Analysis Plan
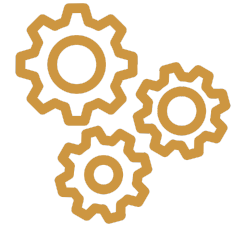
Project name: Increasing set-aside utilization through proactive market research

Project code: 2405

Date finalized: October 9, 2024

Project description

The core focus of this evaluation will be to improve federal contracting outcomes for "socioeconomic small businesses" (SSBs[1]) by encouraging the use of *set-asides,* which limit competition to firms in socioeconomic categories of small businesses. According to the Federal Acquisition Regulation (FAR), all purchases between the micro-purchase ($10,000) and simplified acquisition ($250,000) thresholds must be set aside for small businesses, though not necessarily for SSBs, unless the CO determines there are not at least two small businesses that are competitive in terms of fair market prices, quality, and delivery. Above the simplified acquisition threshold ($250,000), the CO must first consider set-asides for SSBs, unless the CO determines there are not at least two SSBs that are competitive ([FAR 19.203](#)). The process of assessing whether there are at least two small businesses or SSBs that have acceptable capabilities is called "market research."

Despite these strong requirements, SSB set-aside utilization appears to fall below its potential. For example, 2.78% of federal FY23 obligations went to HUBZone businesses, though the target for FY23 was 3% ([SBA Small Business Procurement Scorecard](#)), and not all of this allocated percentage was via a HUBZone set-aside contract. There are various reasons set-asides are not more widely utilized, but a primary barrier may be related to the market research process. The regulations governing market research and set-aside requirements are complex. Conducting market research is cumbersome and inefficient because of the complex and ever-changing nature of data federal contractors must respond to in order to be an acceptable source. **In this project, our objective is to assess whether proactive market research — specifically the proactive identification of SSBs that could fulfill upcoming opportunities — increases the utilization of set-aside contracts.**

We conduct an evaluation among federal contracting officers (COs) posting award opportunities with a "sources sought" or "pre-solicitation" status that have not yet been designated as using an

---

[1] This category includes small businesses certified through one of the [contracting assistance programs](#) run by the Small Business Administration (SBA) and therefore eligible for the set-aside contracting mechanism: women-owned small businesses (WOSB); economically disadvantaged women-owned small businesses (EDWOSB); veteran-owned small businesses (VOSB); service-disabled veteran-owned small businesses (SDVOSB); small businesses certified under the 8(a) business development program; or small businesses located in historically underutilized business zones (HUBZone).

SSB set-aside.[2] Those COs are randomly assigned at the level of the office in which they work to receive either one of two different emails or no email at all.[3] The "Encouragement" email provides information about the requirements to use set-aside contracts, the benefits of doing so, and an encouragement to use set-asides. The "Market Research" email provides all of the information in the Encouragement email plus a list of SSB firms found in a proactive market research process. Each list is specific to a combination of the industry (NAICS code) and good / service (PSC) within which the opportunity falls. Lists are populated as follows:

1. Each contract opportunity is associated with a particular industry (NAICS) and goods/service (PSC).

2. Using the Government-Wide Procurement Equity Tool (GWPET), we compile a list of SSBs who completed federal contracts in this NAICS/PSC area as recently as fiscal year 2021.

3. We cross-reference the list generated in Step 2 against firm registration information in SAM.gov and SSB certification information from SBA's Dynamic Search Tool, removing businesses who do not have active registration or at least one current SBA certification in one of the official SSB categories.

4. We retain lists for all NAICS-PSC combinations that include at least 4 remaining SSBs. Opportunities with other NAICS-PSC combinations (less than 4 remaining SSBs) are excluded from sampling.

5. We share a link with COs that takes them to a spreadsheet hosted online of remaining SSBs for their award opportunity's NAICS-PSC. This list includes basic information about each firm (e.g., name, contact information, SSB certifications). Links to their active vendor profile page on SBA's Dynamic Small Business Search (DSBS) site are also embedded in the list.

Randomization occurs in several phases. **Before randomization**, all offices with SAM.gov postings between FY21 and FY23 are sorted into blocks of three based on their historic procurement volume (i.e., number of postings) and randomized with equal probability to either the business-as-usual (no email), Encouragement, or Market Research groups.[4] Then, for **the first phase of randomization**, we pull contract opportunities data from SAM.gov every two weeks (starting in July 2024 and continuing for eight months). Within each pull, we subset to those that

---

[2] Specifically, we sample opportunities that either have no set aside status, or are only small business set asides. Among opportunities posted between July 22nd and August 8th, 2024, for instance, approx. 91% of otherwise eligible opportunities were in one of these categories. The other 9% included set asides for the SSB categories certified by SBA (the focus of this study), and various other kinds of set asides.

