

# Power Calculations: Research Funding Lottery

Authored by:

McKenzie Leier, Policy Manager at the Science for Progress Initiative at J-PAL Global

Lisa Turley Smith, Research Manager at J-PAL North America

## Scenario

A country's science funding agency has decided to allocate a portion of its budget to a lottery funding system, where proposals are chosen through a random lottery to receive funding. The country gives out \$800 million in grants to fund scientific research each year.

The grant program that will use the partial lottery awards up to \$400,000 grants for unconventional and innovative research ideas. On average a proposal receives \$250,000. The program has approximately \$50 million in grants to award each year. In an average year, the program receives 700 applications and grants awards to roughly 200 applicants. Applicants cannot reapply with the same proposal in following years. The country's government is willing to work with you to run a randomized evaluation to understand the impact of their funding.

The government is not willing to fund any low-quality proposals, and would like to reserve some spots for "must fund" proposals. Based on previous years, they estimate around 500 proposals will meet their minimum standards and 10 proposals will be "must fund" proposals. This will leave around 490 "mid-tier" proposals that can be randomized for funding decisions. They will allow you to run your evaluation for up to 5 years of awarding funding and follow the research output for 5 years after receiving funding. This means that approximately 2,450 proposals will be randomized.

The research question: what is the effect of science funding on a researcher's or research team's scientific output (e.g., publish more papers) five years after receiving funding?

You are now tasked with designing and implementing an evaluation. **Running power calculations as you design an evaluation is critical to know if you will be able to detect a reasonable effect.** For more information on power calculations in general, please refer to [J-PAL's Research Resource on power calculations](#).

The remainder of this document will walk through the specific considerations of power calculations for this scenario.

## Minimum Detectable Effect

Generally, researchers run power calculations to determine the minimum sample size that would **allow them to have enough certainty of some impact of the program**. This is called the minimum detectable effect (MDE). The MDE is often based on prior literature or by determining the smallest effect that would be relevant or interesting.

For example, if in reality there was a 5% increase in publication output, but you were only powered to measure a 10% increase or decrease in output, are you comfortable with potentially finding a null result in your study? Choosing the MDE can be more of an art than a science, but consider what justification you are using for the effect size you choose.

## Sample Size

**How many proposals were eligible to receive funding and able to be randomized?**

In some cases, you are not able to increase your sample size. There is no way to force people to apply to get funding. However, depending on the constraints given by the funding agency, you could increase your sample size giving out more grants for a smaller dollar amount or adjust the “must fund” or “do not fund” tiers. .

Increasing your sample size is the surest way to increase power, enabling you to detect a smaller effect size in your study.

## Standard Deviation of Outcome

**What is the expected variance in the number of publications in the sample?**

To estimate the variance of the outcome variable, search through previous research, observational literature, or publicly-available data. For example, Jacob & Lefgren (2011) estimated the standard deviation of the publications in their study in the untreated group after 5 years was 15.7. This may be a good starting point for this study. You may also be able to estimate the variance through the funder’s operational records.

It is also important to do a gut check and consider if there is anything about this study that would make it very different from the estimates in previous studies. It can be helpful to test a range of standard deviations (or use the highest in the range) to be sure that you will have sufficient sample size in a scenario where the standard deviation of your outcome is larger than expected.

## Timing of Randomization

**At what point in the process should randomization occur?**

It is generally best to randomize later in the process than earlier, although there are often constraints by implementing partners. In the scientific funding process, as with any funding process, not all proposals will meet a minimum threshold for funding. By waiting until after the proposals go through an initial screening process, the study will have more power. Otherwise some proposals who were randomized into the “treatment” group will essentially be noncompliers because they could not receive funding.

## Take-up and Compliance

**What share of applicants, if offered funding, will not accept it (for any reason)? And what share of applicants denied funding (but considered fundable) will still get funding?**

You may be able to get a sense from historical data what share (if any) are offered funding, but do not accept it for whatever reason. Because researchers have applied for the funding, it's reasonable to assume that take-up would be very close to 100%, if not 100%.

There are two possible research questions to consider for proposals randomized into the control group of this experiment (not receiving funding):

- Are you interested in the impact if a proposal doesn't get funding from this particular funding agency?
- Or are you interested in the impact of not getting funded from any source at all?

If interested in the former, compliance in the control group should be 100%. You have less control over the latter research question, but if this is the question of interest, you will likely need a larger sample size to detect an effect to factor in the proposals that receive funding from other sources. You must also consider how you will track and measure this.

## Attrition

**What number of proposals will you not be able to track or only receive partial funding for whatever reason?**

In this scenario, it is unlikely that you would be unable to track the outcomes of the research teams who applied in either treatment or control. However, if payments were allocated over many years and a proposal or research team became ineligible for funding in the middle of the study, this would be considered as partial treatment and often treated as attrition.

## Proportion in treatment

### What proportion of proposals randomized will receive funding?

Ideally, you would be able to randomize proposals with an equal likelihood of being funded or not. The farther you move away from a 50-50 split, the worse power will be. However, total sample size still improves power, and so you should not decrease your sample size to move toward an even split.

## Controlling for Covariates

If you know that certain characteristics impact the outcome, you can add covariates to your model, which can significantly improve power through decreasing the standard deviation of the outcome.

In this example, we would likely want to control for past publications. Past research shows controlling for covariates like past publications significantly increases precision. Based on the effect sizes, published  $r$ -squared, and reported lagged publications in Jacob & Lefgren (2011), we can estimate how the standard deviation will change if we introduce some covariates and full covariates.

