PART B. Collections of Information Employing Statistical Methods

1.   Universe of Sampling Respondent Selection
Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method 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.

The U.S. Small Business Administration (SBA) is interested in deriving estimates of satisfaction and outcomes at the level of each cluster. Due to the relatively small size of each cluster's small business and large organization participation levels, a census survey of each group will be conducted.

A small business participant is defined as one that received services from its cluster in the past fiscal year and thus excludes "inactive" participants in the clusters. A large organization stakeholder in the cluster is defined by the administrator of the cluster but must have participated in some way in the cluster during the previous fiscal year.

Exhibit 1 details the expected small business and large organization survey population size and response rate for each site and overall. Exhibit 1 uses participant data from previous instances of survey implementation for the first seven clusters (which were already involved in the initiative) and uses the average number of participants across the original seven clusters to extrapolate the participant counts of the four clusters that are new to this initiative and will be surveyed for the first time under this request. The response rates used in exhibit 1 have been derived in a similar fashion (known figures from previous year for the first seven clusters and extrapolation for the remaining four), and the estimated number of surveys completed is the product of each cluster's estimated number of participants and estimated response rate.

Exhibit 1. Expected small business and large organization survey population size and response rate, per site and overall

Cluster	Est. number of small business participants	Est. response rate for small businesses	Est. number of surveys completed by small businesses	Est. number of large organization participants	Est. response rate for large organizations	Est. number of surveys completed by large organizations
Cluster 1	68	40%	27	8	38%	3
Cluster 2	34	91%	32	19	90%	18
Cluster 3	38	55%	22	9	33%	3
Cluster 4	267	12%	31	152	9%	14
Cluster 5	34	62%	22	22	46%	11
Cluster 6	52	69%	36	26	77%	20
Cluster 7	28	82%	23	10	80%	8
Cluster 8	75	37%	28	35	31%	11
Cluster 9	75	37%	28	35	31%	11
Cluster 10	75	37%	28	35	31%	11
Cluster 11	75	37%	28	35	31%	11
Total	**821**	**37%**	**305**	**386**	**31%**	**121**

Exhibit 1 does not discuss another source of data utilized in this evaluation: the cluster administrators. They are asked to provide data about services provided via an annual survey, and are interviewed twice a year to obtain qualitative information about the evolution of their organization and operations, etc. Again, there is no sampling, and all eleven cluster administrators will be asked to participate. In past years, all of them have submitted a survey and participated in the two annual interviews, leading to a 100% response rate (see exhibit 2, below). This pattern is expected to hold moving forward.

Exhibit 2. Expected cluster administrators survey population size and response rate, per data collection instrument

Data collection instrument	Total number of cluster administrators	Est. response rate for cluster administrators	Est. number of responses
Cluster Administrator Survey	11	100%	11
Mid-term cluster administrator calls	11	100%	11
Annual cluster administrator calls	11	100%	11

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.

As noted in question 1, no sampling strategy is used in this data collection. Data will be collected from all Regional Innovation Cluster (RIC) Initiative participants. The estimates are intended to be generalizable to the 11 RIC sites, which were selected to fulfill predefined SBA contract criteria but will not necessarily be representative of other regions of the United States or other locations in which the initiative may be deployed in the future. If the response rate falls below the 80 percent threshold included in the Office of Management and Budget (OMB) guidance, nonresponse bias analysis will be conducted and reported using business information collected in the Cluster Administrator Survey, which serves as the survey frame.

One objective of the study is to provide statistically valid estimates of the incremental change in key business outcomes of participants, including revenue and employment growth, as compared to the benchmark sample. To measure these differences, inferential statistics, such as the Student-T test, are used to compare the average difference between the two groups. The study will use a standard threshold of a 5 percent significance level to reject the null hypothesis of no difference. The benchmarking data come from the following sources:
- The Quarterly Census of Employment and Wages (from the Bureau of Labor Statistics), which provides data on the number of employees
- The Dun and Bradstreet (D&B) Business Database, which provides data on both revenue and number of employees

These data sources vary with respect to the frequency with which they are updated, the time periods covered, types of respondents, geographic and industrial granularity, and units of observation. However, the maximum amount of data and the closest match in time period will be used for the dataset.

3.   Describe Methods to Maximize Response Rates and Methods to Deal with Issues of Nonresponse.   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.

To maximize the survey response rate and minimize respondent burden, the following data-collection techniques will be used:

- Surveys will be introduced to the cluster participants via an official introduction e-mail from their respective cluster administrators. This will help legitimize the survey and highlight its importance. The introduction outlines potential benefits to the business

community resulting from the evaluation. After the introduction letters, Optimal will send an e-mail to cluster participants with the direct survey link embedded.

- Optimal will also send up to four reminder e-mails over a period of approximately one month to encourage potential survey respondents to complete the survey before the deadline. The reminders will be sent at different times of the day to increase the probability of reaching respondents at an opportune time.
- Optimal will work with the cluster administrators to replace e-mail addresses on their respective rosters that are invalid or generate a bounce. This will ensure that as much of the survey frame as possible is given the opportunity to respond to the survey.
- The surveys are designed to use skip patterns and simple, often multiple-choice answer options to facilitate the ease with which they can be completed. Surveys can be halted and resumed by participants to provide greater flexibility, and participants are free to skip question(s) that they may not know how to answer or do not want to answer.
- To further maximize the response rate, Optimal plans to provide cluster administrators with a PDF smart form of the survey instruments. These smart forms are electronic, fillable versions of the survey that can be used if participants encounter technical difficulty or do not feel comfortable submitting information online. The clusters can also use these smart forms at their events or in the context of one-on-one counseling to capture data from a greater share of the cluster participants.

4.   Describe Any Tests of Procedures of 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 tests may be submitted for approval separately or in combination with the main collection of information.

The small business and large organization surveys were pretested internally by four individuals to ensure reliability, minimize measurement error, and minimize respondent burden. Feedback from these pretests was also used to identify problematic, difficult, or time-consuming questions and to revise or clarify them. The surveys were also revised based on the feedback received from the cluster administrators. In addition, once the surveys have been implemented on the web and as PDF smart forms, they will be tested by internal research staff for usability and to identify any errors related to their transfer to different mediums.

Pretests on internal research staff were completed to provide a burden estimate for the final instruments. Staff were instructed to complete the survey in one sitting, at a time and place with minimal distractions. The average burden across testers is detailed in exhibit 3.

Exhibit 3. Burden estimates of the large organization survey and small business survey

Instrument	Average burden in minutes
Large organization survey	10
Small business survey	20

5.   Expert Contact Information

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.

The consultants used to design, collect, and/or analyze the information for the agency are from Optimal Solutions Group, LLC. For questions regarding the study or questionnaire design or statistical methodology, contact the appropriate staff member(s) via the information below:

Optimal Solutions Group, LLC
M Square Research Park
5825 University Research Court, Suite 2800
College Park, MD 20740-9998
Telephone: 301-918-7301
E-mail: cluster@optimalsolutionsgroup.com

Specialization	Name	Title	Organization	E-mail address
Program Evaluation	Brittany Borg	Program Analyst	Small Business Administration	Brittany.borg@sba.gov
Public policy	Jennifer Auer	Project and IDIQ Director	Optimal Solutions Group	jauer@optimalsolutionsgroup.com
Labor economist	Mark Turner	CEO	Optimal Solutions Group	mturner@optimalsolutionsgroup.com