.tradesports.com <i>For profit company</i>	Trade in a rich set of political futures, financial contracts, current events, sports and entertainment	Hundreds of thousands of dollars
Economic Derivatives www.economicderivatives.com <i>Run by Goldman Sachs and Deutsche Bank</i>	Large-scale financial market trading in the likely outcome of future economic data releases	Hundreds of millions of dollars
Newsfutures www.newsfutures.com <i>For profit company</i>	Political, finance, current events and sports markets. Also technology and pharmaceutical futures for specific clients.	Virtual currency redeemable for monthly prizes (such as a TV)
Foresight Exchange www.ideosphere.com <i>Non-profit research group</i>	Political, finance, current events, science and technology events suggested by clients.	Virtual currency
Hollywood Stock Exchange www.hsx.com <i>Owned by Cantor Fitzgerald</i>	Success of movies, movie stars, awards, including a related set of complex derivatives and futures. Data used for market research.	Virtual currency.

Accuracy of Prediction Markets

Arguably the most important issue with these markets is their performance as predictive tools. In the political domain, Berg, Forsythe, Nelson, and Reitz (2001) summarize the evidence from the Iowa Electronic Markets, documenting that the market has both yielded very accurate predictions, and also outperformed large-scale polling organizations.

Figure 1 shows data from the past four U.S. presidential elections. The horizontal axis shows the number of days before the election. The vertical axis measures the average absolute deviation between the prices of index contracts linked to the two-party shares of the popular vote for each party, and actual vote shares earned in the election. In the week leading up to the election, these markets have predicted vote shares for the Democratic and Republican candidates with an average absolute error of around 1½ percentage points. By comparison, over the same four elections, the final Gallup poll yielded forecasts that erred by 2.1 percentage points. The graph also shows how the accuracy of the market prediction improves as information is revealed and absorbed as the election draws closer.

Figure 1: Information Revelation Through Time

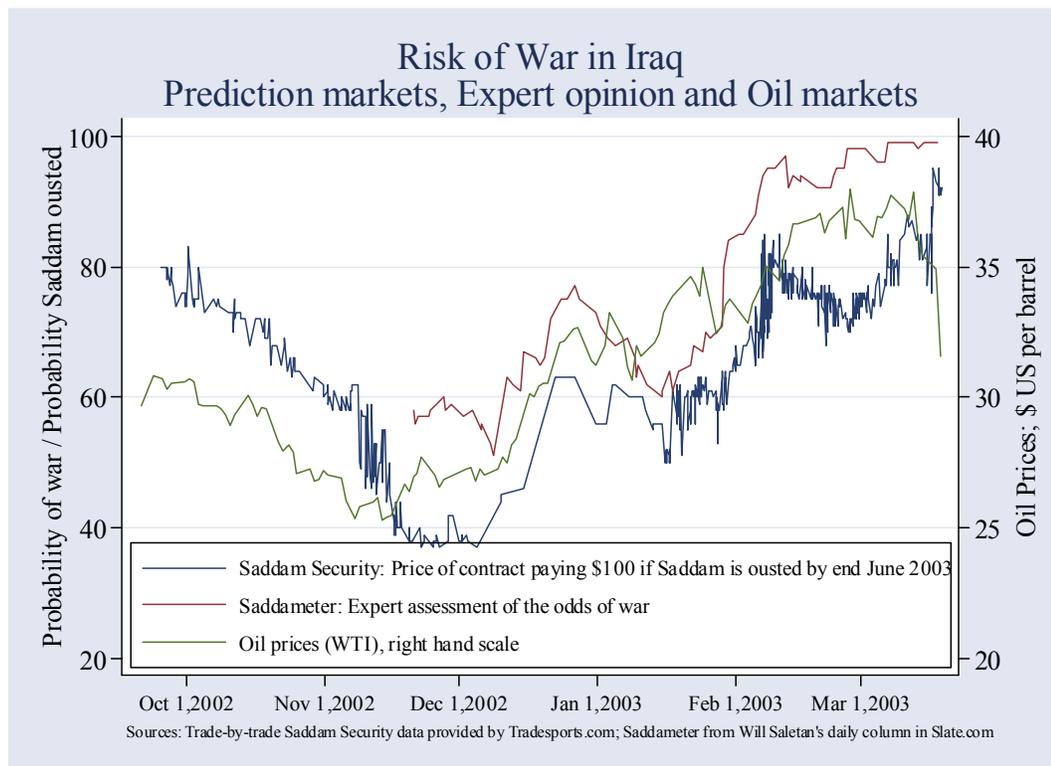


Perhaps more surprising in terms of how well prediction markets can aggregate information is the performance of markets at the level of the individual district. Typically districts are sufficiently small that there is little interest (or funding) for local polling, yet when Australian bookmakers started betting on district-level races, Wolfers and Leigh (2002) report that they were extremely accurate.

That said, comparing the performance of markets with a mechanistic application of poll-based forecasting may not provide a particularly compelling comparison. A more relevant starting point might be to compare the predictions of markets with those of independent analysts. For an example along these lines, consider the “Saddam Security”, which was a contract offered on TradeSports paying \$100 if Saddam Hussein were ousted from power by the end of June 2003.

