

**Using Prediction Markets to Solve Complex Problems**  
– **An Application to the ‘Climate Bet’**

Andreas Graefe  
Institute for Technology Assessment and Systems Analysis  
Forschungszentrum Karlsruhe, Germany  
andreas.graefe@kit.edu

J. Scott Armstrong  
The Wharton School, University of Pennsylvania,  
Philadelphia, PA, USA  
armstrong@wharton.upenn.edu

February 16, 2009

The track record of prediction markets is good. For various types of problems, they have been shown to perform as least as well as alternative forecasting methods. However, the nature of most of the problems for which prediction markets have been used is simple: they require aggregating information and its outcome does not relate to complex decisions being made. Examples include forecasting elections, sports events, or sales figures.

In aggregating dispersed information from a – theoretically unlimited – number of people, prediction markets have the potential to provide social utility. They can contribute to participatory regulation by incorporating the views of a broad public on problems that ask for policy interventions. Such problems are complex as they not only require aggregating information or facts. They also involve people’s values, attitudes, emotions, expectations, fears, commitments, etc. For such problems, achieving consensus becomes highly difficult.

Using the example of the ‘Climate Bet’, we analyze the social utility of prediction markets for complex problems.

**Prediction Markets**

Having already been popular in the late 19<sup>th</sup> century for forecasting election outcomes (Rhode & Strumpf 2004), the launch of the Iowa Electronic Markets (IEM) in 1988, accompanied by

the rise of the internet, finally leveraged prediction markets. Currently, prediction markets are gaining attention in various fields of forecasting. In such markets, participants reveal their estimates by trading stocks whose prices reflect the aggregated group estimate on a particular topic. Based on their individual performance, traders can win money. If a trader thinks the current group estimate is too low (high), he will buy (sell) stocks. Thus, through the prospect of gaining money, traders have an incentive to become active in the group process whenever they expect the group estimate to be inaccurate. For more information about how prediction markets work as well as recent research findings see the Special Interest Group on Prediction Markets at <http://marketsforforecasting.com>.

Since their emergence, numerous studies have been published that demonstrate high forecasting accuracy of prediction markets in the field and in the laboratory. For example, in analyzing 964 polls for the five presidential elections from 1988 to 2008, Berg et al. (2008a) found that the respective market forecasts were closer to the actual election results 74% of the time. This performance was replicated for the 2008 election (Berg et al. 2008b), although Erikson and Wlezien (2008) have shown that this advantage disappeared when comparing the market forecasts to damped polls. Spann and Skiera (2009) analyzed the accuracy of prediction markets for sports forecasting. They found that prediction markets performed equally well to betting odds and were clearly more accurate than experts (tipsters). Others conducted small-scale studies to analyze prediction markets for the field of business forecasting. Chen and Plott (2002) showed that prediction markets were more accurate than internal forecasting mechanisms at Hewlett-Packard for forecasting sales. Spann and Skiera (2003) found that prediction markets appeared to be more accurate than four extrapolations mechanisms for forecasting the use of different services at a German cell phone operator. In conducting laboratory experiments, Graefe and Armstrong (2008) found prediction markets to perform equally well to meetings, nominal groups and Delphi on a quantitative judgment task.

However, all studies available to date refer to rather simple problems that require only aggregation of information or facts. Little is known about prediction markets' performance on more complicated tasks that also involve people's values, attitudes, emotions, expectations, fears, commitments, etc. In being a structured forecasting approach and being able to involve a large number of participants, prediction markets appear to be valuable for dealing with suchlike problems. We use the current discussion on global warming to study the potential of prediction markets.

## **The Climate Bet**

Al Gore has claimed that there are scientific forecasts that the earth will become warmer and that this will occur rapidly. Yet, searches of his book or on the Internet did not reveal any quantitative forecasts or any methodology he relies on. As a result, in June 2007, Scott Armstrong offered Al Gore a bet of \$10,000 on who could best forecast annual mean temperatures over the next ten years. Al Gore declined the bet, citing the reason that he does not bet money.