[3] Approximately 3,197 different offices (identified by Activity Address Codes, AAC) posted procurement opportunities in SAM.gov between the 2021 and 2023 fiscal years. We currently anticipate sampling roughly 1,250 - 1,350 offices over the course of this study.

[4] In the first round of sending intervention emails, the inclusion of Market Research email language in the Encouragement emails resulted in us shifting 114 offices from the Encouragement group to the Market Research group. This creates heterogeneous assignment probabilities by block, whereby a given block's probability of being in the Market Research group will be a function of the number of randomization-eligible awards it entered prior to the first mail merge.

(1) are either a "sources sought" or "pre-solicitation" status, (2) are not SSB set asides (see above), (3) are in one of the retained NAICS-PSC combinations, (4) are not missing any required fields (NAICS, PSC, Office AAC code, CO contact information) and (5) have not yet been canceled or advanced to a later procurement status (e.g., "solicitation") by the time of our pull. Based on the first four data pulls (8 weeks of opportunities data), these screening criteria result in a sample of 6.75% of the opportunities posted in SAM.gov.

In the second phase of randomization, opportunities in each pull that originate from an office that has been randomly assigned to the Encouragement or Market Research email groups are subsequently randomized to receive the associated email (with probability p=0.75) or to not receive an email (with probability p=0.25).[5] Specifically, opportunities are sorted within blocks into groups of four based on the number of SSBs in the market research list for their NAICS-PSC, and one opportunity in each group is randomly chosen not to receive an email. When opportunities appear that are attached to an office that did not have any procurement volume in FYs 21-23, these offices are sorted into blocks of 3 based on office-level procurement volume within that data pull and randomly assigned to one of the three groups with equal probability (and then the second phase proceeds as usual).[6]

Preregistration details

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

Hypotheses

There are two primary hypotheses:

H1: The intervention emails will increase the likelihood that an SSB set-aside contract is utilized for opportunities mentioned in the emails, compared to opportunities not mentioned in an intervention email.

H2: The effect of the Market Research email on set-aside contract utilization will be larger than the effect of the Encouragement email.

There are three secondary hypotheses:

H3 (Officer Spillover): The intervention emails will increase the likelihood that an SSB set-aside contract is utilized for opportunities not mentioned in the emails but associated with officers who received the emails for other opportunities.

H4 (Office Spillover): The intervention emails will increase the likelihood that an SSB set-aside contract is utilized for opportunities not mentioned in the emails but associated with an office

---

[5] As opportunities are sorted (at an office level) into a group that potentially receives emails with a ⅔ probability, setting the email receipt probability equal to ¾ helps ensure that each opportunity has an overall (approx.) 50% chance of receiving an email.
[6] In all cases where offices or opportunities are randomly assigned within groups, to help ensure correct assignment probabilities, we add "dummy rows" when needed to ensure that groups have the intended size. We then drop those dummy rows after the assignment is complete.

that has received at least one email intervention (excluding opportunities from officers who have received at least one email).

H5 (Intermediate): The intervention emails will increase the likelihood that a CO anticipates communicating with a SSB firm.

Data and data structure

This section describes variables that will be analyzed, as well as changes that will be made to the raw data with respect to data structure and variables.

Data source(s):

This evaluation will draw on five data sources: 1) Intervention data, including the email recipient, outcome of the random assignment, and email analytics; 2) Reply data, including information about whether the CO replied to the intervention email and the nature of the reply; 3) SAM.gov opportunities data; 4) Federal Procurement Data System (FPDS) awards data; 5) CO survey data.

These data sources will be merged together to form a dataset for analysis at the contract opportunity level, covering opportunities entered into SAM.gov between June 2024 and March 2025, with corresponding awards data collected until September 2025. There will be one observation per award opportunity: we will not build a panel dataset.

Table 1 provides information about the data sources, variables, and time periods in the resulting dataset. Rows in bold font and highlighted in teal are the basis of primary and secondary outcomes.

