

COMMENTARY

COMMENT ON A “MODEST PROPOSAL” FOR A CRIME PREDICTION MARKET[†]

Thomas E. Feucht* & William J. Sabol**

INTRODUCTION

Henderson, Wolfers, and Zitzewitz conclude their paper, *Predicting Crime*, with what they term a “modest proposal”: an entity, such as the National Institute of Justice (NIJ), should establish a crime prediction market to develop and test the authors’ approach for forecasting crime rates and informing crime policy. They argue that prediction markets, which have been used to predict election outcomes and are used by businesses to forecast sales, costs, and other outcomes, can also be used to forecast crime rates. They argue that prediction markets can be used to provide crime policy decisionmakers with transparent, actionable information that can guide decision-making as well as evaluate the effects of crime policy decisions. The authors state that the market mechanism would provide a centralized location for aggregating information about crime forecasts, and that policymakers can then evaluate the information about future crime rate predictions and thereby make a more informed decision. The authors propose that a federal entity, such as NIJ, should establish the prediction market to overcome potential political battles that might occur at the local level and that a federal role is appropriate because of overlapping jurisdictions.

The proposed initial experiment is straightforward: it would focus on predicting various levels of Uniform Crime Reports (UCR) crimes over the next three years, relying on a group of about 200 invited crime experts who would be endowed with \$200 to purchase prediction contracts in the new markets. Stakes

[†] Findings and conclusions reported here are those of the authors and do not necessarily reflect the official positions or policies of the U.S. Department of Justice. Order of authors is alphabetic.

* Thomas Feucht is Executive Senior Science Advisor at the National Institute of Justice, 810 7th Street NW, Washington, DC, 20531, Thomas.Feucht@usdoj.gov.

** William Sabol is Chief of the Corrections Statistics Program, Bureau of Justice Statistics, 810 7th Street NW, Washington, DC, 20531, William.Sabol@usdoj.gov.

would be modest, and traders could cash in only after accumulating \$500. The top trader each year would receive a non-financial award. The objectives of this experiment would be to learn more about the organization and operation of the market, such as how to write contracts, how much participation to allow, and what stakes to set.

Our comment principally focuses on practical considerations for implementing the authors' modest proposal. We begin by evaluating the proper scope of an experimental crime prediction market. In that part we question whether a national prediction market would be useful for a first experiment. We examine the alternative—a local crime market—as a way of articulating the challenges that would attend any experimental crime prediction market. The scope of the prediction market is linked to several other key considerations that we address in turn: who the right traders are; what the payoff is and how market incentives would work; and how success is to be measured. Finally, we conclude with a discussion of whether prediction markets about crime—national or local—will have a chance to inform crime policy and decision-making.

I. THE QUESTION OF SCALE FOR A CRIME PREDICTION MARKET

One place to begin appraising the potential for crime prediction markets is with the proposed use of UCR crimes as the forecasting objective. A focus on UCR is not unreasonable: the data are well-known; they are relatively reliable and resistant to manipulation by a single, contributing entity; and because they are produced routinely, they provide a convenient target for a prediction market. However, the use of UCR signals a preference for a crime prediction market of national scale, and we are not convinced that this is the right place to begin experimentation. In particular, a UCR experiment may be of little use in assessing a kind of prediction market that may hold greater potential: those with a local focus.

Decision-making about crime is an overwhelmingly local affair. Our nation's spending on crime and justice issues is dominated by state and local jurisdictions.¹ The vast majority of law enforcement agencies and personnel belong to local units of governments, and most persons tried, convicted, sentenced, incarcerated, or supervised by justice authorities are the responsibility of state courts and local justice agencies. The viability of a prediction market will ultimately be determined by the value of the predictive information the market is able to produce. Therefore, it is important to set the scope of the experiment in a way that would maximize its chances for relevance to crime policy decision-making.

1. The Bureau of Justice Statistics compiles data from the Census Bureau's Annual Government Finance Survey and Annual Survey of Public Employment. For 2005, total outlays for the criminal justice system were calculated as \$204.13 billion: local jurisdictions spent \$103.77 billion; states, \$64.95 billion; and the federal government, \$35.41 billion. For more information, see BUREAU OF JUSTICE STATISTICS, JUSTICE EXPENDITURE AND EMPLOYMENT EXTRACTS 2005 (2007), *available at* <http://bjs.ojp.usdoj.gov/index.cfm?ty=pbdetail&iid=1023>.

While forecasting the UCR crime levels may provide the basis for a reasonable game, much like forecasting movie revenues or election victors, a crime prediction market that forecasts UCR crime levels would not generate much useable information for local crime policy decisionmakers. And while the use of UCR data is only proposed for a first experiment, it is not clear to us that such an experiment would help anticipate how other—more locally scaled—markets might operate. That is, the lessons learned about market structure, operations, stakes, and so on may not be easily translated into prediction markets for local markets.

