

Michael Abramowicz

The Politics of Prediction

Innovations Case Discussion: The Policy Analysis Market

Bets are essentially predictions. Whether in financial markets or in explicit betting contexts such as pari-mutuel racing, they have always effectively played that role. The current price of a share of stock provides a prediction about a company's future cash flows, and the pari-mutuel odds reflect a group consensus about the likely outcome of a race. But in neither context is prediction the central purpose. Financial markets exist to promote broad and diversified ownership of enterprises, and pari-mutuel betting has long been a vehicle for gambling and entertainment. An innovation called prediction markets combines the most useful aspects of both approaches—the dynamic trading of financial markets and pari-mutuel's focus on a specific event—to generate group forecasts that are interesting for their own sake.

Prediction markets apparently originated in 1988, when the first Iowa Electronic Markets predicted the winner of the presidential election.¹ The founders of the Iowa Electronic Markets wanted a vehicle that would better explain the workings of financial markets,² and their approach built on the work of experimental economists who had conducted experiments using simplified financial markets in the laboratory.³ Robin Hanson reports having first considered the idea of prediction markets at about the same time as the Iowa Electronic Market was developed, and his revolutionary contribution was to recognize that prediction markets (or “idea futures,” as he then called them) might serve not merely as a test of financial markets, but as a tool in their own right, to produce group consensus forecasts.

Hanson's initial academic article on the subject focused specifically on “science claims,”⁴ but his argument clearly showed that markets could be used to assess virtually any claim about the future that could eventually be verified. The Foresight Exchange, which implemented his vision, includes claims about a wide variety of

Michael Abramowicz is a Professor of Law at the George Washington University Law School. He teaches and does research in areas including intellectual property, civil procedure, administrative law, insurance law, and corporate law. Before coming to GW, Professor Abramowicz served for a year as a visiting assistant professor at Northwestern University School of Law, and three years as an assistant and then associate professor at George Mason University School of Law. He holds a JD from Yale Law School and a BA from Amherst College.

domains, including entertainment, news, science, and politics. Unfortunately, the Foresight Exchange uses play money, because regulatory restrictions continue to limit the application of prediction market technology.

Properly conceived, prediction markets do not compete with polls, statistical analysis, or other approaches for arriving at beliefs about an uncertain future. Rather, they are a tool for aggregating the views of people who may have used sophisticated methodologies, such as the tools of econometrics, to make individual estimates. Prediction markets provide financial incentives for the best-situated individuals to apply the best available tools to predictive problems, and to test the depth of conviction of those who have done detailed analyses themselves, as well as those who have studied the work and reputations of such expert analysts. Thus they can effectively identify a consensus position.

INSTITUTIONAL FAILURES IN THE FUTUREMAP DEBACLE

Though prediction markets have performed accurately in several contexts, many casual observers still see them as more toy than tool when it comes to the design and implementation of public policy. Hanson's account of the Defense Department's abortive FutureMAP program (specifically, the aspect called the "Policy Analysis Market") provides insights into an innovation arriving in the wrong place, at the wrong time.

Why did FutureMAP fail? First, as Hanson astutely theorizes, politicians and the public were antagonistic to the program because it appeared to violate certain taboos. The program allowed some to benefit from deaths associated with terrorist attacks—which seemed morally unacceptable. Of course, many economic practices (such as the sale of annuities) set up some people to benefit from others' deaths.

But this was just the start. Even if Hanson and the other designers of the program had carefully excluded the possibility that predictions could be made directly on whether terrorist attacks would occur or even on lives more generally, other concerns may have doomed the effort. To many, gambling itself is taboo and using outcomes from gambling as a basis for policy is politically unthinkable.

Of course, taboos may sometimes capture an important philosophical intuition. What is the chance, however, that such philosophical arguments will be debated seriously by the relevant policymakers? And will policymakers seriously consider any relevant economic or institutional arguments in the presence of a taboo? By its very nature, political taboo presents the risk of ridicule to any politicians who attempt to engage in serious analysis. And so, decisions are based on superficial arguments rather than serious ones.

The taboo problem can be seen as related to a more serious problem with political systems: the distortionary effect of the *availability heuristic*.⁵ Individuals will overestimate the probabilities of outcomes that are salient in public discourse and the media. For example, the public overestimates risks that are highly publicized, such as the dangers of hazardous waste, and public policy responds exces-

The Politics of Prediction

sively, meanwhile failing to respond to other problems. Similarly, if any aspect of a proposed policy contains a potentially salient taboo, it will receive extra attention in initial news reporting. For example, the fact that FutureMAP might have allowed betting on terrorist acts initially received far more attention than any of FutureMAP's positive aspects: its possible applications, or the economic and institutional justifications for using it. The public's judgment of the program will depend more on the salient argument than on the less available ones.