Table 1. Variables in the merged analysis dataset

Name	Category	Data source	Description
Office ID	Intervention	Intervention Dataset	AAC code associated with the office
Recipient ID	Intervention	Intervention Dataset	Anonymized unique ID field associated with the CO
Opportunity ID	Intervention	Intervention Dataset	Anonymized unique ID field associated with the opportunity
Office Intervention	Intervention	Intervention Dataset	Categorical indicator for office-level assignment: 0 no email (business-as-usual); 1 Encouragement; 2 Market Research
Office Encouragement	Intervention	Intervention Dataset	Binary indicator for whether or not the office was assigned to the Encouragement group
Office Market Research	Intervention	Intervention Dataset	Binary indicator for whether or not the office was assigned to the Market Research group

Opportunity Intervention	Intervention	Intervention Dataset	Binary indicator for opportunity-level assignment: 0 No email; 1 Email
Opportunity Encouragement	Intervention	Intervention Dataset	Binary indicator for whether or not the opportunity was assigned to the Encouragement group
Opportunity Market Research	Intervention	Intervention Dataset	Binary indicator for whether or not the opportunity was assigned to the Market Research group
Office Agency ID	Intervention	Intervention Dataset	Unique ID field associated with the office's agency (likely L1 code)
Procurement Volume	Intervention	Intervention Dataset	Count variable for the number of contract opportunities associated with the office in FY23
Block ID	Intervention	Intervention Dataset	Unique ID field associated with the randomization block (at office-level)
Opportunity Set-Aside Designation	Intervention	Intervention Dataset	Categorical indicator for the pre-intervention set-aside designation (e.g., no set aside, to be determined, small business but not SSB set aside, etc.)
Market Research SSB Count	Intervention	Intervention Dataset	Count variable for the number of SSBs identified in the proactive market research process (by design, the minimum is 4)
Market Research SSB Diversity	Intervention	Intervention Dataset	Diversity index capturing how many and to what extent different SSB categories are represented
Email Date	Intervention	Intervention Dataset	Date email was sent
Reply Date	Reply	Replies Dataset	Date of response to the intervention email
Email Bounce Back	Reply	Replies Dataset	Binary indicator for whether or not the email bounced back
Email Auto-Reply OOO	Reply	Replies Dataset	Binary indicator for whether or not the sender received an out-of-office auto-reply in response to the email
Email Auto-Reply Non-OOO	Reply	Replies Dataset	Binary indicator for whether or not the sender received an auto-reply that was not about an OOO notice
Email Reply Providing Information	Reply	Replies Dataset	Binary indicator for whether the CO replied to the email providing information

Email Reply Requesting Information	Reply	Replies Dataset	Binary indicator for whether the CO replied to the email requesting information
Link Visit	Analytics	Analytics Dataset	Binary indicator for whether or not the OES-hosted link was visited NOTE: *The link variables will be collected for the market research links*
Link Visit Timing	Analytics	Analytics Dataset	Date link was first visited
Link Visit Count	Analytics	Analytics Dataset	Count variable for the number of times the market research link was visited
Link Duration	Analytics	Analytics Dataset	Integer indicating the cumulative number of minutes spent on the link, across all visits
Agency	Opportunity	SAM.gov	L1 code
Division	Opportunity	SAM.gov	L2 code
Office	Opportunity	SAM.gov	L3 code
Opportunity Posting Date	Opportunity	SAM.gov	Date opportunity was posted in SAM.gov or was moved to its current phase (i.e., "sources sought" or "pre-solicitation")
Stage of Opportunity	Opportunity	SAM.gov	Categorical indicator for whether in sources sought or pre-solicitation phase
NAICS	Opportunity	SAM.gov	Industry code
PSC	Opportunity	SAM.gov	Product and service code
Solicitation Set-Aside Status	Opportunity	SAM.gov	Binary indicator for whether the opportunity is designated as an SSB set-aside once entered as a solicitation in SAM.gov
Award Set-Aside Status	Award	FPDS	Binary indicator for whether the award is designated as an SSB set-aside
SetAsideUtilization	**Outcomes**	**N/A**	**Binary indicator for whether solicitation set-aside status or award set-aside status (or both) is equal to 1**
Contractor SSB Certifications	Award	FPDS	Set of binary indicators noting the SSB certifications of the awarded contractor (8(a), WOSB, EDWOSB, HUBZone, VOSB, SDVOSB)
Consistent CO	Award	FPDS	Binary indicator for whether the CO included in the randomization is listed in FPDS
Number of COs	Award	FPDS	Integer designating the number of COs associated

			with the opportunity across opportunity and award phases
CO Demographics	CO Covariates	Survey	Demographics (gender, race / ethnicity category, high-level place of work)
CO Work History	CO Covariates	Survey	Work history (years of procurement experience, years of federal government experience, industry experience)
CO Market Research Experience	CO Covariates	Survey	Market research approach (authority / independence over market research process, which non-CO officials are involved)
CO Set-Aside Experience	Outcomes	Survey	Anticipated set-aside utilization plans
CO SSB Experience	**Outcomes**	**Survey**	**Current contact with SSBs, anticipated communication with SSBs**
CO Intervention Experience	Outcomes	Survey	CO direct/indirect exposure to and perceptions of intervention emails

Outcomes to be analyzed:

There are two outcomes to be analyzed. `SetAsideUtilization` will provide the outcome variable in our tests of H1, H2, H3, and H4. `SSBComm` will provide the outcome variable in our test of H5.