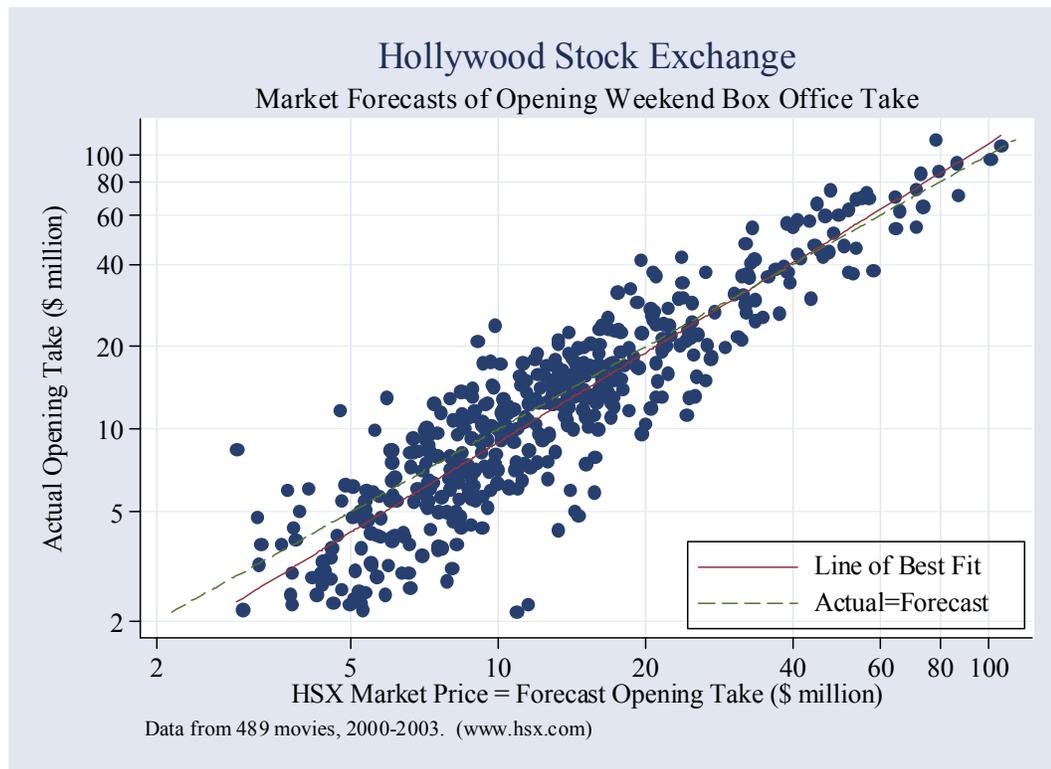
Figure 2 shows that the price of this contract moved in lockstep with two other measures: expert opinion as shown by an expert journalist’s estimate of the probability of the United States going to war with Iraq; and oil prices, an obvious barometer of political strife in the Middle East.

Figure 2: The Saddam Security



In a corporate context, the Hollywood Stock Exchange predicts opening weekend box office success, and Figure 3 shows that these predictions have been quite accurate. Further, this market has been about as accurate at forecasting Oscar winners as an expert panel (Pennock, Lawrence, Giles and Nielsen, 2003).

Figure 3: Predicting Movie Success



Some firms have also begun to experiment with internal prediction markets. An internal market at Hewlett-Packard produced more accurate forecasts of printer sales than did the firm’s internal processes (Chen and Plott, 2002). Ortner (1998) described an experiment at Siemens in which an internal market predicted that the firm would definitely fail to deliver on a software project on time, even when traditional planning tools suggested that the deadline could be met. While the Hollywood markets have drawn many participants simply on the basis of their entertainment value, the HP and Siemens experiences suggested that motivating employees to trade was a major challenge. In each case, the firms ran real money exchanges, with only a relatively small trading population (20-60 people), and subsidized participation in the market, by either endowing traders with a portfolio or matching initial deposits. The predictive performance of even these very thin markets was quite striking.

In another recent prediction market, traders in “Economic Derivatives” predict the likelihood that economic data released later in the week will take on specific values. The traditional approach to aggregating forecasts is to simply take an average or a “consensus

estimate” from a survey of 50 or so professional forecasters. We now have data from the first year of operation of these markets.

Table 3 analyzes these early outcomes, comparing average market and consensus forecasts of three variables: total nonfarm payrolls data released by the Bureau of Labor Statistics; retail trade data (excluding autos) released by the Bureau of the Census; and business confidence as measured by the Institute for Supply Management’s survey of manufacturing purchasing managers. The market-based predictions of these economic indicators are always extremely close to the corresponding “consensus” forecast, and hence the two estimates are highly correlated. There are no statistically (or economically) meaningful differences in forecast performance – measured as either the correlation with actual outcomes, or in terms of average absolute forecast errors. That said, this early sample is sufficiently small that precise conclusions are difficult to draw.

**Table 3: Predicting Economic Outcomes
Comparing Market-Aggregated Forecasts with Consensus Surveys**

	Non-Farm Payrolls	Retail Trade (ex Autos)	ISM Manufacturing Purchasing Managers’ Index
	(Monthly change, ‘000s)	(Monthly change, %)	
Panel A: Correlations			
Corr(Market, Consensus)	0.91	0.94	0.95
Corr(Consensus, Actual)	0.26	0.70	0.83
Corr(Market, Actual)	0.22	0.73	0.91
Panel B: Mean absolute error			
Consensus	71.1	0.45	1.10
Market (empirical)	72.2	0.46	1.07
Market (implied expectation)	65.7	0.34	1.58
Panel C: Standard deviation of forecast errors (Standard error of forecast)			
Consensus	99.2	0.55	1.12
Market (empirical)	97.3	0.58	1.20
Market (implied expectation)	81.1	0.42	1.96
Sample size	16	12	11

Notes: “Market” = market-implied mean forecast from <<http://www.economicderivatives.com>>

“Consensus” = average of around 50 forecasters from <<http://www.briefing.com>>

“Actual” = Preliminary estimates from original press releases (BLS, Census, ISM).

