Evaluation of the National Science Foundation's Partnerships for International Research and Education (PIRE) Program

Volume 2: Supplementary Materials

GS-10F-0086K, Order No. NSFDACS13T1393

Final Report

November 20, 2015

Prepared for:

John Tsapogas
Cassidy Burke
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230

Submitted by:

Abt Associates
55 Wheeler Street
Cambridge MA 02138

In Partnership with:

Thomson Reuters Scientific
Abt SRBI

Introduction to Volume 2

This is the second of three volumes of the final report for the Evaluation of NSF's Partnerships for International Research and Education (PIRE) program.

- Volume 1 includes an Executive Summary, an overview of the PIRE program and the objectives of the evaluation, a brief explication of study methods, and main findings.
- Volume 2 includes additional material: technical details on the study methods and survey response rates, supplementary information about the research outcomes of the PIRE program, and a comparison of PIRE participants' employment outcomes to national trends for science and engineering (S&E) degree recipients.
- Volume 3 includes copies of the study surveys and a cross-walk of undergraduate major fields of study mapped to primary research fields (or primary field for graduate degree) used in Volume 1, Exhibits 2.3.

1. Constructing the Counterfactual

The selection of an appropriate comparison group was critically important because a key objective of the evaluation was to explore the extent to which international research collaborations (of the scale supported by PIRE awards) produce meaningful differences in the nature, quality or quantity of research produced relative to research without international partners or in which international collaboration is a less critical component. To fulfill the evaluation's research questions, we needed to construct a comparison group that would allow project-level analyses—that is, a comparison of the group of PIRE awards to a counterfactual group of awards—and participant-level analyses—a comparison of outcomes for individual participants in these awards. This section begins by discussing the types of comparison groups we considered and the procedures we used, and provides characteristics of the comparison group that resulted.

1.1 Rationale for the Comparison Group

The purpose of a comparison group is to represent the counterfactual—namely, what would have occurred in the absence of the PIRE program. In a random assignment study, the control group represents this counterfactual, and differences in outcomes between the program and control groups provide an estimate of the effect of the program on those outcomes. Because the PIRE program makes awards on the basis of merit, it was not feasible to randomly assign some PIRE proposals to "award" status (i.e., a treatment group) and others to a "non-award" status (i.e., a control group). Moreover, award decisions for past cohorts of the program had already been made when the evaluation was planned. Although a regression discontinuity (RD) design was contemplated, PIRE proposals do not receive a continuous score that is compared to an exogenous "cutoff" to make funding decisions, a key requirement for an RD design. Rather, each proposal receives three or more categorical scores, from one (lower merit) to five (highest merit); subsequently, panelists discuss the rated proposals and place each into one of three categories (highly competitive, competitive, or not competitive). From these categories, NSF program officers recommend which proposals to fund, often taking into account additional factors such as geographic and disciplinary variation across the portfolio of funded projects, and characteristics of the institutions and PIs.

Another potential quasi-experimental design (other than an RD design) was also deemed inappropriate, namely a design where unfunded PIRE proposals were matched to funded PIRE projects using propensity score matching (PSM) techniques to control for selection bias. For the evaluation of the PIRE program, this method was not feasible. PSM techniques require data on a large number of pre-treatment (in this case, pre-PIRE award) characteristics of the members of the treatment and comparison groups to model selection and develop propensity scores. Data on pre-award characteristics for funded and unfunded PIRE proposals are limited. Moreover, for the PIRE program, it is far from clear what measures *at the project level* exist for project characteristics before the award decision is made—the project does not exist as a measurable entity until the award decision has been made. For example, the PIRE award itself could affect which individuals come together to form the research group that ultimately engages in research, produces publications, and collaborates with foreign partners. In the absence of the PIRE funding, those individuals did not form an already established group for which characteristics can be measured. At the participant level also, it is unclear what group of individuals would form the counterfactual for an unfunded PIRE proposal: without PIRE funding, what individual postdoctoral fellows, graduate students, or undergraduate students could be identified as those who "would have participated" in PIRE and could therefore comprise a valid comparison group? After considering the above evaluation designs, the current approach was adopted as the most feasible design. We describe limitations of this comparison group in Volume 1.

1.2 Constructing the Comparison Group for PIRE

For each of the 59 PIRE projects awarded in the 2005, 2007, 2010 and 2012 cohorts, we attempted to identify a similar NSF project that did not require an international collaboration, but which was a similar to its corresponding PIRE projects along other key criteria, following a series of iterative steps.

1.2.1 Methods

In the first step, we generated a list of candidate awards for each PIRE award by searching NSF award data for all active and expired awards from 2000 through 2014. We first classified each PIRE award based on the NSF directorates which provided co-funding, or from which the PIs would likely have sought funding in the absence of the cross-directorate PIRE program. (For example, to identify a comparison project for a PIRE award focused on mitigating the effects of earthquakes on human populations, we would have searched non-PIRE awards within the GEO, ENG and SBE directorates.) For this first step, we applied the first five applied the key criteria shown in Exhibit 1.1

Exhibit 1.1: Criteria for Matching Comparison Projects to PIRE project

Criteria for Comparison Project	Definition
Award amount	Award amount must be within 20 percent of the PIRE project award amount. Award amount includes any supplemental funding awarded.
Duration	Award duration must be within 1 year of the duration of the PIRE project. Duration includes funded or no-cost extensions.
Award effective date	Effective date (start date) must be within 1 year of the PIRE project's award effective date.
Award expiration date	Expiration date (end date) must be within 1 year of the PIRE project's award expiration date.
Continuing grant	Priority was given to continuing grants, as all PIRE awards are continuing grants.
At least two participating institutions	Award must involve collaboration across two or more institutions (U.S. or foreign).
Limited to research awards	Research must be primary focus of grant. Awards where the primary focus was the training of graduate students (e.g., IGERT, GK-12) or purchase of equipment or instrumentation (e.g., Shipboard Scientific Support Equipment) are excluded.
Same or similar disciplinary focus	Disciplinary focus of the research must overlap that of the PIRE award. This criterion was met by examining award abstracts, the post-graduate scientific or engineering backgrounds of senior personnel, and the academic departments or similar institutional units of employment or affiliation of senior personnel.
Must not *require* an international collaboration	Funding program must not *require* an international collaboration as a pre-condition for award. Programs in which international collaboration was encouraged, optional, or not mentioned are eligible.
Graduate student involvement	At least one graduate student participant must be involved in the comparison project (all PIRE projects have graduate students).
No prior PIRE award	Lead PI must not be the PI or co-PI of a PIRE project.

Next, using award abstracts, proposals, and annual and final reports, we screened each candidate project further to ensure that:

- The project was funded by an NSF program whose primary purpose was research (i.e., in contrast to programs focused on training, workforce diversity or equipment or infrastructure;
- The project included at least two different participating institutions (U.S. or foreign);
- The project included at least one graduate student;

- The PI was not a current or former PIRE PI;
- The applicable program solicitation(s) either encouraged applicants to propose an international collaboration; mentioned international collaboration as an option; or did not mention international collaboration at all. [1]

In addition, we carefully scrutinized the candidate's award's abstract, annual reports, and the CVs of PIs, co-PIs and senior personnel, to identify candidate projects where the research area(s) was similar to the PIRE award. To make this determination, we looked for evidence from award abstracts, reports, and CVs of senior personnel that the comparison project would require types of knowledge or expertise similar to that of the PIRE project. If multiple candidates that met these criteria were identified, the comparison project that best matched the research focus of the corresponding PIRE award was selected. Finally, if no comparison award was eligible, we attempted to generate new candidates (i.e., the first step), by including additional directorates (if likely to yield viable new candidates) in the search query and:

- Allowing the start and end dates of candidate projects to be within 24 months of the PIRE award (n=10 comparison projects);
- Allowing the duration of candidates to be within 24 months of the total PIRE duration (n=7 comparison projects);
- Allowing the candidate award amount to differ by up to 25 percent of PIRE amount (n=5 comparison projects); and
- Including standard grants as an eligible award type (n=10 comparison projects).

We identified a matching project for 55 of the 59 PIRE projects from a population of approximately 1,500 projects that met step one criteria. Below, we summarize the resulting characteristics of the matched pairs of projects. For four of the PIRE awards, no matched comparison project was found. To protect the identity of the selected comparison projects and personnel, we withheld the award numbers, project titles, institutions and other identifying characteristics from NSF.[2]

1.2.2 Results

Exhibits 1.2 through 1.6 summarize the characteristics of the comparison group of projects and highlight key similarities and differences to the corresponding PIRE projects. To ensure that no one in or outside of NSF can identify the selected comparison awards, we present aggregated data across the 55 matches rather than individual characteristics of each match.

Award amount

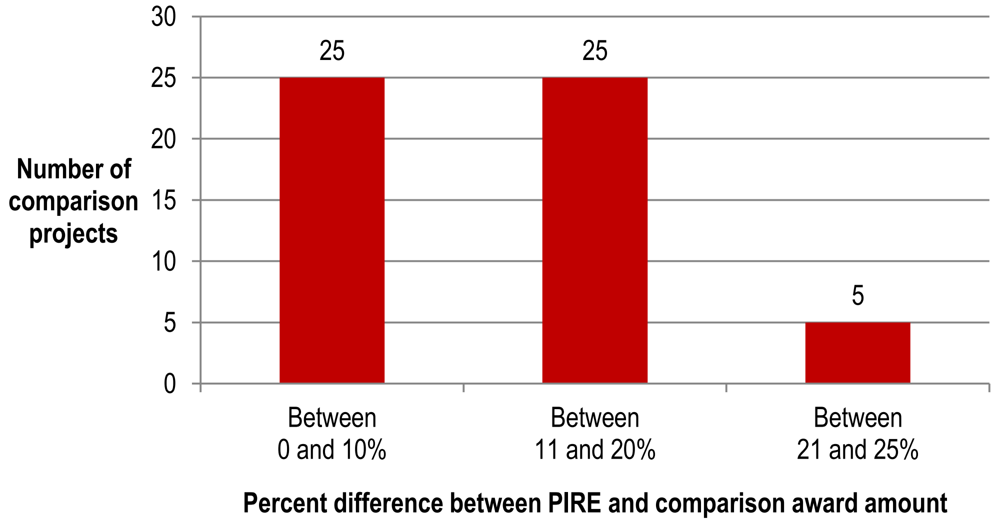
The mean award amount for the 55 matched PIRE awards was $3.3 million and for comparison awards, $3.1 million. Fifty of the matched comparison awards (91 percent) received funding that was within 20 percent of the amount received by their PIRE counterpart; just five comparison awards had funding that differed by more than 20 percent (but less than 25 percent) from their PIRE counterpart (Exhibit 1.2)

---

[1] An NSF program that disallowed international collaborations would have been ineligible but no such program was encountered.

[2] The purpose of the comparison group was to evaluate PIRE against a counterfactual, not to provide direct comparisons of the outcomes of any individual comparison project to its matched PIRE project. Withholding the comparison project identities helped ensure participation in the evaluation by comparison project personnel.

Exhibit 1.2: Percentage Difference in Award Amount Between PIRE and Matched Comparison Awards



Award onset and duration

The majority of comparison projects (N=48) had a total award duration that was within plus or minus 12 months of the corresponding PIRE project's duration (Exhibit 1.3). The remaining seven lasted up to 24 months more or less than the PIRE project. Nearly half of the comparison awards began within 8 months of the PIRE award's starting date (what NSF calls the "award effective date"). Another 18 comparison projects began within one year of the PIRE award.

Exhibit 1.3: Comparison Project Duration and Award Start Date Relative to PIRE

	Percent	N
Comparison Project Duration		
Within 0–12 months of PIRE project's duration	87%	48
Within 13–24 months of PIRE project's duration	13%	7
Award Effective Date		
Within 8 Months of PIRE Award effective date	49%	27
Within 9–12 Months of PIRE Award effective date	33%	18
Within 13–24 Months of PIRE Award effective date	18%	10

Program emphasis on international collaboration

All comparison projects met the criterion that international collaboration could not be a required component. Most comparison projects were funded by programs in which the solicitation did not mention international collaboration at all (31 projects, 56 percent; Exhibit 1.4). Sixteen comparison projects were selected in which the program solicitation encouraged international collaboration, and an additional eight where such collaboration was mentioned as an option for PIs to consider.