- `SetAsideUtilization`: A binary indicator for whether the opportunity was ever classified as an SSB set-aside, either in SAM postings for this opportunity later in its procurement lifecycle or in this opportunity's entries in FPDS (if it was ultimately awarded)

- `SSBComm`: The CO-level[7] response to the following Likert-scale question, to be included on the Qualtrics survey of COs (to be reviewed and approved by GSA's Office of Customer Engagement):

    Question: How likely is it that you will communicate with a small business owner certified as HUBZone, 8(a), women-owned small business, or service-disabled veteran-owned small business regarding a potential contract in the next six months?

    Answer categories:

    1. Highly likely

---

[7] We anticipate being able to match survey responses to the sample of COs using name and email. If we cannot match exactly, we can potentially match based on string similarity using some measure of "edit distance" (e.g., Levenshtein distance). However, we will not be able to assess match potential until we have the survey data, and we also may not have a large sample of matched COs. In this event, we will aggregate to the office level and calculate the mean office-level response to this question.

2. Somewhat likely

3. Uncertain

4. Somewhat unlikely

5. Highly unlikely

Imported variables:

As discussed above, the variables in this evaluation are imported from five different data sources. Depending on the level of analysis, data will be merged and transformed based on Opportunity ID, CO ID, and Office ID.

Data exclusion:

We will include opportunities that appeared in SAM.gov in the sources sought or pre-solicitation phases between July 8, 2024, and March 30, 2025. We will exclude opportunities that are in NAICS-PSC combinations for which we could not generate a viable market research list of SSBs, could not identify a CO contact or their email address, and opportunities that have already been set-aside for an SSB.

Treatment of missing data:

For three reasons, we anticipate missing data for our primary outcome. First, not all opportunities will have later postings in SAM.gov, and not all opportunities turn into awards. Second, we anticipate being unable to match all opportunities to awards (see Limitations section below for more discussion of this issue). Observations missing outcome data will be treated as attrited from our sample. Our power calculations have allowed for significant attrition of this kind.[8] Third, we anticipate that not all COs included in our sample will complete the survey, which affects the outcome variable we intend to use to test H5, as well as some of our exploratory analysis.

Descriptive statistics, tables, and graphs

- A table reporting descriptive statistics and balance tests for the two email and business-as-usual groups, including date of email. We will test the significance of the difference between each of these alongside an omnibus F-test of all of the differences.

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:

Tests of Primary Hypotheses:

To test **H1, H3, and H4**, we will estimate the following model using OLS:

---

[8] Preliminary tests of our outcome measurement protocol using pre-intervention data suggested attrition rates of roughly 20-33%. We allowed for attrition of this magnitude in preliminary power simulations.

$$SetAsideUtilization_{ijk} = \beta_0 + \beta_1 OppEmailed_i + \beta_2 COEmailed_k + \beta_3 OfficeEmailed_j + X_{ijk}\delta + \varepsilon_{ijk}$$

where $SetAsideUtilization_{ijk}$ represents the primary outcome of interest, an indicator variable for whether opportunity i associated with office j and managed by CO k was designated as an SSB set-aside. $OppEmailed_i$ is an indicator variable equal to 0 if the opportunity was randomly assigned to the business-as-usual condition and 1 if the opportunity was randomly assigned to be discussed in either an Encouragement or Market Research email. Therefore, $\tau_{H1} = \beta_1 + \beta_2 + \beta_3$ represents the coefficient of interest to test **H1**, or the difference in the likelihood of an SSB set-aside between the intervention and comparison groups. Our null hypothesis is that $\beta_1 + \beta_2 + \beta_3 = 0$.

$COEmailed_k$ is an indicator variable equal to 0 if CO k associated with opportunity i never received an email about any opportunity during the intervention period, and 1 if CO k received at least one Encouragement or Market Research email about an opportunity during the intervention period. Therefore, $\tau_{H3} = \beta_2 + \beta_3$ represents the coefficient of interest to test **H3**, or officer-level spillover effects from one opportunity to another for CO k. Our null hypothesis is that $\beta_2 + \beta_3 = 0$.