Interestingly, these markets yield not just a point estimate for each economic indicator, but involve a menu of ten to twenty winner-takes-all contracts as to whether the

indicator will take on specific values. This family of contracts reveals an approximation to the full probability distribution of market expectations. Consequently we can calculate the level of uncertainty surrounding specific point estimates. That is, because the market prices give weight to a range of outcomes beyond the central estimate, the weight given to outcomes further from the central estimate yields estimates of uncertainty. One measure of uncertainty is the expected absolute forecast error (although calculations using standard deviation provide the same qualitative results). The market-based assessments of uncertainty are shown in the last line of panel B. Comparing these implied expectations with outcomes in the first two rows of panel B suggests that the market-based assessments of uncertainty are of about the right magnitude. Finally, one can compare the implied standard errors of the forecasts with the reported standard errors of the statistics that the market is attempting to forecast. For instance, the Census Bureau reports that the change in retail trade is estimated with a standard error of around 0.5 percent, while the standard error implied by the prediction market is 0.42 percent. Taken literally, this suggests that the market believes that it is less uncertain about the Census Bureau estimate than the Census Bureau is.⁴ Such results suggest either that the statistical agencies' errors are predictable, that their standard error estimates are (slightly) upwardly biased, or that traders are overconfident.

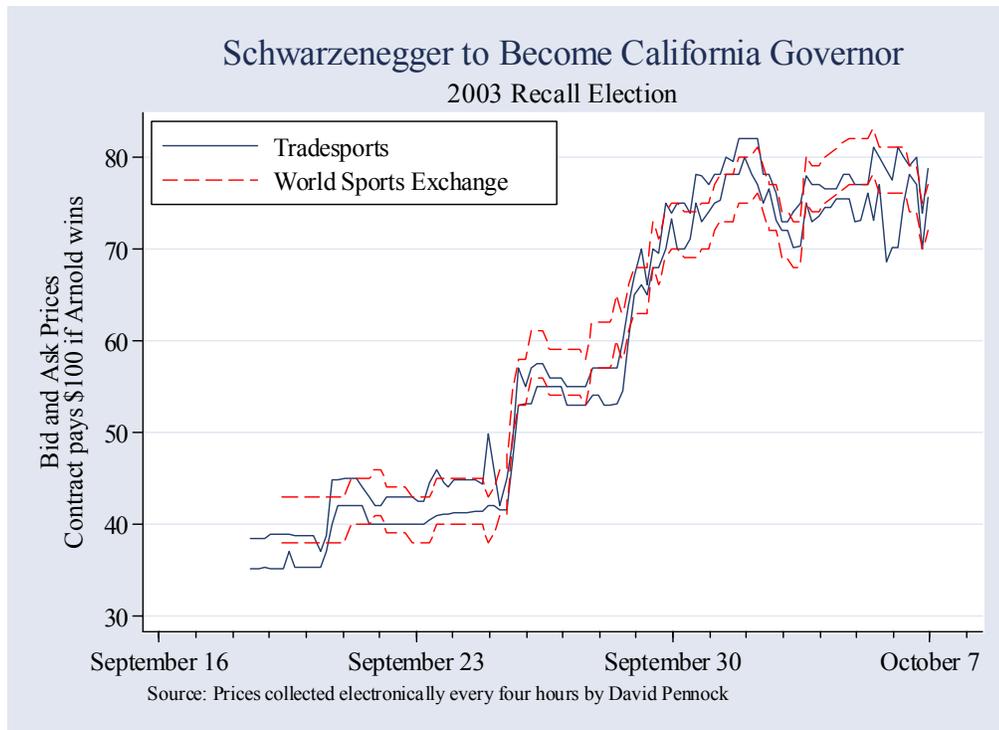
Possibilities for Arbitrage

Prediction markets appear to present few opportunities for arbitrage. There are several ways of looking for arbitrage opportunities: whether prices for similar contracts can be arbitrated across different exchanges or different securities; whether predictable patterns in the movement of the prices allow for arbitrage; and whether arbitrageurs might be able to exploit predictable deviations from rationality.

⁴ A similar comparison can be made for non-farm payrolls, although the inference is less direct. The BLS estimates that their final estimate of the change in non-farm payrolls has a standard error of around 64,000, while the preliminary estimate is more uncertain. The BLS has yet to estimate a standard error for their preliminary estimates, but the root mean squared error of the preliminary estimate relative to the final estimate is around 50,000. If the revision to the preliminary estimate and the subsequent error in the revised estimate were uncorrelated, this would imply a standard error for the preliminary estimate of about 81,500. Comparing these numbers with the average standard error of the market forecast of 81,100 suggests that the market is about as sure of the advance estimate as the BLS.

