

A conversation with Philip Tetlock on February 12, 2014

Participants

- Philip Tetlock — Professor of Psychology and Management, University of Pennsylvania; Co-leader of the Good Judgment Project
- Alexander Berger — Senior Research Analyst, GiveWell

Note: These notes were compiled by GiveWell and give an overview of the major points made by Professor Tetlock.

Summary

GiveWell spoke to Professor Tetlock to learn more about opportunities for philanthropy in forecasting. Conversation topics included the Good Judgment Project, efforts to depolarize political debate, macroeconomic forecasting tournaments, and establishing the credibility of tournaments.

The need for accountability in forecasting

Political pressure often leads pundits to try to avoid accountability, such as by using vague measures of uncertainty and complex conditional statements.

In the intelligence community, there is an "accountability ping-pong" dynamic in which political pressure leads the community to adjust their threshold for reporting potential threats up or down in response to the most recent public mistake. That is, after a notable failure to foresee a coming event (i.e. a false negative), the threshold for predicting events will be lowered, resulting in more false positives. This reduces accuracy in the long run.

The Good Judgment Project (GJP)

The Good Judgment Project (GJP) is a research study organized as part of the Intelligence Advanced Research Projects Activity (IARPA) Aggregative Contingent Estimation (ACE) forecasting tournament. The IARPA ACE forecasting tournament is an attempt to learn how to improve the accuracy and precisions of intelligence estimates by avoiding "accountability ping-pong" dynamics.

The Good Judgment Project is a large project, especially by comparison with standard psychology or political science research projects. Testing hypotheses on the drivers of accuracy in forecasting is expensive because it requires many forecasters and a large research staff. It is less expensive, though not cheap, to just run a forecasting tournament.

Funding for future work

GJP is looking for funding to continue its work, as the IARPA tournament funding will end in the summer of 2015. For GJP to reach its goals, much more funding is needed.

IARPA might continue some forms of forecasting research, but it is unlikely that government agencies will continue to support GJP. Funding for further work is more likely to come from the private sector than the public sector. It may be possible to develop into a for-profit venture: some companies are interested in forecasting tournaments as a way to improve planning and could potentially fund this kind of work.

Depolarizing political debate

Forecasting tournaments could help depolarize political debates by inducing each side to make testable predictions about policies. Tournaments could be structured similarly to the Good Judgment Project (GJP) but with a narrower focus. Tournaments could address large questions, such as “What are Chinese geopolitical intentions?” by asking smaller questions with observable, near-term indicators, such as “How many Japanese people will die as a result of Sino-Japanese conflict in the next X years?” If one side consistently won a forecasting tournament over time, that side would gain credibility in the debate. Forecasting tournaments could be organized for both foreign and domestic policy topics.

GJP hopes to include efforts to depolarize political debates in its long-term research agenda. To have an impact, such a project would need to be done on a large scale over a long time period. It is difficult to find funders to commit to projects of this scope and duration. The project would also require influential sponsors and high-profile participants who would be genuinely interested in the truth and able to encourage their colleagues to join. A bipartisan set of think tanks might be the appropriate institutional home for such an effort.

Other areas for forecasting

Techniques used by GJP could be applied to other areas of government. For example, the National Institutes of Health (NIH) could use forecasting tournament methods in its grantmaking process to assess the potential long-term impact of research proposals and to improve the allocation of funding.

GiveWell asked about the possibility of forecasting tournaments around macroeconomic policy decisions.

A forecasting tournament could provide a helpful framework for resolving debates in macroeconomic policy, such as disagreements over the impacts of various Federal Reserve policies. Generally speaking, it would be valuable to track the predictions made by prominent liberal and conservative economists who disagree about particular policies to see if one side is consistently more accurate. The Brookings Institution and the American Enterprise Institute, which has credibility among conservatives, might be good institutions to host such a forecasting tournament.

Establishing the credibility of tournaments

People are more likely to participate in forecasting tournaments if they perceive their credibility to be at stake, and if the tournaments are perceived as non-partisan and credible. Tournaments could potentially establish credibility by hosting modest but high-visibility demonstration projects with policy relevance, or by being associated with prominent individuals. Demonstration projects would probably not be very expensive.

Prediction market-style tournaments could help generate public interest in forecasting. Pundits have become more cautious about dismissing quantitative models after Nate Silver's accuracy in predicting 2012 election outcomes.

Short- vs. long-term forecasting

Accuracy decreases rapidly with longer-term predictions. Professor Tetlock's research shows that it is difficult for experts to make good predictions about outcomes five or more years into the future. Predictions beyond the lifespan of the person making them may be entertaining but should not be treated as credible. There's little or no evidence that people are capable of reliable non-trivial multi-decade predictions, and there is a strong track record of such predictions failing.

There are some exceptions to this general principle, such as stock markets, which are volatile in the short term but relatively predictable in the long term, but such exceptions are unusual.

All GiveWell conversations are available at <http://www.givewell.org/conversations/>