The general objective of the challenge was to promote the proper use of science in formulating public policy. This involves such things as full disclosure of forecasting methods and data, and the proper testing of alternative methods. A specific objective was to develop useful methods to forecast global temperatures. In particular, it was hoped that other competitors would join to show the value of their forecasting methods.

Al Gore was invited to select any currently available fully disclosed climate model to produce the forecasts (without human adjustments to the model's forecasts). Scott Armstrong's forecasts would have been based on the naive (no-change) model; that is, for each of the ten years of the challenge, he will use the most recent year's global mean temperature as the forecast for each of the years in the future. The naïve model is a commonly used benchmark in assessing forecasting methods. It is a strong competitor when uncertainty is high or when improper forecasting methods have been used.

Specifically, the challenge would have involved making forecasts for ten weather stations that are reliable and geographically dispersed. An independent panel composed of experts agreeable to both parties would have designated the weather stations. Data from these sites would have been listed on public web sites along with daily temperature readings and, when available, error scores for each contestant.

The criteria for accuracy would have been the average absolute forecast error at each weather station. Averages across stations would have been calculated for each forecast horizon (e.g., for a six-year ahead forecast). Finally, simple unweighted averages would have been made of the forecast errors across all forecast horizons. For example, the average across the two-year ahead forecast errors would have received the same weight as that across the nine-year-ahead

forecast errors. This unweighted average would have been used as the criterion for determining the winner.

It was aimed at starting the bet at the beginning of 2008. The full story can be reviewed at the <http://climatebet.com>.

### **Who Would Have Won? – Solving the Climate Bet with Prediction Markets**

The question about who would win this bet is important as it can promote the proper use of science for informing public policy decisions. Thus, it will be assumed that Al Gore and Scott Armstrong had made a gentleman's bet involving no money and that the ten years of the bet had started as scheduled on January 1, 2008.

To answer this question, a prediction market has been launched at <http://hubdub.com> on January 28, 2009. We have chosen hubdub.com as it is an independent and neutral prediction market platform that is open for everyone to participate. It offers numerous bets on different topics, some of which are also related to global warming. All questions have been created by hubdub.com users. In using play-money instead of real-money, it does not face concerns associated with gambling. Furthermore, the use of play-money should help to attract a larger number of participants as risk-averse people might not be willing to invest real-money.

The terms of the original bet were altered for simplicity, and transparency. Thus, the criterion was based on the UAH satellite measures on global mean temperature. An alternative would be to use the Hadley series. In addition, the bet was simplified to use only a single starting point; that is, January 1, 2008.

The contract specifications of the bet are shown in Figure 1. A screenshot of the prediction market page, accessible at <http://tinyurl.com/gore-armstrong-bet>, is provided in Appendix 1. The page explains the contract specifications and shows the current market forecast as well as comments provided by users. In addition, an automatically generated list of news related to the topic is shown in the right column of the page.

Figure 1: The Climate Bet Prediction Market at hubdub.com – Contract Specifications

### Who will win the "Climate Bet" – Al Gore or Wharton Professor Scott Armstrong?

**Background:** In June 2007, Wharton Professor Scott Armstrong offered Al Gore a bet of \$10,000 on who could best predict global mean temperature over the next ten years. Al Gore declined the bet, citing the reason that he does not bet money (the full story can be reviewed at <http://theclimatebet.com>).

Now, assume that Armstrong and Gore had made a gentleman's bet (no money) and that the ten years of the bet started on January 1, 2008.

- Armstrong's forecast was that there would be no change in global mean temperature over the next ten years.
- Gore did not specify a method or a forecast. Nor did searches of his book or on the Internet reveal any quantitative forecasts or any methodology he relies on. He did, however, imply that the global mean temperature would increase at a rapid rate - presumably at least as great as the IPCC's 1992 projection of 0.03°C-per-year; thus. The IPCC's 1992 projection is to be taken as Gore's forecast.