$OfficeEmailed_j$ is an indicator variable equal to 0 if office j never received an email about any opportunity during the intervention period, and 1 if office j received at least one Encouragement or Market Research email about an opportunity during the intervention period. Therefore, $\tau_{H4} = \beta_3$ represents the coefficient of interest to test **H4**, or office-level spillover effects from one opportunity to another for office k. Our null hypothesis is that $\beta_3 = 0$.

X_{ijk} represents a vector of pre-intervention control variables, included to increase the precision of the estimates. These control variables will be:

- Office-level procurement volume (*one of the blocking variables*);
  - For offices in our historical FY21-FY23 data, this will be procurement volume (number of opportunities) across those years. For offices only observed in our sample timeframe, we will count the number of opportunities observed for this office within our final sample.
  - To accompany this composite measure of procurement volume, all models that include it will also control for a binary indicator for offices that are not in our historical procurement data. In effect, this is similar to the "missing indicator" method for handling missing data in control variables.

- Binary indicator for whether an office had at least one contract awarded in FY21-23 data

- Time (number of days) between the opportunity posting date and the intervention send date;

- Number of SSBs on the market research list for the NAICS-PSC associated with the opportunity (*the other blocking variable*);

- Binary indicator for whether the list included at least one SSB in all certification categories;

- Binary indicator for whether the opportunity was "sources sought" status (vs. "pre-solicitation") when it entered our dataset;

- Binary indicator for whether an opportunity was posted as a small business (but not SSB) set aside (the residual category is opportunities without any initial set aside status);

- Binary indicator for whether an opportunity had a "group code" (the first one/two digits of a PSC, representing broader product/services categories) instead of a full PSC; and

- Email round fixed effects

For reference, we will compute CR2 standard errors clustered by office, as this is the unit of random assignment. Due to the relatively high number of clusters, we do not anticipate that clustering will substantially change the standard errors themselves. However, we plan to perform statistical inference for our analyses using a design-based "*randomization inference*" strategy (see our section on Inference below).

To test **H2**, we will estimate the same model as for **H1**, but we will replace $OppEmailed_i$ with

$Encouragement_{ij}$ (coefficient β_E) and $MarketResearch_{ij}$ (coefficient β_{MR}). We will then test the difference between their coefficients: the difference between the effect of being sent an Encouragement vs a Market Research email on the likelihood of an SSB set-aside. Specifically, using the **H2** model, we will estimate the following difference in coefficients: $\tau_{H2} = \beta_{MR} - \beta_E$.

Using one-sided randomization inference, we will then test the null hypothesis that the effect of Market Research emails is not greater by comparing the observed τ_{H2} to it's simulated randomization distribution (more detail in our Inference section below), estimating the proportion of times values in this distribution are greater than or equal to τ_{H2} (i.e., more positive). Formally, with R values in the randomization distribution, indexed by r, this p-value is:

$$p_{H2} = \frac{1}{R}\sum I(\tau_{H2}^r \geq \tau_{H2})$$

Tests of Secondary Hypotheses:

To test **H5**, we will restrict our sample to the set of COs that were associated with a sampled opportunity for whom we have survey data, and then we will estimate the following model using OLS:

$$SSBComm_k \ = \ \beta_0 \ + \ \beta_1 COEmailed_k \ + \ X_j\delta \ + \ \varepsilon_k$$

where $SSBComm_k$ represents the outcome of interest. This is the response from CO k to a Likert-scale survey question about the anticipated likelihood of communication with SSBs. CO $Emailed_k$ is as defined above. Therefore, $\tau_{H5} \ = \ \beta_1$ represents the coefficient of interest, or the difference in the anticipated likelihood of communication with SSBs between the intervention and comparison groups, and our null hypothesis is that $\beta_1 \ = \ 0$. X_k represents a vector of the office-level pre-intervention control variables that are included in the test of **H1**. Observations will be clustered by agency.

Confirmatory Analyses:

We will estimate effects on one primary outcome: `SetAsideUtilization`. We will apply inverse probability weight to each observation depending on their probability of receiving the intervention condition that they did receive (email, no email, etc.). The opportunity-level probability of receiving any email is 0.5, implying that the weights will be equivalent for all observations. However, due to an error in the first round of emails (see Footnote 4), the probability of receiving the Market Research email is higher than the probability of receiving the Encouragement email, conditional on receiving any email. Moreover, this difference in probabilities is not constant by block, so it must be accounted for. With this in mind, the block-level probability of receiving a Market Research email is 0.25*(number of offices actually assigned to MR after the first-round correction/number of offices originally assigned to MR). We can define the other block-level probabilities (Encouragement or business-as-usual) similarly. For example, if hypothetically we have 900 offices, each condition will have 300 offices assigned to them. If, in practice, 400 offices are assigned to MR and only 200 offices are assigned to Encouragement, the probability of receiving a MR email is 0.25*(400/300)=0.333 and the probability of receiving a nudge email is 0.25*(200/300)=0.167. As a result, the weight applied to each MR observation will be 1/0.333=3 and the weight applied to each nudge observation is 1/0.167=6.

