Student Final Project

Instructor: Dr. Glen Weyl

ECO406

1/15/19

Using the Liberal Radical Mechanism to Fund Replication Studies

Research and development generates ideas, and these ideas become public goods if the results are open source. Publicly available scientific research, like many other public goods, is generally underfunded because it generates significant positive externalities. The Liberal Radical (LR) mechanism was introduced to remedy the problem of the underfunding of public goods by determining and provisioning funding at the theoretically optimum level. The LR mechanism allocates an amount of funding proportional to the square of the sum of the square roots of individual contributions to a project (Buterin, Hitzig, and Weyl 2018).

We propose to use the LR mechanism to fund scientific research. Under normal conditions, each contribution represents information revealing the private valuation of the project to each respective individual. For this use case, the project would be a well-specified program of scientific study to be funded and individuals would be considering their expected private valuation of the products of the study, which would be mandated to be open source if funded through this mechanism. Researchers would propose and raise money for studies through a centralized portal, which we will build as a proof of concept. The other side of this portal would allow for contributions. A separate LR mechanism would be used for each project. A system administrator would approve these projects, as discussed in Buterin, Hitzig and Weyl (section 4.4).

Since funding for science is a large problem space, we propose to focus in particular on funding replications of existing studies. Replications are important because they scrutinize established (and sometimes highly cited) studies and keep scientists honest and accountable. According to Dreber et al. (2015) the "costs associated with irreproducible preclinical research alone are about $28 billion a year in the United States" and "a lack of reproducibility threatens to undermine the validity of statistical hypothesis testing as it is currently practiced." The norm of replicating scientific studies has the additional benefit of encouraging good research practices so that studies are robust and therefore more likely to replicate. Many journals now require data, code, and other products of research outside of the final publication to allow replication as a result of the push for more replications and a general call for research reproducibility (a short summary of the state of research credibility of economics can be found in Roodman (2018)). Psychology and many other disciplines have had 'replication crises' during which a subset of researchers in the field have called attention to evidence that research in their scientific discipline is not consistently reproducible.

Like science research in general, replications are underfunded. Standard economic reasons discussed above, academic politics, and defensiveness on the part of researchers who are authors of existing studies contribute to underfunding. Using the LR mechanism to fund replication studies mitigates both of these problems. LR mechanisms are explicitly designed to fund projects to their economic optimum levels, subject to certain technical constraints. The decentralized and secret contributions of the LR mechanism (practical implementation to be described in detail below) circumvents political constraints to some extent by allowing researchers to express that replication studies are valuable to them without having to put their name on the funding they contribute and incurring a reputation cost or other penalty.

Although the community that benefits from replication studies is broader than the community of scientific professionals, they would likely be the main participants in this system of funding. People in this community of scientists would have to be able to estimate the private value of the replication study to them to participate. When functioning correctly, the LR mechanism induces individuals to report their private valuation of the project truthfully. Other than the salient obstacles to the proper functioning of the mechanism mentioned in Buterin, Hitzig, and Weyl, which are problems even in contexts where the outcome of a project is well-defined and certain to happen if the project is funded, there are additional considerations with replication funding. Since there is uncertainty about the outcome of the project (successful replication, replicated with caveats, results not replicated, replication found something interesting that neither contradicts nor confirms the results of the original study, replication plan shifted, etc.) an individual's dollar-equivalent utility (called 'value' previously but made explicit here for conceptual clarity) is uncertain and will be realized only after funding is finalized and the project is completed.

There are two main sources of uncertainty that can be actively reduced in the case of replications studies: uncertainty from under-specification of the project and uncertainty from lack of information that exists but is not yet known to relevant agents. Under-specification of the project will affect contributor's abilities to determine their private valuations of the study being carried out. A detailed proposal for a replication study would therefore be required to start a project and open it up to funding. The proposal would need to include contingency plans for various likely obstacles and follow best practices used by established pre-registration sites, such as the Center for Open Science (https://cos.io/prereg/). Some baseline requirements would include: the paper (or other scientific product) to be replicated, their expected results, whether

there has been any progress on the project already, and the proposing scientists' initial predictions. This would help potential contributors evaluate the (private) value of the study to them, conditional on the study being carried out. An appropriate minimum level of funding, set uniquely for each individual proposed project, would have to be reached in order to go forward.

Information relevant to the project would reduce uncertainty about the outcome and consequences of the study. This, in turn, would reduce the risk for an individual funder contributing their expected private valuation. Due to risk aversion, this expected dollar-equivalent utility for an individual would actually increase with the reduction of uncertainty even when it does not change the individual's estimation of the most likely outcome and its consequences. Uncertainty may therefore deter potential contributors from participating because they see other, better opportunities to invest in. Information that shifts the probability distribution assigning likelihood to expected outcomes is also valuable since it makes estimations of the value of a replication more precise. This is a good thing independently, but it is even more important because this mechanism would be a repeated game and people would cease to participate if they felt they were systematically misled or incorrect about likely outcomes.