**Settlement date:** January 1, 2018

**Settlement details:** The criterion will be the mean absolute errors of Armstrong's and Gore's annual forecasts for the ten year period, with the errors to be measured against the UAH global temperature record (<http://vortex.nsstc.uah.edu>). The win goes to the smallest mean absolute error.

Figure 2 shows how market participants predicted the outcome of the bet in mid-February. Scott Armstrong was expected to win with a probability of 72% (vs. Al Gore 28%). Now, if one thinks Gore's chances are higher, one would buy the contract – and the price will go up. If one thinks they are lower, one would sell – and the price would go down. Thus, through the process of trading, one reveals information to the market for which one expects to win (play) money. In other words, one has an incentive to become active whenever one thinks the current forecast is wrong.

Figure 2: Forecast of the Climate Bet Market at hubdub.com



After signing up at hubdub.com, each participant receives an initial amount of H\$1000 Hubdub play-money dollars, which can be used to trade on every question available at the

market platform. Participants win if they are right and lose what they staked if they are wrong. The more Hubdub dollars they have, the higher their ranking compared to other participants.

### **Early Results**

#### *Results of the first year of the climate bet*

- Based on the Hadley measures of land temperatures<sup>1</sup>, the actual change for 2008 was  $-.08^{\circ}\text{C}$ . Thus, in absolute terms Armstrong missed by  $.08^{\circ}\text{C}$ , while Gore missed by  $.11^{\circ}\text{C}$ . Advantage Armstrong by  $0.03^{\circ}\text{C}$ .
- Based on the UAH data<sup>2</sup>, the actual change for 2008 was  $0.05$ . Thus, Armstrong's error was  $0.05$  where as Gore's was  $0.02$ . Advantage Gore by  $0.03^{\circ}\text{C}$ .

Given both measures, the first year could be interpreted as a draw. However, since the bet will be judged based on the UAH data, Gore won the first year.

#### *Forecast from Climate Bet Prediction Market*

At the launch of the market, the chances for both Al Gore and Scott Armstrong to win the bet were set to 50%. With the first trade, Scott Armstrong was predicted as the winner. At the time of writing, the market predicted a chance of 73% for Armstrong to win, involving 85 predictions. In seeing Armstrong's chances almost consistently above 60%, this forecast has remained relatively stable since the start of the market.

### **Discussion**

The results suggest that involving prediction markets for solving questions like the Climate Bet can be valuable for informing policy decisions. Participants expect Armstrong to win the bet, a result that conforms to the historical benchmark analysis. Given the way how the issue of global warming is discussed in the media and motivates policy decisions, this result seems surprising. Based on the early forecast, the public's perceptions of this issue appear to differ. Nonetheless, concerns might be raised on using prediction markets for suchlike problems.

#### *Who are the participants?*

Since the Hubdub market is open for everyone, it is unclear who is participating. Most likely, participants from either camp – sceptics as well as proponents of global warming – might

---

<sup>1</sup> HadCRUT3: <http://hadobs.metoffice.com/hadcrut3/>

<sup>2</sup> <http://vortex.nsstc.uah.edu/>

populate the market. However, the possibility that one camp might dominate can not be excluded, particularly at this early stage where the market involves only a small number of predictions. The market forecasts might gain reliability with an increasing number of participants.

However, people might have concerns with involving unskilled amateurs in the decision-making process by using prediction markets. In general, these concerns appear to be unwarranted:

- Experts do not have value in forecasting change, particularly in situations involving high uncertainty. This has been shown for forecasting future political and economic events (Tetlock 2006) as well as conflict situations (Green & Armstrong 2007). Although expertise beyond a minimum level was found to lead to more accurate forecasts, additional expertise did not improve accuracy. In fact, there was some evidence that accuracy might even decrease with increasing expertise.
- In executing uninformed trades so called ‘noise traders’ provide additional liquidity. According to assumptions of rational models of liquidity provision, this enhances incentives for other participants to get involved and informed – and to reveal their information through trading. In response, it is often assumed that market forecasts might become even more accurate. This is supported by earlier research, which has shown that prediction markets provide accurate forecasts in spite of biased traders (Forsythe et al. 1999) or even intentional attempts to manipulate market prices (see below).

