The ARPA Approach to Screening Prospective Energy Technology Research Investments

Proposal for the 2020 Research Conference on Economics of Innovation in the Energy Sector

Anna P. Goldstein

Background. The important role of publicly funded R&D in energy innovation is widely acknowledged. Technology cost reductions are needed for both clean energy supply technologies and energy-efficient end-use technologies. Within the US Department of Energy (DOE), which is the world's largest public spender on energy R&D, a majority of applied R&D and demonstration funds are directed toward lowering costs of existing technologies. However, progress in existing technologies alone will not be sufficient to reach ambitious emissions-reduction targets. New enabling technologies such as long-duration storage, low-carbon fuels, and carbon capture face prohibitively high capital requirements and/or long timelines for commercialization.

In 2009, the Advanced Research Projects Agency – Energy (ARPA-E) was established as a new funding office within DOE to help overcome long-term, high-risk barriers to energy innovation. ARPA-E was designed as an ARPA model agency to fill the "white space" left by other agencies and explore potential new breakthrough technologies [1].

One key aspect of the ARPA model is its approach to the uncertainty inherent to R&D. It breaks with the common practice of research funding agencies, which tend to distribute funds within groups of proposals on the basis of average peer review scores, in order from highest to lowest. There is mixed evidence in the literature on how well this conventional peer review performs at selecting the best proposed projects. One notable study found that some variation in research outcomes among funded NIH proposals could be explained by scoring percentile [2], while others have found that this effect is limited to the top percentiles [3]. Furthermore, a number of scholars have put forth the critique that peer review is biased against high-risk research; Boudreau et al. [4] found that reviewers are systematically biased against proposals with greater novelty.

ARPA-E uses an alternative method of selecting projects, where proposals are chosen for funding based on the judgment of individual expert staff members, rather than ranking external review scores. These staff members are scientists and engineers hired as short-term program managers; they seek external opinions on funding decisions without being bound to act on them. The use of individual discretion has been endorsed by some and criticized by others, but there have been no quantitative studies pointing to either its effectiveness or lack thereof. It is generally difficult to access data on unfunded projects, as funders do not make these review scores and comments available. Furthermore, much of the quantitative research on project selection that does exist focuses on US biomedical research support, specifically at the NIH, with rather than energy-relevant research programs.

The proposed paper will help fill this gap by providing the first empirical description of project selection in an ARPA model agency. I will apply advanced analytical methods to the question of which projects are selected when program staff are empowered to use their judgment rather than rely on ranking of peer review scores.

Preliminary evidence. Through a consulting arrangement with the National Academies during their evaluation of ARPA-E [5], I was able to construct a dataset of all ARPA-E proposals during the 2009-2015 timeframe. This dataset contains de-identified information on funded and unfunded applicants, in addition to measures of research output (publications, patenting, and market engagement activities).

Analysis of these data so far have revealed a number of interesting trends [6]:

(1) Approximately half of ARPA-E projects are "promoted", i.e. selected despite low review scores
(2) Proposals are more likely to be selected if reviewers disagree on the quality of the proposal, particularly if the proposal has at least one champion
(3) "Promoted" projects perform equally well on average compared to non-"promoted" projects on short-term research productivity metrics

Proposed methodology. Preliminary results show that project selection at ARPA-E is driven by factors other than peer review scores, thus introducing the opportunity to overcome risk aversion among peer reviewers. The first aim of this paper is to fully describe the relationship between review scores, funding decisions, and outcomes among funded projects. Beyond measuring the average trends, I will also explore whether the set of "promoted" projects differ from high-scoring projects in their distribution of outcomes, i.e. are there more failures in addition to more big wins?

Interviews with program managers indicate that some of their decisions are influenced heavily by review comments, but it is not known how this influence works in practice. The second aim of this paper is to use text-based analytical methods to explore comments from external reviewers, which are recorded in the dataset of proposals. Natural language processing will be used to identify common themes among funded and unfunded projects. Specifically, review comments will be processed using Latent Dirichlet Allocation (LDA) to generate topic models, and the presence and prevalence of key topics will be regressed against the funding decision for the proposal. I will also explore the use of machine learning to determine whether the selections of the ARPA-E program staff can be reproduced out-of-sample by a supervised model.

Impact. The outcome of this study will be a thorough quantitative description of how project selection relates to the upfront risk tolerance of a public research funder. As investments in energy R&D scale up worldwide, some have suggested ARPA-E as a model for other governments to replicate. This study will improve the ability of policymakers and analysts to make evidence-based decisions regarding research management strategies, thereby maximizing program impact and accelerating the transition to a sustainable global energy system.

References

[1] P. Azoulay, E. R. H. Fuchs, A. P. Goldstein, and M. Kearney, "Funding Breakthrough Research: Promises and Challenges of the 'ARPA Model,'" *Innov. Policy Econ.*, vol. 19, pp. 69–96, 2019.

[2] D. Li and L. Agha, "Big names or big ideas: Do peer-review panels select the best science proposals?," *Science*, vol. 348, no. 6233, pp. 434–438, 2015.

[3] F. C. Fang, A. Bowen, and A. Casadevall, "NIH peer review percentile scores are poorly predictive of grant productivity," *Elife*, vol. 5, pp. 1–6, 2016.

[4] K. J. Boudreau, E. Guinan, K. R. Lakhani, and C. Riedl, "Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance and Resource Allocation in Science," *Manage. Sci.*, 2016.

[5] National Academies of Sciences Engineering and Medicine, *An Assessment of ARPA-E*. 2017.

[6] A. P. Goldstein and M. Kearney, "Uncertainty and Individual Discretion in Allocating Research Funds," 2017.