WHEN INNOVATION MEETS THE POLITICAL PROCESS

If in fact the public does not want policy determined by gambling, let alone by gambling on human lives, then the political process functioned quite effectively in this case. The summary judgment rendered on FutureMAP by the political system might be considered evidence of procedural efficacy, rather than analytic laziness. Legislators have limited time, and the public has limited attention. If more attention had been paid to FutureMAP, less would have been available for other issues.

But this argument is fundamentally flawed: it ignores the potential public interest in identifying and implementing effective policies, whether or not these policies are unpopular at first. After all, a major justification for representative government, in contrast to a direct democracy in which the electorate chooses policies through plebiscites, is that the public does not have the time to become informed on all issues. If our representatives merely appeal to the first responses of the electorate, our indirect democracy will not perform much better than direct democracy.

One response to this dilemma would be for a random sample of the public to become relatively well informed about a new program, such as FutureMAP. For example, Peter Dienel has developed a procedure he calls citizens' panels, in which randomly selected citizens evaluate policies.⁶ Would such a group have eventually come to support the program, at least in some modified form? There is no way of knowing for sure. And can we let policy outcomes depend on drawing small random groups of citizens (or even of judges or legislators) to make decisions? If we did, policy would effectively become random, because the law of small numbers implies that policy outcomes might depend on which citizens happen to be drawn for a project. A group of twenty randomly selected citizens, for example, might purely by chance reach a different conclusion than another such group. Education is also important here: obtaining a representative informed sample would require educating a large enough sample of the public, perhaps for months or even years, on every policy issue. About the best we can do is rely on government agencies, but the agendas of bureaucrats may deviate substantially from the those of the public. To the extent that public opinion is represented in the political branches of government, it probably provides a beneficial corrective to technocratic governance more often than it causes harm.

Michael Abramowicz

RESOLVING THE TRADEOFF BETWEEN
DEMOCRATIC BREADTH AND ANALYTIC DEPTH

Those who try to design institutions in modern democracies encounter a fundamental tradeoff: between the representativeness inherent in a process and the quality of information on which the process relies. In other words, democratic “breadth” competes with analytic “depth.”

Different institutions resolve this tradeoff in different ways. A court, for example, is not at all a representative body, but certainly provides analytic depth; judges generally make decisions after carefully considering a great deal of information, including counterarguments offered by both parties. Various institutional features of the judicial process mitigate against the lack of representativeness. Judges, for example, are bound by precedent, as well as by rules of procedure.

Administrative agencies reflect a similar approach to the tradeoff. In general, agency officials will carefully consider relevant scientific evidence. Although agencies are not very representative (reflecting the ideological leanings of the president), a variety of institutional features, including judicial review, constrain ideological action. Finally, legislatures reflect a slightly different balance, perhaps achieving more representativeness, but less depth, with greater potential for issues to be resolved by either the mentality of the mob or by special interests.

On the whole, the arrangements that have evolved in each institution may well reflect sensible resolutions of the tradeoff between representativeness and information, along with other tradeoffs inherent in democratic government. Even if the political system was heavy-handed with FutureMAP, failing to recognize its potential benefits and exaggerating its costs, the system more broadly does produce a balance between false positives (policies that are enacted even though they should not be) and false negatives (policies that are rejected even though they should be enacted). I would suggest we count FutureMAP as a false negative, because I would hazard that in a hypothetical world with a much better informed legislature and much better informed public, the program would have survived. Still, any attempt to remove those imperfections by reducing representativeness might lead to many false positives and false negatives in other arenas. If, for example, legislatures met like courts in panels of three to consider each issue in more detail, there might be a better chance that FutureMAP would have been enacted, but also a good chance that many programs that ought not be enacted would be enacted simply as a result of the random selection of legislators.

A FUTURE FOR “IDEA FUTURES”?

So far, my argument would seem to imply that—the title to Hanson’s paper notwithstanding—prediction markets have little future in public policy. After all, if policymakers ignored readily available arguments on behalf of FutureMAP, would they not do the same with the results of prediction markets? It is hard to avoid the conclusion that even if the FutureMAP project had continued, policy-

The Politics of Prediction

makers might have blithely ignored its assessments of U.S. policy options that contradicted their agendas. What is the use of improving the aggregation of information unless doing so also improves decision-making?

