

B. Collections of Information Employing Statistical Methods

The agency should be prepared to justify its decision not to use statistical methods in any case where such methods might reduce burden or improve accuracy of results. When the question “Does this ICR contain surveys, censuses or employ statistical methods” is checked, “Yes,” the following documentation should be included in the Supporting Statement to the extent that it applies to the methods proposed:

- 1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.*

1. Survey of I-Corps Team members

The survey of I-Corps will include all teams who participated in an I-Corps program and completed the Time 1 survey (cleared under 3145-0238) and agreed to be contacted for Time 2. As NIH I-Corps teams started their program at a more advanced stage, this will be the first time they are contacted, I-Corps teams at NIH will be contacted for the first time. We include all previous I-Corps teams and anticipated awardees that will go through the program. A total of 2,000 I-Corps team members will be surveyed.

2. Survey of non-I-Corps PIs

This survey of non-I-Corps PIs uses statistical methods to determine the sample universe and the resulting estimate of PIs that have explored commercialization of their NSF-based product.

The potential respondent universe for the survey of non-I-Corps PIs includes PIs who received awards from five divisions that commonly “feed” the I-Corps Teams pipeline. The awards must have been received between 2009 and 2013. The five divisions have awarded over 11,000 projects in this time period. A detailed analysis of these awards yielded 7,897 unique awards. We excluded from the total number of awards: (1) PIs who received I-Corps Teams awards; (2) duplicate records for PIs who received more than one award from these divisions in the time period in question; and (3) PIs who only received awards that focus on organizing a conference, symposium, workshop, or another similar activity that does not involve conducting research. Exhibit A-4 shows the distribution of unique PIs per division.

Exhibit A-4. *Number of Awards per Top Divisions Feeding into the I-Corps Program*

NSF Division	Total # of PIs who received an award between 2009–2013
Chemical, Bioengineering, Environmental, and Transport Systems (CBET)	1,824
Division of Materials Research (DMR)	1,492
Civil, Mechanical, and Manufacturing Innovation (CMMI)	1,986
Division of Computer and Network Systems (CNS)	1,711
Division of Computing and Communication Foundations (CCF)	884
Total	7,897

We will limit the non-I-Corps PI sample to 7,897 PIs. The Time 2 survey of I-Corps PIs is based on a census of all I-Corps Teams during the period covered by the study PIs.

3. Case Study In-Depth Interviews

The case studies are not intended to yield statistical estimates and thus use qualitative methods to select cases that will provide illustrative examples of what teams have gone through during their efforts to explore commercialization.

For the 10 I-Corps Teams in the case study, we will work to select the case study PIs and teams from the 584 I-Corps Teams that will have completed the program by the end of 2015. We anticipate the selection of the projects/ Teams to follow a variety of criteria that are hypothesized to impact the Teams' experience with I-Corps. These criteria include:

- Geographical location of home institution;
- Seniority of PI or other demographic of interest (e.g., teams led by women and minorities);
- Readiness for commercialization; and
- Research domain.

For the 10 non-I-Corps PIs, we will rely on the results of the survey. For this reason, we anticipate that 720 respondents will be the possible universe from which we will select the non-I-Corps PIs. We will analyze responses to identify teams with a range of commercialization experiences. We will select the PIs, when possible, according to similar criteria used for the I-Corps Teams.

2. *Describe the procedures for the collection of information including:*
 - * *Statistical methodology for stratification and sample selection,*
 - * *Estimation procedure,*
 - * *Degree of accuracy needed for the purpose described in the justification,*
 - * *Unusual problems requiring specialized sampling procedures, and*
 - * *Any use of periodic (less frequent than annual) data collection cycles to reduce burden.*

We developed a three-step process to identify the comparable non-I-Corps projects using extant data and data collected from the comparison group survey:

1. **Select NSF directorates and programs.** To identify the appropriate comparison group, we need to understand which directorates and programs provided primary funding for research activities the I-Corps projects have undertaken. While NSF specifies participating directorates in this RFQ, we can refine the potential comparison group to reflect a similar distribution of PIs across participating directorates as those reflected in I-Corps (or in other programs, if desired). We may also catalogue the research fields and the funding lineage of I-Corps projects to use as selection criteria for the comparison group.
2. **Screen projects within selected directorates and programs.** In the first part of the survey, we will ask PIs of all projects funded by the selected programs/directorates about their motivation to commercialize their research products and experience with entrepreneurial trainings. These questions will allow us to determine the potential comparable projects with which we should follow up. The screening questions are not a separate instrument, but are designed to determine which PIs will be asked to complete the remainder of the instruments. Appendix B shows the questions to be used for the screening survey.
3. **Validate comparability of matched projects.** We will cross-validate survey responses with information extracted from Electronic Jacket (e-Jacket) and consolidate the final project sample. Eventually, the two groups of projects will be comparable in four ways:
 1. The two groups represent NSF directorates and programs in a similar way;
 2. Project PIs in both groups are motivated to commercialize their funded research;
 3. The matched projects have funded research stemming from similar fields or to be applied in similar fields; and
 4. The research products developed by the matched projects show a similar level of maturity or readiness for commercialization.