Exhibit 1.4: Comparison Program Solicitation's Emphasis on International Collaboration

Comparison Program's Emphasis on International Collaboration	Percent	N
Not mentioned in solicitation	56%	31
Encouraged in solicitation	29%	16
Optional in solicitation	15%	8

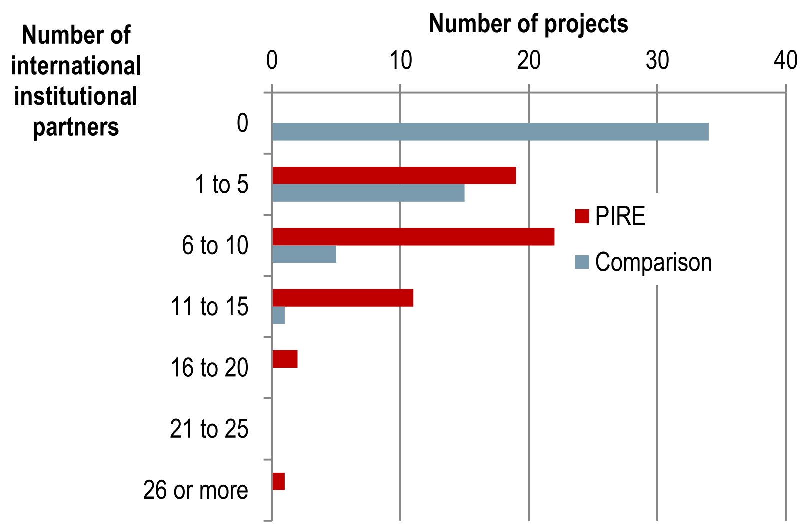
Comparison projects were selected from across NSF's 7 directorates. The MPS directorate had the largest share of comparison awards (36 percent) with GEO and BIO each contributing 26 percent and 18 percent, respectively; Exhibit 1.5). Five of the comparison awards were from ENG programs. Because there were fewer than five comparison projects in each of the remaining directorates (CISE, SBE and EHR), we combined these awards into a single row in the exhibit.

Exhibit 1.5: Directorate of NSF Program Funding Comparison Projects

Directorate of NSF Program Funding Comparison Projects	Percent	N
Mathematical and Physical Sciences (MPS)	36%	20
Geosciences (GEO)	26%	14
Biological Sciences (BIO)	18%	10
Computer and Information Sciences & Engineering (CISE), Social Behavioral and Economic Sciences (SBE), Education & Human Resources (EHR)	11%	6
Engineering (ENG)	9%	5

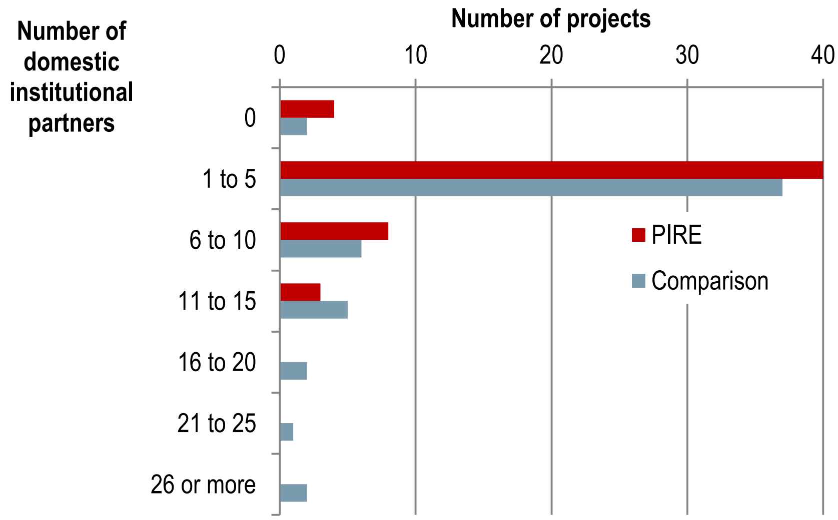
We required each comparison award to have at least two participating institutions. On average, PIRE projects had 4 domestic and 8 international institutional partners relative to 6 domestic and 2 international institutional partners for comparison projects. Exhibits 1.6a and 1.6b show the distributions of the number of international and domestic institutional partners, respectively, across PIRE and comparison projects.

Exhibit 1.6a: Distribution of the Number of International Institutional Partners Across PIRE and Comparison Projects



Sources: PI-submitted annual and final reports to NSF.

Exhibit 1.6b: Distribution of the Number of Domestic Institutional Partners Across PIRE and Comparison Projects



Sources: PI-submitted annual and final reports to NSF.

Exhibit 1.7 shows the number of domestic and international partners in each matched pair of projects (in order of the number of international partners among PIRE projects). In 49 of the 55 matched pairs, the PIRE project had more international institutional partners than its matched comparison project. In just 6 pairs did the comparison project have an equal or greater number of international partners than the PIRE project. For partnerships with domestic (U.S.-based) institutions, in 29 of the pairs, the comparison project had a greater number of domestic partners than its corresponding PIRE project. However, in 26 pairs, the PIRE project had an equal or greater number of domestic (institutional partners).[3]

3 Partnering institutions were identified from PI-reported organizational partners and the total unique number of partners across annual and final reports were counted. PIs may have had different criteria for what organizations to include (or not) as partners.

Exhibit 1.7: Number of Domestic and International Institutional Partners in Each Matched Pair of PIRE and Comparison Projects

Pair	Number of Domestic Partners		Number of International Partners	
	PIRE	Comparison	PIRE	Comparison
1	0	4	1	0
2	0	27	1	2
3	1	1	1	9
4	3	1	1	0
5	9	2	1	1
6	1	1	2	9
7	4	3	2	0
8	5	2	2	1
9	1	1	3	0
10	1	5	3	1
11	2	3	3	0
12	3	1	3	0
13	3	3	3	0
14	3	3	3	0
15	4	1	3	1
16	5	26	3	0
17	4	7	4	1
18	2	1	5	0
19	2	22	5	0
20	3	0	6	7
21	4	2	6	0
22	4	2	6	0
23	4	4	6	0
24	4	6	6	2
25	5	14	6	3
26	7	1	6	0
27	1	2	7	0
28	1	4	7	2
29	3	17	7	6
30	4	5	7	5
31	1	3	8	0
32	1	3	8	0
33	2	4	8	0
34	2	0	8	1
35	3	4	8	0
36	4	13	8	0
37	2	5	9	10
38	4	3	9	0
39	4	8	9	0
40	10	7	9	0
41	12	17	9	0
42	1	3	11	0
43	10	13	11	0
44	10	13	11	0
45	4	5	12	4
46	6	9	12	1
47	15	3	14	0
48	0	2	15	0
49	0	2	15	5
50	1	11	15	13
51	3	1	15	0

Pair	Number of Domestic Partners		Number of International Partners	
	PIRE	Comparison	PIRE	Comparison
52	7	4	15	0
53	7	6	17	2
54	11	5	17	0
55	3	3	26	0

Sources: PI-submitted annual and final reports to NSF.

2. Survey Response Rates

This chapter describes the initial survey samples (drawn from personnel listed in PI-submitted annual and final project reports to NSF), ineligibility exclusions, and response rates for each of the six surveys fielded for the study:

- Principal investigator (PI) survey (for lead PI and co-PIs);
- Postdoctoral survey;
- Graduate student survey;
- Undergraduate student survey;
- Foreign senior investigator (FI) survey; and
- Institutional administrator survey.

Exhibit 2.1 summarizes the final response rates (separately for PIRE and comparison group PIs, postdocs and graduate students).

Exhibit 2.1: Sample Size and Response Rates for Study Surveys

	PIRE		Comparison	
Respondent Group	Sample Size	Response Rate	Sample Size	Response Rate
Lead PIs	59	71%	55	62%
All PIs	293	58%	205	56%
Postdocs	211	55%	235	50%
Graduate students	531	54%	705	47%
Undergraduates	735	40%	NA	NA
Foreign investigators	251	47%	NA	NA
Institutional administrators	59	78%	NA	NA

Sources: Annual reports; study surveys.

Notes: Sample size is equal to the number of eligible respondents. The response rate indicates the percentage of this sample that completed a survey. Because it is impossible to verify that the intended recipient has received an emailed survey invitation, response rates include sample members who may not have received the survey invitation.

Below we describe how we determined these response rates for each respondent group.

2.1 Principal Investigators

2.1.1 Sampling Frame

We identified a total of 533 PIs and co-PIs from NSF extant data (proposals and annual and final reports). Of these:

- 112 were unique lead PIs and 421 were unique co-PIs;
- For 2 participants who were lead PIs on multiple PIRE projects, we invited each to complete two surveys, one for each project, resulting in a total of 535 survey invitations; and
- 5 co-PIs were affiliated with foreign institutions at the time of award participation and were removed from the sample file, resulting in 530 unique records in the PI survey sample.

We included the census of these records in the survey sample.

2.1.2 Ineligibles

During the survey field period, we determined that 32 co-PIs were ineligible for the PI survey:

- 24 were incorrectly identified in annual reports as a co-PI when they either had a different role or no role on the project (14 from PIRE projects, 10 from comparison projects);
- 5 were actually foreign senior investigators; we removed these from the PI sample and added them to the FI sample (all 5 were PIRE project participants);
- 2 were deceased at the time of survey fielding (one from a PIRE project, one from a comparison project); and
- 1 from a PIRE project indicated that he had not participated in the project after the proposal was submitted to NSF.

We removed these ineligibles, leaving a total of 498 eligible records in the sample.

2.1.3 Response Rates

Of the 498 eligible PIs, we classified 284 as respondents, a response rate of 57 percent. We based this classification on examination of the proportion of survey items that each sample member completed:

- 276 completed at least 70 percent of the items in the survey and were classified as respondents.
- 8 had completed less than 70 percent of the survey overall but at least 55 percent of the items identified as providing data for key outcomes and were classified as respondents.
- The remainder either did not initiate the survey or completed an insufficient proportion of items.

Separate analyses of the response rates for the subset of lead PIs indicated an overall response rate of 66.7 percent (76 of 114 Lead PIs).

2.2 Postdoctoral Researchers

2.2.1 Sampling Frame

We identified a total of 463 U.S.-based postdoctoral participants from NSF extant data (proposals and annual and final reports) and from information we requested from project PIs (i.e., we asked PIs to identify any participants not included already in annual reports and to provide a contact email). We included the census of these records in the survey sample.

2.2.2 Ineligibles

During the survey field period, we determined that 17 sample members were ineligible for the postdoctoral survey:

- 9 had been misclassified and were removed from the postdoc sample and added to the graduate student sample (6 PIRE, 3 comparison);
- 4 project participants indicated that they had not participated as a postdoctoral researcher, but in some other role (e.g., research associate; all 4 from PIRE projects);
- 4 had never participated in the project or had limited participation (1 PIRE, 3 comparison; one had served as a teaching assistant for less than one month; had taught a summer field course but had declined a postdoctoral role on the project).

We removed these ineligibles, leaving a total of 446 eligible records in the postdoctoral sample.

2.2.3 Response Rates

Of these 446 eligible postdocs, we classified 232 as respondents, a response rate of 52 percent. We based this classification on examination of the proportion of survey items that each sample member completed:

- 223 completed at least 80 percent of the items in the survey and were classified as respondents.
- An additional 9 completed at least 50 percent of the items identified as providing data for key outcomes and were classified as respondents.
- The remainder either did not initiate the survey or completed an insufficient proportion of items.

2.3 Graduate Students

2.3.1 Sampling Frame

We identified a total of 1,230 unique U.S.-based graduate student project participants from NSF extant data (proposals and annual and final reports) and from information we requested from project PIs. During the survey field period we identified an additional 9 individuals who had been misclassified as part of other survey samples; these were added to the graduate sample. This resulted in a total of 1,239 sample records. We included the census of these records in the survey sample.

2.3.2 Ineligibles

During the survey field period, we determined that 3 sample members were ineligible for the graduate student survey:

- 2 had been misclassified and were removed from the graduate student sample and added to the undergraduate student sample (they had first participated in the project as undergraduate students). Both were PIRE participants.
- 1 comparison project participant was not enrolled as a graduate student at the time of participation in the project.

We removed these 3 ineligibles, leaving a total of 1,236 eligible records in the graduate student sample.

2.3.1 Response rates

Of these 1,236 eligible graduate students, we classified 616 as respondents, a response rate of 50 percent. We based this classification on examination of the proportion of survey items that each sample member completed:

- 601 respondents completed at least 75 percent of the items in the survey overall and at least 50 percent of the items identified as providing data for key outcomes and were classified as respondents;
- An additional 15 completed less than 75 of the survey items overall, but more than 50 percent of the items identified as providing data for key outcomes and were classified as respondents;
- Six respondents completed at least 75 percent of the items in the survey overall, but less than 50 percent of the items identified as providing data for key outcomes and were classified as non-respondents.

2.4 Undergraduate Response Rates

2.4.1 Sampling Frame

We selected a probability-proportional-to-size sample of 770 undergraduates from the population of 897 U.S.-based undergraduates identified in NSF extant data (proposals and annual and final reports) and from information we requested from project PIs. These 897 undergraduate participants were from 50 of the 59 PIRE projects, the sample of undergraduates per PIRE project is proportional to the number of undergraduates per PIRE project in the population.

2.4.1 Ineligibles

Of the 770 undergraduates in the initial sample, prior to fielding the survey we removed as ineligible 8 who were undergraduates based at a foreign institution.