#### *Can the market be manipulated?*

The question about who is participating inevitably raises the question of manipulation. Manipulation is a common and often raised concern when using prediction markets and has been cited as one of the reasons for the dismissal of the policy analysis market in 2003 (see Textbox 1). However, most empirical studies to date showed that attacks on result accuracy have not been successful historically (Rhode & Strumpf 2004), in the laboratory (Hansen et al. 2006), or in the field (Camerer 1998). Only one study reports successful manipulation of prices at the IEM (Hansen et al. 2004). In reviewing studies of price manipulation, Wolfers

and Zitzewitz (2004) concluded that, besides a short transition phase, none of the known attacks had a noticeable influence on the prices.

**Textbox 1: The DARPA policy analysis market (PAM)**

From 2001 to 2003, the Defense Advanced Research Project Agency (DARPA) of the U.S. government sponsored the FutureMAP project, also known as the Policy Analysis Market (PAM). The original goal of this project was to improve existing intelligence institutions by predicting military and political instability around the world, how the U.S. would affect such instabilities, and vice versa.

Later, the focus was narrowed to predict five parameters for each of eight nations in the Middle East: military activity, political instability, economic growth, U.S. military activity, and U.S. financial involvement. In addition, traders should predict additional parameters like U.S. GDP growth, world trade, or total U.S. military casualties.

On July 28, 2003, shortly before the scheduled start of PAM on September 1, two Democratic Senators held a press conference accusing the U.S. Department of Defense to plan a “terror market” for people to bet on terrorist events. Instantly, the topic caught the interest of the media. During the next two days, 128 media articles were published, most of them casting a damning light on PAM. Not surprisingly, PAM was cancelled immediately.

Later, Hanson (2007), who was involved in the project, conducted a statistical news analysis on more than 600 media articles that mentioned PAM. He found that more informed articles favored PAM. Yet, the political decision to dismiss PAM was made and it is unlikely that it will be reversed anytime soon. For a review of the origin and development of the project, see Hanson (2007).

*Can play-money work?*

People might have concerns that play-money markets are less accurate. Since participants do not stake real money, they might not be willing to reveal their true information. Thus far, two studies analyzed the relative performance of play-money and real-money markets. For sports events, Servan-Schreiber et al. (2004) could not identify differences in accuracy, while Rosenbloom and Notz (2006) found real-money markets to be more accurate for non-sports events. Thus, the answer to this question remains open.

Nonetheless, to address the current problem of the Climate Bet, play-money markets appear to be feasible. In using play-money, one can attract a large number of participants that might not be able – or willing – to invest real-money. Furthermore, people might not be willing to invest real-money in a long-term bet but rather look for short-term and more lucrative investments. In addition, using real-money for complicated problems like global warming might even be counter-productive. In particular, financially well-equipped groups that have a strong interest in the market outcome might be able to manipulate the results.

*Isn't the time frame to long?*

Prediction markets are usually used to forecast events in the near future in order to maintain participants' motivation. Yet a trader's interest does not necessarily decrease because of long time horizons and prediction markets have shown to perform well for long-term forecasting. The play-money market Foresight Exchange (FX)<sup>3</sup> has been active since 1994 and has managed to establish a lively community of traders involved in predicting events decades in the future. In analyzing results from the FX, Pennock et al. (2001) found that prediction market prices strongly correlated with outcome frequencies.

### **Future Work**

Future work will aim at addressing some of the concerns raised in the discussion. This involves marketing the Climate Bet prediction market on various websites and blogs. The goal is to increase the number of participants in order to obtain more reliable results. Furthermore, we observe possibilities to launch a real-money market in addition to the current play-money market. Finally, we plan on investigating the influence of providing additional information to participants. Currently, no information is given in addition to the contract specifications. It will be interesting to observe how further information (like the results of the first year of the bet or of the historical benchmark analysis) will influence the market predictions.