Statistical Power

Based on the number of I-Corps projects at the end of 2013, we assume an equal sample size of 255 projects for the comparison group to achieve a balanced sample between the I-Corps projects and non-I-Corps projects in the comparison group. With a total sample of 510 projects (255 I-Corps projects and 255 comparison group projects), we will be able to detect a program effect of 0.13 standard deviations or higher for a mean comparison using matched pair t-test and 0.22 or higher for regression-based analyses, which is a reasonable range for a program effect. Based on the number of I-Corps teams estimated to have completed the training workshops by early 2015 (more than 500), the precision would be substantially greater. The actual number of completed projects selected for the treatment and comparison groups would depend on the number having completed the longitudinal survey 1 year after their training.

3. *Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.*

The Time 2 longitudinal survey will be open for up to 3 months. The survey of non-I-Corps PIs will be open for 1 month to give PIs ample time for participation. Following a modified Dillman et al. (2009) method, we will use a strategy that includes encouragement from NSF, repeated invitations, and personalization.¹ Specifically, we will use the following strategy to encourage non-I-Corps PIs to participate:

- **Pre-survey notification from NSF.** NSF will send an initial e-mail to all participants who will receive the Web survey. This e-mail will add credibility to the e-mail solicitation that will come from MSG a day later in order to encourage participation and increase the response rate. MSG will provide a draft of the e-mail to NSF.
- **E-mail invitation from contractor to PIs.** We will e-mail an invitation to participate in the survey to the participants. We will personalize the e-mail invitations by addressing each PI by name by using MS Word's mail merge tool. In addition, we will also include the name of the NSF contact and a survey administrator contact so that participants may ask questions about the data collection. This technique increases response rates (Dillman et al., 2009). For non-I-Corps PIs, the e-mail invitation will include general information about the I-Corps Evaluation project, the importance of PI participation, a statement that participation is voluntary and data are confidential, and a link to the survey.
- **Personalized reminder.** A personalized reminder e-mail will be sent to participants approximately two weeks after the initial invitation to encourage anyone who has not yet participated and to thank those who completed the survey (Dillman et al., 2009).
- **Final reminder.** A final reminder e-mail will be sent to participants 48 hours before the close of the survey to again remind those who have not yet participated and to thank those who completed the survey.

MSG will ensure confidentiality of the data by maintaining a login and password interface for each user. Once data are submitted, the project will be saved in database format for storage and retrieval by either standard report formats or ad hoc database queries.

4. *Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of test may be submitted for approval separately or in combination with the main collection of information.*

¹ This method is associated with higher response rates than surveys that are administered without repeated invitations and personalization.

The research team conducted a cognitive test of the screener survey with eight PIs from I-Corps and non-I-Corps projects in 2015. The test enabled us to assess participants' understanding and clarity of the questions, response burden of survey items, and to measure the time it took respondents to complete the survey. Again in 2016, we received approval to recruit, pre-test and conduct cognitive interviews with PIs, as the instrument had been modified for the I-Corps participants. This enabled us to have a total of 36 completed surveys and 18 cognitive interviews completed. We revised the instrument to incorporate the findings of the pre-tests and interviews.

Appendix A shows the survey instrument and skip patterns for I-Corps and non-I-Corps participant groups. We also show the two instruments separately in Appendices B (I-Corps respondents) and C (non-I-Corps PIs).

5. *Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.*

NSF has contracted with MSG to conduct the Evaluation of I-Corps Teams Program MSG has subcontracted with Westat for all survey data collection. MSG will be responsible for data analysis related to the PI survey, while Westat will analyze the data collected in the semi-structure interviews.

Daniel Geller, Ph.D., Director of Evaluation Services, Manhattan Strategy Group
E-mail: dgeller@manhattanstrategy.com
Office: 301-828-1348

Marilia Mochel, Project Manager, Manhattan Strategy Group
E-mail: mmochel@manhattanstrategy.com
Office: 301-828-1512

Ying Zhang, Ph.D., Senior Researcher, Manhattan Strategy Group
E-mail: yzhang@manhattanstrategy.com
Office: 301-828-1346