In democratic societies, prediction markets will start to matter for public policy not when they are understood and believed by politicians and technocrats, but when they are understood and believed by the public. Imagining a hypothetical world in which FutureMAP is functioning well and the public understands prediction markets, one can imagine a news report noting the market's anticipated reaction to a U.S. policy decision. Imagine this one, for example: the administration is confident that its new agreement with Saudi Arabia will increase economic stability there, but the markets project increased economic volatility following the administration's announcement. In such a policy world, policymakers would find it at least a bit harder to manipulate statistics to support their own favored policies for the wrong reasons. At the very least, this might lead to more honesty about the true reasons for policy decisions, and a negative market assessment of a potential policy course might even lead government officials to avoid that policy course.

Thus the case for prediction markets for public policy has an unfortunate chicken-and-egg quality: prediction markets may be beneficial to public policy once the public has come to appreciate them, because they will make it easier for the public to assess government policy. But the public will likely come to appreciate the markets only after a great deal of exposure to them. Even if FutureMAP had succeeded, even the most educated members of the public would likely not have allowed prediction market forecasts to affect their personal evaluations of policy options until the government had tried many more experiments with prediction markets. This bind helps explain why I would bet against widespread government use of prediction markets any time soon.

Still, over the very long term, these markets may be a good bet. As a rule, promising technologies tend to earn public trust. With prediction markets, experimentation by business seems likely to outpace experimentation by government. But even business experimentation will be far too slow, because anyone can free-ride off successful experiments, while everyone can ridicule unsuccessful ones. But eventually success in the private sector may lead to selective experiments in the public sector. Even before the public learns to appreciate their virtues, some politicians who care genuinely about policy outcomes might come to rely on prediction markets, perhaps even privately-sponsored ones, as a way to facilitate their own assessments of policies. We may overestimate the political changes that can occur in ten or twenty years but underestimate the possibilities of transformation in fifty years or a century.

THE PROMISE OF PREDICTION

In some distant future, prediction markets may be so well established the government will rely on them whenever it engages in an explicitly predictive task. It might seem that widespread use of prediction markets would not work, because they

would not be zero-sum games. The traditional structure of a prediction market depends on traders disagreeing with one another about the future. Especially in contexts where individuals may have inside information, and where particular topics provoke little inherent interest, few trades may occur, and the market's forecasts may be unreliable. But prediction markets can be designed with automated market makers, so that traders can buy and sell against the "house," subsidized by the government. In another significant contribution to the prediction market movement, Hanson devised the "market scoring rule" that is becoming the most common form of automated market maker.⁷ With this and other automated market makers, the price at which transactions with the house occur depends on the current market forecast, and so individuals can trade whenever they believe that the current market forecast is wrong.

It is easy to imagine extensions to FutureMAP for a wide range of policy contexts. One of the project's most powerful aspects was its embrace of combinatorial markets, which would have allowed users to assess the effect of any government policy option on any variable of interest, such as economic output, or health. In any policy context where policy options might be expected to have a measurable effect on certain variables, these markets could be useful.

And yet it might still seem that prediction markets, thus described, would not have done much in the context of the FutureMAP controversy itself. Of course, if the markets were widespread, then the successes of those markets might have suppressed the sense of taboo associated with them. But imagine some future controversy concerning some other new technology that tends to lead to feelings of taboo but could also improve information aggregation and policymaking. How could a prediction market help? Maybe it would only help where the effect of the new technology is particularly dramatic. Even if FutureMAP had succeeded, after all, it would likely have had only a very slight effect on any measurable outcome, such as the probability of a terrorist attack occurring. That does not mean that the program was useless, but rather that many government initiatives may produce benefits (or costs) that are real but difficult to measure. A tiny decrease in the probability of terrorist attacks might be well within the margin of error of a prediction market, yet be quite valuable from a policy standpoint.

In such contexts, a potential solution is to use what I have called a "normative prediction market": a market that forecasts a subjective evaluation of a policy. For example, imagine this commitment: in a few years, we will select someone at random to evaluate a technology being considered today, write a report, and rate the technology on some numeric scale. The prediction market will forecast that rating, whose only purpose will be to resolve payouts in the prediction market, not to actually determine policy. The prediction market forecast of the rating, rather than the rating itself, would determine whether the technology is implemented, or it would contribute to a determination made through more traditional means.