Since the ultimate goal of a prediction market is to produce actionable predictions, we believe the experimental work should begin where the action is: at the local level. In the end, however, it is not simply the case that since “crime is local,” a local crime prediction market is the easy solution. At least three inter-related issues complicate the matter: (1) who should participate in the market; (2) how market incentives might work to improve prediction; and (3) how success might be measured. These are significant challenges to resolve for either a local or national crime prediction market. Following a brief discussion of each of these issues, we revisit the issue of the scope of markets and the potential for crime prediction markets to actually inform crime policy and decision-making.

II. WHO GETS TO PLAY IN THE PREDICTION MARKET? QUESTIONS ABOUT EXCLUSIVENESS AND INCLUSIVENESS

Lessons learned from other prediction markets—especially ones where the market reached the wrong prediction, such as the election market for the 2008 New Hampshire Democratic presidential primary²—suggest that a successful prediction market requires a large and diverse pool of traders. A large pool is required to limit the effect that a single trader may have on the overall market outcome. Additionally, prediction markets require a diverse pool to generate sufficient variation in the bets. Henderson, Wolfers, and Zitzewitz propose a first experiment for predicting UCR levels that would be limited to about 200 criminologists and chiefs of police. We wonder whether a national crime prediction market (like the one the authors propose) could actually attract a significantly diverse group of traders. Our suspicion is that interest in national UCR data probably does not extend beyond police chiefs, criminologists, and other experts. Thus, a national UCR prediction market could generate a large pool of expert traders, but diversity will be harder to come by.

If relevant crime policy information is local in nature, as we suggest, then a first experiment might consider how large and diverse a pool of participants could be used for a crime prediction market for a local area like a neighborhood in Denver, the city of Cleveland, or greater Miami-Dade County. Local crime data and predictions about local crime would have greater salience than national UCR data for homeowners, local merchants, and other community residents, and even for criminals themselves. The question becomes, would local prediction markets attract sufficiently large and diverse pools of traders?

2. Daniel Gross, *Why Were The Political Futures Markets So Wrong About Obama and Clinton?*, SLATE, Jan. 8, 2008, <http://www.slate.com/id/2181745>.

Clearly, any crime prediction market would be best served by a diverse group of traders with diverse information sources on which to base their predictions and bets. One might expect that a large metropolitan area could easily sustain markets of hundreds of participants. Yet even in a community with a diverse population, it may be difficult to achieve adequately diverse traders, with diverse sources of information and strategies for forming predictions. The key questions are how local residents obtain their information and form their views about future crime and whether a jurisdiction would contain sufficient variation among local residents in predicting future crime rates to generate a robust distribution of outcomes.

We note that local crime data may be subject to greater relative error and easier manipulation (compared to national data like UCR), and crime reports in local newspapers or other local media may be more vulnerable to spurious factors like unbalanced crime reporting in the media. With local crime data, there is no corollary to the national compilation of UCR data that can help eliminate errors. It may be hard for potential local prediction market traders to have confidence in the reliability—and inherent predictability—of these data. Conceivably, a diverse urban populace could possess diversity in information *sources* about crime.

We also question whether the pool of local crime experts knowledgeable about and willing to bet on local crime predictions would be sufficiently large. A variety of experts generating predictions based on different analyses of data using different prediction models *as well as* other non-expert traders with unique knowledge would result in the greatest diversity. For local crime prediction markets, the number of experts may be quite small, and this would not bode well for developing a successful prediction market.

In addition, it seems important that the numbers of experts and other traders be carefully balanced in order to avoid an “affiliation bias,”—where the direction of the market is biased as a result of too many traders who belong to a particular group or share a common view. Frankly, we do not know what the right equation of experts and “others” is, but it is reasonable to assume that it is not an experts-only market. It is possible to imagine a structure of incentives that could entice greater participation among one sort of trader or the other, thus diminishing the potential for affiliation bias. Calibrating and maintaining differential incentives, however, merely complicates the value and payoff structure of the market, as discussed below.

Clearly, the challenge is to find a jurisdiction that has good, reliable crime data, is cosmopolitan enough to provide sufficient expert traders, and is provincial enough to attract and retain the right balance of non-expert traders. This is no mean task.

