This document details the planned data analysis approach to examine the factors influencing the outcome of ARPA-E projects, and serves as a means of distinguishing between hypothesis generation and hypothesis testing [1,2]. Full coding of the project outcomes has not yet been completed.

Introduction

Since it began funding projects in 2009, ARPA-E has used two methods to solicit new project ideas. Designed programs are crafted by agency staff, in conjunction with expert workshops, and specify a specific technology area, along with performance and cost targets. Project proposals are then solicited to meet those objectives. The agency also uses OPEN programs, which allow for proposal submissions for any energy technology, and without cost or performance targets specified by expert panels. Now, 10 years after the agency began awarding funding, over 450 projects have been completed, and we can use the outcomes of those projects to assess how effective experts are at establishing technology performance and cost targets that align with market demand/adoption of technology.

Aims/Hypotheses

1) Generate a dataset of outcomes for completed ARPA-E projects. The outcomes will focus on whether the companies involved are still following the technology targets outlined during the program solicitation/at the funding stage (persist), whether they've pivoted to different goals or technological approaches (pivot), or if they have left the market entirely (perish). Project statuses will be assigned by (at least) two independent coders, who will then reconcile their analysis.
2) Assess whether there is a significant difference in the frequency of the projects funded through OPEN solicitations or directed programs that are continuing to pursue the initial targets specified (persist), pivoted from initial objectives (pivot), or have left the market entirely (perish). If the expert-derived targets used in the structured programs are better than the market-derived objectives from OPEN programs, the proportion of projects persisting with original strategy should be higher, and the proportion of projects that perish should be lower.
3) (Data pending) Assess whether there is a significant difference in the frequency or amount of follow-on funding that projects funded through OPEN solicitations or directed programs are able to obtain. If the expert-derived targets used in the structured programs are better than the market-derived objectives from OPEN programs, the proportion of projects that obtain, and the follow-on funding amounts obtained by projects funded through designed programs should be higher than for OPEN projects.

Null Hypothesis:

1. The distribution of projects funded through OPEN and designed programs between the three groups (persevere, pivot, perish) is equivalent.
2. (Data pending) The distribution of project follow-on funding that projects funded through OPEN and designed programs is equivalent.

Coding

Coding is a general term for conceptualizing qualitative data to systematically determine core concepts or outcomes [3]. Here, projects will be coded independently and then reconciled amongst coders. We borrow from innovation literature to identify three possible categories of project success, adhering to the original targets and technology outlined in the ARPA-E award (persisting), making significant changes to the technology or market goals since the ARPA-E award, but still pursuing a technology in the field (pivoting), or exiting the related energy technology market entirely (perishing).[4,5] We will provide a calculation of the Cohen's kappa of agreement between coders before reconciliation.

Sample

Project information for the coding analysis is drawn from the publicly available information about projects and awards provided by ARPA-E. We only include projects which were completed, or would have been completed according to their original timeline, on or before January 1, 2019. For each program, we collect the award amount, project start and end dates, technology categories, project status, and location (state) of the program awardee. Within this sample, there are 465 projects, 32 Designed programs, and 3 calls for OPEN proposals. Table 1 summarizes the number of projects included in the sample. Within this sample, there are two possible project statuses: cancelled and alumni. Cancelled projects had their awards terminated ahead of the project timeline. Figure 1 shows the distribution of project awards for the OPEN programs and designed programs included in our sample. Awarded projects are also sorted into technology categories. Table 2 lists the number of projects from our sample awarded to each category, along with the total ARPA-E award amounts per technology category. ARPA-E has also been collecting information about the follow-on funding that projects receive, whether from private or other public sources. If possible, we will obtain this data and use it as an outcome in addition to the coding analysis.