In testing H5, we will estimate effects on one secondary outcome: `SSBComm`.

Exploratory analysis:

Exploratory analysis will take five primary forms. The first set explores differential effects by subgroups. The second set of exploratory analyses explores analytics associated with the Market Research links to understand the experience of the Market Research email more comprehensively. The third set of exploratory analysis explores intermediate outcomes to gather some information about possible causal mechanisms. The fourth provides CO-level and office-level equivalents of a few of our key analyses as a robustness check. Finally, the fifth provides a diagnostic check related to an assumption underlying our interpretation of τ_{H1}.

Differential effects

We will analyze heterogeneity in the pooled email intervention's effects based on the variables described below. We will do this by interacting the variables below with an indicator for exposure to either the Encouragement or Market Research emails and examining the effect on the primary outcome.

- Opportunity Phase: We will examine heterogeneous effects based on whether an opportunity was in "sources sought" or "pre-solicitation" when it entered our study. We expect "sources sought" opportunities to have larger effects because we will have intervened at an earlier phase in the procurement process for such opportunities.

- Agency: We will examine heterogeneous effects based on whether the opportunity was associated with an office within an agency that received an A+ score on the [FY23 SBA Agency Scorecard](). We expect opportunities posted by high-performing agencies to exhibit larger effects.

- Pre-Intervention Set-Aside Designation: We will examine heterogeneous effects based on whether, upon entry into our study, the opportunity did not show a set-aside status at all versus showing a (not SSB) small business set-aside. We expect opportunities designated as (not SSB) small business set-asides to have larger effects.

- SSB Set-Aside Status Measurement Source: As described above, our primary outcome variable - `SetAsideUtilization` - will draw from multiple data sources to capture whether an opportunity (or its subsequent award) was ever designated for an SSB set-aside contract (or awarded under this mechanism). We will explore whether the results for our primary hypotheses (H1 and H2) vary across data sources: SAM.gov (opportunities) vs. FPDS (awards). By necessity, these exploratory analyses will exclude opportunities with outcomes that could only be measured using the other source.

- Analysis Plan Posting Date: As we had access to outcomes data (including Google Analytics, SAM.gov, and FPDS data) prior to posting the Analysis Plan publicly, we will explore whether effects vary including and excluding pulls from before the Analysis Plan is posted.

Google analytics

For those in the Market Research group, we are collecting analytics on interaction with the CO-specific market research links via Google Analytics. We will compile descriptive statistics on these analytics, including the proportion of links that received at least one visit, average number of visits, average time (in minutes) spent on the link, and time between email send date and first visit. We will also examine how these statistics covary with the primary set-aside utilization outcome.

Intermediate outcomes

Using the main regression model and analytic approach specified above for our test of H1 where possible, and modified analytical strategies in other cases, we will explore the emails' effects on intermediate outcomes, as follows:

- Change in set-aside status across the opportunity-specific entries in SAM.gov, to understand at what point the set-aside status changed[9]

- Appearance in FPDS as an award (versus never getting to that phase)

- SSB set-aside designation upon entry in SAM.gov as an opportunity with either "sources sought" or "pre-solicitation" status.

  - Note that such opportunities will be removed from our sample for randomization and intervention, so this will be an out-of-sample exploratory analysis

  - Here, we will also provide descriptive statistics for change over time in the proportion of otherwise eligible opportunities entering SAM.gov as SSB set asides.