III. WHAT IS THE PAYOFF? QUESTIONS ABOUT VALUATION AND WORTH

Results from other prediction markets suggest that the stakes involved have to be real stakes—that something of value must be derived from the prediction market. In the proposed crime prediction market, an important part of the market payoff is the non-financial reward that accrues to the successful

trader—the one whose predictions prove most correct. This non-financial reward is likely to have different value for different traders. For experts, like criminologists or police chiefs, disclosing who was the best crime predictor would be a form of professional prestige and recognition. At the same time, this intangible good may be of little value in drawing non-expert participants into the market. For example, this same acknowledgment may not be a strong enough incentive to draw local residents or merchants into the market. Compared to other small-stakes markets, like local lotteries, 50/50 fundraiser drawings or other sweepstakes and prize raffles, the already modest tangible payoff of a crime prediction market is not sufficiently enhanced for non-expert traders by these intangible benefits. It is, of course, an empirical question.

But, disclosure could also stifle participation. Expert traders and others who perform badly on a consistent basis may not wish to be identified, and such disclosures may become a disincentive to participation. This would be disastrous for the market since bad predictions are a necessary ingredient for the overall performance of the market.

Beyond its potential for incentivizing traders, disclosure could be important for market transparency, as Henderson, Wolfers, and Zitzewitz point out. It is also essential if one of the market's goals is to identify prediction strategies and models that perform better (and worse) than the market average.

The disclosure dilemma poses a unique challenge. While we would prefer a market that discloses information about participants' bets—much like mutual fund ratings—we doubt that a crime prediction market that discloses participants' outcomes would generate more enthusiastic participation overall compared to a market that did not disclose.³

Even if a large, diverse, balanced, and adequately motivated group of traders could be assembled to participate in an experimental crime prediction market, it is not clear to us how to get around the “endowment effect” of the sponsored experiment. In the modest proposal, the initial traders would be endowed with a small, \$200 grant that they would use for trading, and they could cash in only after realizing a certain amount, such as \$500. However, as Kahneman and his collaborators point out, in games in which traders are initially endowed with money, the traders might be willing to accept contracts that greatly exceed their willingness to pay because they do not have to put up their own money.⁴ One could argue that the endowment effect would be less among experts, like police chiefs and criminologists. This uncertainty does not strengthen our confidence in either an expert-only national crime prediction market or in a more diverse local market.

3. Allowing participants to choose to reveal themselves does not solve this problem if, as we suggest, one purpose of the market is to evaluate the performance of the experts. Moreover, if expert traders are allowed to choose to reveal themselves, then market watchers will not have full information about high and low performers. And, requiring disclosure leads to the participation problem that we have outlined.

4. Daniel Kahneman et al., *Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias*, J. ECON. PERSP., Winter 1991, at 194–97.

IV. MEASURING SUCCESS

The proposal for an experiment in crime prediction markets begs the question of evidence: when and how will we know whether the experiment was a success? It seems clear to us that achieving a meaningful measure of success will require an effort much larger than the modest proposal sets forth.

Given the issues presented here, several experiments and a period of several years would be required to provide information about the conditions that are necessary to create an efficient market. Substantial resources would have to be committed to crime prediction markets for several years, running several experiments in different localities in order to decide whether crime prediction markets can “work.” And even then, evidence that the crime prediction market was efficacious would be elusive.

The key question about crime prediction markets is whether they can provide accurate crime forecasts that are useful to crime policy decisionmakers. If a single manifestation of a crime prediction market fails, it is not conclusive proof that crime prediction markets are a bad idea, nor would one success prove the opposite. But would a second failure (or success) lead to a firmer conclusion? Would multiple failures lead us to jettison the prediction market enterprise, or would they merely signal that the prediction markets were not organized properly?

Evaluating the efficacy of a prediction market is similar to testing hypotheses derived from theories. As described succinctly by Mark Blaug, no single experiment can lead to the rejection of an underlying theory. Summarizing Popper and Lakatos, Blaug points out that as tests fail and evidence builds against a hypothesis, its proponents develop immunizing stratagems to protect the core axioms of the theory, generating reasons why an experiment was flawed or finding ways to explain discordant results.⁵

Testing any particular manifestation of a prediction market is analogous to testing hypotheses from a theory. It is impossible to design a single test that would determine the efficacy of prediction markets because the success (or failure) of a single experimental prediction market would not prove that the underlying idea is good (or without merit). How much effort should be applied to establishing and testing an experimental crime prediction market before declaring victory (or accepting defeat)? Ultimately, the only acceptable empirical answer to this question will be obtained through the accretion of a preponderance of evidence—something that we think will take slightly more time and more resources than the authors’ “modest proposal” suggests. A federal funder for an experiment in crime policy prediction markets would be wise to develop a priori a stratagem for accumulating evidence and a stratagem for ending the period of experimentation.

5. MARK BLAUG, *THE METHODOLOGY OF ECONOMICS, OR, HOW ECONOMISTS EXPLAIN* 17–18 (1980).

V. THE POLICY UTILITY OF PREDICTION MARKETS

In the end, the most compelling reason to experiment with crime prediction markets is that they might make a valuable contribution to crime policy and decision-making. Henderson, Wolfers, and Zitzewitz suggest that a national crime prediction market using UCR data has a motivated and interested national policy audience, but we are not so confident. We suggest that local crime policy decision-making is where the action really is—but even there, we are not optimistic about the potential contribution of crime prediction markets.