### **Conclusion**

We launched a real-money prediction market at hubdub.com to address the question of who will win the Climate Bet between Al Gore and Scott Armstrong. The goal of this project is to analyze the social utility of prediction markets for solving complex problems. Early results show that the market predicts Scott Armstrong to win the bet – a result that seems surprising given the perception of global warming that is communicated by the media.

### **References**

- Berg, J., Nelson, F. & Rietz, T. A. (2008a). Prediction Market Accuracy in the Long Run, *International Journal of Forecasting*, 24, 285-300.  
Berg, J., Nelson, F. D., Neumann, G. R. & Rietz, T. (2008b). *Was There Any Surprise About Obama's Election?*, Available at [http://www.forecastingprinciples.com/PM/images/articles/berg\\_obama\\_surprise.pdf](http://www.forecastingprinciples.com/PM/images/articles/berg_obama_surprise.pdf).

---

<sup>3</sup> <http://www.ideosphere.com>

- Camerer, C. (1998). Can Asset Markets be Manipulated? A Field Experiment with Racetrack Betting, *Journal of Political Economy*, 106, 457-482.
- Chen, K.-Y. & Plott, C. R. (2002). *Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem*. In: Social Science Working Paper No.1131, California Institute of Technology, Pasadena. Available at <http://www.hss.caltech.edu/SSPapers/wp1131.pdf>.
- Erikson, R. S. & Wlezien, C. (2008). Are Political Markets Really Superior to Polls as Election Predictors?, *Public Opinion Quarterly*, forthcoming.
- Forsythe, R., Rietz, T. A. & Ross, T. W. (1999). Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets *Journal of Economic Behavior & Organization*, 39, 83-110.
- Graefe, A. & Armstrong, J. S. (2008). *Comparing Face-to-face Meetings, Nominal Groups, Delphi and Prediction Markets on an Estimation Task*, Working paper. Available at [http://www.forecastingprinciples.com/PM/images/articles/graefe\\_armstrong\\_method\\_comparison.pdf](http://www.forecastingprinciples.com/PM/images/articles/graefe_armstrong_method_comparison.pdf).
- Green, K. C. & Armstrong, J. S. (2007). The Ombudsman: Value of Expertise for Forecasting Decisions in Conflicts, *Interfaces*, 37, 287-299.
- Hansen, J., Schmidt, C. & Strobel, M. (2004). Manipulation in Political Stock Markets - Preconditions and Evidence, *Applied Economics Letters*, 11, 459-463.
- Hanson, R. (2007). The Policy Analysis Market - A Thwarted Experiment in the Use of Prediction Markets for Public Policy, *Innovations: Technology, Governance, Globalization (MIT Press)*, 2, 73-88.
- Hanson, R., Oprea, R. & Porter, D. (2006). Information Aggregation and Manipulation in an Experimental Market, *Journal of Economic Behavior & Organization*, 60, 449-459.
- Pennock, D. M., Giles, C. L. & Nielsen, F. A. (2001). The real power of artificial markets, *Science*, 291, 987-988.
- Rhode, P. W. & Strumpf, K. S. (2004). Historical Presidential Betting Markets, *Journal of Economic Perspectives*, 18, 127-141.
- Rosenbloom, E. S. & Notz, W. (2006). Statistical Tests of Real-Money versus Play-Money Prediction Markets, *Electronic Markets*, 16, 63-69.
- Servan-Schreiber, E., Wolfers, J., Pennock, D. M. & Galebach, B. (2004). Prediction Markets: Does Money Matter?, *Electronic Markets*, 14, 243 - 251.
- Spann, M. & Skiera, B. (2003). Internet-based Virtual Stock Markets for Business Forecasting, *Management Science*, 49, 1310-1326.
- Spann, M. & Skiera, B. (2009). Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters, *Journal of Forecasting*, 28, 55-72.
- Tetlock, P. E. (2006). *Expert Political Judgment: How good is it? How can we know?*, Princeton University Press.
- Wolfers, J. & Zitzewitz, E. (2004). Prediction Markets, *Journal of Economic Perspectives*, 18, 107-126.