During the survey field period, we determined that an additional 27 were ineligible for the undergraduate survey:

- 11 indicated that they had not been an undergraduate when they first participated in PIRE;
- 7 indicated that they had never worked on the PIRE project; and
- 1 indicated that he was not the individual we were seeking.

After removing these ineligibles, the sample of eligible undergraduates was 735.

2.4.1 Response Rates

Of these 735 eligible undergraduate students, we classified 290 as respondents, a response rate of 40 percent. We based this classification on examination of the proportion of survey items that each sample member completed:

- 287 respondent to more than 85 percent of the items overall and at least 50 percent of the items identified as providing data for key outcomes and were classified as respondents; and
- An additional 3 sample members completed less than 85 percent of the items overall, but at least 50 percent of the items identified as providing data for key outcomes and were classified as respondents.

2.5 Foreign Investigator Response Rates

2.5.1 Sampling Frame

We identified a total of 261 unique foreign senior investigators from NSF extant data (proposals and annual and final reports) from information we requested from project PIs (i.e., we asked PIs to identify any foreign senior investigators not included already in annual reports and to provide a contact email). We included the census of these records in the survey sample.

2.5.2 Ineligibles

During the survey field period, we determined that 10 sample members were ineligible for the FI survey:

- 6 had never participated in the project;

- 2 reported limited participation in the project;[4] and
- 2 reported having participated but not as a foreign senior investigator.

We removed these ineligibles, leaving a total of 251 eligible records in the sample.

2.5.1 Response Rates

Of these 251 eligible FIs, we classified 117 as respondents, a response rate of 47 percent. We based this classification on examination of the proportion of survey items that each sample member completed:

- 104 completed at least 75 percent of the survey overall and at least 50 percent of the items identified as providing data for key outcomes;
- 1 completed two-thirds of the survey and at least 50 percent of the items identified as providing data for key outcomes; and
- 12 completed less than two-thirds of the survey but at least 50 percent of the items identified as providing data for key outcomes.

2.6 Institutional Administrator Response Rates

2.6.1 Sample Construction

We identified potential respondents for the institutional administrator survey as follows:

1. As part of the PI survey, we asked the lead PIs of each of the 59 PIRE awards in the evaluation to identify two administrators at their institution of higher education (IHE) who could answer questions about the PIRE award and their institution's practices for supporting international engagement by faculty and students.
2. If the lead PI did not identify two administrators, we conducted web searches for administrators in appropriate campus offices at the institution (e.g., Office of Global Engagement, Office of Research or the Provost or Vice Provost responsible for overseeing the institution's external research portfolio).
3. For each institution, we designated each administrator as "primary" or "secondary."
4. If the primary administrator did not respond to the survey, we invited the secondary administrator to complete the survey.
5. As part of the survey we allowed respondents to identify an alternate administrator who could best respond to the survey items; if the primary or secondary did so, we invited this alternate to complete the survey.

2.6.2 Selecting Respondents for the Analysis Sample

Our target response rate was one completed institutional administrator survey per PIRE award. That is, the target response rate was N=59. From the initial sample file of 118 administrators (two per PIRE award), we identified a total of 72 survey records from 59 awards (0 to 3 surveys completed per project; see Exhibit 2.2).

4 One was an advisor to undergraduates and another was identified by the project's PI as not having a role in research; neither of these was deemed to fit the definition of a foreign senior investigator.

Exhibit 2.2: Number of Eligible Administrators per PIRE Project

Number of Surveys Received per PIRE Project	Number of PIRE Projects	Total Number of Surveys Received
0	2	0
1	43	43
2	13	26
3	1	3
	Total = 59	Total = 72

Sources: PI submitted annual reports; internet web searches; Institutional Administrator sample files
Notes: The target sample was one response per PIRE award (N=59). A convenience sample was constructed by soliciting nominations of two administrators from lead PIs (we designated one as a primary respondent) and by conducting internet searches of institutional web pages when no nominations were received. If a primary respondent did not reply to the survey, we contacted the secondary respondent. If a respondent nominated an alternate administrator, we invited that person to complete the survey.

Since the unit of response for the survey is the PIRE award (not an individual administrator), we implemented a series of decision rules to select one survey respondent for each award:

- If none of the eligible administrators invited to complete the survey award opened the survey web link, we selected the primary administrator as that award's eligible invited sample member.
- If only one administrator from an award opened the survey, we kept that record in the sample;
- If two administrators from an award opened the survey but only one had navigated through the entire survey and submitted their survey as final, we kept this record in the sample;

After applying these rules, there were three awards with multiple records remaining. Here are the specific decision rules applied:

- One award had two incomplete surveys but each was an eligible survey sample member. The primary administrator answered one item and the secondary administrator answered two items. We retained the sample member who completed two items.
- For one PIRE award, there were three potential administrators, one who was nominated by a primary administrator, one who screened out of the survey, and one who had an incomplete survey. We retained the sample member who submitted an incomplete survey.
- For the third PIRE award, one administrator screened out of the survey, and another administrator submitted two responses to the survey. We retained the most recent survey received.

The final sample of eligible respondents was N=59 (one per PIRE award). We then examined these responses for completeness to determine the final response rate.

2.6.3 Response Rates

Of these 59 eligible responses, we classified 46 as respondents, a response rate of 78 percent. We based this classification on examination of the proportion of survey items that each sample member completed:

- 41 respondents completed at least 90 percent of the survey items; and
- 5 respondents completed less than 90 percent of the survey overall but at least 60 percent of the items identified as providing data for key outcomes.

3. Analysis Methods

3.1 Analyses of Research Outcomes

In Volume 1, we report three types of analyses of publication outcomes (quality, quantity, and international collaborations) at the project level (i.e., for projects as a whole including all journal articles reported by the PI as resulting from the project), and at the participant level (i.e., analyzed separately for each of three types of project participants: PIs, postdocs, and graduate students):

- Descriptive analyses of key publication outcomes across PIRE projects (Research Question 1) or PIRE participants (Research Question 3).
- Descriptive comparative analyses to show *differences* in research outcomes for PIRE and comparison projects (Research Question 2) and participants (Research Question 4):
  - At the project level, these descriptive comparative analyses included only the subset of 45 matched pairs of projects where each project within each pair had at least one journal article identified in the Web of Science.
  - At the participant level, these analyses included only participants from within matched project pairs who had at least one post-onset and at least one pre-onset publication indexed in the Web of Science.
- Impact analyses to test the impact of PIRE on two research outcomes—the number of publications and the field normalized citation impact (field NCI) for post-onset publications.[5]

Before conducting these analyses, we first matched references and participants listed in PIs' annual or final reports to NSF to Web of Science records:

- For each of the 59 PIRE and 55 comparison projects, bibliographic information for journal articles (i.e., authors, year of publication, article title, journal name and volume) was identified from lists of publications that PIs submitted to NSF in annual reports. These records were then submitted for comparison by Thomson Reuters Scientific staff to records in the Web of Science databases.
- To identify participants' publications, we collected the names of PIRE and comparison projects' PIs, postdocs and/or graduate students, and when possible, email addresses and/or a bibliographic record for a selected "seed publication" when no email address was available in NSF data. Thomson Reuters attempted to find publications on which participants were authors by matching the participant's name and/or email address and using seed publications when available to disambiguate individuals with identical or similar names. Searches covered publications published in 2000 or later.

We present the results of matching references or participants to the Web of Science below.

3.1.1 Identifying PIRE Projects' References in the Web of Science

From the annual reports submitted by PIs to NSF, we collected all publications reported to date (by July 2014) including peer-reviewed journal articles, books, chapters in edited volumes, and manuscripts in preparation, under review, or in press. Only papers that were published in journals

5 Post-onset articles are those published after the participant had begun participating in the PIRE or comparison project. Pre-onset articles are those published prior to this onset of participation. (See Section 2.1.3: Data Sources, Bibliometric Data.)

were submitted as records to Thomson Reuters (books, chapters in edited volumes, and manuscripts in preparation, under review, or in press were excluded).

From 51 of the 59 PIRE projects, a total of 1,648 records were submitted and 1,399 (85 percent) were matched to articles in the Web of Science (Exhibit 3.1). Articles listed in PIs' annual reports that had publication dates preceding the award onset year were excluded. For the remaining eight PIRE projects, either the PI's available annual reports to NSF included no published peer-reviewed journal articles (n=6 projects), or none of the records submitted to Thomson Reuters were identified in the Web of Science (n=2 projects). Five of the PIRE projects with no reported peer-reviewed journal publications were from the 2012 PIRE cohort and one was from the 2010 cohort (Exhibit 3.2).

Exhibit 3.1: PI-Reported Journal Articles Submitted and Matched to Articles in the Web of Science

	N of Records Submitted to Thomson Reuters[1]	N of Articles Identified in Web of Science	Percent Identified
PIRE (N=51 projects)	1,648	1,399	85.0%

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science.

Notes:

[1] Records submitted were restricted to papers published in journals (whether or not the journal was peer-reviewed was not known, but the Web of Science indexes only peer-reviewed journals) with a publication date no earlier than the year that the PIRE award began. Other publications reported by PIs such as books, chapters in edited volumes, and manuscripts in preparation, under review, or in press were excluded since bibliometric indicators were not available for these types of publications.

Exhibit 3.2: Number of PIRE Projects with No Reported Publications or No Articles Identified in Web of Science

Cohort	Year of Award	N of projects with Missing Bibliometric Data	Reasons
2	2007	1	1 project with no records matched to Web of Science[1]
3	2010	1	1 project with no publications reported[2]
4	2012	6[3]	5 projects with no publications reported 1 project with no records matched to Web of Science

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science.

Notes:

[1] This PIRE project was a computer and information sciences (CISE) research initiative and in these disciplines scholarly dissemination relies less on peer-reviewed journals than in other fields.

[2] This PIRE project was not intended to produce peer-reviewed journal articles, but rather, focused on contributing other resources data repositories, software, and similar products for dissemination.

[3] These 6 projects began in October or November of 2012 or January of 2013 and at the time bibliometric data retrieval began in June, 2014, had limited time to produce journal articles (a maximum of 19 months from award start).

An analysis of the unmatched articles (n=249 articles) showed that more than half (53 percent) were published in journals not indexed in the Web of Science. For the remaining 47 percent (n=116), the article was published in a journal indexed in the Web of Science but either the journal title or the article title from the PI's annual report could not be matched to an existing Web of Science record (Exhibit 3.3).

Exhibit 3.3: PIRE Project Articles Not Identified in Web of Science

	N	%
Articles published in journals not indexed in Web of Science	133	53%
Articles published in journals indexed in Web of Science but not matched[1]	116	47%
Total of articles not identified	249	100%

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science
Notes:
[1] The article was not matched either because the journal title matched a journal in the Web of Science exactly but the article was not found (PIRE: n=95) or the journal title reported by the PI was similar to, but not an exact match for, a journal indexed in the Web of Science (n=21).

3.1.2 Identifying PIRE and Comparison Projects' References in the Web of Science

For analyses comparing PIRE and comparison group projects' research outcomes, it was necessary to restrict the analytic sample to a subset of 45 matched pairs of PIRE and comparison projects:

- A matched comparison project was identified for 55 of the 59 PIRE projects;
- Of these, 7 PIRE projects reported no journal articles or none of the articles reported were matched to Web of Science data. The corresponding comparison project for each of these 7 PIRE projects was likewise dropped;
- Among the comparison projects, 2 reported no journal articles or none of the articles reported were matched to Web of Science data. The corresponding PIRE project for these 2 comparison projects were dropped;
- In one PIRE-comparison pair, both the PIRE and comparison project reported no journal articles or none of the articles reported were matched to Web of Science data; this pair of projects was dropped.

Exhibit 3.4 summarizes the results.

Exhibit 3.4: Description of Projects Excluded from Comparative Bibliometric Analysis

Reason for Exclusion	PIRE	Comparison	Total
Number of projects before exclusions	59	55	114
No matched comparison project	-4	NA	-4
PIRE project in matched pair has no matched articles in Web of Science	-7	-7	-14
Comparison award in matched pair has no matched articles in Web of Science	-2	-2	-4
Both PIRE and comparison award in matched pair have no matched articles in Web of Science	-1	-1	-2
Number of matched pairs of projects	45	45	90

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science.
Notes: Of the three comparison projects with no articles in the Web of Science, two were funded under the CISE directorate and the third project's main research product is a data repository.

Among the 45 matched PIRE and comparison projects, there was no statistical difference between the PIRE and comparison group percentage of articles matched ($p>0.30$). For these 45 pairs of projects, 86 percent of the PIRE projects' submitted records and 85 percent of the comparison projects' records were matched to Web of Science data.

Exhibit 3.5: PI-Reported Journal Articles Submitted and Matched to Articles in the Web of Science

	N of Records Submitted to Thomson Reuters[1]	N of Articles Found in Web of Science[2]	Percent Matched
PIRE (N=45 projects)	1529	1320	86.3%
Comparison (N=45 projects)	1328	1128	84.9%
Total	2857	2448	85.7%

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science

Notes

1 Records submitted were restricted to papers published in journals (whether or not the journal was peer-reviewed was not known, but the Web of Science indexes only peer-reviewed journals). Other publications reported by PIs such as books, chapters in edited volumes, manuscripts in preparation, under review, or in press were excluded since bibliometric indicators were not available for these types of publications.