- Email saturation: Number of opportunities and COs emailed in office j over the intervention period[10]

- Questions surrounding perceptions of SSBs

  - Note that these outcomes will be analyzed only among those COs who respond to the survey

- Responses to the emails (bounce backs, OOOs, information provided/requested)

  - Note that the analysis of these outcomes would compare only the Market Research vs. Encouragement groups, omitting the business-as-usual group

- If data can be accessed, interaction with the profile pages of SSBs on SBA's DSBS

Office-level and CO-level effects

We are interested in exploring patterns of effects at the office and CO levels. First, to understand how the intervention emails may affect office-level patterns of set-aside utilization, and to partially mitigate our limited ability to measure SSB set aside use at the opportunity level (see [Treatment of Missing Data](#) above and [Limitations](#) below), we will perform alternative office-level analyses of **H1** and **H2**. We will construct an office-level dataset with *three outcome measures*: (1) the proportion of SSB set aside awards in an office across our study time frame (number of SSB set asides/number of awards), if possible excluding awards from GSA procurement schedules; (2) the total dollar value of SSB set asides in an office across our study timeframe; and (3) the number of "pre-solicitation/sources sought" opportunities for this office that entered SAM.gov as SSB set asides. We will regress each outcome on a vector of office-level controls along with an indicator for whether an office was sent at least one intervention email (**H1**), or indicators for being sent at least one Encouragement Email and at least one Market Research email (**H2**). We plan to include the following controls: office-level historic procurement volume (i.e., the counts used for email assignment), a dummy for offices not found in our FY21-FY23 data, observed office-level

---

[9] Not all postings that enter our sample will have "later" (i.e., post-pre-solicitation) notices in SAM.gov. This does not necessarily mean the opportunity was canceled without leading to an award.

[10] As the level of email saturation is a function of procurement patterns rather than a feature of the randomization design, we analyze the effect of email saturation only as an exploratory analysis.

procurement volume during our study, and an office-level average of the number of market research list SSBs for each opportunity's NAICS-PSC.

Second, to understand how the intervention emails - specifically, receiving a certain number (i.e., "dose") of emails - affect CO-level set-aside utilization, we will use random permutations of office/opportunity-level email assignment to estimate each CO's probability of receiving their observed number of intervention emails. Specifically, we will calculate the proportion of times CO k receives their observed number of emails, $NumEmails_k$, across permutations of email assignment. We call this $ProbEmails_k$.[11] This yields an unadjusted IPW weight for CO k of $IPW_k = \frac{1}{ProbEmails_k}$. We will then "stabilize" these weights by estimating the proportion of all COs receiving $NumEmails_k$, denoted $AllCOsNumEmails_k$, and finally estimating $IPW_k^s = \frac{AllCOsNumEmails_k}{IPW_k}$.[12] We will then run the main model to test H1-H4 but with $NumEmails_k$ as the primary intervention variable rather than the binary indicator for receiving an intervention email (or not), and with observations weighted by IPW_k^s.

Interpretation of H1

Our current plan for constructing an estimate of τ_1 based on our primary confirmatory model (see [Statistical Models & Hypothesis Tests](#)) rests on an assumption that spillover to opportunities mentioned in intervention emails is negligible–i.e., there are only substantively meaningful spillover effects for not-emailed opportunities. This assumption is plausible, but it may be incorrect. If it is incorrect, then our estimate of τ_{H1} described above instead captures the combined effects of mentioning an opportunity in an intervention email (**H1** above) *AND* the average net spillover to emailed opportunities, which we'll call τ_{H1}^{alt}.[13] While τ_{H1}^{alt} is still policy relevant, it differs from the quantity outlined in **H1.** And even in this situation, β_1 would only equal the quantity described in **H1** if observed spillovers to emailed and not-emailed opportunities were the same, which is unlikely.

To ensure we are interpreting our estimates correctly, we will conduct a diagnostic test for whether we can provide *evidence against* spillover to emailed opportunities (see the next paragraph). If we cannot provide sufficient evidence against this type of spillover, we will update

---

[11] If the weights we propose are not estimable (i.e., many COs have unique numbers of emails), we will employ a modified strategy considering ranges of *NumEmails.*

[12] Stabilizing these weights helps ensure that a few COs with rare email counts do not have an outsized influence. For a similar stabilization strategy in other settings, see: Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5), 550-560. This is also discussed in: Gelman, A., & Meng, X. L. (Eds.). (2004). *Applied Bayesian modeling and causal inference from incomplete-data perspectives.* Chapter 7: The propensity score with continuous treatments. John Wiley & Sons.

[13] Note that in this situation, our estimates of within-CO and within-office spillover to opportunities not mentioned in emails (the quantities targeted by H3 and H4) would still be correct.

our interpretations accordingly. In that event, we will also perform a robustness check fitting a more complex variant of our confirmatory model that estimates separate spillover effects for emailed and not-emailed opportunities.[14] But in the interests of preserving statistical power, we do not plan to estimate separate spillovers to emailed and not-emailed opportunities in our confirmatory tests.