Our proposed solution to discover and reveal information for all community members is to start a prediction market on the outcome of the replication. Current popular areas of application for prediction markets include politics and sports (like markets at PredictIt.com). Prediction markets as arranged in our system would consist of buying and selling contracts of a binary outcome, a standard way to design such a market. These markets only directly provide information about the result of the study: it either will or will not fulfill the criteria for a successful replication of the target study. This covers only one dimension of many possible degrees of uncertainty, but it is an important one. The information discovery mechanism would

be a significant step towards encouraging participation by mitigating uncertainty without an inordinately large increase in complexity. Since findings that are unlikely to replicate probably, on average, yield the most value when executed, the results of a binary prediction market for replication will be useful. The criteria for a successful replication would have to be formulated individually for each study and would have to be extremely clear. Otherwise, there will be a greatly increased risk that the prediction market does not resolve and the associated contracts will not pay out.

A properly functioning prediction market would not only provide reliable information about the probability of success of a replication, but it would also give community members aware of the market reason to seek information, do research, and think deeply about the likely outcome of the replication. Dreber et al. (2015) have shown that prediction markets are well-suited to predicting the results of replication studies, finding that "prediction markets set up to quantify the reproducibility of 44 studies published in prominent psychology journals … predict the outcomes of the replications well and outperform a survey of market participants' individual forecasts." The authors also observe that there has not been much work done on potential applications of prediction markets, even though they "aggregate private information on reproducibility [at low cost], and can generate and disseminate a consensus among market participants" (Dreber et al., 2015).

One potential concern, however, is that many projects and their associated prediction markets will not resolve by design: the market will have to be dissolved if a project is not funded to the threshold at which the money is disbursed. This non-resolution concern will be addressed below in the discussion on subsidization. Another concern is that prediction markets are much less effective outside of psychology, a field in which initial credence in results is often low and

that is known for past methodological troubles. Future research on prediction markets and practical experience from the growing platforms that facilitate them will provide a clearer picture of efficacy by discipline.

Prediction markets have several necessary preconditions for them to function well. These include being well-defined, quickly resolved, probable resolution, limited hidden information (no insider trading), and sources of disagreement/interest ("Subsidizing Prediction Markets," 2018). Prediction markets that are well-defined will sell contracts that pay out in the largest set of possible circumstances, preferably 100% of the time. Writing airtight contracts that constitute well-defined markets has been a problem for many new prediction market platforms, so close attention will have to be paid to the definition of replication and the technical terms of the contract. The problems of quick and probable resolution cannot be avoided in the use case of making predictions about replications, but their impacts can be mitigated. Participants will have to be compensated so that the risk of non-funding (when the project does not meet its threshold) and risk of long payout time (when the study occurs over a long time) are overwhelmed by the subsidy to the market. Some markets will be more expensive than others to subsidize, depending on how severe the two risks are.

Preventing insider trading on markets when participants in the market are highly connected through social ties is normally a significant challenge. Wolfers and Zitzewitz (2004) puts it succinctly: "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." Although it is advantageous that the relevant replication will not have been carried out yet, scientists proposing will not be able to participate in the prediction market

associated with their project (enforcement would have to be strict). One possible risk is that previous, unpublished work kept private by a researcher is used to skew a prediction market. However, this sort of work is very costly, both in terms of researcher time and money. If people are induced to operate their own mini-experiments prior to the closing of the prediction market, this is a good thing and does not truly qualify as insider trading. Even in the case where someone spends enough time on it to dominate the market and drive other participants out when the others realize they have material and substantial information that generates a significant advantage in the market, this is still a god thing because then there will have been substantial research done—that is the point of this entire mechanism.

The final precondition for functional prediction markets is sources of disagreement and interest. The markets we propose to set up are likely inherently interesting to scientists who work in a field related to the proposed replication, and there is much disagreement in the scientific community. The fact that scientists will often think they are right or have special knowledge (the latter is true, but can frequently lead to overconfidence) will, in theory, translate into trade volume for these markets. The existence of sources of disagreement and interest usually leads to markets with enough individual participants and enough capital invested. Insufficient participation is necessary because systematic bias is more likely when there are fewer participants. Insufficient positions taken/contracts created means that there may not be enough transactions to converge to a point closer to the true replication probability. Information aggregation does not occur to any significant extent without many participants and many contracts created.

An interesting side observation: participating in the prediction market encourages participation in funding the associated proposed replication because it increases the likelihood of the market paying out.

Given that there are many cases where the prediction market does not resolve and a long timeline for payout, each prediction market would have to be subsidized to some extent. There are several forms this could take. The money could come exclusively from a central funder, like a university, or be sourced in a partially decentralized fashion like the other two layers of the funding mechanism (e.g., another LR mechanism for funding the prediction market). This, however, may add too much complexity to be worthwhile for a system with applications at a smaller scale. What would the technical details for subsidizing a prediction market be? Options include paying for initial entry, providing liquidity by taking the other side of some trade, multiplying ultimate winnings by some factor, or some combination of the above options. There is not much technical literature on the subject of subsidizing prediction markets, but reviewing what does exist suggests a few ways forward. Wolfers and Zitzewitz (2004) mention endowing traders with a portfolio or matching initial account contributions. "Subsidizing Prediction Markets" emphasizes the importance of using subsidy money to be a market maker that provides liquidity directly (there are limits that can be set to protect the market maker from unlimited downside).