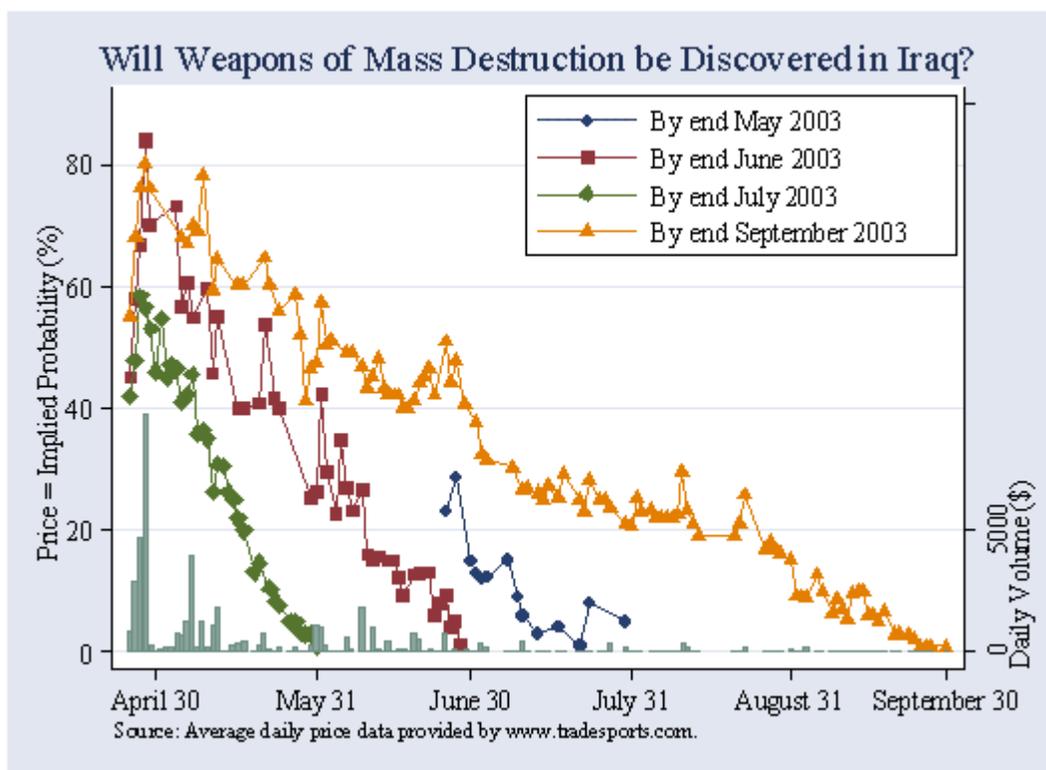
Figure 4 shows the bid and ask prices on a contract that paid \$100 if Schwarzenegger was elected California's governor in 2003, sampling data on bid and ask prices from two online exchanges every four hours. While both sets of data show substantial variation, they co-move very closely, and opportunities for arbitrage (when the bid price on one exchange is higher than the ask on another), are virtually absent.

Figure 4: 2003 California Gubernatorial Election



The pricing of families of related securities tends to be internally consistent. For example, Figure 5 shows the prices of several securities launched by Tradesports that paid off if weapons of mass destruction were found in Iraq by May, June, July, or September 2003. Their prices moved closely together in a way that suggests that the prices of each contract digested similar information at close to the same time.

Figure 5: Will Weapons of Mass Destruction be Discovered in Iraq?



In most cases, the time series of prices in these markets does not appear to follow a predictable path and simple betting strategies based on past prices appear to yield no profit opportunities; for example, Leigh, Wolfers and Zitzewitz (2003) demonstrate this point for the aforementioned Tradesports “Saddam Security”. However, there is also some evidence that this small-scale market responded to news about Iraq with a slight lag relative to deeper financial markets. Tetlock (2004) surveys a wide range of data from Tradesports, finding that their financial contracts are largely priced efficiently.

Prediction markets do seem to display some of the deviations from perfect rationality that appear in other financial markets. There is substantial evidence from psychology and economics suggesting that people tend to overvalue small probabilities and undervalue near certainties. For example, there is a well-known “favorite-longshot bias” in horse races (for example, discussed in this journal by Thaler and Ziemba, 1988), in which bettors tend to overvalue extreme longshots and thus receive much lower returns

for such bets, an effect which is offset by somewhat higher (albeit still negative) returns for betting on favorites. The “volatility smile” in options refers to a related pattern in financial markets (Bates, 1991; Rubenstein, 1994), which involves over-pricing of strongly out-of-the-money options, and under-pricing of strongly in-the-money options (relative to their future values or their ex-ante values from the Black-Scholes option pricing formula).⁵ These experiences suggest that it is likely that prediction markets will also perform poorly at predicting small probability events.

A somewhat vivid example of this kind of miscalibration comes from financial variables that trade on Tradesports. Table 4 reports the bid and ask prices in the prediction market for a contract that will pay \$100 if the Standard and Poor’s 500 index finishes 2003 in a certain range. For comparison, one can look at the actual prices of December S&P options traded on the Chicago Mercantile Exchange. We used the method discussed in Leigh, Wolfers, and Zitzewitz (2003) to translate the financial market prices into prices for a security comparable to the Tradesports contract. Comparing Tradesports prices with the actual option prices in Chicago suggests that the extremely unlikely (high and low) outcomes for the S&P 500 are relatively over-priced on Tradesports. In fact, the price differences implied a (small) arbitrage opportunity that persisted for most of summer 2003 and has reappeared in 2004. Similar patterns existed for Tradesports securities on other financial variables like crude oil, gold prices, and exchange rates. This finding is consistent with the longshot bias being more pronounced on smaller-scale exchanges.

⁵ Ait-Sahalia, Wang, and Yared (2001) argue that the conclusion of miscalibration is less clear cut in this context because these prices may be driven by small likelihoods of extreme price changes.

**Table 4: Price of Standard and Poor's Future Price Securities
Tradesports vs. Actual Prices from the Chicago Mercantile Exchange**

S&P level at end of 2003	Price on Tradesports		Estimated price from actual December S&P options
	Bid	Ask	
1200 and over	2	6	2.5
1100 to 1199	11	16	13.2
1000 to 1099	28	33	33.3
900 to 999	25	30	30.5
800 to 899	14	19	13
700 to 799	3	8	5
600 to 699	4	7	2
Under 600	5	8	1
S&P level on July 23, 2003	985		

Notes: Prices are market close, July 23, 2003. Prices given in first two columns are for a security that pays \$100 if S&P finishes 2003 in given range. Prices in third column are estimated from actual option settlement prices using the method in Leigh, Wolfers, and Zitzewitz (2003), adjusting for the 13 day difference in expiry date.

