Appendix I: Approach to Matching

The study design requires the construction of project- and participant-level comparison groups. First, a matched comparison group of non-PIRE, NSF-funded projects that do not require an international collaboration but which are similar to PIRE awards will be constructed. From each set of matched projects, participant-level comparison groups will be formed.

This appendix describes the criteria for matching PIRE and comparison projects and for matching participants within each set of matched projects.

Project Matches

To construct the matched comparison group of NSF-funded projects, candidate comparison projects funded by non-PIRE NSF programs will be identified and matched to each PIRE project included in the evaluation on the basis of the following criteria:

- The PI of the comparison project must not also be (or have been) PI of a PIRE project;
- The comparison project duration (award period) must be within 1 year of the duration of the PIRE project (e.g., if the PIRE award is 5 years in duration, the comparison award must be 4, 5 or 6 years in duration);
- The comparison project award amount must be within 20 percent of the PIRE project award amount;
- The comparison project must be funded by an NSF program that does not *require* an international component (projects in which international collaboration is not required but is encouraged, optional, or not mentioned explicitly meet this criterion);
- The comparison project must have at least one graduate student participant;
- The comparison project must have research as its primary focus (PIRE projects have research as a primary focus);
- The comparison project must engage expertise in scientific or engineering disciplines that overlap at least partially with those involved of the PIRE project (i.e., the NSF directorate in which the program funding the comparison award is housed must be commensurate with at least one of the NSF directorates that corresponds to the PIRE award and the scientific or engineering disciplines in which comparison project personnel have expertise must be the same as the expertise of at least one member of the PIRE project team);[1]

[1] NSF awards are typically made from programs within divisions that sit within directorates; for programs like PIRE that can cut across programs, divisions or directorates, NSF has, *a priori*, identified which directorates each of the 59 PIRE awards corresponds to – this was done partly to facilitate identification of proposal reviewers with the necessary expertise to judge the merits of the proposal.

- The comparison project must include at least two institutions (U.S. or foreign) total (by definition, a PIRE award includes at least one U.S. and at least one foreign institution).

If multiple matches for a PIRE project are identified, we will select one or two projects at random for inclusion in the comparison group.)

Participant Matches

Within matched projects, participants in PIRE and comparison projects will be matched using a greedy matching distance technique. Matches will be based on a pre-project measure of one of the key research outcomes for the study (e.g., the number of publications prior to project participation; the average citation measure for publications prior to project participation).

There are many multivariate matching techniques for identifying potential matches for each PIRE participant. One frequently used technique, propensity score matching (PSM), would require a large sample of PIRE participants to create reliable propensity scores. Additionally, to create propensity scores the evaluation would require a large number of variables from existing data than are unavailable on participants prior to data collection. Thus, we will use another matching technique, "greedy matching" to match participants within each PIRE-comparison project pair. Unlike PSM techniques, greedy matching is not sensitive to sample size. "Greedy" matching begins by (1) randomly sorting the PIRE participants and the comparison project participants; (2) then, the first PIRE case in the list is matched to the closest comparison participant—namely, the one most similar to the PIRE participant. Next, (3) the second PIRE participant is matched to its closest comparison participant among those remaining; and (4) this process repeats until all PIRE participants have been matched.

Here, "similarity" between a PIRE and comparison participants is defined as the average distance per year between the PIRE and comparison participants' score on a pre-test (e.g., the number of publications prior to project participation).

For example, using the number of publications as the pre-test measure, the average distance per year between the i^{th} PIRE participant and the j^{th} comparison participant is defined as

$$D_{ij}=\frac{\sum_{n=1}^{N} X_n^1 - X_n^0}{N}$$; where $X^1=\{ X_1^1, X_2^1, X_{3,}^1 \ldots X_N^1 \}$ is the number of publications across N years for PIRE participants and $X^0=\{ X_1^0, X_2^0, X_{3,}^0 \ldots X_N^0 \}$ is the number of publications across N years for Comparison participants.

We will use a SAS macro called "gmatch[2]" to calculate distances and find matches for each PIRE participant. Within each matched set of projects, the SAS gmatch macro will first sort the PIRE participants in a random order. Starting with the first PIRE participant in the randomly ordered list, it will identify the comparison participant who had the smallest average distance to the PIRE participant's pretest measure. This comparison participant will be identified as a match and removed from the list of potential matches for subsequent PIRE participants. The routine will then move on to the next PIRE participant on the list, and from the remaining set of comparison participants, identify the comparison participant with the smallest average distance to the PIRE participant. This process will continue to the end of the list of PIRE participants. The macro will then return to the top of the list of PIRE participants and find the next best match, and so on.

The "gmatch" matching routine will allow for the PIRE participants to have as few as zero matched comparison participants if none were found who satisfied the distance criterion and, as many as five matches, if all five potential comparison participants met the distance criterion. In general, the number of potential matches per PIRE participant within project is obtained by dividing the number of comparison participants within project by the number of PIRE participants, and rounding down to the whole number value. For example, if in project 1, there were 30 PIRE and 56 potential comparison participants, the number of comparison participants matched to each PIRE participant could range from a minimum of zero to a maximum of 1 (56/30 = 1.87, rounding down = 1).

[2] Kosanke, J., and Bergstralh, E. (2004). Match cases to controls using variable optimal matching. http://www.mayo.edu/research/departments-divisions/department-health-sciences-research/division-biomedical statistics-informatics/software/locally-written-SAS-macros