To provide evidence on the presence of spillover to emailed opportunities, we will perform a Two One-Sided equivalence Test (TOST). This test evaluates whether we can reject a null hypothesis of "meaningful" spillover to emailed opportunities, where "meaningful" effects are those outside a pre-specified equivalence region. We consider effects of average spillover exposure ranging from -0.3pp to +0.3pp to be sufficiently small that we will not change our planned interpretation.

To implement the TOST, we will first estimate spillover to emailed opportunities following a similar strategy to our CO-level dosage test above. We will subset to emailed opportunities in offices that posted at least two opportunities in our data, and calculate each remaining opportunity's accumulated "dosage" of potential spillover exposure (number of emails this CO received for other opportunities, or that other COs in this CO's office received): $AccumDosage_{ijk}$. Using random permutations of email assignment, we will estimate the probability of each opportunity's observed spillover dosage, and use these probabilities to construct IPWs. We will then regress $SetAsideUtilization_{ijk}$ on $AccumDosage_{ijk}$ using weighted least squares to estimate the marginal effect of an additional dosage of spillover. Finally, we will use this estimate as a basis for our equivalence test (constructing a 90% randomization inference CI for it using test inversion, and then following the approach to approximating a TOST outlined by Rainey, 2014).[15]

Inference criteria, including any adjustments for multiple comparisons:

For all OLS models, we will calculate CR2 standard errors for reference, with clusters defined at the office level. However, as our primary statistical inference method, we will calculate two-tailed p-values using randomization inference with 10,000 draws from the randomization distribution of each estimator under the sharp null hypothesis of no effect for any unit. Specifically, we will calculate p-values for each effect, τ_x, as follows, where $R = 10000$ and r indexes draws from its randomization distribution:

$$p_X = \frac{1}{R}\sum I(|\tau_X^r| \geq |\tau_X|)$$

We will reject the null hypothesis if $p < 0.05$. The only exception is for **H2**, where we will estimate a one-tailed p-value instead as described above.

We have pre-specified one test for each primary and secondary hypothesis, corresponding to the estimands τ_{H1} to τ_{H5} above. The null hypotheses for these tests are independent in the sense that

---

[14] Specifically, we will replace *COEmailed* with a binary indicator for whether any <u>other opportunity</u> posted by this CO received an email, replace *OfficeEmailed* with a binary indicator for whether any opportunity for <u>another CO</u> in the same office received an email, and then interact each substituted term with *OppEmailed*.
[15] Rainey, C. (2014). Arguing for a negligible effect. American Journal of Political Science, 58(4), 1083-1091.

rejecting one does not necessarily imply rejecting another. For instance, evidence that the program of sending emails increases SSB set aside use (H1) does not require that the Market Research emails be more effective (H2). Put differently, we think of these hypotheses as speaking to separate claims we seek to evaluate using our data.[16] Therefore, we treat each hypothesis as being in its own family and will not implement a correction for multiple hypothesis tests.

Limitations:

There are three primary limitations of the study. _First_, we have had to balance logistical feasibility against behavior modification potential - intervening later in the solicitation process than the set-aside decision for many opportunities. For instance, some of the CE interviews suggest that the decision is sometimes already made even by the "pre-solicitation" or "sources sought" stages. This is a known limitation, and we plan to consider intervening earlier in follow-up evaluations.

Second, we may not be able to match all opportunities to their resulting award, or to SAM.gov postings for later status, due to both data incompleteness and a limited time period between the end of the intervention and our analysis.[17] We expect a higher match rate for opportunities we sample earlier in our intervention time frame, since we will be able to catch awards that took up to, e.g., a year to be finalized. With around a year of awards data, we estimate a higher match rate closer to 80%. But 66.62% is our current best guess for the match rate we should expect towards the end of our intervention time frame. We do not currently expect this rate to differ across groups (i.e., we don't expect the intervention to influence the various factors that might shape when matches are/aren't available).

Third, this is a bundled intervention (see Table 1), so we will not be able to distinguish between the effects of the different elements of the intervention email (except for across the Nudge vs. Market Research versions).

---

[16] Rubin, M. (2024). Inconsistent multiple testing corrections: the fallacy of using family-based error rates to make inferences about individual hypotheses. Methods in Psychology, 100140.

[17] Based on a minority of eligible opportunities from 2022 that also had later "award notices" posted in SAM.gov (likely non-representative), we cautiously estimate that a majority of opportunities take up to four or five months to be awarded, and some take up to a year or more. For opportunities towards the beginning of our timeframe, we anticipate having over a year of awards data to observe, but for opportunities towards the end of our timeframe, we may have less than six months to observe awards data.