Another behavioral bias reflects the tendency of market participants to trade according to their desires, rather than their objective probability assessments. Strumpf (2004) provides evidence that certain New York gamblers are more likely to bet on the Yankees, while Forsythe, Reitz and Ross (1999) provide evidence that individual traders buy and sell in political markets in a manner correlated with their party identification. Even so, as long as marginal trades are motivated by profits rather than partisanship, prices will reflect the assessments of (unbiased) profit motive. Thus far there is little evidence that these factors yield systematic unexploited profits. (For instance, systematically betting either for or against the New York Yankees during the 1990s would have yielded losses in either case.)

A further possible limitation of prediction market pricing arises if speculative bubbles drive prices away from their likely outcomes. Traditional markets may be subject to bubbles because of constraints on short selling and because investors will be reluctant to commit a large share of their wealth to an arbitrage opportunity, since if the mispricing does exist, it may get worse before it gets better (Shleifer and Vishny, 1997). Since prediction markets typically impose no restrictions on short selling, and the markets are sufficiently small-scale that it is unlikely that informed investors will be

capital-constrained, the scope for bubbles might be more limited. It is almost impossible to make any serious attempt at describing the frequency of bubbles in the data we have so far. However, through September 2003 we suspected a bubble in the Tradesports security on whether Hillary Clinton would win the Democratic nomination. Our suspicions were based on her public statements that she was not a candidate, and the tenor of discussion among traders, which seemed to indicate that trading was being driven by expectations of future price movements rather than by fundamentals. Equally, these high prices may have reflected betting from those with access to campaign insiders who knew more about her state of mind than we did.

Empirically, the best that we can say is that the performance of past markets at predicting the future has been, on average, pretty good, whether or not specific markets were in some cases distorted by biases or bubbles. Laboratory experiments hold out the possibility of learning more about bubbles, as it is possible for the experimenter to know the “true price,” and hence to observe deviations. Plott and Sunder (1982 and 1988) have set up extremely stylized examples in which bubble-like behavior occurs in simple prediction markets. At the same time, bubbles in experimental markets often burst and give way to more rational pricing.

Can Event Markets Be Easily Manipulated?

The profit motive has usually proven sufficient to ensure that attempts at manipulating these markets were unsuccessful. There have been several known attempts at manipulation of these markets, but none of them had much of a discernible effect on prices, except during a short transition phase.⁶ For example, Wolfers and Leigh (2002) report candidates betting on themselves at long odds to create a “buzz”, while Strumpf (2004) placed random \$500 bets on the Iowa Electronic Markets to trace their effect. In a similar vein, Camerer (1998) attempted with little effect to manipulate betting on horse races by canceling large wagers at the last moment. Clearly the extent to which markets are manipulable depends – at least in part – on how thin the markets are.

⁶ Rhode and Strumpf document that attempts at manipulation in early twentieth century political markets were typically unsuccessful.

It was feared that the DARPA markets would create the opportunity for a terrorist to profit from an act of terrorism or an assassination. This concern may have been misplaced, both because the proposed markets were unlikely to have included terrorism or assassination contracts in the first place, and because the small scale of these markets means that terrorists would not have been able to earn much relative to the presumed going rate for an assassination. An alternative view holds that such trade is actually a good thing to the extent that trading ultimately reveals previously secret information about the intentions of terrorist groups. That said, if terrorists are sophisticated enough to place bets in futures markets, surely they can do so with standard futures contracts on oil prices, by selling short stock in insurance companies or the entire stock market, and the like. Indeed, rumors have circulated widely that there was unusual trading in options on United and American Airlines stock in the week prior to the attacks of September 11, 2001. A careful analysis by Poteshman (2004) found little evidence to support these rumors, suggesting that if terrorists did profit from their actions, they neither left a noticeable footprint nor needed a prediction market to do so.

3. Market Design

The success of prediction markets, like any market, can depend on their design and implementation. Some of the key design issues include how buyers are matched to sellers, the specification of the contract, whether real money is used, and whether a diversity of information exists in a way that provides a basis for trading. We consider these in turn.

In most prediction markets, the mechanism that matches buyers to sellers is a continuous double-auction, with buyers submitting bids and sellers submitting asking prices, and with the mechanism executing a trade whenever the two sides of the market reach a mutually agreeable price. However, the new prediction markets in announcements of economic statistics operate more like the parimutuel systems that are common in horse-race betting. In a parimutuel system, all of the money that is bet goes into a common pot and is then divided among the winners (after subtracting transaction costs). Many prediction markets are also augmented by market makers who announce willingness to buy and sell at a certain range of prices; similarly, most sports bets are

placed with bookmakers who post prices. Finally, while these mechanisms are relatively useful for simple markets, Hanson (2003) has proposed the use of market scoring rules to allow for simultaneous predictions over many combinations of outcomes. Instead of requiring separate markets for each combination of possible outcomes, traders effectively bet that the sum of their errors over all predictions will be lower.

