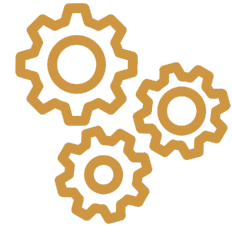


## Analysis Plan

Project Name: Increasing Access to Minnesota Child Care Stabilization Base Grants

Project Code: 2209-B

Date Finalized: May 2, 2023



---

### Project Description

This evaluation is part of the Office of Evaluation Sciences (OES) [American Rescue Plan Act of 2021](#) (ARP) portfolio. The ARP was designed to address immediate needs related to the pandemic, with a specific focus on addressing historically disparate outcomes across race, class, and geography that were further exacerbated by the pandemic. As federal programs are innovating and finding new ways to achieve these goals, the OES [portfolio of evaluations](#) will measure whether ARP-funded interventions are working as intended and share lessons learned.

In support of the [ARP Equity Learning Agenda](#), OES is working with agency partners to better understand how to improve awareness, access, and allocation of ARP programs and resources, focusing on ARP programs with equity goals. This set of evaluations will be intentional and strategic in building evidence to understand the role of ARP programs and supported interventions in improving outcomes for historically underserved populations.

This project aims to identify methods to enhance equitable access to child care grants among child care providers in the state of Minnesota. The pandemic highlighted the instability of the child care market and put additional financial burdens on child care providers. ARP allocated approximately \$24 billion for child care stabilization grants that the Department of Health and Human Services (HHS)'s Administration for Children and Families (ACF), working with states, territories, and tribes, provides as subgrants to child care providers.

Specifically, this project aims to answer the following questions:

- To what extent does a communication bundle of proactive phone calls, text messages, and behaviorally-informed emails impact awards of Minnesota's Child Care Stabilization Base Grant (CCSBG) among eligible providers who have never or infrequently taken-up the program?
- To what extent does a communication bundle of text messages and behaviorally-informed emails impact awards of CCSBG among eligible providers who have never or infrequently taken-up the program?
- Is the effect of the communication bundle of text messages and behaviorally-informed emails on awards different than the effect of a communication bundle of proactive phone calls, text messages, and behaviorally-informed emails on awards?

## Preregistration Details

This Analysis Plan will be posted on the OES website at [oes.gsa.gov](https://oes.gsa.gov) before outcome data are analyzed. In addition, this project will be preregistered in the AEA RCT Registry at <https://www.socialscienceregistry.org/>.

## Hypothesis

### Primary

- A communication bundle including personalized emails, texts, and proactive phone outreach impacts awards of CCSBG.
- A communication bundle including personalized emails and texts impacts awards of CCSBG.
- The effect of a communication bundle including emails, texts, and proactive phone outreach is the same as a communication bundle including only emails and texts.

### Exploratory

*Subgroup effects and difference in effects by subgroups:*

- A communication bundle including personalized emails, texts, and proactive phone outreach impacts awards of CCSBG based on provider type (i.e., for family child care homes or child care centers).
- A communication bundle including personalized emails, texts, and proactive phone outreach impacts awards of CCSBG for providers based on their past application behaviors before the start of our evaluation (i.e., having never applied or infrequently applied).
- A communication bundle including personalized emails, texts, and proactive phone outreach impacts awards of CCSBG for providers based on their location (i.e., urban or rural).
- A communication bundle including personalized email and texts impacts awards of CCSBG based on provider type (i.e., for family child care homes or child care centers).
- A communication bundle including personalized emails and texts impacts awards of CCSBG for providers based on their past application behaviors before the start of our evaluation (i.e., having never applied or infrequently applied).
- A communication bundle including personalized emails and texts impacts awards of CCSBG for providers based on their location (i.e., urban or rural).

*Treatment on the Treated Analysis (LATE/CACE):*

- Speaking to a Child Care Aware representative on the phone impacts awards of CCSBG.
- Award of CCSBG impacts center closure for providers.

## ***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):**

- Child care provider licensing, CCSBG application history, and contact information data;
- Monthly list of likely eligible providers;
- Monthly text message data that captures message timing, successful/unsuccessful delivery, and opt outs among other measures;
- Monthly email data that captures message timing, delivery, bounceback, etc.;
- Operational data from phone call data that captures elements of the conversation, including whether the provider answered, phone call order, talking points discussed, length of call, etc.; and
- Monthly application and award data, including provider information, and award amounts.

### **Outcomes to Be Analyzed:**

The primary outcome is:

- A monthly dichotomous indicator equal to one if the provider was awarded funding for CCSBG for that license ID; and, zero otherwise.