Among the unmatched articles, Exhibit 3.6 shows that roughly half were produced by PIRE (49 percent) and the comparison group (51 percent). Likewise, just over half of the articles that were published in journals not indexed in the Web of Science were from PIRE projects (52 percent). However, of the articles that were published in Web of Science journals but not matched, less than half (45 percent) come from PIRE projects. These findings suggests some caution in interpreting differences in research outcomes between PIRE and the comparison group: for example, if the unmatched articles in the comparison group also had systematically higher (or lower) field normalized citation impacts than the unmatched articles from the PIRE group, the differences in research outcomes reported below could be biased.

Exhibit 3.6: Articles Not Identified in Web of Science

Project[1]	Articles Not Identified		Articles Published in Journals Not Indexed in Web of Science		Articles Published in Journals Indexed in Web of Science	
	N	%	N	%	N	%
PIRE (N=45 projects)	207	55%	102	61%	105	50%
Comparison (N=45 projects)	171	45%	65[2]	39%	106	50%
Total	378	100%	167	100%	211[3]	100%

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science.

Notes:

1 Nine projects (6 PIRE, 3 comparison) did not report any journal articles in references listed in PI-submitted annual reports. For an additional 2 PIRE projects, none of the listed references matched Web of Science records.

2 Includes 62 articles in journals not indexed in the Web of Science and 3 articles where no journal name was provided.

3 Of the 211 articles published in journals indexed in the Web of Science the PIRE articles were published in 72 unique journals and the comparison group articles in 84 unique journals. The article was not matched either because the journal title reported by the PI was similar to, but not an exact match for, a journal indexed in the Web of Science (PIRE: n=18; Comparison: n=13) or the journal title matched a journal in the Web of Science exactly but the article was not found (PIRE: n=87; Comparison: n=93).

3.1.3 Identifying PIRE Participants in the Web of Science

To examine the research outcomes for PIRE participants, we restricted analyses reported in Volume 1 to articles co-authored by a PIRE participant (PI, postdoctoral researcher, or graduate student) and published in or subsequent to a participant's first year of participation in the PIRE project (i.e., post-onset publications). These publications could include articles or reviews that did not result directly from the PIRE project. Whenever possible, both participant name and email address were submitted to Thomson Reuters. When no email address was found in annual reports, an internet search for an email address for the participant was conducted; in addition, publications reported to NSF were searched for any on which the participant was a co-author. Thomson Reuters used a combination of email addresses and/or "seed" publications to disambiguate authors with identical or similar names. Disambiguation of individuals with similar names is more successful for lower-frequency names, names with a known seed publication, and/or names with an email address associated with the individual.

A participant was considered to be identified if one or more articles published after the participant had begun participating in the PIRE project listing the individual as an author was found in a journal indexed in the Web of Science. Articles could have been published before or after the individual's onset of participation in the project (PIRE or matched comparison project). Post-onset articles may not have resulted from participation in the project. For participants without a reported participation start date (i.e., from the study survey), the PIRE award effective date was used as the participation start date.

Not all individuals were found to have a *post-onset* publication in the Web of Science, that is, a publication in the year of or after the start of participation in the project (Exhibit 3.7). Note that articles published post-onset may have resulted from activities other than participation in the PIRE project.

Exhibit 3.7: Number of PIRE Participants with at Least One Post-Onset Publication Matched to Web of Science Data

Participant Type	N of Participants Submitted[1]	N of Participants with at Least One Article Found in Web of Science[2]	Percent Participants Matched to Web of Science
PIs	307	265	86%
Postdocs	209	135	65%
Graduate students	467	181	39%

Sources: Annual reports submitted by PIs to NSF; Web of Science.

Notes:

[1] Whenever possible, both participant name and email address were submitted to Thomson Reuters. When no email address was found in annual reports, an internet search for an email address for the participant was conducted; in addition, publications reported to NSF were searched for any on which the participant was a co-author. Thomson Reuters used a combination of email addresses and/or "seed" publications to disambiguate authors with identical or similar names. Disambiguation of individuals with similar names is more successful for lower-frequency names, names with a known seed publication, and/or names with an email address associated with the individual.

[2] A participant was considered to be identified if one or more articles published after the participant had begun participating in the PIRE project listing the individual as an author was found in a journal indexed in the Web of Science. Articles could have been published before or after the individual's onset of participation in the project (PIRE or matched comparison project). Post-onset articles may not have resulted from participation in the project. For participants without a reported participation start date (i.e., from the study survey), the PIRE award effective date was used as the participation start date.

3.1.4 Identifying PIRE and Comparison Participants in the Web of Science

To compare research outcomes of PIRE participants to that of their counterparts in the comparison group of projects, analyses were restricted to articles co-authored by a PIRE or comparison project participant (PI, postdoctoral researcher, or graduate student) and classified into one of two time periods:

- Post-onset: those published in or subsequent to a participant's first year of participation in the PIRE or comparison project (note that these publications could include articles or reviews that did not result directly from the PIRE or comparison project); and

- Pre-onset: those published in the years prior to a participant's first year of participation in the PIRE or comparison project.

Only participants with at least one pre-onset and at least one post-onset publication were included in descriptive comparative or impact analyses. (Analysis samples were further restricted by other considerations discussed below). Exhibit 3.8 shows the resulting number of participants in each group.

Exhibit 3.8: Number of PIRE, Comparison Group Participants with at Least One Pre-Onset and at Least One Post-Onset Article Identified in the Web of Science

Participant Type	N of Participants Submitted		At Least One Pre-Onset AND at Least One Post-Onset Article in Web of Science[1]			
			PIRE		Comparison	
	PIRE	Comparison	N	Percent	N	Percent
PIs	307	213	210	68%	171	80%
Postdocs	209	235	67	32%	76	32%
Graduate students	467	710	40	9%	71	10%

Sources: Annual reports submitted by PIs to NSF; Web of Science.
Notes:
[1] Whenever possible, both participant name and email address were submitted to Thomson Reuters. When no email address was found in annual reports, an internet search for an email address for the participant was conducted; in addition, publications reported to NSF were searched for any on which the participant was a co-author. Thomson Reuters used a combination of email addresses and/or "seed" publications to disambiguate authors with identical or similar names. Disambiguation of individuals with similar names is more successful for lower-frequency names, names with a known seed publication, and/or names with an email address associated with the individual.
[2] A participant was considered to be identified if one or more articles published in 2000 through 2014 listing the individual as an author was found in a journal indexed in the Web of Science. Articles could have been published before or after the individual's onset of participation in the project (PIRE or matched comparison project). Post-onset articles may not have resulted from participation in the project.

3.1.5 Descriptive Comparative Analyses for Project- and Participant-Level Research Outcomes

In Volume 1, we report descriptive comparative analyses to show *differences* in research outcomes for PIRE and comparison projects (Research Question 2) and participants (Research Question 4). A comparison project was identified for 55 of the 59 PIRE projects. However, at the project level, these descriptive comparative analyses included only the subset of 45 matched pairs of projects *where each project within each pair had at least one journal article identified in the Web of Science*. Ordinary least squares (OLS) regression analyses included the outcome as the dependent variable; independent variables included a group indicator (1=PIRE project; 0=comparison project) and dummy variables to account for the matched pairs of projects.

For the project-level comparisons we used OLS regression models to fit data in which the outcomes were dependent variables and the independent variables included a dichotomous PIRE indicator (1=PIRE project, 0=comparison project) and dummy variables that indicated the matched pairs of projects:

$$Y_i = \beta_0 + \beta_1 (PIRE_i) + \sum_{j=1}^{J-1} \beta_{1+j} Pair_j + \varepsilon_i$$

where

Y_i is the outcome of interest for project *i*.

$PIRE_i$ is 1 if project *i* is a PIRE project and 0 if a comparison project.

$Pair_j$ is 1 if project *i* is in pair *j*, is 0 otherwise (*j*=1,2,…,54).

ε_i is the residual for the i^{th} project, assumed to be normally distributed with mean = 0 and variance = σ^2.

$\hat{\beta}_0$ is the model-adjusted mean value of the outcome for the comparison projects.

$\hat{\beta}_1$ is the difference between the mean value of the outcome of the PIRE and comparison projects.

$\hat{\beta}_0 + \hat{\beta}_1$ is the model-adjusted mean value of the outcome for the PIRE projects.

At the participant level, we conducted a descriptive comparative analysis of the percentage of foreign affiliations among co-authors of participants' post-onset publications. These analyses included only participants from within matched project pairs *who had at least one post-onset and at least one pre-onset publication indexed in the Web of Science*. The analysis used OLS regression models that included a group indicator (1=PIRE project; 0=comparison project) and dummy variables representing the matched pairs of projects; these participant-level analyses also controlled for the pre-onset measure of the outcome (e.g., percentage of foreign institutions per publication per participant):

$$Y_i = \beta_0 + \beta_1 (PIRE_i) + \sum_{j=1}^{J-1} \beta_{1+j} Pair_j + \varepsilon$$

where

Yi is the outcome of interest for participant *i*.

$PIRE_i$ is 1 if participant *i* is a PIRE participant and 0 if a comparison participant.

$Pair_j$ is 1 if project *i* is in pair *j*, is 0 otherwise (*j*=1,2,…,54).

ε is the residual, assumed to be normally distributed with mean = 0 and variance = σ^2.

$\hat{\beta}_0$ is the model-adjusted mean value of the outcome for the comparison participants.

$\hat{\beta}_1$ is the difference between the mean value of the outcome of the PIRE and comparison participants.

$\hat{\beta}_0 + \hat{\beta}_1$ is the model-adjusted mean value of the outcome for the PIRE participants.

Exhibit 3.9 shows the pre-onset average percentage of foreign institutions per article for PIRE and comparison group participants. Although postdocs appear similar prior to participating in the project, among both PIs and graduate students, PIRE participants had higher average percentages than their counterparts in the comparison group. Regression analyses controlled for these pre-onset differences.

Exhibit 3.9: Average Percentage of Foreign Institutions per Article for PIRE and Comparison Participants' Pre-Onset Articles

Participant Type	Average Percentage of Foreign Institutions per Pre-Onset Article	
	PIRE	Comparison
PIs	18.0%	13.5%
Postdocs	39.2%	38.9%
Graduate students	35.6%	19.9%

These regression analyses do not allow conclusions about the underlying cause of any observed differences. That is, any difference between PIRE and comparison outcomes could result from factors other than PIRE. Nevertheless, these analyses provided a test of the statistical significance of observed differences. Unless reported otherwise, regression analyses controlled for the matched pairs of PIRE-comparison projects. Because we matched projects on criteria such as award amount, duration, award start and end years, and research area(s), controlling for these pairs effectively controls for all such matching criteria.

3.1.6 Impact Analyses Of Participant-Level Research Outcomes

To estimate the effect of PIRE on participants' research outcomes, we used a quasi-experimental, comparative interrupted time series (CITS) analysis (i.e., a difference in difference model with multiple pre and post time points), to compare the research outcomes for participants in PIRE to those of participants in the comparison group of other NSF projects. We ran separate analyses for each group (PIs, postdocs and graduate students) for two outcomes:

- Annual number of published journal articles; and
- Average field NCI per year.[6]

For each outcome, PIRE participants were matched to corresponding participants from comparison group projects by limiting the analysis sample such that the two groups were equivalent at baseline—that is, before they began participating in the project—on the number of publications and the average field NCI.[7] (Participants were also matched on the starting year of participation.) We describe the analysis samples resulting from establishing baseline equivalence below.

Only participants from a matched pair (N=55) of PIRE and comparison projects and who had at least one pre-onset and at least one post-onset publication were included in the CITS analyses. Because the CITS model accounted for the matched project pairs, some participants in a PIRE or comparison project who had at least one publication were dropped if there were no participants of that type in the matched project. Consider, for example, a matched pair of projects, $PIRE_{i=1}$ and $COMP_{i=1}$ where $COMP_{i=1}$ had three postdoctoral researchers and $PIRE_{i=1}$ had one, and the $PIRE_{i=1}$ postdoc had no publications during the post-onset period. This $PIRE_{i=1}$ postdoc was dropped from the analysis sample (because there was no observed data for the post-onset period), and, as a result, the three postdocs in $COMP_{i=1}$ were also dropped because they had no counterpart in the matched PIRE project. This limitation of the CITS model reduced the analytic sample sizes for each participant type.

The analyses included four years of bibliometric data prior to participation onset and multiple years of post-onset data to estimate the impact of PIRE. By including data on participants in comparison

6 Average field NCI per year was calculated by summing the field NCI per article for each of the years after the article's publication, and dividing by the number of articles with a non-missing field NCI.