# Appendix

## Appendix 1: The Climate Bet Prediction Market at hubdub.com – Screenshot

The screenshot shows the hubdub.com website interface. At the top, the logo 'hubdub.com' is displayed with the tagline 'Predict the News'. Navigation links include 'Home', 'Leaderboards', 'Forums', and 'My Hubdub'. A search bar and a 'Questions' dropdown menu are visible. The user's profile 'agraefe' is shown with a net worth of H\$1,220, cash of H\$380, and predictions of H\$840. The main content area features a question titled 'Who will win the "Climate Bet" – Al Gore or Wharton Professor Scott Armstrong?' with a 'Forecast history %' chart showing J. Scott Armstrong at 72% and Al Gore at 28%. The question suspends in 8 years. A 'Background' section explains the bet's history, and a 'Settlement details' section outlines the criteria. Social sharing options for Digg, Delicious, Reddit, and Facebook are provided. A 'Predictions (78)' section is partially visible at the bottom.

**hubdub.com** Predict the News

agraefe | My account | Invite friends | Logout | About | Help | Edition: UK | US | All

Search [ ] Questions [v] Go

Home | Leaderboards | Forums | My Hubdub | New question

Politics | Sport | Entertainment | World | Business | Technology | Science | General

All questions » Science » Environment Created Thu 29th Jan 3:25pm EST by agraefe

Copy question as new

### Who will win the "Climate Bet" – Al Gore or Wharton Professor Scott Armstrong?

Forecast history %

Participant	Percentage
J. Scott Armstrong	72%
Al Gore	28%

Question suspends in 8 years

Suspend date: Mon 1st Jan 2018 2:59am EST (8 years to go) details

Activity: H\$20,825 | Predictions: 78 | Comments: 5

**Background:** In June 2007, Wharton Professor Scott Armstrong offered Al Gore a bet of \$10,000 on who could best predict global mean temperature over the next ten years. Al Gore declined the bet, citing the reason that he does not bet money (the full story can be reviewed at <http://theclimatbet.com>).

Now, assume that Armstrong and Gore had made a gentleman's bet (no money) and that the ten years of the bet started on January 1, 2008.

- Armstrong's forecast was that there would be no change in global mean temperature over the next ten years.
- Gore did not specify a method or a forecast. Nor did searches of his book or on the Internet reveal any quantitative forecasts or any methodology he relies on. He did, however, imply that the global mean temperature would increase at a rapid rate - presumably at least as great as the IPCC's 1992 projection of 0.03°C-per-year; thus. The IPCC's 1992 projection is to be taken as Gore's forecast.

**Settlement details:** The criterion will be the mean absolute errors of Armstrong's and Gore's annual forecasts for the ten year period, with the errors to be measured against the UAH global temperature record (<http://vortex.nsstc.uah.edu>). The win goes to the smallest mean absolute error.

Request settlement | Flag

Post on Digg | Delicious | Reddit | Facebook

Get widget! | Challenge friends

**Predictions (78)**

#### Related News

- Australian bushfires rage - Premium content Nature 16 hours ago
- The Christopher Booker prize: our man in Michigan Guardian Unlimited 2 days ago
- Latest contender for Christopher Booker prize Guardian Unlimited 2 days ago
- Bank tips almost vertical recovery Glasgow Herald 2 days ago
- Emergency talks called on 'harder' climate change The Age 5 days ago

#### Related tags

green temperature forecast global warming public policy

#### New in Science » Environment

- Will any nation other than Japan prevent the Steve Irwin from leaving its port by sea during 2009?
- When will the first report of the erosion of the Wilkins Ice shelf occur?
- Will the U.S. Forest Service be abolished and save the country over 2 billion dollars?