Exploratory outcomes are:

- Whether the provider applied in a given month;
- The monthly amount of funding received, imputed to zero if they did not apply;
- Whether the provider is eligible for funding in a given month;
- Whether a site's license ID is not active in a given month (3 months after the intervention for the LATE analysis); 0 otherwise; and
- A dichotomous indicator for whether the center was active (i.e., a center opened with an active license) in a given month (including 3 months after the end of the intervention).

Additional exploratory operational measures captured from the operational data for phone calls include:

- Whether the CCA staff caller reached the provider, left a voicemail, left a non-voicemail message with another person, or received a wrong/disconnected number notice
- Whether the caller left a voicemail or was unable to because the provider's voicemail was full or not set up
- If a provider was reached, whether they were available to talk or requested a call back at a different time
- Call duration
- Call content (i.e., talking points discussed)
- Whether the called provider indicated they were familiar with the CCSBG program
- Whether the called provider indicated they had applied that month
- Whether the call was flagged for a follow-up with another CCA staff member
- Whether the provider indicated they did not wish to be contacted again

- Open ended call notes/flags from the CCA caller
- When the call occurred, including when during the monthly application cycle, time of day, and day of week.

### **Imported Variables:**

#### Experimental design variables

Our randomization blocks are created from the combination of three mutually exclusive categorical variables:

- A categorical measure for the number of times a center/license ID has applied for CCSBG in the six application windows between August 2022 and January 2023. This measure includes three (3) categories:
  - Applied 0 times;
  - Applied 1-4 times; or
  - Applied 5 times.
- A categorical measure for provider type. This measure includes two categories:
  - Family child care home; or
  - Child care center or certified child care center.
- A categorical measure for variation of the opt out text sent:
  - Standard opt out text variation; or
  - Transparent default opt out text variation.

Thus, we randomize within 12 block combinations (3 application count categories x 2 provider types x 2 versions of opt out text).

#### Additional covariates

- An indicator for having opted-out of text messages for the CCSBG take-up intervention, which occurred prior to delivery of the communications bundle (if available, prior to the start of the take-up intervention);
- An indicator for whether a provider never applied in the six months prior to the beginning of the study;
- An indicator for whether the provider is a family child care home; and
- An indicator for whether the provider is located in a rural zip code.

#### Random assignment

- An indicator for whether the provider was assigned to the email and text message group;
- An indicator for whether the provider was assigned to the phone call group;
- A continuous measure for the phone call batch (groups of 50 phone numbers) in which the provider was randomly assigned. This value is missing for providers not assigned to the phone call group.

#### Take-up of the intervention

- An indicator for successful contact by phone for the take-up intervention.

## Transformations of Variables:

### Outcomes measured each month for each license-id

- **Awarded CCSBG** - This will be coded as 1 if the grant amount paid is greater than \$0; 0 otherwise (including missing).
- **Applied for CCSBG** - This will be coded as 1 if the variable applied is 1; 0 otherwise (including missing).
- **Amount awarded** - This will be coded as the amount paid if a grant was awarded; 0 otherwise. If this is out of range of possible values, and the correct information cannot be obtained from DHS, this will be top/bottom coding using a cutoff mean  $\pm 2$  sd).
- **Likely eligible** - This will be coded as 1 if an email with an application link was sent for that License ID in a given month; 0 otherwise.
- **Site closure** - This will be coded as 1 if the site's license ID is not active in a given month (3 months after the intervention for the LATE analysis); 0 otherwise.

### Random assignment

- **Email and text only**: An indicator that is 1 if the provider was randomly assigned to the email and text message group and not to the phone call group or to a phone call group where the phone call batch had not been shared in that month or a prior month; 0 otherwise.
- **Phone, email, and text**: Randomly assigned to a phone call batch that has been shared in that month or a prior month.

### Weighting

Among the sample of provider clusters the probabilities of assignment to the text group (comprising either the text and email communication bundle or the text, email, and phone call bundle) was 50%. Assignment to the text and email group did not differ based on provider characteristics. Some providers in the text and email group also were randomly assigned to the text, email, and phone call group. This probability differed based on the frequency of their applications to CCSBG prior to the start of the evaluation. The overall probability of assignment to the phone call group (for later reference, defined as  $p_{\text{phone group}}$ ) was:

- 43% of provider clusters that had applied zero times (N=675 License IDs);
- 14% of provider clusters that had applied one to four times (N=135 License IDs); and
- 17% of provider clusters that had applied five times (N=90 License IDs).

Additionally, we randomly assigned the order in which phone calls were made. Among provider clusters assigned to the phone call group, each provider was randomly assigned to a phone call batch of 50 phone numbers (i.e., Batch 1 - 18). We refer to the last batch shared in month  $t$  below as  $batch_t$ .

