"Building Programme Evaluation Into the Design of Public Research-Support Programmes" (Jaffe 2002)

The Question: What are the major challenges of evaluating the impact of research funding programmes? What features of research grant programmes would allow for more credible evaluation of their impacts?

The Challenges: This article highlights two major issues with evaluating the effectiveness of research-support programmes. Programme impact is the average effect of programme funds on the entities receiving the grants. The first problem is "selectivity," or "selection bias." Funding agencies actively seek to fund the most promising projects and the best teams. Thus, the projects with the highest chance of success also have the highest chance of getting funding. As a result, winning a grant may not *cause* subsequent research successes; rather, the promising nature of the project resulted both in winning the grant and in the project's ultimate successes.

A second problem is "additivity." How can research success be attributed to any one source of funding, since projects often receive support from multiple sources? Research grants may not increase total research spending if winners simply substitute the grants for other research resources, i.e. the grant 'crowds out' other funds. Conversely, winning a prestigious grant may signal that a project is worthwhile, resulting in 'crowding in' of funds from many sources to a single project. Understanding the full impact of a particular grant program requires an understanding of the broader funding environment.

The Lessons: Grant agencies can award funds in ways that enable more reliable after-the-fact evaluations of programme impact.

The first strategy to combat selectivity is randomization. An agency can identify a group of potential awardees and pick winners randomly within that group. If winners are truly chosen randomly, then any difference in subsequent output between winners and non-winners should be due to receiving the grant, rather than inherent differences in project potential. To address concerns about not funding highly promising projects, there could also be multiple groups of potential awardees with different random probabilities of winning an award. For example, projects in the 'high priority' group could have a higher random chance of winning than those in a 'marginal' group. The average programme impact would then be a weighted average of the treatment effects for the two groups.

The second strategy is a regression discontinuity (RD) design. This strategy takes advantage of a discontinuity in the probability of funding at some threshold level of project quality. For example, all projects may be given an overall quality score. Then, a cutoff score is established such that all projects with scores above that cutoff will be funded. Assuming that applications just above or just below the cutoff are very similar, differences in outcomes for these applicants will be due to the grant itself rather than unobserved differences in applicant or researcher quality. These differences in outcomes show the impact of the research funds specifically for projects with quality around that particular cutoff value. With additional assumptions, it is possible to estimate the average treatment effect of research funds for all funded projects. This evaluation approach can be used as long as the cutoff score strongly predicts the probability of grant funding. If there are a few decisions that deviate from the cutoff rule for reasons unrelated to project quality (for instance, increasing the diversity of grant awardees), this regression discontinuity approach is still valid.