Impact analysis for investments in science and technology

I. Introduction

The recent stagnation in federal funding has raised important concerns about the advancement of science.[1] While, on aggregate, private foundations support a small share of scientific endeavors,[2] they have been increasingly active in addressing funding shortages.[3]

As with any other fund-granting institution, private foundations must ensure that their resources are spent as productively as possible. Evaluating the returns to one's investments may be difficult because the process of scientific discovery is inherently uncertain and researchers may go down many "wrong" paths before making a breakthrough discovery. Funders of scientific research face the challenge of supporting scientists as they take risks and make mistakes while limiting the extent to which they waste valuable resources on projects with a low likelihood of success.

Striking the right balance between these often opposing goals is difficult and funders should expect that this process will be iterative. Furthermore, the "right" balance will depend on the goals of the funder. For instance, some institutions may focus on supporting projects in underexplored research areas, others could focus on the career development of young scientists or targeted minorities.

II. Goals and outline

This document is meant to introduce potential scientific funders to best practices for structuring scientific assessments so that their impact can be evaluated in the future. It is organized into three sections:

1. Developing goals and expectations: many organizations, public, for profit, and non profit, are involved in scientific funding. Here, we discuss various factors a new funder may want to consider when setting the focus of their organization.

2. Designing for evaluation: the impact of science funding can be evaluated scientifically. This section provides a practical guide to how funders can design their process so that its impact can be readily evaluated using scientific techniques (e.g. randomized trials or control groups).

3. Tracking appropriate outcomes: depending on one's goals, there are many ways to measure scientific output. Here we discuss metrics for publication, training, and commercial output, and provide suggestions on which may be the most relevant for particular organizations.

---

[1] http://www.nature.com/nmeth/journal/v13/n7/full/nmeth.3923.html.
[2] http://www.nature.com/nchembio/journal/v4/n9/full/nchembio0908-519.html#B4.
[3] https://www.insidephilanthropy.com/home/2016/5/18/just-how-much-are-private-funders-giving-to-basic-science-re.html

III. Developing goals and expectations

First and foremost, investing in scientific research requires patience and a tolerance for failure. Imagine a $1 million project with a miniscule 0.00001% chance of curing cancer. In practice, relatively few individual organizations have the risk tolerance to spend $1 million on a single investment that will fail 99.999% of the time.

Yet imagine if there are 200,000 such potential projects, all with a 0.00001% chance of success. If their probabilities of success are uncorrelated, then these projects collectively represent a $200 billion dollar investment with an 87% chance of curing cancer. Because of the enormous social value of finding a cure for cancer, almost everyone would agree that this is a worthwhile portfolio of investments. And yet the risk aversion of individual organizations makes it less likely that any of the component $1 million dollar investments would get made in the first place.

As this simple example illustrates, it is important for funders to think of their scientific investments as part of a broader social portfolio of investments made by government, private sector, and other non-profit entities. While the success rate of individual projects may not be high, the impact of a portfolio of diverse approaches to solving big problems can be both high and also relatively low risk.

The key to making this happen is to invest in high potential projects that would not be otherwise supported by traditional funding institutions. In our above example, diversity is key because if one invested in 200,000 identical projects, one would simply loose $200 billion dollars 99.999% of the time. Diversification may involve funding projects across different research disciplines or projects adopting different methodologies. It may also involve targeting the career development of different types of researchers. For example, funders may think about:

1. Providing funding to scientists and ideas in areas that are not targeted by major funders such as the National Institutes of Health or the National Science Foundation.
2. Similarly, these organizations have low rates of funding and funders can step in and provide support not necessarily for the best scientists who have many funding options, but for highly qualified scientists that fall just short of the funding threshold.
3. Support for early stage ideas that enable scientists to get traction and obtain larger sources of funding
4. Support for translational areas that help science make it out of the lab.

IV. Designing for Evaluation

As with any other investment, funders of scientific research should rightly want to understand the impact of their resources. This provides an opportunity to build on strengths in their existing funding model and to improve on weaknesses.

Yet evaluation is difficult without some initial planning. Imagine if a foundation provides a grant to a scientist and 2 years later she has trained 2 graduate students, and published 10 additional articles, several of them in prominent journals. Yet, in order to assess the impact of this grant, it is not simply enough to count. Rather, one needs to understand what her research outcomes would have been like had she not received any support: funding has an impact if enables to do more or better research than in the absence of that funding. This is analogous to the challenge that scientists face when assessing, for example, the impact of a medical treatment: how do we know whether the patient got better because of the treatment or because of something else?

Rejected applicants as a control group: In medicine, scientists address this challenge by comparing outcomes for treated patients with outcomes for a control group of similar patients who were not treated. Funders of scientific research can do the same thing by collecting data on similar scientists who were not funded.

The easiest way to do this is to keep track of applicants who are similarly qualified but not selected due to resource constraints. When it comes time to evaluate the impact of funding on research output, one should collect data on research output from *both* of these groups. If unfunded scientists produce less output, one can infer that funding has an impact on the careers of recipients.

As a practical matter, applicants who have been rejected will likely be less responsive to requests for data than funded applicants. To get around this, one should ask for enough information *at the time of application* to identify research outcomes later on. In the United States, this can be as simple asking for an ORCID, which is a unique digital identifier most scientists already have. When it comes time to evaluate, data on publications can be linked via ORCIDs for both funded and unfunded applicants. This also has the benefit of automating data collection for funded applicants as well.[4]

[Are there IDs we can ask for that allow linkage to non-publication related outcomes?]