For a prediction market to work well, contracts must be clear, easily understood, and easily adjudicated. For example, we don't see contracts like "Weapons of Mass Destruction are not in Iraq," but rather contracts specifying whether such weapons will have been found by a certain date. This requirement for clarity can sometimes turn out to be complex. In the 1994 U.S. Senate elections, the Iowa markets proposed what looked to be a well-specified market, with contracts paying according to the number of seats won by each party. The day after the election (and while votes were still being counted in some jurisdictions), Senator Richard Shelby (D-Alabama) switched sides to become a Republican. As another example, in the course of Ortner's (1998) internal prediction market on whether a software project would be delivered to the client on schedule, the client changed the deadline.

One intriguing question is how much difference it makes whether prediction markets are run with real money or with some form of play-money. Legal restrictions on gambling have led some groups like NewsFutures.com to adopt play-money exchanges, with those who amass the largest play-fortunes eligible for prizes. Prices on play and real-money exchanges are not linked by arbitrage: in August 2003, for example, George W. Bush was a 67 percent favorite to win reelection on real-money exchanges, but was a 50-50 bet on NewsFutures. However, we do not yet have sufficient comparative data to know the extent to which money makes predictions more accurate. Indeed, it has been argued that the play-money exchanges may even outperform real-money exchanges because "wealth" can only be accumulated through a history of accurate prediction. In a suggestive experiment, Servan-Schreiber, Wolfers, Pennock and Galebach (2004) compared the predictive power of the prices from real-money and play-money exchanges over the 2003 NFL football season, finding that both yielded predictions that were approximately equally accurate. Interestingly, both sets of prices also outperformed all but a dozen of 3,000 people in an online contest, and also easily

outperformed the average assessments of these “experts.” One practical advantage of play money contracts is that they offer more freedom to experiment with different kinds of contracts. On play-money exchanges, such as Foresight Exchange, one often sees quite loosely worded “contracts” such as that a “scientific study will conclude that astrology is a statistically significant predictive method to describe an individual’s personality traits.”

Even well-designed markets will fail unless a motivation to trade exists.⁷ Most prediction markets are not large enough to allow hedging against specific risks. However, the play-money exchanges and sports gambling industry both suggest that it may be possible to motivate (small-scale) trading simply through the thrill of pitting one’s judgment against others, and being able to win a monetary prize may sharpen this motivation. Trade also requires some disagreement about likely outcomes. Disagreement is unlikely among fully rational traders with common priors. It is more likely to occur when traders are overconfident in the quality of their private information or in their ability to process public information or when they have priors that are sufficiently different to allow them to agree to disagree.

These insights suggest that some prediction markets will work better when they concern events that are widely discussed, since trading on such events will have higher entertainment value and there will be more information on whose interpretation traders can disagree. Ambiguous public information may be better in motivating trade than private information, especially if the private information is concentrated, since a cadre of highly informed traders can easily drive out the partly informed, repressing trade to the point that the market barely exists. Indeed, attempts to set up markets on topics where there are insiders with substantial information advantages have typically failed. For instance, the Tradesports contracts on the next Supreme Court retirement or the future of the papacy have generated very little trade despite the inherent interest in these questions. Trade can also be subsidized either directly, or indirectly by adding noise trades into the market, which provides the potential to profit from trading.

Finally, the power of prediction markets derives from the fact that they provide incentives for *truthful revelation*, they provide incentives for research and *information*

⁷ The inflation futures market on the Coffee, Sugar, and Cocoa Exchange is a case in point; this market generated little volume, ultimately failing.

discovery, and the market provides an algorithm for *aggregating opinions*. As such, these markets are unlikely to perform well when there is little useful intelligence to aggregate, or when public information is selective, inaccurate, or misleading. Further, the weights that markets give to different opinions may not be an improvement on alternative algorithms where the accuracy of pundits is directly observable. For example, the public information on the probability of weapons of mass destruction in Iraq appears to have been of dubious quality, so it is perhaps unsurprising that both the markets were as susceptible as general public opinion to being misled.

4. Making Inferences from Prediction Markets

How might economists use the results from prediction markets in subsequent analysis? The most direct form of inference involves using these predictions directly. For instance, in their experiments at Hewlett-Packard, Chen and Plott (2002) elicited expectations of future printer sales, which were of direct interest for internal planning purposes.

Some analyses have tried to link the time series of expectations elicited in prediction markets with time series of other variables. For instance, in Leigh, Wolfers, and Zitzewitz (2003), we interpreted movements in the Saddam Security as an index for the risk of war, and interpreted the comovement with the oil price shown in Figure 2 as a causal relationship, concluding that war led to a \$10 per barrel increase in oil prices. A similar analysis suggested that equity prices had built in a 15 percent war discount. Applying a similar methodology, Slemrod and Greimel (1999) linked the price of a Steve Forbes security in the 1996 Republican primary market with a rising interest rate premium on municipal bond prices because Forbes's signature issue was a "flat tax" that would have eliminated the tax exemption for municipal bond interest. As with any regression context, one must be cautious before inferring that these correlations reflect causation, and must consider the issues of reverse causation, omitted variables, statistical significance, functional form, and the like.

It seems quite possible to design prediction market contracts so that they would bring out the connection between an event and other variables. For instance, in 2002 we

could have floated two securities, one paying $\$P$ if Saddam were ousted in a year (where P is the future oil price), with the purchase price refunded otherwise, and another that paid $\$P$ if Saddam remains in power, again refunding the purchase price. The difference in the equilibrium price of these two securities can be interpreted as the market's expectation of the effect of ousting Saddam on oil prices. This inference does not require researchers to wait until sufficient variation in the political situation has accrued for a regression to be estimated. Moreover, changes in the market's beliefs about how ousting Saddam would affect oil prices can be directly measured through such a conditional market.