For example, in a variation I call *predictive cost-benefit analysis*, the numeric scale might be dollars, and we might commit that an official of a relevant administrative agency will conduct a traditional, detailed cost-benefit analysis ten years

The Politics of Prediction

hence. The particular result of that cost-benefit analysis might depend on who happens to be selected to conduct that analysis and the ideological tendencies of the president ten years from now. For some policies, Republicans and Democrats might be expected to have systematically different views. The prediction market participants, however, would not know in advance who that individual will be, and will thus attempt to forecast a rating based on an average of anticipated ideological views.

This approach provides a kind of virtual representativeness. Whatever the beliefs of the prediction market traders, they will have incentives to apply a moderate political standard to the problem. At the same time, the decision should be very well informed. Eventually, in developing a full report, the agency officials can be expected to consider complex and subtle arguments. Anticipating this, the prediction market participants will have an incentive to consider those arguments as well. Of course, the larger the subsidy provided to the market, the more carefully they will consider the arguments. With a sufficiently large subsidy, prediction market participants will even have some incentive to release their own models and arguments, hoping to influence the agency official to make a cost-benefit assessment that will reward their original forecast.

This discussion suggests that, superficial appearances aside, Hanson's case has more to do with the theory of representative democracy than it does with strategies for national security. The case—along with the related literature on prediction markets—provides some indication of how prediction markets might be incorporated into our decision-making institutions, such as courts, administrative agencies, and legislatures. It might at first seem that law is only occasionally predictive in nature. From this perspective, individual legal decision-makers might rely on prediction market forecasts, in the same way that they might rely on many other types of data, but prediction markets do not perform tasks that are integral to the enterprise of law as a whole. Representative government is a tool for facilitating democratic participation through voting, but votes may legitimately involve the expression of preferences, rather than merely predictions of which party's platform would more likely succeed. On the other hand, prediction markets can be used to forecast the average subjective assessments of citizens or their representatives, and those very assessments reflect anticipated preferences and beliefs. Ultimately, these markets can be structured to perform or to simulate the core functions of our democratic institutions in a way that overcomes the tradeoff between representativeness and information.

CONCLUSION

Relative to their potential, prediction markets still remain in their infancy, in law as in other domains. The FutureMAP project represented a recognition that government policy frequently depends on predictive assessments. In many cases, government makes predictions only implicitly, and even when its predictions are explicit, we must worry that the prediction may be offered to support an agenda

being pursued largely for independent reasons. While prediction markets should be useful for any organization that seeks to improve its aggregation of information, they are particularly critical for government, because they make decisions more transparent. In the long term, as a scalable technology, normative prediction markets could accomplish tasks traditionally reserved to institutions like legislatures and administrative agencies, and do so in a way that is more representative and better informed.

-
- 1 Hanson describes some precursors to modern prediction markets. An important precursor to the Iowa Electronic Markets was betting markets on earlier presidential elections. Paul W. Rhode & Koleman S. Strumpf, "Historical Presidential Betting Markets," *Journal of Economic Perspectives* 18 (2004): 127.
 - 2 Robert Forsythe and the other creators of the market used the experience to assess how effectively financial markets aggregate information. Robert Forsythe et al., "Anatomy of an Experimental Political Stock Market," *American Economic Review* 82 (1992): 1142. [This second note should say who the creators are. Forsythe et al? If so, say so!]
 - 3 Vernon L. Smith, "Markets as Economizers of Information: Experimental Examination of the 'Hayek Hypothesis,'" *Economic Inquiry* 20 (1982): 165.
 - 4 Robin Hanson, "Could Gambling Save Science? Encouraging an Honest Consensus," *Social Epistemology* 9 (1995): 3.
 - 5 Timur Kuran & Cass R. Sunstein, "Availability Cascades and Risk Regulation," *Stanford Law Review* 51 (1999): 683.
 - 6 See Peter C. Dienel, "Contributing to Social Decision Methodology: Citizen Reports on Technological Projects," in *Social Decision Methodology for Technological Projects* (C. Vlek & C. Cvetkovich eds. London: Springer, 1989, 253. For an American analogue, see Ronald F. Wright, "Why Not Administrative Grand Juries?," *Administrative Law Review* 44 (1992): 465.
 - 7 Robin Hanson, "Combinatorial Information Market Design," *Information Systems Frontiers* 5 (2003): 107.