

Analysis Plan

Evaluation Name: Assessing Procurement Readiness for Small Manufacturing Firms

Evaluation Code: 2603

Date Finalized: August 19, 2026

Project Description

This evaluation is designed to build evidence in response to the first key question in [SBA's FY2026 Annual Evaluation Plan](#): **What are effective strategies to encourage small business manufacturers' participation in the federal marketplace?**

The SBA's Made in America Manufacturing Initiative aims to foster more opportunities for public and private investments in support of manufacturing in America and increase participation in the federal marketplace, among other goals. However, increasing participation in the federal marketplace requires procurement-ready firms, yet the SBA cites procurement readiness as a bottleneck for small businesses participation in the federal marketplace.¹ Lack of authoritative information on what it takes to deliver on a federal contract can prevent competitive firms from submitting a bid or lead uncompetitive firms to submit a bid, causing inefficiencies and diluting the market by doing so. Results from [OES's FY25 descriptive evaluation](#) found that the rate of manufacturing awards to small businesses can be very different across sectors, suggesting that firms in certain sectors are more competitive than in others. In 2022, SBA developed a Procurement Readiness Assessment Tool (PRAT) with the aim of attracting more procurement-ready firms to bid on contracts as a part of its [HUBZone Early Engagement Initiative](#).

This evaluation has two objectives. The first is to use a randomized evaluation to evaluate whether a generalized PRAT affects firms' self-awareness of their own procurement readiness and confidence in the firm's ability to succeed in the federal marketplace. The second is to assess whether the generalized PRAT is helpful for identifying business strengths, areas for improvement, and next steps. This objective is exploratory and will be assessed through survey feedback about the PRAT.

Together, we have four research questions for this study. Research Questions 1 and 2 will be evaluated through confirmatory hypothesis testing. Research Questions 3 and 4 will be assessed through exploratory descriptive analysis.

Research Question 1 (confirmatory): Is self-awareness of procurement readiness different among firms that report it after receiving the PRAT compared to firms that report it before?

Research Question 2 (confirmatory): Is confidence in their ability to succeed in the federal marketplace different among firms that report it after receiving the PRAT compared to firms that report it before?

Research Question 3 (exploratory): Do firms perceive the PRAT to be helpful in identifying a firm's strengths, areas for improvement, and next steps for the firm to become more procurement ready?

Research Question 4 (exploratory): What are the PRAT's implementation strengths and weaknesses (e.g., dropoff rates for completing the tool and length of time to complete the tool)?

Preregistration Details

This Analysis Plan will be posted on the OES website at oes.gsa.gov before outcome data are analyzed.

¹ Stakeholder conversations, September 2025.

Hypotheses

The primary confirmatory hypotheses assess whether the PRAT affects firms' certainty in their federal procurement readiness and confidence in their ability to succeed in the federal marketplace:

H1: Firms that receive the PRAT before answering the self-assessment questions will have different levels of certainty in their procurement readiness than firms that receive the PRAT after answering the self-assessment questions.

H2: Firms that receive the PRAT before answering the self-assessment questions will have different confidence levels than firms that receive the PRAT after answering the self-assessment questions.

We also have one secondary exploratory hypothesis related to the helpfulness of the PRAT:

H3: Firms' reported helpfulness of the PRAT tool is not equal to the neutral score (middle score) 3.

Data and Data Structure

Data Source(s)

We used the following datasets to construct the evaluation sample:

- SAM.gov entity information for manufacturing businesses firms (NAICS 31x-33x; did not need to be primary): Data included active/inactive registration status, last updated registration, purpose of registration status, NAICS codes, and Unique Entity Identifier (UEI)
- USA Spending contract awards from FY2023 to present: Data included UEI and contract awards, including "latest performance date" (i.e., contract end date) for each award
- SBA's Small Business Search manufacturing firms (NAICS 31x-33x; did not need to be primary NAICS): Data included emails, NAICS codes, and UEI

We will use the following datasets to monitor intervention implementation:

- GovDelivery bulletin report - outreach message [July 16, 2026]
- GovDelivery bulletin report - reminder 1 [July 22, 2026]
- GovDelivery bulletin report - reminder 2 [July 29, 2026]
- GovDelivery bulletin report - reminder 3 [August 5, 2026]

We will use the following dataset to construct outcome measures:

- Qualtrics dataset [August 14, 2026]

In addition, we may use the following dataset to construct and explore various measures, though we do not plan to pre-specify analyses of these outcomes:

- GovDelivery engagement dataset [August 14, 2026]

There are two notes regarding the Qualtrics and GovDelivery datasets. First, note that these will be time bound to only include data from Qualtrics and GovDelivery engagement from the full launch. Second, note that we will be unable to exactly match across these two datasets, since no unique ID connects entries in the two. We may attempt fuzzy matching across these two datasets based on email address, but this match process is not part of our pre-specified analytical strategy.