With no covariates, our standard deviation would be 15.7 papers published over the course of 5 years. By adding some covariates ( $r^2 = 0.29$ ), our standard deviation is reduced to 13.2. By adding full covariates ( $r^2 = 0.62$ ), our standard deviation is reduced to 9.68.

---

## Power Calculation Inputs

**Significance level: 0.05**

**Power: 0.8**

It is standard practice to have 80% power and a 5% significance level.

**Take-up (in treatment): 95%**

**Take-up (in control): 0%**

**Attrition: 0%**

We expect that all proposals offered funding will accept, but leave some room for a few to reject funding for whatever reason. We expect to be able to track all outcomes across time and for no research teams to receive partial funding.

## Minimum Detectable Effect: 1.25 additional papers in 5 years

We want to be able to detect a difference in publishing 1.25 more additional papers over the course of 5 years, based on previous quasi-experimental research by Benavente et al 2012 of Chile's FONDECYT program and Jacob & Lefgren (2011).

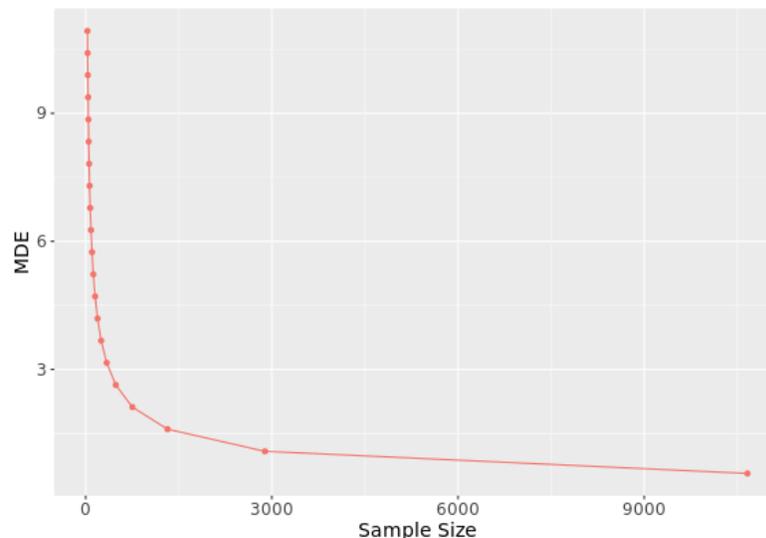
## Standard Deviation of Outcome: 9.68 papers in 5 years

Based on Jacob & Lefgren (2011), we have estimated the standard deviation of the outcome to be 15.7 papers in 5 years, but we intend to add full covariates to our model, thus we will run our power calculations with a standard deviation of 9.68 papers in 5 years.

## Proportion in Treatment: 0.408

Of the 700 proposals, about 500 are expected to be fundable and of these, 10 will be “must fund” proposals that the government agency does not want to randomize, leaving 490 to be randomized. With a \$50 million pool and an average grant size of \$250,000 there will be sufficient funds to award funding to about 200 proposals. ( $200/490 = 0.408$ )

**To detect an MDE of 1.25, the required sample size is 2160.**



To vary these numbers, or to run your own preliminary power calculations, check out [J-PAL's Power Calculator](#). See also our sample STATA and R code on the [J-PAL Research Resources page](#).

**Given you can run this experiment over the course of 5 years, giving you an expected total sample of 2,450 unique proposals, you are sufficiently powered at 80% given the above parameters.**

## Sensitivity Checks

Given that power calculations involve a lot of assumptions, it is wise to perform sensitivity analyses to test how power changes with any changes to key assumptions.

---

Example 1: Let's say that the government changes their mind about the number of "must fund" proposals after initial review. They now have 35 proposals that are "must fund" and they do not want to randomize. This change leaves 465 proposals to be randomized each year, changing the proportion in treatment to 0.430 (200/465) and your expected sample size over 5 years to 2,325.

**The increased proportion in treatment means that your new required sample size is 2,128, a slight decrease from the previous estimate. This means you would still have sufficient power for the study.**

Example 2: Next let's assume that a political change happens and the agency now has only \$30 million to give away each year. They plan to maintain a grant size of \$250,000. This means that only 120 proposals can be funded each year. The proportion in treatment would change to 0.245 (120/490).

**The required sample size is now 2,546, greater than the 2,450 expected unique proposals over 5 years. This would mean you would be slightly underpowered to detect a change of 1.25 papers in 5 years.**

Example 3: Finally, let's say that we determine it makes more sense to only use partial covariates. It's important to understand if the study would still be powered with a larger standard deviation since this is fairly unknown. As a reminder, the standard deviation would be 13.2 with partial covariates.

**The required sample size is now 4,016, much greater than the 2,450 expected unique proposals over 5 years. If this is closer to the true standard deviation, your experiment would be very underpowered to detect an effect of 1.25 papers over 5 years.**

---

For any power calculations, researchers should consider which assumptions may change and perform sensitivity analyses for best and worst case scenarios.

## References

- Barnett, A., Blakely, T., Liu, M., Garland, L., & Clarke, P. (2024). The impact of winning funding on researcher productivity, results from a randomized trial. *Science and Public Policy*, 51(6), 1042–1050. <https://doi.org/10.1093/scipol/scae045>
- Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of Public Economics*, 95(9), 1168–1177. <https://doi.org/10.1016/j.jpubeco.2011.05.005>