Very few of these contingent markets have been constructed, though this year's Iowa Electronic Market on the 2004 presidential election is instructive. Table 5 shows the prices of a series of contracts which are standard index contracts that pay a penny for each percentage of the two-party popular vote won by each party, but are contingent in that the contract pays out only if the democratic nominee is also successfully predicted. These contracts pay nothing if the nominee is not correctly predicted.

**Table 5: Contingent Markets: 2004 Presidential Election
Contracts pay according to Vote Share, Conditional on the Democratic Nominee**

Contract pays conditional on specific Democratic Candidate	Democratic Candidate Vote Share (Contract price, \$)	Republican Vote Share against this Candidate (Contract price, \$)	Implied Prob. this Candidate wins the Nomination	Expected Share of Popular Vote if Nominated
	<i>A</i>	<i>B</i>	$C=A+B$	$D=A/C$
John Kerry	\$0.344	\$0.342	68.6%	50.1%
John Edwards	\$0.082	\$0.066	14.8%	55.4%
Howard Dean	\$0.040	\$0.047	8.7%	46.0%
Wesley Clark	\$0.021	\$0.025	4.6%	45.7%
Other Dems	\$0.015	\$0.017	3.2%	46.9%

Notes: Columns A and B show the prices of contracts that pay a penny for each percentage of the two-party popular vote won by Democrats or Republicans respectively, conditional on picking the winner of the Democratic nomination. (Contracts pay \$0 if the selected candidate does not win the Democratic nomination.)

Source: Closing prices as of January 29, 2004, Iowa electronic markets.

Because the Democratic and Republican shares of the two-party vote must sum to one, a portfolio containing contracts tied to both the Democratic and Republican vote

shares, but conditional on Kerry winning the nomination, will definitely pay \$1 if Kerry wins the primary, and \$0 otherwise. Implicitly then, this market embeds a winner-takes-all market on the Democratic primary race, and adding the prices shown in columns A and B yields the prices of these synthetic securities that represent the probability that any specific candidate wins the Democratic nomination (shown in column C). The final column calculates the implied expected vote share for each candidate, if that candidate were to win the nomination, by deflating the cost of the Democratic vote share contract conditional on that candidate by the probability of that candidate actually winning the nomination. Hanson (1999) has called these contingent markets “decision markets”, arguing that these expectations should be used to guide decision-making. As such, delegates to the Democratic convention interested in selecting the strongest candidate would simply compare the ratios in the final column and accordingly vote for John Edwards. Berg and Reitz (2003) make a related argument using data from the 1996 Republican nomination race.

While we are optimistic that these data on contingent prediction markets can be used to inform decision-making, some care is required. In making statements about the comovement of two variables, social scientists have long struggled to distinguish correlation from causation, and these decision markets do not resolve this issue.

One could imagine that traders hold a frequentist view of probability and that they price the securities in Table 5 by simply inventing hundreds of possible scenarios, and prices simply reflect average outcomes across these scenarios. An econometrician running regressions based on these hundreds of scenarios would note a robust correlation between Edwards winning the nomination and the Democrats winning the presidency. But a careful econometrician would be reluctant to infer causation, noting that there are important “selection effects” at play, as the scenarios in which Edwards wins the nomination are not random. For example, the markets may believe that Edwards will not win the nomination unless Southern Democrats become energized, but if this does happen, it is likely that Edwards will win both the nomination and the presidency. Alternatively, with Kerry viewed as the likely nominee, Edwards may be perceived as a possible nominee only if he shows himself to be a politician of extraordinary ability, overcoming Kerry’s early lead in the delegate count. If so, it also seems likely that a

candidate of such extraordinary ability would win the general election. Or Edwards might be perceived as thin-skinned, and likely to drop out of the race if it appears that the Democratic party is unlikely to win the White House. As such, the relatively high price of the Edwards-Democratic security may reflect either something about Edwards' ability or the selection effects that lead him to win the nomination.

Just as econometricians often deal with selection effects by adding another equation that explicitly models the selection process, there is a prediction market analogy - floating another contract that prices the variables driving the selection of Democratic candidates. For example, adding a contract that pays off if a candidate drops out of the nomination race early would allow an assessment of the extent to which prices of contingent contracts are being driven by that specific selection mechanism, thereby yielding a more accurate indication of candidate ability. But since many key traits of candidates may be unobservable or difficult to capture in a contract that would attract trading, it may be impossible to rely fully on contingent markets to guide voters to the candidate with the greatest vote-winning potential.⁸

These relatively simple contingent markets, as well as more complex combinatorial markets, are as yet virtually untested and remain a useful focus for further research. There may be important and interesting applications in domains where selection problems are minimal.

5. Innovative Future Applications?

Prediction markets are extremely useful for estimating the market's expectation of certain moments. Simple market designs can elicit expected means or probabilities, more complex markets can elicit variances, and contingent markets can be used to elicit the market's expectations of covariances and correlations, though as with any estimation context, further identifying assumptions are required before a causal interpretation can be made. The research agenda on these markets has reflected an interplay between theory, experiments, and field research, drawing on scholars from economics, finance, political

⁸ Furthermore, it is worth noting that the incentives to manipulate a contract rise with its use in decision making, and the apparent failure of past manipulation attempts mentioned above do not guarantee that it would fail in this context.

science, psychology, and computer science. This research program has established that prediction markets provide three important roles: 1) incentives to seek information; 2) incentives for truthful information revelation; and 3) an algorithm for aggregating diverse opinions. Current research is only starting to disentangle the extent to which the remarkable predictive power of markets derives from each of these forces.