7 The criterion for baseline equivalence was that difference in means between the two groups was less than .25 standard deviation for each pre-participation measure.

projects with the opportunities to publish research occurring at the same time as their matched PIRE participants, the CITS approach controls for the effects of any factors external to PIRE that might occur at the same time as participation in PIRE and that could also affect the outcomes of interest (e.g., global economic trends in research and development). By including data on the participants prior to the onset of their participation in a project (PIRE or comparison), the CITS design controls for persistent (time-invariant) characteristics of projects that might differentially affect outcomes (i.e., persistent, time-invariant differences between participants). For example, participants' fields of research likely affect both pre- and post-participation outcomes, and it is reasonable to assume that the effect on these outcomes is stable over time. Thus, our impact analysis models included controls for differential effects of different research disciplines on PIRE outcomes. In sum, using data for PIRE participants and matched comparison individuals both prior to and after project onset allows us to control for persistent, project-specific factors that could explain any observed differences, thus reducing the number of plausible alternative explanations for observed effects.[8]

Of three interrupted time-series analysis models we considered (a non-linear baseline model, a linear baseline model, and a baseline mean model) we used baseline mean models for each participant group. A baseline mean model is the simplest and least risky of the three models; by including multiple years of pre-treatment data (four year), the mean is more precise than models using just one year of baseline data. Moreover, a baseline mean model avoids assumptions necessary in linear and non-linear models that can lead to especially large errors in impact estimates if the model assumptions are invalid (Bloom, 2003).

For a CITS design, pre-participation outcomes must be statistically independent of post-participation outcomes. To ensure that the "pre-onset" and "post-onset" field NCIs for a given article were statistically independent, for each of an identified participant's publications found in the Web of Science, Thomson Reuters calculated a customized "NCI per year" for each year beginning with 2000 and ending with 2013. This NCI per year used only the citations of a target article that were made in citing articles published in the indicated year. (For example, an article's NCI in 2005 could differ from its NCI in 2013). In contrast, using a single NCI measure, such as one collected in 2013, would have introduced non-independence between the pre-onset measure of an author's average field NCI and a post-onset average field NCI.

The interrupted time series models took the form:

$$Y_{jkl} = \beta_0 + \alpha_{ok} + \beta_1(trtgrp_{kl}) + \beta_2(trtyr_{jkl}) + \beta_3(trtgrp_{kl} * trtyr_{jkl}) + \beta_4(matchgrp_1) + ...$$
$$+ \beta_{4+m-1}(matchgrp_{m-1}) + \varepsilon_{jkl}$$

Where:

Y_{jkl} is the outcome measure for year *j* on participant *k* in the *l*th project.

$trtgrp_{kl}$ is 1 if participant *k* in the *l*th project is a PIRE participant and 0 if participant *k* in the *l*th project is a comparison participant.

$trtyr_{jkl}$ is 1 if year *j* is after participant *k* started to participate in the PIRE program, and 0 otherwise.

[8] The CITS design used here assumes that the effect of these (time-invariant) characteristics on study outcomes for a given participant is the same both before and after the onset of participation in the project (whether the project is a PIRE or comparison project).

$matchgrp_1$ is 1 if participant belongs to the first matched pair of projects, and 0 otherwise.

… … additional dummies for additional matched pairs.

$matchgrp_{m-1}$ is 1 if participant belongs to the (*m*-1) matched pair of projects, and 0 otherwise (the dummy for the M^{th} matched pair is omitted from the model).

and where the coefficients are interpreted as follows:

$\hat{\beta}_0$ is the adjusted mean for comparison participants in pre-PIRE years (adjusted for model covariates, i.e. the dummies for the matched pairs of projects).

$\hat{\beta}_0 + \hat{\beta}_1$ is the adjusted mean for PIRE participants in pre-PIRE years.

$\hat{\beta}_1$ is the adjusted average difference between the pre-PIRE mean of PIRE participants and comparison participants.

$\hat{\beta}_2$ is the adjusted comparison participant difference between means of pre-PIRE years and means of post-PIRE years.

$\hat{\beta}_2 + \hat{\beta}_3$ is the adjusted PIRE participant difference between means of pre-PIRE years and means of post-PIRE years.

$\hat{\beta}_3$ is IMPACT (the effect of PIRE) = ($(\beta_2 + \beta_3) - \beta_2$).

α_{ok} is a random intercept for participant *k*

$$\alpha_{ok} = \delta_{00} + \mu_{0k}$$

where μ_{0k} is normally distributed with a mean 0 and standard deviation τ .

ε_{jk} is the usual error participant-year specific error term.

Analysis samples for impact analyses

To establish baseline equivalence for each PIRE and comparison participant group (PIs, postdocs, graduate students) we first matched participants on the year that their participation in the project began. We then examined the distributions of the pre-onset number of publications per year and average field NCI per year, and tested the PIRE and comparison samples for baseline equivalence. If the standardized mean difference for an indicator was >.25 standard deviations, we dropped participants who were apparent outliers in the distribution for that indicator. Exhibit 3.10 summarizes the baseline equivalence for each participant type for each baseline measure. Details for each participant group are below.

Exhibit 3.10: Standardized Mean Differences Between PIRE and Comparison Participants For Pre-Onset Characteristics

	Number of Participants		Pre-Onset Standardized Differences		
	PIRE	Comparison	Start Year	N of Publications per Year	Average Field NCI per Year
PIs	256	199	0.15	0.11	*0.51*
Postdocs	118	102	0.17	0.04	0.13
Graduate students	172	297	0.20	0.11	0.04

- PIs: 6 PIs with 9 or more publications per year in the pre-onset period were identified as outliers and dropped from the analysis sample. Each project within a pair had to include at least one participant in the analysis sample in order to include participants from that pair of projects. As a result of dropping these 6 PIs with extreme values, we had to drop an additional 4 PIs who were no longer in a matched project pair. The remaining 455 PIs (256 PIRE and 199 comparison) in 54 matched project pairs were well-matched on the annual number of pre-onset publications (the difference in means between the two groups is less than .25 standard deviations). We were unable to balance the groups on pre-onset average field NCI per year, as the distributions were too dissimilar. The resulting analysis sample of PIs had a mean standardized difference on average field NCI per year of .51.
- Postdocs: Using the same method described for PIs, we dropped from the analysis sample 16 postdocs with extreme values (average field NCI per year ≥ 4 or annual publications ≥ 2 per year). As a result of dropping these postdocs with extreme values, we had to drop an additional 63 postdocs who were no longer in a matched project pair. The remaining 220 postdocs (102 PIRE and 118 comparison) in 34 matched project pairs were well-matched on the annual number of pre-onset publications and the average field NCI per year (the differences in means between the PIRE and comparison groups were less than .25 standard deviations).
- Graduate students: We dropped from the analysis sample 11 graduate students with extreme values (average field NCI per year ≥ 3 or annual publications ≥ 3 per year). As a result of dropping these graduate students with extreme values, we had to drop an additional 29 students who were no longer in a matched project pair. The remaining 469 students (172 PIRE and 297 comparison) in 45 matched project pairs were well-matched on the annual number of pre-onset publications and the average field NCI per year (the differences in means between the PIRE and comparison groups were less than .25 standard deviations).

3.2 Analyses of Participant Experiences

In order to compare PI, postdoctoral and graduate student participant experiences in PIRE and comparison projects, we fit OLS regression models in which the outcomes were dependent variables and the independent variables included a dichotomous PIRE indicator (1=PIRE project, 0=comparison project) and dummy variables that indicated matched pairs of projects.

To determine whether or not additional covariates were included in the OLS models, we first identified the set of control variables that had statistically significant associations with the outcome ($p<0.20$ criterion) after controlling for other statistically significant control variables, namely, the PIRE indicator and dummy variables that indicated matched pairs of projects in the model. This was accomplished using backwards elimination with forward checking.[9] In this method, all of the control variables, the PIRE indicator, and the dummy variables that indicated matched pairs of projects in the model were entered as predictors in the model. The control variable with the largest non-significant value was dropped from the subsequent model. We repeated this step until the only control variables that remained in the model met the $p<0.20$ criterion. In the forwards checking step, each of the previously eliminated control variables was checked by adding each one to the model with only the significant predictors. In this step, each variable had a chance to get back into the model.

For PIs and postdocs, the only covariates that remained after applying this backwards elimination with forward checking process were the PIRE indicator and the dummy variables representing the matched pairs of projects. For graduate students, the models also included a covariate for the graduate student's primary field of study for their graduate degree.

The models for PIs and postdocs were specified as follows:

$$Y_i = \beta_0 + \beta_1 (PIRE_i) + \sum_{j=1}^{J-1} \beta_{1+j} Pair_j + \varepsilon$$

where

Y_i is the outcome of interest for participant *i*.

$PIRE_i$ is 1 if participant *i* is a PIRE participant and 0 if a comparison participant.

$Pair_j$ is 1 if project *i* is in pair *j*, is 0 otherwise (j=1,2,…,54).

ε is the residual, assumed to be normally distributed with mean = 0 and variance = σ^2.

$\hat{\beta}_0$ is the model adjusted mean value of the outcome for the comparison participants.

---

9 Backwards elimination methods are attractive because they are often used and familiar. But use of this method using the conventional $p<0.05$ criterion has been criticized because the selection criterion tends to favor covariates with strong relationships to the outcome, but may omit important confounders (i.e., variables that have a weaker relationship to the outcome, but have a strong relationship to the predictor variable of interest). Maldonado and Greenland (1993) evaluated a backwards elimination strategy and a change-in-estimate strategy using simulated data from a Poisson regression model. They found that the p-value based method performed adequately when the alpha levels were higher than conventional levels (0.20 or more), and found that the change-in-estimate strategy performed adequately when the cut point was set to 10 percent. However, their data, generated from a Poisson model, and their analysis model, with only a single covariate in addition to the key exposure variable, are very different from the models used here. Budtz-Jorgensen et al. (2007) compared several covariate selection strategies including backwards elimination and change-in-estimate. They looked at the backwards elimination strategy with three p-value cut-off levels, 0.05, 0.10, and 0.20, and, following the recommendation of Maldonado and Greenland (1993) used a 10 percent criterion for the change-in-estimate method. They found that although the change-in-estimate strategy did an adequate job of identifying confounders and keeping them in the model, it sometimes threw out variables that were correlated with the outcome, but were not confounders. Therefore, this method threw out variables that, if retained, would have reduced the residual error and reduced the standard error of the exposure coefficient (thus increasing the power to detect exposure effects— exposure effect is analogous to our key predictor of interest). Although they found that backwards elimination with a $p<0.05$ criterion was unsuited for confounder identification, they found that when the p-value criterion was set to $p<0.20$, backwards elimination resulted in a reduction of residual error variance and did not throw out important confounders. They recommended the backwards elimination strategy with a $p<0.20$ criterion over the change-in-estimate strategy.

$\hat{\beta}_1$ is the difference between the mean value of the outcome of the PIRE and comparison participants.

$\hat{\beta}_0 + \hat{\beta}_1$ is the model adjusted mean value of the outcome for the PIRE participants.

The models for graduate students were the same as above but included a covariate for the graduate student's primary field of study for their graduate degree:

$$Y_i = \beta_0 + \beta_1 (\text{PIRE}_i) + \sum_{j=1}^{J-1} \beta_{1+j} Pair_j + \beta_{56} (\text{Field}_i) + \varepsilon$$

where

Field_j is the discipline of the i^{th} participant.

4. Research Outcomes: Supplementary Findings

Volume 1, Chapter 3 reports the main findings for research outcomes. Here we report results that supplement those in Volume 1.

4.1 Foreign Contributions to Publications Citing or Cited By PIRE and Comparison Project Articles

Two of the research outcomes we examined in Volume 1, Chapter 3 concerned the percentage of foreign institutions represented among authors per paper for (1) citations of PIRE and comparison group articles and (2) papers cited by the PIRE and comparison groups' articles. Exhibit 4.1 repeats the key finding of the regression-adjusted means for 45 matched pairs of PIRE and comparison projects.

Exhibit 4.1: Citations of, and Citations by, PIRE or Comparison Projects' Articles

	Mean PIRE	Mean Comparison	Difference	p-Value[3]	Standard Error
Percentage of non-U.S. institutions on articles citing project articles (PIRE, comparison group)[1]	65.5%	53.7%	11.7	< .01**	2.12
Percentage of non-U.S. institutions on articles cited by project articles (PIRE, comparison group)[2]	54.2%	47.9%	6.3	< .01**	1.55

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science

Notes:

1 Across papers citing one or more PIRE or comparison project-produced publications, the percentage of authors' institutions per citing publication that were located outside the U.S.

2 Across papers that were cited by one or more PIRE or comparison project-produced publications, the percentage of authors' institutions per cited publication that were located outside the U.S.

3 ** indicates that the difference was statistically significant at p<.01.