Table 1: Number of projects funded through ARPA-E by program type and project status

	Program Type		
Project Status	OPEN	Directed	All
Active	28	231	259
Alumni	100	326	426
Cancelled	14	31	45
Recently Cancelled	2	10	12
Total	144	598	742
Included in analysis	**114**	**357**	**471**

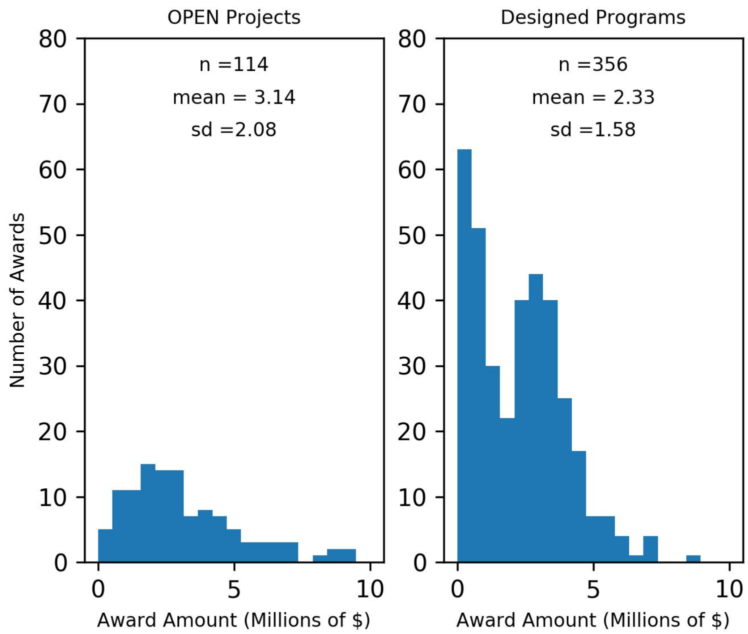

Figure 1: Histograms of award amounts for OPEN and designed programs.

Table 2: Projects and total awards by technology category

Technology Category	Number of projects	Total Amount Awarded over all projects (millions)
Transportation Fuels	81	$257
Distributed Generation	71	$165
Resource Efficiency	67	$166
Storage	60	$146
Transportation Storage	49	$115
Electrical Efficiency	36	$77
Building Efficiency	34	$86
Grid	28	$70
Manufacturing Efficiency	26	$48
Centralized Generation	7	$29
Transportation Vehicles	7	$13
Transportation Network	5	$15

Preliminary Power Analysis
Power analysis is a means of assessing the likelihood of an experiment will have a statistically significant result. Here, we use a chi square test to determine whether the frequency of projects that persevere, pivot, and perish is different for OPEN and designed programs. The power of an experiment is dependent on the effect size, the sample size, and the significance criterion, although any of these parameters is also a function of the other three parameters [6]. Here, the sample size is constrained by the number of completed projects. We assume a significance criterion (α) of 0.05 (meaning that the probability of a Type 1 error, or falsely concluding that the frequency of the three outcomes is unequal between the two groups, is 0.05). Figure 1 shows how the power of the experiment varies with the sample size, given different effect sizes. We see that for the number of projects already completed, the experiment has a high power (>0.8) even for small (<0.2) effect sizes.