Refining the control group: The approach described above is the simplest way to conduct basic impact evaluation. Its validity, however, depends on how similar rejected applicants are to funded applicants. If they are substantially less qualified, then we may expect them to have worse research outcomes that funded applicants, even if the latter group were not funded: this would tend to overstate the impact of funding.

One can get around this in two ways. The first is to randomizing who gets funding. This is akin to randomization in medical trials or A/B testing in business settings. Randomization ensures that “treatment” and “control” scientists are comparable. Funders may understandably not want to randomly allocate their scarce funds to unqualified scientists. One way to address this concern is to first exclude those applicants who are deemed unaccepted and then randomize within the set of acceptable applicants.

[4] It is important to use the same process to identify outcomes for funded and unfunded applicants. If one uses a more in-depth process to identify research for funded applicants, then one may find more research for this group and mistakenly attribute this difference in output to the impact of the funding when, in reality, it was only a difference in data collection approaches across the two groups.

Another approach makes use of basic data on the characteristics of all applicants. A starter list may be as follows: ORCID, highest education, year of graduation, undergraduate and graduate institution, prior funding history, and keywords describing primary fields of research.[5] These variables can be used to make sure that one is comparing comparable treatment and control scientists: ones who differ only in finding and not in terms of publication history, education, and research area. One may also ask applicants to provide a list of peer scientists at the time of application, though keeping in mind that self-reported peers may not be true peers.

Impacts on research areas: So far we have discussed evaluation in terms of understanding the impact of funding on a specific scientist. Funders, however, may be interested in understanding their impact on a field of research as a whole. In this view, it is not enough to compare treatment and control outcomes at the level of an individual scientist. This is because two applicants may have similar ideas: if funding enables one scientist to publish her results ahead of another, that yields a big impact from the perspective of her individual output, but it may not yield as large an impact her field because that research idea would have been published regardless. In order to assess the impact of funding on an entire area, one can still apply the same techniques as the ones described above, but focusing on fields rather than individuals as the unit of "treatment." For example, if one decides to focus funding on translational research in diabetes, one may compare the number of new clinical trials in diabetes to those in other similar disease areas.

V. Tracking appropriate outcomes

One can assess the effectiveness of a grant by evaluating the difference in research outcomes associated with funded and non-funded groups.

Before discussing some potential outcomes to consider, we note that it is important to keep in mind that the outcomes one tracks become the incentives for one's grant recipient. Programs that only track publications, for instance, will provide recipients with an incentive to publish, but may not necessarily give them incentives to commercialize their work. By contrast, counting patents may provide incentives for recipients to patent unimportant work.

Publication and citation based measures: The most common way to assessing research productivity is to count publications and citations. Researchers have found a positive correlation between these measures of output and breakthrough discoveries.[6] With ORCIDs, it is easy to link researchers to their academic output and one may consider looking at publications, publications

---

[5] For example, for life scientists, one may want to ask for MeSH (Medical Subject Headings) keywords. For chemical scientists, one may want to ask for CAS (Chemical Abstracts Service) keywords.

[6] See, for instance, Lawani, S.M. and Bayer, A.E., 1983. Validity of citation criteria for assessing the influence of scientific publications: New evidence with peer assessment. *Journal of the American Society for Information Science*, *34*(1), pp.59-66.

in top journals, or high impact publications.[7] This should be considered a basic part of any impact evaluation.

[Any studies on correlation between bibliometric outcomes and other things?]

Commercial or applied impact: A weakness of publication-based measures is that they may fail to capture the impact of a scientist or research program outside of academia. If a funding organization's goals are focused on applied impact, one may want to consider other metrics such as patents generated, new clinical trials, incorporated start ups, etc.

[What is the best practice ID for funders to collect here?]

Career outcomes: Funders may be interested in supporting creative individuals rather than specific projects, in which case impact assessments should include measures of career traction or influence. For example, one can examine the career evolution of funded scientists (i.e. whether they are appointed professors or, if they are already professors, whether they are promoted to the rank of associate professors or full professors) and the quality of their placement. Funders may also be interested in the career outcomes of young researchers working in a funded scientist's laboratory. In this case, impact assessments should include measures of career traction or influence of these researchers. Information regarding the scientists' careers could be retrieved from LinkedIn or the universities' websites.

the number of students the researcher trains, whether he or she obtains additional grants or is promoted.

[This is harder to collect for applicants. Do we have best practice suggestions for this?]

VI. Additional resources and concluding thoughts

Private foundations have an important role to play in stimulating scientific research, especially where public funding may be limited due to budget, vision, or scope. This document provides an introduction to the broad issues that funders face in setting up organizations that can be evaluated and improved upon.

[Can we provide a list of additional resources here?]

Finally, we also stress that the impact of private foundations will only be felt if these organizations make a sustained commitment to funding science over time, and across projects. This means being failure tolerant and aiming to provide a continuous source of funding, rather than one-off funds. This is because it is difficult to rely on the success of any particular grant recipient; rather, it is the accumulation of support for an area that attracts talent and increases the probability that a few projects will have an enduring impact.

---

[7] One way to measure high impact is to examine a cohort of articles in the same field published in the same year and say that a publication is "high impact" if it is cited in the top 1% of 5% of its cohort.