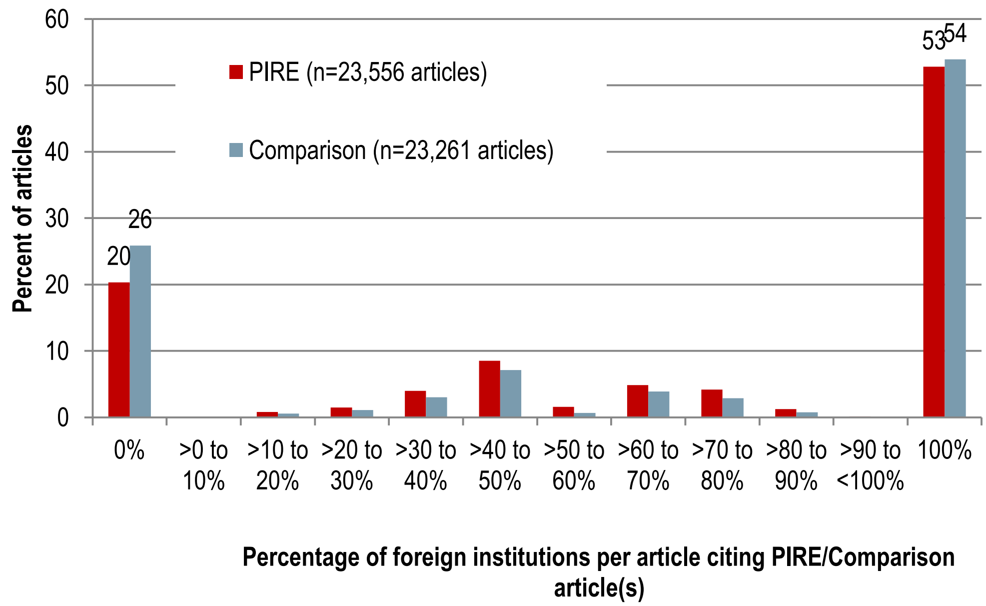
As a supplement to these results, we note that the distributions of the percentage of foreign institutions per article were strongly bi-modal. We first discuss the distribution of the percentage of foreign institutions per article citing PIRE or comparison project articles.

4.1.1 Foreign Institutions on Papers Citing PIRE or Comparison Project Articles

Across papers citing PIRE-produced articles, the proportion of foreign institutions per paper tended to be either 0 or 1 (or nearly 0, nearly 1). Exhibit 4.2 shows the percentage of articles (on the y-axis) for which the proportion of foreign institutions of authors was 0, between 0 and 0.1, between 0.1 and 0.2, etc., up to 1 (along the x-axis). The distribution has a peak at 0 (meaning none of the authors' institutions were foreign) and a more pronounced peak at 1 (meaning that all of the authors' institutions were foreign). Although Exhibit 4.1 shows that the average percentage of foreign institutions on articles citing PIRE articles (65.5 percent) is higher than that for comparison projects' articles (53.7 percent), note the following from Exhibit 4.2:

- 20 percent of PIRE articles, compared to 26 percent of comparison group articles, were cited by papers with no coauthors from a foreign institution; and
- 54 percent of PIRE articles, compared to 53 percent of comparison group articles, were cited by papers where 100 percent of the coauthors were from foreign institutions.

Exhibit 4.2: Distribution of the Proportion of Foreign Institutions per Article among Articles Citing PIRE or Comparison Project-Produced Research Article(s)[1]



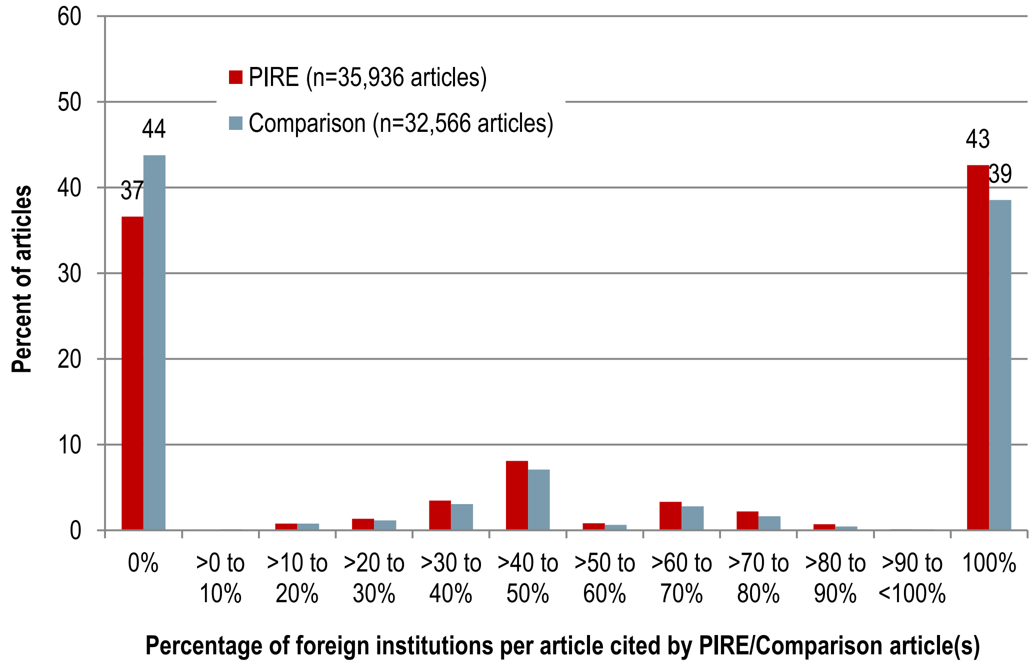
Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science.
Notes:
[1] Across papers citing one or more PIRE or comparison project-produced publications, the percentage of authors' institutions per citing publication that were located outside the U.S.

4.1.2 Foreign Institutions on Papers Cited by PIRE or Comparison Project Articles

Across papers cited by PIRE or comparison project-produced articles, the proportion of foreign institutions per paper tended to be either 0 or 1. Exhibit 4.3 shows the percentage of articles (on the y-axis) for which the percentage of foreign institutions of authors was 0, between 0 and 10 percent, between 10 and 20 percent, etc., up to 100 percent (along the x-axis). The distribution has a peak at 0 percent (meaning none of the authors' institutions were foreign) and a more pronounced peak at 100 percent (meaning that all of the authors' institutions were foreign). Although Exhibit 4.1 shows that the average percentage of foreign institutions on articles cited by PIRE articles (54.2 percent) is higher than that for articles cited by comparison projects' articles (47.9 percent), note the following from Exhibit 4.3:

- 37 percent of articles cited by PIRE research articles, compared to 44 percent of articles cited by comparison group research articles, were papers with no coauthors from a foreign institution;
- 43 percent of articles cited by PIRE articles, compared to 39 percent of articles cited by comparison group articles, were papers with more than 90 percent foreign institutions represented among the authors.

Exhibit 4.3: Distribution of the Proportion of Foreign Institutions per Article among Articles Cited by PIRE or Comparison Project-Produced Research Article(s)[1]



Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science.
Notes:
[1] Across papers that were cited by one or more PIRE or comparison project-produced publications, the percentage of authors' institutions per cited publication that were located outside the U.S.

4.2 Pending Publications, Conference Papers and Other Research Products by PIRE and Comparison Group Participants

In Volume 1, Chapter 3 we reported findings from analyses of published journal articles co-authored by PIRE and comparison PIs, postdocs and graduate students. In addition, we collected self-reported data from survey respondents in these participant groups on the following other types of research products:

- Pending publications (research reports in press, under review or in preparation);
- Invited talks;
- Conference abstracts, posters or presentations;
- Patents; and
- Other products disseminated to the scholarly community or public (e.g., software).

Exhibit 4.4 summarizes the average number of pending publications, invited talks, conference abstracts, posters or presentations and other research products, as well as the average percentage of foreign co-contributors across each type of product for each respondent group. Exhibit 4.5 summarizes the results for patents. PIRE and comparison group differences were not tested for statistical significance.

Exhibit 4.4. Publications in Preparation, Other Research Products and Percent Foreign Coauthors: PIs, Postdoctoral Researchers and Graduate Students

	PIs				Postdocs				Graduate Students			
	Mean Number		% Foreign		Mean Number		% Foreign		Mean Number		% Foreign	
	PIRE	Comparison	PIRE	Comparison	PIRE	Comparison	PIRE	Comparison	PIRE	Comparison	PIRE	Comparison
Peer-reviewed research reports, in press[1]	2.8	1.7	65%	35%	0.6	0.7	59%	46%	0.7	0.7	66%	17%
Research reports under review[2]	0.8	1.0	69%	33%	0.5	0.5	54%	41%	0.2	0.3	52%	23%
Manuscripts in preparation[3]	2.1	2.3	77%	28%	1.3	1.3	66%	39%	1.0	1.0	61%	19%
Invited talks[4]	9.1	7.0	34%	10%	2.9	2.1	28%	19%	1.2	1.2	29%	10%
Conference abstracts, posters or presentation[5]	11.8	11.5	58%	25%	2.8	3.5	54%	25%	2.6	3.7	51%	17%
Other products disseminated to the scholarly community or public[6]	1.0	1.1	65%	11%	0.4	0.3	52%	27%	0.3	0.5	52%	23%

Sources: Principal Investigator Survey, Items D3a, D3b, D4a; Postdoctoral survey, Items C2a, C2b, C3; Graduate student survey, Items C2a, C2b, C3.

Notes:

1 Research reports in press: Mean: PI, PIRE N=168, missing=2; Comparison N=114. PD, PIRE N=115, missing=7; Comparison N=112, missing=9. Grad, PIRE N=269, missing=16; Comparison N=326, missing=5. Percent foreign: PI, PIRE N=74; Comparison N=43, missing=2. PD, PIRE N=35; Comparison N=30, missing=1. Grad, PIRE N=61, missing=23; Comparison N=92, missing=11.

2 Research reports under review: Mean: PI, PIRE N=167, missing=3; Comparison N=114. PD, PIRE N=117, missing=5; Comparison N=110, missing=11. Grad, PIRE N=265, missing=20; Comparison N=327, missing=4. Percent foreign: PI, PIRE N=75; Comparison N=53, missing=4. PD, PIRE N=39; Comparison N=24, missing=2. Grad, PIRE N=47, missing=13; Comparison N=64, missing=14.

3 Manuscripts in preparation: Mean: PI, PIRE N=169, missing=1; Comparison N=114. PD, PIRE N=117, missing=5; Comparison N=110, missing=11. Grad, PIRE N=272, missing=13; Comparison N=326, missing=5. Percent foreign: PI, PIRE N=125, missing=1; Comparison N=89, missing=4 PD, PIRE N=70, missing=1; Comparison N=58, missing=4. Grad, PIRE N=126, missing=22; Comparison N=167, missing=16.

4 Invited talks: Mean: PI, PIRE N=169, missing=1; Comparison N=113, missing=1. PD, PIRE N=115, missing=7; Comparison N=112, missing=9. Grad, PIRE N=280, missing=5; Comparison N=328, missing=3. Percent foreign: PI, PIRE N=135; Comparison N=95. PD, PIRE N=67; Comparison N=51. Grad, PIRE N=121, missing=5; Comparison N=113, missing=4

5 Conference abstracts, posters, presentations: Mean: PI, PIRE N=169, missing=1; Comparison N=113, missing=1. PD, PIRE N=115, missing=7; Comparison N=111, missing=10. Grad PIRE N=280, missing=5; Comparison N=325, missing=6. Percent foreign: PI, PIRE N=146; Comparison N=97. PD, PIRE N=81; Comparison N=79, missing=2. Grad, PIRE N=220, missing=8; Comparison N=261, missing=13.

6 Other products: Mean: PI, PIRE N=164, missing=6; Comparison N=113, missing=1. PD, PIRE N=114, missing=8; Comparison N=111, missing=10. Grad, PIRE N=272, missing=13; Comparison N=320, missing=11. Percent foreign: PI, PIRE N=54, missing=1; Comparison N=37. PD, PIRE N=27; Comparison N=15, missing=2. Grad, PIRE N=47, missing=13; Comparison N=64, missing=14.

Fewer than 5 percent of PIs, postdocs or graduate students reported that they had contributed to any patents (Exhibit 4.5). Three PIRE PIs indicated that some proportion of their patents had foreign contributors and one PIRE and one comparison graduate student each reported that their patents had foreign contributors.