Outcomes to Be Analyzed

We will measure outcomes on August 14, 2026, three weeks after the first email was sent and one week after the third reminder was sent.

For Research Questions 1-2 we are interested in two primary outcomes: 1) self-awareness of federal procurement readiness ("Self-Awareness"), which we measure using a Likert-scale question about certainty in self-assessment of procurement readiness; and 2) self-confidence in ability to succeed in the federal marketplace ("Confidence"), which we measure using a Likert-scale question about confidence. We study these outcomes in a confirmatory analysis.

1. Outcome Measure 1: Self-Awareness
 - a. Definition: Certainty in self-assessment of procurement readiness
 - b. Measure: In the Self-Assessment Questions, the respondent is asked: "How certain are you about your firm's procurement readiness?"
 - i. Not at all certain
 - ii. Somewhat uncertain
 - iii. Neither certain nor uncertain
 - iv. Somewhat certain
 - v. Very certain
 - c. Code: Integer variable ranging from 1 to 5
2. Outcome Measure 2: Confidence
 - a. Definition: Firm self-confidence in their ability to succeed in the federal marketplace
 - b. Measure: In the Self-Assessment Questions, the respondent is asked: "How confident are you that your firm could be successful in the federal marketplace?"
 - i. Not at all confident
 - ii. Somewhat not confident
 - iii. Neither confident nor unconfident
 - iv. Somewhat confident
 - v. Very confident
 - c. Code: Integer variable ranging from 1 to 5

For Research Question 3, we are interested in two indicators of PRAT helpfulness: whether the PRAT helped a firm identify strengths and areas of improvement ("Improvement"); and whether the PRAT helped a firm identify next steps to becoming more procurement ready ("Next Steps").

1. Exploratory Indicators: Helpfulness Index created by averaging the below two outcome variables
 - a. Definition: How helpful the PRAT is perceived to be by firm representatives
 - b. Indicator 3: "How helpful did you find the tool in identifying your firm's strengths and potential areas for improvement?"

- i. Very helpful
 - ii. Helpful
 - iii. Neither helpful nor unhelpful
 - iv. Unhelpful
 - v. Very unhelpful
- c. Indicator 4: “How helpful did you find the tool in identifying next steps for your firm to become more procurement ready?”
 - i. Very helpful
 - ii. Helpful
 - iii. Neither helpful nor unhelpful
 - iv. Unhelpful
 - v. Very unhelpful
- d. Code: Integer variable ranging from 1 to 5, calculated using the average score across the two Likert-scale helpfulness indicators

For Research Question 4, we are interested in the following indicators related to dropoffs while completing the PRAT (“Drop Off”), and length of time to complete the tool (“PRAT Time”):

1. Indicator 5: “Drop Off”
 - a. Measure: Number of questions completed in the PRAT (before dropoff, if this occurs)
 - b. Code: Integer from 0 to total number of PRAT questions
2. Indicator 6: “PRAT Time”
 - a. Measure: Number of seconds spent on the PRAT
 - b. Code: Continuous from 0 to positive infinity

We also anticipate analyzing the following indicators, though we will not pre-specify our expectations and analytical strategy:

- Email open rate
- PRAT link click rate
- Response to PRAT question about intent to pursue federal contracting: “Does your business include federal contracting in your growth strategy?”
- Response to PRAT question about first self-awareness question: “How procurement ready is your business for federal contracting?”
- Verified SAM.gov registration: We will check new registrants/activity on SAM.gov to identify any firms that went from inactive to active in their registration or if there were any new registrants.
- Open-ended text about what the firm thinks are its next steps in pursuing federal contracts
- Open-ended text about how the tool could have been more useful for the firm
- Comparison between PRAT score and self-assessed readiness score
- Patterns and trends in the PRAT results:
 - Areas where low procurement-ready firms tend to be weak
 - Areas where low procurement-ready firms tend to be strong

- Areas where high procurement-ready firms tend to be weak
- Areas where high procurement-ready firms tend to be strong
- All descriptives can be further separated by region, sector, and other baseline covariates

Imported Variables

There are no imported variables in this project. All variables to be analyzed are in the outcomes dataset discussed above: the Qualtrics dataset to be downloaded on August 14, 2026.