Prediction markets doubtlessly have their limitations, but they may be useful as a supplement to the other relatively primitive mechanisms for predicting the future like opinion surveys, politically appointed panels of experts, hiring consultants, or holding committee meetings. We are already seeing increasing interest in these markets in the private sector, with the experiments at Hewlett-Packard now being supplemented with new markets on pharmaceuticals and the likely success of future technologies on NewsFutures.

DARPA's ill-fated attempt at establishing a Policy Analysis Market ultimately failed. However, it seems likely that private-sector firms will continue to innovate and to create new prediction markets, so policymakers will still be able to turn to prediction markets run by firms like Tradesports, Net Exchange, Incentive Markets and NewsFutures. It may be a sensible political outcome to have these event markets run by publicly-regulated, private-sector firms. Nonetheless, to the extent that the valuable information generated by trade in these markets is not fully internalized into the profits earned by these private firms, prediction markets may be underprovided.

6. References

- Aït-Sahalia, Yacine, Yubo Wang, and Francis Yared (2001), "Do Options Markets Correctly Price the Probabilities of Movement of the Underlying Asset?" *Journal of Econometrics*, 102, 67-110.
- Athanasoulis, Stefano, Robert Shiller, and Eric van Wincoop (1999), "Macro markets and financial security," *Economic Policy Review*, 21-39.
- Bates, David (1991) "The Crash of '87: Was It Expected? The Evidence from Options Markets," *Journal of Finance* 46(3), 1009-44.
- Berg, Joyce and Thomas Rietz (2003), "Prediction Markets as Decision Support Systems", *Information Systems Frontiers*, 5(1).
- Berg, Joyce, Robert Forsythe, Forrest Nelson and Thomas Rietz (2001) "Results from a Dozen Years of Election Futures Markets Research," forthcoming in Charles Plott and Vernon Smith (Eds) *Handbook of Experimental Economic Results*.
- Chen, Kay-Yut and Charles Plott (2002), "Information Aggregation Mechanisms: Concept, Design and Field Implementation for a Sales Forecasting Problem."
- Camerer, Colin (1998), "Can Asset Markets be Manipulated? A Field Experiment with Racetrack Betting," *Journal of Political Economy* 106(3).
- Forsythe, Robert, Thomas Rietz and Thomas Ross (1999), "Wishes, Expectations and Actions: Price Formation in Election Stock Markets," *Journal of Economic Behavior and Organization*, 39, 83-110.
- Hanson, Robin (1999) "Decision Markets," *IEEE Intelligent Systems*, 14(3), 16-19.
- Hanson, Robin (2003) "Combinatorial Information Market Design," *Information Systems Frontiers*, 5(1), 105-119.
- Heckman, James J. (1979) "Sample Selection Bias as a Specification Error," *Econometrica* 47(1), 153-61.
- Leigh, Andrew, Justin Wolfers and Eric Zitzewitz (2003), "What do Financial Markets Think of War in Iraq?," *NBER Working Paper 9587*.
- Leigh, Andrew and Justin Wolfers (2002), "Three Tools for Forecasting Federal Elections: Lessons from 2001", *Australian Journal of Political Science* 37(2).
- Looney, Robert (2003), "DARPA's Policy Analysis Market for Intelligence: Outside the Box or Off the Wall", *Strategic Insights*, Vol. II Issue 9, September 2003.

- Ortner, Gerhard (1998), "Forecasting Markets – An Industrial Application", *mimeo*, Technical University of Vienna.
- Pennock, David, Steve Lawrence, C. Lee Giles, and Finn Arup Nielsen (2001), "The Real Power of Artificial Markets," *Science*, 291, 987-988.
- Plott, Charles and Shyam Sunder, (1982), "Efficiency of experimental security markets with insider information: An application of rational-expectations models," *Journal of Political Economy*, 90(4), 663-98.
- Plott, Charles and Shyam Sunder, (1988), "Rational Expectations and The Aggregation of Diverse information in Laboratory Security Markets," *Econometrica*, 56, 1085-1118.
- Poteshman, Allen (2004), "Unusual Option Market Activity and the Terrorist Attacks of September 11, 2001", *mimeo*, University of Illinois at Urbana-Champaign.
- Rubenstein, Mark (1994), "Implied Binomial Trees," *Journal of Finance*, 49, 771-818.
- Rhode, Paul and Koleman Strumpf (2004), "Historical Presidential Betting Markets", *Journal of Economic Perspectives*, Spring 2004.
- Servan-Schreiber, Emile, Justin Wolfers, David Pennock and Brian Galebach (2004), "Prediction Markets: Does Money Matter?", *Electronic Markets* 14:3, Fall 2004.
- Shiller, Robert (2003), *The New Financial Order: Risk in the Twenty-first Century*, Princeton University Press: Princeton, NJ.
- Shleifer, Andrei and Robert Vishny (1997), "The Limits of Arbitrage," *Journal of Finance* 52(1), 35-55.
- Slemrod, Joel and Timothy Greimel (1999), "Did Steve Forbes Scare the Municipal Bond Market?", *Journal of Public Economics*, 74(1), 81-96.
- Spann, Martin and Bernd Skiera (2003), "Internet-Based Virtual Stock Markets for Business Forecasting", *Management Science*, 49, 1310-1326.
- Strumpf, Koleman (2004), "Manipulating the Iowa Political Stock Market", *mimeo*, University of North Carolina.
- Tetlock, Paul (2004), "How Efficient are Information Markets? Evidence from an Online Exchange", *mimeo*, Harvard University.
- Thaler, Richard H., and William T. Ziemba. 1988. "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries," *Journal of Economic Perspectives*, Vol. 2, No. 2. Spring, pp. 161-174.