Exhibit 4.5 Distribution of the Number of Patents Reported by PIs, Postdocs and Graduate Students

	Number of PIs		Number of Postdocs		Number of Graduate Students	
Number of patents reported	PIRE	Comparison	PIRE	Comparison	PIRE	Comparison
0	163	109	114	109	269	309
1	3	2	0	2	4	14
2	1	2	0	0	0	0
3	0	0	0	0	1	0
4	0	0	0	0	0	0
5	1	0	0	0	0	0
	Number of PIs		Number of Postdocs		Number of Graduate Students	
% of patents with a foreign co-inventor	PIRE	Comparison	PIRE	Comparison	PIRE	Comparison
0%	2	4	NA	2	4	12
20%	0	0	NA	0	0	0
40%	1	0	NA	0	0	0
60%	0	0	NA	0	0	0
100%	2	0	NA	0	1	1

Sources: *Principal Investigator Survey, Item D4a; Postdoctoral survey, Item C3; Graduate student survey, Item C3.*
Notes: Number of patents: PIs, PIRE: missing=2, Comparison: missing=1. Postdocs: PIRE: missing = 8, Comparison: missing =10. Graduate students: PIRE: missing=11, Comparison: missing =8.
Percent foreign: PIs, PIRE: missing=0, Comparison: missing=0. Postdocs: PIRE: missing = 0, Comparison: missing =0. Graduate students: PIRE: missing=0, Comparison: missing=1.

4.3 PIRE Projects with Partners in Developed versus Developing Countries

Of interest to NSF was whether there were differences in the quantity or quality of research (or the extent to which publications included foreign authors) for PIRE projects depending on the economic status of the nations in which foreign partners were located. That is, NSF wanted to know: Do PIRE projects that collaborate solely with individuals (or institutions) in developed nations have different research outcomes than PIRE projects that collaborate with individuals in developing nations?

4.3.1 Methods

Using NSF administrative data on the 59 PIRE projects included in the evaluation, we identified each country in which foreign collaborating personnel and institutions were located. Next, using OECD's Development Assistance Committee (DAC) data on countries receiving Official Development Assistance, we determined whether each country was a developed or developing nation. (The DAC list of ODA recipients is based on the gross national income per capita determined by the World Bank, but excludes the G8, EU members and countries scheduled for entry into the EU. This list includes nations identified by the United Nations as Least Developed Countries.)

We classified each PIRE project into two mutually exclusive categories:

A. All foreign partners were in developed nations;
B. Not all foreign partners were in developing nations.

We examined the relationship PIRE projects' classification (i.e., into Group A or Group B) and four research outcomes. For each outcome, we tested whether research outcomes for PIRE projects in group A (all foreign partners were in developed nations) differed from those of other PIRE projects.

To test for differences in the number of articles produced by PIRE projects in group A compared to PIRE projects in group B, we used a t-test at the project level to compare the means:

$$T = \frac{\bar{y}_d}{SE_p}$$

Where T is the *t*-score.

$\bar{y}_d$ is the mean difference between group A and group B.

SE_p is the pooled standard error:[10]

$$SE_p = \sqrt{\left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right)}$$

Where:

n_1 is the number of projects in group A.

n_2 is the number of projects in group B.

s_1 is the standard deviation of the outcome of projects in group A.

s_2 is the standard deviation of the outcome of projects in group B.

To examine the relationship between field NCI, journal NCI, percentage of foreign authors and percentage of foreign institutions on articles and PIRE projects' classification we used two-level hierarchical linear model where articles were nested within projects and the group indicator appeared in level-2 (i.e., the project level):

Level-1 Model: Article Level

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

where

Y_{ij} is the outcome for the i^{th} article in the j^{th} project.

β_{0j} is the mean of the outcome for articles at the j^{th} project.

[10] We failed to reject the null hypothesis, H_0: the variance of group A = variance of group B, so we used the pooled variance for the t-test.

ε_{ij} is the random effect representing the difference between article *ij*'s outcome and the predicted mean outcome for project *j*. These residual effects are assumed normally distributed with mean 0 and variance σ^2.

Level-2 Model: Project Level

$\beta_{0j} = \gamma_{00} + \gamma_{01}(X_j) + \mu_{0j}$

where

X_j is 1 if project *j* is in group A and 0 if in group B.

γ_{00} is the mean outcome for group B.

γ_{01} is the difference between group A and group B means.

μ_{0j} is the deviation of project *j*'s mean from the grand mean—this effect is assumed normally distributed with mean 0 and variance τ^2.

Combining the article and project level equations give the following single equation model:

$Y_{ij} = \gamma_{00} + \gamma_{01}X_j + \mu_{0j} + \varepsilon_{ij}$

4.3.2 Results

Exhibit 4.6 summarizes the results. There were no statistically significant differences between research outcomes for PIRE projects working only with foreign partners from developed nations and projects working with foreign partners from developing (or developing and developed) nations.

Exhibit 4.6: Research Outcomes for PIRE Projects With Foreign Partners from Developed or Developing Nations

Indicator	Foreign Partners in Developed Nations Only	Foreign Partners in Developing (and Developed) Nations	Foreign Partners in Developed Nations Only	Foreign Partners in Developing (and Developed) Nations	Difference	Standard Error of the Difference	P-value
	N of PIRE Projects with Least One Article Identified In Web of Science		Mean Number of Articles				
Number of published articles[1]	21 projects	30 projects	28.0	27.1	0.9	6.7	.895
	N of Articles Identified with the Indicator		Mean across Articles				
Field Normalized Citation Impact (NCI)[2]	556	783	1.67	1.93	-0.26	0.27	.329
Journal NCI[2]	571	797	1.11	1.42	-0.31	0.17	.072
Percentage of foreign authors on articles (published 2009 or later)[2]	516	637	28.5%	33.0%	-4.5%	4.3%	.292
Percentage of foreign institutions on articles[2]	587	811	36.2%	36.0%	0.1%	4.3%	.977

Sources: Annual reports submitted by PIs to NSF; Thomson Reuters' Web of Science; http://www.oecd.org/dac/stats/historyofdaclistsofaidrecipientcountries.htm; http://unohrlls.org/about-ldcs/

Notes:

1 Six PIRE projects reported no publications in their annual reports; for 2 additional projects, none of the reported references were identified in the Web of Science.

2 Analyses were limited to articles where the indicator was available for that article.

5. Participant Experiences: Supplementary Material

In Volume 1 of the report, we report extensive findings comparing PIRE participants to participants in a matched comparison group of other NSF projects. Here, we take a different approach and compare PIRE postdoctoral, graduate and undergraduate participants to national trends for selected employment outcomes using data from two nationally representative surveys:

- The 2013 Survey of Doctoral Recipients (SDR); and
- The 2010 National Survey of Recent College Graduates (NSRCG).

The outcomes of interest for these analyses were:

- Current employment status;
- Type of expertise job requires;
- Employed at an institution of higher education;
- Faculty rank;
- Tenure status; and
- Years since earning tenure.

Although we identified and selected a sub-sample of respondents from these surveys who had similar characteristics as PIRE participants, because no matching was possible, the results of this exercise must be viewed as exploratory and suggestive only. We use these data to show a national "benchmark" against which to compare PIRE participants, with the caveat that there are a multitude of potential reasons, aside from participation in the PIRE program, for any observed differences. Primary among these reasons is that former PIRE participants differed from the sub-samples of SDR/NSRCG respondents even before participation in the PIRE project.

5.1 Methods for Comparing PIRE Participants to National Trends

To compare PIRE postdoctoral, graduate student and undergraduate participants' career outcomes to nationally representative trends, we used data from the 2013 SDR and the 2010 NSRCG. Only former PIRE participants were included in benchmarking analyses of employment outcomes: survey respondents who were still participating in an active PIRE project were not asked about employment outcomes.

From the SDR, we selected a sub-sample of respondents with characteristics similar to the PIRE postdoctoral participants. From the NSRCG, we selected two sub-samples of respondents, one with characteristics similar to former PIRE graduate students, and one with characteristics similar to former PIRE undergraduate students.

After selecting samples that were as similar as possible on selection criteria, we conducted a chi-square or t-test to determine whether or not there were significant differences between the PIRE and SDR (or NSRCG) samples. Next, we used stepwise regression models (see Section 3.2), each with an employment outcome as the dependent variable and each including demographic characteristics as control variables to test the significance of any observed differences between the PIRE and national groups.

5.2 Former PIRE Postdoctoral Participants vs. Doctoral Degree Holders

5.2.1 Comparing Former PIRE Postdocs to SDR 2013 Respondents

Former postdoctoral PIRE participants were assumed to have completed their doctorate prior to the start of participation in PIRE. Furthermore, we assumed that postdoctoral participants had completed their doctorate in the same field they reported as their primary research field on the survey. Only respondents who had completed their postdoctoral appointment were included (51 current PIRE postdoctoral participants were excluded). The resulting sample included 66 respondents.

The sample of 2013 SDR respondents (n=30,696) was subset to include only those respondents who had received their doctorate within 20 years of the PIRE survey (February 2015), who were not currently employed in a postdoctoral position, and who reported that their field of research was in the biological or physical sciences, computer sciences, engineering, mathematics or statistics, psychology, economics, political science, sociology or another social science field (839 respondents who reported that their field was "health or related field" were dropped because no PIRE postdocs reported any health-related field). The resulting SDR sub-sample included 15,754 respondents.

Although respondents to the 2013 SDR who were currently employed as a postdoc were excluded, it was not possible to identify a sub-sample of SDR respondents who had held a postdoctoral appointment in the past. This is an important difference between the PIRE and SDR samples: the sample of former PIRE postdoctoral participants is, by definition, a group that has held at least one postdoctoral appointment after completing their doctorate; in contrast, the SDR sub-sample includes individuals who held a postdoctoral appointment and those who had never held a postdoc. Holding a postdoctoral position is likely to be strongly correlated with the employment outcomes described below.

5.2.2 Characteristics of PIRE Postdocs and SDR 2013 Sub-Samples

Exhibit 5.1 shows demographic and other characteristics of the PIRE and SDR samples. Note that there were statistically significant differences between the two groups for membership in a group traditionally under-represented in S&E (i.e., a greater percentage of PIRE postdocs), citizenship status (more PIRE postdoctoral participants than the national average were non-U.S. citizens with temporary visas), and disability status (PIRE postdocs were less likely to be disabled than the national average). In addition, former PIRE postdocs had, on average, received their doctorate more recently than the sub-sample; and the fields of research among PIRE postdocs differ from the fields of research represented in a national sub-sample of doctorates.

Exhibit 5.1: Characteristics of PIRE Postdocs & SDR 2013 Sub-samples

	PIRE (N=66)		SDR (N=15,754)		
	N	%	N	%	p-Value
Demographic Characteristics					
Gender					0.9198
Female	25	40%	6,775	39%	
Male	38	60%	8,979	61%	
Minority status					0.0316
Not under-represented minority	49	79%	12,854	91%	
Under-represented minority	13	21%	2,900	9%	
Citizenship					0.0001
U.S. citizen	32	51%	12,742	80%	
Non-U.S. citizen with permanent visa	10	16%	2,178	15%	
Non-U.S. citizen with temporary visa/other non-U.S. citizen	21	33%	834	5%	
Disability status					0.0081
Not disabled	56	90%	12,581	80%	
Disabled	6	10%	3,173	20%	
Other Characteristics					
Average number of years since doctorate	7.57		10.53		<0.0001
Primary field of research					<0.0001
Biology & agricultural/environmental studies	12	19%	3,971	25%	
Computer science & math	8	13%	1,448	9%	
Chemistry, physics, & earth science	30	48%	2,407	16%	
Engineering	7	11%	3,229	22%	
Social science	6	10%	4,699	28%	

Sources: Postdoctoral Survey, Evaluation of NSF's PIRE program; SDR 2013.
Notes: Number missing, PIRE: Gender: 3; Minority status: 4; Citizenship: 3; Disability: 4; Number of years since doctorate: 0; Field of research: 3.

5.2.3 Employment Outcomes

Exhibit 5.2 shows that former PIRE postdocs were less likely to be employed at the time of the study survey (February 2015) than the 2013 SDR sub-sample; were less likely to report that their job duties required the expertise of a bachelor's degree in a field other than S&E or social sciences; if employed, were more likely to be employed at a college, university or other institution of higher education; and had earned tenure more recently (most likely due, in part, to having received their doctoral degrees more recently as well).

Exhibit 5.2: Comparisons of PIRE Postdoctoral Participants To a Sub-Sample of SDR 2013 Respondents

	PIRE		SDR		
	N	%	N	%	p-Value
Total Sample Size	66		15,754		
Current employment status					0.0018
Unemployed	13	21%	806	5%	
Employed	49	79%	14,948	95%	
Of those Employed:					
Job duties require the expertise of a bachelor's degree or higher in…					
Engineering, computer science, math or the natural sciences	44	90%	10,720	73%	0.0761
Social sciences	7	14%	4,977	32%	0.4339
Other field	5	10%	4,186	27%	0.0093
Employed at an institution of higher education	35	71%	6,687	43%	<0.0001
Of those employed at an institution of higher education:					
Faculty rank					0.8783
Professor	2	6%	838	17%	
Associate professor	7	22%	1,765	33%	
Assistant professor	17	53%	2,266	40%	
Instructor, lecturer, or other	6	19%	602	11%	
Tenure status					0.7034
Tenured	8	29%	2,196	45%	
Not tenured	20	71%	2,869	55%	
Average number of years since tenure (among tenured respondents)	2.38		5.93		0.0124

Sources: Postdoctoral Survey, Evaluation of NSF's PIRE program; SDR 2013.
Notes: Number Missing (PIRE): Employment: 4.
Participants who responded that tenure and faculty ranks were not available at their institutions or positions were excluded from the analysis. P-values were calculated using a regression with control variables selected though backward and forward step-wise regression

5.3 Former PIRE Graduate Student Participants vs. Recent College Graduates

5.3.1 Comparing Former PIRE Graduate Students to NSRCG 2010 Respondents

Former PIRE graduate student participants who reported that their highest degree was a master's or bachelor's degree were included; those indicating that they had received a doctoral degree were excluded (n=108 excluded). Only those not currently enrolled in graduate school were included in the PIRE sample (n=32 excluded) and who had completed their degree within four years of the survey date (n=43 excluded; the NSRCG only invites respondents who have completed college within three years of the survey date). Finally, data on demographic and degree characteristics were missing for three PIRE graduate students. The resulting sample included 30 respondents.