Transformations of Variables

The outcome variables are fully described in the section “Outcomes to Be Analyzed.” No further transformation is necessary.

Transformations of Data Structure

Our final dataset will be at the firm level (one observation per firm), with separate variables for each outcome. We anticipate that this will be equivalent to an email-level dataset. In the sample construction phase of the project, for each email address that had multiple firms, we randomly selected one firm to be mentioned in that email address’ PRAT outreach email. However, it is possible for a given individual to fill out the PRAT multiple times for multiple businesses. In this case, the individual’s treatment assignment could have been different across their entries, and any observations beyond the first are effectively treated anyway regardless of the Qualtrics-designated assignment, since these $n > 1$ observations are always post-PRAT exposure. We will only know that this has occurred if the person enters their email address at the end of the Qualtrics survey in order to receive the PRAT report via email. Let us call this variable *PRAT_Report_Email*. In the event that we see multiple observations for one email address as identified in *PRAT_Report_Email*, we will retain all observations for the main analysis (i.e., treat these observations as independent). We will run a secondary analysis with only the first observation for each email address included.

Data Exclusion

We will exclude all observations in the Qualtrics dataset from before the full launch (July 16, 2026). This will exclude test participants and the approximately 1,000 contacts that received the email via a soft launch to test operations and functionality. We will also exclude all observations in the Qualtrics dataset that are blank (i.e., the contact clicked on the PRAT link in the email but did not engage with the PRAT), those that declined to consent to the PRAT on the first screen (i.e., clicked “No”), and those that dropped off before they could be randomized to treatment (i.e., self-assessment after PRAT) or control (self-assessment before PRAT). Our outcomes dataset will therefore include all observations that consented to take the PRAT and were randomized to treatment or control.

Treatment of Missing Data

Randomization occurs within the Qualtrics interface, and our analysis dataset will only include those who were randomized to treatment or control. The primary intent-to-treat analysis will include all observations in the analysis dataset. For observations that were randomized but subsequently declined to answer one (or more) of the questions that provide our outcomes data, we will drop this observation from the analysis of that outcome variable. For each outcome variable, we will report Manski bounds and Lee bounds, to assess how different the missing outcome values would have to be to yield meaningful bias (i.e., a false positive or false negative).

It is not unreasonable to expect that missing primary outcomes data will be more common in the treatment group. In the treatment group, the self-assessment questions — which form the basis of the primary outcomes — came after the PRAT, whereas in the control group they came before the PRAT. In the event we observe this differential missingness, we will probe its nature in two ways:

- a) We will subset observations with missing self-assessment questions to those with non-missing PRAT feedback questions, which occurred *after* the self-assessment questions in both treatment and control. After subsetting we will examine whether missingness on the self-assessment questions is correlated with treatment assignment. A significant result on this analysis would indicate that the experience of taking the PRAT before vs. after answering the self-assessment questions changed the likelihood of answering those questions.
- b) We will subset observations with missing answers beginning with the self-assessment questions (i.e., all answers including and after the self-assessment questions are blank), which indicates the contact dropped off at some point. We will then regress the time of drop off (measured in seconds in one specification and question number in another) on treatment assignment, to see if the timing of dropoff is associated with treatment assignment. A significant result on this analysis would indicate that the experience of taking the PRAT before vs. after answering the self-assessment questions changed the timing of dropoff.

Descriptive Statistics, Tables, and Graphs

We will report:

1. A balance table (means, p-values) comparing firms assigned to treatment vs. control on pre-treatment characteristics, including self-reported NAICS, number of employees, and zip code..
2. Descriptive statistics for intervention implementation, including email addresses suppressed in GovDelivery, delivery failure rate, email open rate, and PRAT click rate for the initial outreach message and all reminders.
3. The confirmatory coefficient table (estimates, SEs, p-values).
4. Exploratory coefficient tables (estimates, SEs, p-values, where applicable)..
5. Plots of means and confidence intervals for the outcomes analyzed for the confirmatory hypotheses

Statistical Models and Hypothesis Tests

This section describes the statistical models and hypothesis tests that will make up the analysis — including any follow-ups on effects in the main statistical model and any exploratory analyses that can be anticipated prior to analysis.