A key issue not addressed in the modest proposal is that of translating forecasts from the prediction market into policy and action.⁶ Accuracy of information is not sufficient in and of itself to compel action by decisionmakers. Moreover, in dealing with forecasts of future outcomes, accuracy is not known until after the decisions have been made, and like all good forecasts, the accuracy of the estimates comes with uncertainty.

One compelling advantage of crime prediction markets over the use of expert panels, commissions, and consultants that Henderson, Wolfers, and Zitzewitz point out—and with which we agree—is that they can be constructed in ways that enhance the transparency of the information that goes into making policy decisions. For example, a website could show the forecasts and local residents could assess whether local decisionmakers did anything in response to the forecasts. If the decisionmakers did not respond to what turned out to be accurate forecasts, then local residents could take steps to hold them accountable for failing to act.

That stylized example, however, only addresses the advantages of a crime prediction market and not the disadvantages. Because the prediction market produces predictions with uncertainty, a decisionmaker would have to make decisions based on the predictions with the same uncertainty. The risk for local decisionmakers is that the prediction market forecast will be wrong (perhaps for reasons related to stakes, diversity, or number of traders), and they will have to explain their decision to act based on a market mechanism that aggregates opinion from an unknown group of traders, rather than relying on their own knowledge or input from other named experts. Aggregating opinion and knowledge is a potential strength of a prediction market, but it also presents a unique challenge to the decisionmaker: holding accountable those others on whose predictions a decision would be based.

We feel strongly that the point of creating a crime prediction market would be to produce actionable data for policymakers. Merely demonstrating that a market can predict crime more accurately than any individual expert would not address the key issue of how knowledge gets translated into action by policymakers. In preparing these comments, we reviewed prediction markets in

6. The exception is the authors' suggestion for "event studies." The potential value of these studies to inform policy is a second-order function of the prediction market mechanism. We have no comments on this possible use of prediction markets since it would be premised on a market first solving the challenge of simply making correct predictions about crime.

other areas. Our review was not comprehensive, but we were nevertheless struck by a number of cases in which the main goal of the prediction market was either essentially a game (such as betting on elections), or a device for identifying winners (such as the influenza markets), and how few cases we could find in which the information from the prediction market was used in decision-making. Additionally, many examples of prediction markets used in decision-making are associated with business decisions, where information about the prediction markets is closely held, making it difficult to see how the market information gets translated into action.

CONCLUSION

Our concerns about the organization and structure of prediction markets are rooted in the need to know how the information generated from the prediction market might be used in the policy decision-making process. A single national crime prediction market has, we believe, rather limited utility for informing crime policy and decision-making, and a single local crime market may unfortunately not be much better. If the purpose of crime prediction markets is to inform local crime policy, then no single prediction market would be adequate; rather, many local crime prediction markets would have to be developed. Would a successful experiment with a local prediction market provide sufficient and compelling reasons for other local markets to spring up on their own? Would a national UCR-based experiment provide a greater boost to the development of local markets? Uncertainty about whether a first, federally-funded experiment would be likely to spur further growth in prediction markets makes it difficult to judge the worth of the initial federal investment. Experience with the diffusion of ideas in other areas of crime policy—where successful diffusion often began from an innovation with local, not national roots—leads us to favor a local prediction market for the first experiment.

The basic question of the relevance of market predictions for crime policy and decision-making may seem to be a concern that is secondary to finding out simply whether or not a market can predict crime in the first place. We think that because crime prediction markets would be new to the field, it would be important to confront the “so what” question of the relevance of market predictions for actual crime policy and decision-making. There would be much to learn about how decisionmakers evaluate the information, how they deal with its advantages and disadvantages, and how they obtain political cover for decisions that turn out to be less than optimal. These important questions certainly complicate any consideration for establishing and evaluating crime policy prediction markets, but this is where the real worth of a crime prediction market will be tested and perhaps proven.

The article by Henderson Wolfers, and Zitzewitz helps to illuminate the difficult challenges of establishing, testing, and improving practical models for informing crime policy decision-making. Their proposal for an experimental crime prediction market is intriguing and imaginative. It is sure to spark considerable attention in a field where prediction and prevention have become the currency of the day.

On balance, we think that the modest proposal for a federally funded experimental crime prediction markets is indeed more substantial than modest. The proposal for an experiment with fairly constrained parameters raises a number of questions about scope, participation, and other potential complications that should be addressed if a first modest experiment in crime prediction markets is to lead to more than modest results.