The sample of 2010 NSRCG respondents (n=12,326) was subset to include only those respondents whose highest degree was a bachelor's or master's degree (any doctoral recipients were excluded), who were not currently enrolled in a graduate school, and who reported that the field of research for their *most recent degree* was in the biological or physical sciences, computer sciences, engineering, mathematics or statistics, psychology, economics, political science, sociology or another social science field (1,269 respondents who reported that their field was one of several "health" fields were dropped, as were 46 who reported that their field was business, managerial, humanities or communications, because no PIRE graduate students reported that they were pursuing a graduate degree in any such field).[11] The resulting NSRCG sub-sample included 5,738 respondents.

5.3.1 Characteristics of former PIRE graduate students and NSRCG 2010 sub-samples

Exhibit 5.3 shows demographic and other characteristics of the PIRE and NSRCG sub-samples. There were statistically significant differences between the two groups for disability status and highest degree received (a greater percentage of former PIRE graduate students had completed a master's degree than respondents in the NSRCG sub-sample), and former PIRE graduate students had received their highest degree more recently, on average, than those in the NSRCG sub-sample. The distribution of primary fields of research among PIRE graduate students also differs from the distribution of fields of national sub-sample of recent college graduates' most recent degree.

Exhibit 5.3: Characteristics of Former PIRE Graduate Students & NSRCG 2010 Sub-Samples

	PIRE (N=30)		NSRCG (5,738)		
	N	%	N	%	p-Value
Demographic Characteristics					
Gender					0.1307
Female	15	50%	2,784	36%	
Male	15	50%	2,954	64%	
Minority status					0.5150
Not under-represented minority	26	87%	3,377	83%	
Under-represented minority	4	13%	2,361	17%	
Citizenship					0.6255
U.S. citizen	26	87%	4,972	88%	
Non-U.S. citizen with permanent visa	2	7%	255	3%	
Non-U.S. citizen with temporary visa/other non-U.S. citizen	2	7%	511	9%	
Disability status					0.0015
Not disabled	29	97%	4,775	84%	
Disabled	1	3%	963	16%	
Other Characteristics					
Highest degree received					<0.0001
Master's or professional	28	93%	2,851	27%	
Bachelor's	2	7%	2,887	73%	
Average number of years since highest degree	1.74		2.05		<0.0001
Primary field of research					0.0019
Biology & agricultural/environmental studies	7	21%	649	20%	
Computer science & math	2	6%	851	23%	
Chemistry, physics, & earth science	10	30%	527	5%	
Engineering	12	36%	2,872	30%	
Social science	2	6%	839	22%	

Sources: Graduate Student Survey, Evaluation of NSF's PIRE program; NSRCG 2010.

5.3.1 Employment Outcomes

Exhibit 5.4 shows that a smaller proportion of former PIRE graduate students were employed at the time of the study survey (73 percent) relative to the NSRCG sub-sample (89 percent), but this difference was not statistically significant. Former PIRE graduate student participants were more likely to report that their current job required expertise in an S&E field (95 percent, relative to 66

[11] Because the PIRE graduate survey did not ask respondents for the field of their highest degree, but rather, the field of research for the degree they were pursuing at the time of participation in the PIRE project, we opted to use NSRCG respondents' reported field for their most recent degree.

percent of the NSRCG sub-sample), and were more likely to be employed at a college, university or other institution of higher education (41 percent vs. 6 percent). (The NSRCG does not include questions about faculty rank or tenure status.)

Exhibit 5.4: Comparisons of Former PIRE Graduate Participants To a Sub-Sample of NSRCG 2010 Respondents

	PIRE		NSRCG		p-Value
	N	%	N	%	
Total Sample Size	30		5,738		
Current employment status					0.0766
Unemployed	8	27%	573	11%	
Employed	22	73%	5,165	89%	
Of those Employed:					
Job duties require the expertise of a bachelor's degree or higher in…					
Engineering, computer science, math or the natural sciences	21	95%	3,877	66%	<0.0001
Social sciences	3	14%	955	19%	0.7393
Other field	1	5%	919	18%	0.0508
Employed at an institution of higher education	9	41%	346	6%	0.0010

5.4 Former PIRE Undergraduate Participants vs. Recent College Graduates

5.4.1 Comparing former PIRE undergraduate participants to NSRCG 2010 respondents

The sample of former PIRE undergraduate participants (n=278) was limited to those who: had completed their bachelor’s degree at the time of the survey (42 excluded); whose highest degree was a bachelor’s or master’s degree (15 excluded); who had received this degree within four years of the survey date (13 excluded); and who were not currently enrolled in school (123 excluded). One additional undergraduate who did not report a field of study for highest degree earned was excluded from the sample. The resulting sample included 91 former PIRE undergraduates.

The sample of 2010 NSRCG respondents (n=12,326) was subset to include only those respondents whose highest degree was a bachelor’s or master’s degree (any doctoral recipients were excluded), who were not currently enrolled in a graduate school, and who reported that the field for their highest degree was in the biological or physical sciences, computer sciences, engineering, mathematics or statistics, psychology, economics, political science, sociology or another social science field, one of several “health” fields, or a business, managerial, humanities or communications field. The resulting NSRCG sub-sample included 6,685 respondents.

5.4.2 Characteristics of Former PIRE Undergraduate Participants and NSRCG 2010 Sub-Samples

Exhibit 5.5 shows demographic and other characteristics of the PIRE and NSRCG sub-samples. A greater percentage of former undergraduate PIRE participants were male than respondents in the NSRCG sub-sample. On average, former undergraduate PIRE participants had completed their highest degree more recently than the respondents in the NSRCG sub-sample, and the distribution of fields for this highest degree differed from the distribution of fields in the NSRCG sub-sample. These differences were each statistically significant.

Exhibit 5.5. Characteristics of Former PIRE Undergraduate Participants vs. Recent College Graduates

	Former PIRE Undergraduates (N=91)		NSRCG (N=6,685)		
	N	%	N	%	p-Value
Demographic Characteristics					
Gender					0.0199
Female	37	41%	3,512	53%	
Male	54	59%	3,173	47%	
Minority status					0.9594
Not under-represented minority	74	82%	3,992	82%	
Under-represented minority	16	18%	2,693	18%	
Citizenship					0.2544
U.S. citizen	86	95%	5,870	91%	
Non-U.S. citizen with permanent visa	1	1%	285	3%	
Non-U.S. citizen with temporary visa/other non-U.S. citizen	4	4%	530	7%	
Disability status					0.4277
Not disabled	73	80%	5,565	84%	
Disabled	18	20%	1,120	16%	
Other Characteristics	**N**	**%**	**N**	**%**	**p-Value**
Highest degree received					0.1172
Master's or professional	23	25%	3,352	33%	
Bachelor's	68	75%	3,333	67%	
Average number of years since highest degree	1.76		2.06		<.0001
Primary field of research					<.0001
Biology & agricultural/environmental studies	30	33%	649	13%	
Computer science & math	8	9%	851	15%	
Chemistry, physics, & earth science	13	14%	527	3%	
Engineering	24	26%	2,872	19%	
Social science	16	18%	1,786	50%	

Sources: Undergraduate Student Survey, Evaluation of NSF's PIRE program; NSRCG 2010.
Notes: Number missing, PIRE: Minority status: 1

5.4.1 Employment Outcomes

Exhibit 5.6 shows that former PIRE undergraduates were equally likely to be employed at the time of the survey, but were more likely to hold a job whose duties required expertise in S&E and less likely to hold a job that required expertise in a field other than social sciences or S&E (such as a job in business, health, or education). No other differences were statistically significant.

Exhibit 5.6: Comparisons of Former PIRE Graduate Participants To a Sub-Sample of NSRCG 2010 Respondents

	PIRE		NSRCG		
	N	%	N	%	p-Value
Total Sample Size	**91**		**6,685**		
Current employment status					0.9549
Unemployed	12	13%	637	9%	
Employed	79	87%	6,048	91%	
Of those Employed:					
Job duties require the expertise of a bachelor's degree or higher in…					
Engineering, Computer Science, Math or the Natural Sciences	61	77%	4,160	52%	<.0001
Social Sciences	19	24%	1,208	21%	0.0664
Other Field	28	35%	1,493	36%	0.0057

Definitions and Acronyms

Continuing grant. An NSF award in which a full amount is authorized but released in annual sub-amounts pending satisfactory progress. Differs from a standard grant in which all funds are released at once at the start of the award.

FI = Foreign senior investigator. A non-U.S.-based participant in an NSF award who is equivalent to a co-principal investigator in the U.S.

Field normalized citation impact (field NCI). The ratio of the number of citations of the target paper to the "world average baseline" which Thomson Reuters defines as the average number of citations per paper for all reviews or research articles published in the same year as the target paper in a journal within the same journal category as the journal in which the target article was published. Information about journal categories can be found here:
http://ip-science.thomsonreuters.com/mjl/scope/scope_scie/ and here:
http://wokinfo.com/essays/journal-selection-process/

Journal normalized citation impact (journal NCI). Also known in some literatures as the ratio of observed to expected citations, is the ratio of number of citations of the target paper to the average number of citations per article for all reviews and research articles published in the same year and in the same journal as the target paper.

Percentage of foreign authors per article is the ratio of the number of authors based at non-U.S. institutions to the number of unique authors of the article. For example, for an article with 10 co-authors, 4 of whom are based in two institutions outside the U.S. and 6 of whom are based at three institutions within the U.S., the article's percentage of foreign authors is 40 percent. This indicator is available only for articles published in 2009 or later. For articles published earlier than 2009, Thomson Reuters cannot disambiguate multiple authors at the same institution. Instead, the percentage of institutions listed in author information that are outside the U.S. is used as a proxy for the relative contribution of non-U.S. co-authors to the article. See the definition for an article's percentage of foreign institutions, below.

Percentage of foreign institutions per article is the ratio of the number of unique non-U.S. institutions out of the number of unique institutions listed in author information per article. Because Thomson Reuters cannot disambiguate multiple authors at the same institution for articles published earlier than 2009, the ratio of non-U.S. institutions is used as a proxy for the relative contribution of non-U.S. researchers to the article. For example, a paper with 10 co-authors based in two institutions outside the U.S. and three institutions within the U.S., the article's percentage of non-U.S.-based institutions is 40 percent (two out of five institutions total are non-U.S.).

Percentage of foreign institutions per article for articles citing PIRE/comparison publications. Across papers citing one or more PIRE or comparison project-produced publications, the percentage of authors' institutions per citing publication that were located outside the U.S.

Percentage of foreign institutions per article for articles cited by PIRE/comparison publication(s) Across papers that were cited by one or more PIRE or comparison project-produced publications, the percentage of authors' institutions per cited publication that were located outside the U.S..

PI = Principal investigator. A senior-level, U.S.-based researcher on an NSF award, either the lead principal investigator or a co-principal investigator.

Pre-onset. The period of time prior to a PIRE or comparison project participant's first year of participation in the project.

Post-onset. The period of time including and after a PIRE or comparison project participant's first year of participation in the project.

S&E. Science and engineering. Used to refer collectively to the scientific or engineering fields typically supported by NSF.

STEM: Science, technology, engineering and mathematics. Similar to S&E, this acronym is used in the literature to refer collectively to the set of fields in these four major areas. Sometimes it includes fields not directly supported by NSF (e.g., biomedical fields).

Sponsoring institution. The home institution of the lead principal investigator on an NSF award.

References

Bloom, H.S. (2003). Using “Short” Interrupted Time Series-Analysis to Measure the Impacts of Whole-School Reforms. *Evaluation Review* 27(1), 3–49.

Budtz-Jorgensen, E., Keilding, N., Grandjean, P., & Weihe, P. (2007). Confounder identification in environmental epidemiology: Assessment of health effects of prenatal mercury exposure. *Annals of Epidemiology*, 17(1), 27–35.

Maldonado, G., & Greenland, S. (1993). Simulation study of confounder-selection strategies. *American Journal of Epidemiology* 138(11), 923–936.