Statistical Models

The confirmatory statistical model will estimate intent-to-treat effects of assignment to treatment using ordinary least squares. For both primary outcomes, it is a linear regression model of the specification:

$$Y_i = \alpha + \beta T_i + \gamma_s + \epsilon_i$$

Y_i is the outcome for firm i , T_i is firm i 's treatment assignment, γ_s are NAICS category fixed effects (manufacturing NAICS, non-manufacturing NAICS, and unknown NAICS), and ϵ_i is the error term. The primary estimate is the coefficient on treatment assignment β . The model includes NAICS category fixed effects because randomization was stratified based on this variable, as measured via self-report in the

second question in Qualtrics. Besides NAICS category fixed effects, the model does not include any pre-treatment covariates.

Confirmatory Analyses

Outcome	Test
Research Question 1 (confirmatory): Is <u>self-awareness of procurement readiness</u> different among firms that report it after receiving the PRAT compared to firms that report it before? H1: Firms that receive the PRAT before answering the self-assessment questions will have different levels of certainty in their procurement readiness than firms that receive the PRAT after answering the self-assessment questions.	
Let $Y_i^{Awareness}$ be an integer measuring how certain a firm is in their answer to the self-awareness question about procurement readiness.	$\beta \neq 0$
Research Question 2 (confirmatory): Is <u>confidence in their ability to succeed in the federal marketplace</u> different among firms that report it after receiving the PRAT compared to firms that report it before? H2: Firms that receive the PRAT before answering the self-assessment questions will have different confidence levels than firms that receive the PRAT after answering the self-assessment questions.	
Let $Y_i^{Confidence}$ be a Likert scale measuring how successful a firm feels they can be in the federal marketplace.	$\beta \neq 0$

Exploratory Analysis

Research Question 3 (exploratory): Do firms perceive the PRAT to be <u>helpful</u> in identifying a firm's strengths, areas for improvement, and next steps for the firm to become more procurement ready? H3: Firms' reported helpfulness of the PRAT tool is not equal to the neutral score (middle score) 3.	
Let $Y_i^{Helpfulness}$ be an index of two 5-level Likert scale questions measuring how helpful a firm found the PRAT tool (1=very unhelpful, 2= unhelpful, 3=Neither unhelpful nor helpful, 4= helpful, to 5=very helpful). This score will be calculated by averaging the two Helpfulness indicators: one about identifying areas of strength and improvement; and one about identifying next steps. Higher values indicate higher perceived helpfulness. We will also report the histogram of Helpfulness index scores.	$E[Y_i^{Helpfulness}] \neq 3$

Research Question 4 (exploratory): What are the PRAT's <u>implementation strengths and weaknesses</u> (e.g., dropoff rates for completing the tool and length of time to complete the tool)?	
This analysis is not based on a hypothesis and estimand but rather on descriptive statistics about dropoff rates, time to completion, user feedback, reported technical glitches, and other usage information as it arises.	N/A

Inference Criteria, Including any Adjustments for Multiple Comparisons

We will conduct two-sided tests with a significance level of $\alpha = 0.05$. We will not adjust for multiple comparisons since we only have one primary test for each hypothesis.

Limitations

- **Limited evaluation period:** While participation in the federal marketplace involves long term planning and timelines, we are limited to a two-month evaluation period. However, we will preserve the treatment assignment lists so that in the future, SBA and/or OES could analyze the evaluation outcomes or additional outcomes (such as awards) with a longer time horizon.
- **Attrition:** There are two types of attrition that affect the sample size.
 1. The first type is attrition at recruitment, whereby businesses in the sample unsubscribed to the GovDelivery topic and therefore did not receive or stopped receiving the emails ("bulletins"). This reduced our sample from 85,539 contacts at full launch to 81,013² by the second reminder.
 2. The second type is attrition associated with engagement with the PRAT. In the design document power calculations, we assumed a 10% takeup of the PRAT. This assumed rate actually bundled attrition at four stages:
 - 1) The rate of clicking on the PRAT link
 - 2) *Among the subset of X contacts who clicked in stage 1:* The rate of engaging with the PRAT
 - 3) *Among the subset of X contacts who engaged in stage 2:* The rate of consenting to the PRAT
 - 4) *Among the subset of X contacts who consented at stage 3:* The rate of providing outcomes data in the PRAT

We will report PRAT attrition across these four stages, including whether attrition in Stage 4 differed across treatment and control.

Link to An Analysis Code/Script

The analysis code script is in development and can be provided upon request.

² While 80,644 contacts received the second reminder, this was after removing 369 contacts who had already completed the PRAT. Therefore, the remaining (i.e., non-attrited) sample at this stage was 81,013 (80,644+369).