There are a few alternatives for information discovery and aggregation, though none seem as well suited to the task of providing information to use in private valuations necessary for the LR mechanism as prediction markets. Polls with incentives don't have the same dynamic where people can continuously update their responses with ease in response to new information and the expectation indicated by the poll. Atanasov et al. (2017) found that prediction polls

slightly outperform prediction markets in some use cases for science forecasting, but the prediction polls used in the paper required weighting responses using an algorithm that takes into account the credibility of each respondent. This likely would not be possible for our application, at least initially.

It appears creating a prediction market associated with each LR project is the best available option currently. Beyond simply giving users of the system an idea of which projects to contribute to and information to pin down their private valuations, the LR mechanisms and prediction markets could also interact in a more formalized way. One option is to highlight projects with large prediction market participation on the website, particularly ones that appear to have a less than 50% chance of replication (the best range for promotion may be the area where there is doubt that the study will replicate but that the odds are not so low that there is widespread agreement). These studies are more likely to be important, highly cited, and valuable to replicate. A second way that the state of the prediction market might directly interact with the associated LR mechanism is featuring project with especially high trade volume, and even announcing additional subsidies for these sorts of projects. This would have the added benefit of generating buzz, much like promotions do. Another idea is to open the prediction markets up to the public even as funding is limited to a certain community. This may be a good policy because identity management is important for the LR mechanism but not for prediction markets, generally.

The environment in which the system for funding replication studies is situated would be a determining factor for many specific implementations details. For example, a replication funding mechanism geared towards members of a professional scientific organization would function somewhat differently from one set up for internal use at a university. Identity

management would be one of the most important first steps towards setting up in any environment, since LR mechanisms depend on it. In the case of scientific organizations, member accounts could be used for identity management purposes, while similar arrangements at universities would likely use pre-existing identification systems such as Princeton's netID.

Much of the work to be done in setting up a mechanism using the LR mechanism would be in lobbying relevant institutions and promoting the usefulness of the mechanism. Since replications are often contentious, there would need to be some measure of anonymity in contribution, as discussed above, and decision-makers would have to be convinced that adopting such a system would be beneficial to their field or department. Publicizing the mechanism would be important, since people who don't know about it can't contribute. A serious, highly relevant consideration for any mechanism of funding is legal issues such as liability and regulation. The Commodity Futures Trading Commission (generally, as some institutions received waivers) regulates any publicly available prediction markets, and it would be important to look into the legal implications of this (example of CFTC mandate: https://www.cftc.gov/PressRoom/PressReleases/pr6423-12). Technical details of setting up prediction markets would have to be specified precisely, as there are many possible designs.

The website we have designed and implemented is a demonstration of concept for funding science with an LR mechanism. It has the benefit of being easily updated so participants can react to the current funding level of the project. Implementation complexities include the algorithm that updates the price paid out by the contract (Dreber et al. (2015) provides a sketch of such an algorithm) and the particulars of market making with the subsidy money. We hope the website demonstrates, in a small way, the feasibility of using the LR mechanism to fund science, particularly replication studies.

Works Cited

Atanasov, Pavel, Phillip Rescober, Eric Stone, Samuel A. Swift, Emile Servan-Schreiber, Philip Tetlock, Lyle Ungar, Barbara Mellers. 2017. "Distilling the Wisdom of Crowds: Prediction Markets vs. Prediction Polls." *Management Science*. Vol. 63, No. 3.

Buterin, Vitalik, Zoë Hitzig, and E. Glen Weyl. 2018. "Liberal Radicalism: Formal Rules for a Society Neutral among Communities." *SSRN*. https://ssrn.com/abstract=3243656.

Dreber, Anna, Thomas Pfeiffer, Johan Almenberg, Siri Isaksson, Brad Wilson, Yiling Chen, Brian A. Nosek, and Magnus Johannesson. 2015. "Prediction markets in science." *Proceedings of the National Academy of Sciences*. Dec 2015, 112 (50) 15343-15347.

Roodman, David. "Hookworms and Malaria and Replications, Oh My!" Replication Network. https://replicationnetwork.com/2018/12/21/hookworms-and-malaria-and-replications-oh-my/ (retrieved 1/9/19).

Wolfers, Justin and Eric Zitzewitz. 2004. "Prediction Markets." *Journal of Economic Perspectives*. Vol. 18, No. 2, pp. 107-126.

'Zvi.' "Subsidizing Prediction Markets." LessWrong. https://www.lesswrong.com/posts/AeKS2m6uLM8RYfvND/subsidizing-prediction-markets (retrieved 1/10/19).