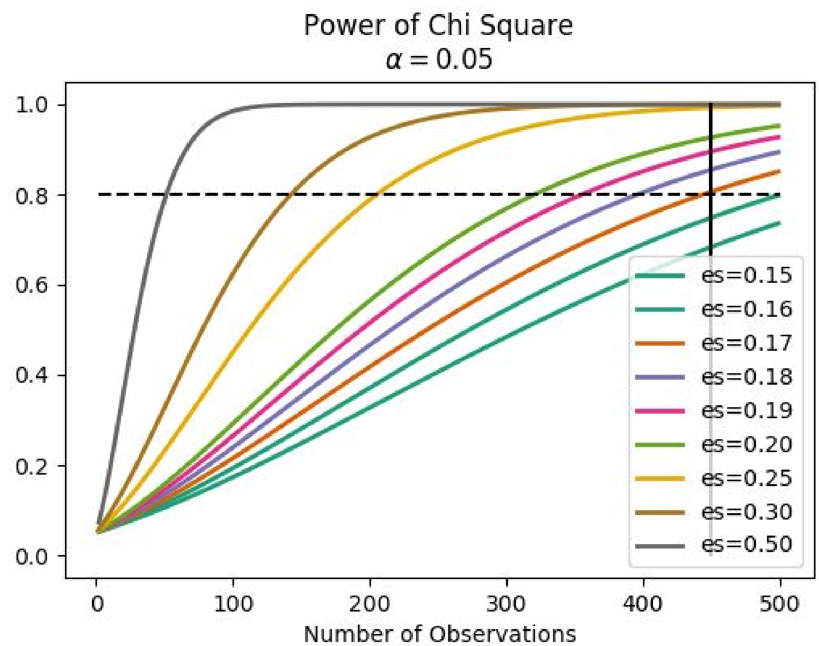

Figure 2: Power analysis results. For even relatively small effects, there is an 80% probability that the chi-square test will correctly identify two different distributions of project outcomes for the OPEN and designed projects.

Model construction
Beyond the type of program the project is funded through, other factors could impact the outcome of a project. We plan to use logistic regression to measure the probability of each of the three discrete outcomes. Our models account for fixed effects associated with the year the project was funded, and random effects including the awardee type (whether a for-profit or non-profit organization) and primary technology category. For the award amount variable and follow-on funding variable, which are continuous, we will also test for potential transformations of the variable using GAM and Stukel's test. Table 3 lists the variables included in each model we plan to examine.

We'll use logistic regressions based on the discrete outcomes for different projects, examining attributes like whether the program was solicited through an OPEN or designed program, the initial award amount, whether the awardee was a for-profit or non-profit entity, the technology category, and whether the awardee had diverse partners (i.e. if the primary awardee was a for-profit company with non-profit partners). We include checks on fixed effects for the starting year of the project.

Pending the availability of the follow-on funding data, we will also use this data as an outcome. The exact model specification will depend on the distribution of follow-on funding amongst the projects. If the follow-on funding is well-distributed amongst the total number of projects included in the sample, we will use a linear regression model, considering the same models specified in Table 3. However, if there are many projects that received no follow-on funding, a linear model will not be an accurate predictor of the relationship between our model and the amount of follow-on funding received. Instead, we would then use a logistic regression with the discrete outcomes of received or did not receive follow-on funding, again examining the variables specified in Table 3.

Table 3: Regression models. Logistic regressions will be repeated for each of the 3 possible project outcomes (persist, pivot, perish), and pending data availability, the frequency (logistic) or amount (linear) of follow-on project funding. For each model, we'll estimate the coefficients, standard errors, and significance.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	B (SE)	B (SE)	B (SE)	B (SE)	B (SE)	B (SE)
x_{OPEN}	B (SE)	B (SE)	B (SE)	B (SE)	B (SE)	B (SE)
$x_{AWARDAMOUNT}$		B (SE)				B (SE)
$x_{AWARDEETYPE}$			B (SE)			B (SE)
$x_{TECHCATEGORY}$				B (SE)		B (SE)
$x_{PARTNERS}$					B(SE)	B (SE)
Fixed Effects-Starting Year	Y	Y	Y	Y	Y	Y
N	X	X	X	X	X	X

Acknowledgements
Many thanks to Casey Canfield for her discussions on model design and registration.

References

1. Nosek, B. A. & Lakens, D. Registered reports: A method to increase the credibility of published results. *Soc. Psychol.* **45**, 137–141 (2014).
2. Miguel, E. *et al.* Social science. Promoting transparency in social science research. *Science* **343**, 30–31 (2014).
3. Strauss, A. *Qualitative analysis for social scientists*. (Cambridge University Press, 1987).
4. Crilly, N. ‘Fixation’ and ‘the pivot’: balancing persistence with flexibility in design and entrepreneurship. *International Journal of Design Creativity and Innovation* **6**, 52–65 (2018).
5. Roland M. Mueller, K. T. Design Thinking vs. Lean Startup: A Comparison of Two User-Driven Innovation Strategies. in *Leading Innovation Through Design: Proceedings of the DMI* 151–161 (2012).
6. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. (Lawrence Erlbaum Associates, 1988).