All analysis will use inverse probability weights to account for the differential probability of assignment to treatment which will vary by month and number of applications to CCSBG prior to

the start of the evaluation. The probability of assignment to each of our three treatment statuses is as follows:

- $P(\text{Control}) = \frac{1}{2}$ , which is constant for all months  $t$ .
- For a cluster assigned to treatment in month  $t$ , the cluster can either be assigned to the text and email only treatment ( $\text{Text}_t$ ) or the text, email, and phone treatment ( $\text{Phone}_t$ ). For a given center in month  $t$ , these probabilities are given by the following:
  - $P(\text{Phone}_t) = p_{\text{phone, group}} \cdot \frac{50 \cdot \text{batch}_t}{900}$
  - $P(\text{Text}_t) = 1 - P(\text{Phone}_t) - P(\text{Control})$

For example, a cluster who had applied zero times before the evaluation would have the following probability of being assigned to a phone call in the fourth batch or earlier:

$$P(\text{Phone}_t) = p_{\text{phone, group}} \cdot \frac{50 \cdot \text{batch}_t}{900} = 0.43 \cdot \frac{50 \cdot 4}{900} = 0.43 \cdot \frac{200}{900} = 0.0955$$

Finally, let  $p_{it}$  represent the probability of an individual cluster being assigned to the treatment it was actually assigned to in month  $t$ . Let our treatment set  $T_{it} = \{\text{phone}, \text{text}, \text{control}\}$  be the set of three possible treatment assignments, where every cluster  $i$  is assigned to a single treatment in a single month  $t$ . Then, we create weights  $\lambda_{it}$  for our regression models using the following specification:

$$\frac{1}{P(\text{Phone}_t)} \text{ if } T_{it} = \text{phone (the unit } i \text{ is in a treatment condition } \text{phone} \text{ in month } t), \text{ and}$$

$$\frac{1}{P(\text{Text}_{it})} \text{ if } T_{it} = \text{text (the unit is in a treatment condition } \text{text} \text{ in month } t), \text{ and}$$

$$\frac{1}{1 - P(\text{Phone}_{it}) - P(\text{Text}_{it})} \text{ if } T_{it} = \text{control (the unit is in a control condition in month } t). \text{ As the probability of assignment to the control group is 50\% for all units (and only the probabilities of being assigned to the text or phone call vary), this weight will be } \frac{1}{1 - 0.5} = 2 \text{ for all control observations.}$$

### Transformations of Data Structure:

A given provider may have more than one license ID (e.g., multiple child care centers that are associated with the same individual provider). In addition, two providers may share the same email or phone number. Finally, there may be more than one center/license ID at a given physical address. To account for this, we will cluster providers according to whether they share the same contact information, or physical address for a center. Outcomes will be collected at the license ID level, so that there is one observation per month for each license ID.

**Data Exclusion:**

No observations will be excluded from the sample.

**Treatment of Missing Data:**

At the beginning of each month's application period, we will receive a list of eligible providers from DHS. Ineligible providers will not be sent an email or text message, in the month that they are ineligible. However, ineligibility may be correlated with treatment assignment, e.g., providers may be more likely to renew their license if they heard about CCSBG after being sent a text message. We will include these data in our study to estimate an intent-to-treat effect.

A License ID that is missing from the application data will be coded as having not applied in that month.

***Descriptive Statistics, Tables, & Graphs***

We will create the standard OES bar chart showing the probability of being awarded CCSBG in any given month for each condition (no additional communications, text and email, and text, email, and phone), with 95% confidence intervals.

***Statistical Models & 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:****Primary Model**

Our primary estimating model is a weighted least squares regression that measures the effect of the text and email communications and the effect of the text, email, and proactive phone call on award of CCSBG in each month. This is analogous to a Horvitz-Thompson estimator using propensity score weights  $\lambda_{it}$  above, since the true randomization weights are known. This specification will take the form:

$$(1) y_{it} = \beta_0 + \beta_1 \text{text\_email\_only}_{it} + \beta_2 \text{phone}_{it} + X_1 \text{optout}_i + X_2 \text{rural}_i + \phi_i + M_t + e_{it}$$

where  $i$  indexes license ID,  $t$  indexes month, and:

- $y_{it}$  is a binary indicator for award of CCSBG funding in month;
- $\text{text\_email\_only}_{it}$  is a binary indicator for assignment to be sent the email and text treatment and not been assigned to a phone number batch that was shared;

- $phone_{it}$  is a binary indicator for assignment to be sent the email and text treatment and to a phone call batch that was shared in that month or any prior month;
- $optout_i$  is a binary indicator for opting out of the CCSBG text message communications for the CCSBG take-up evaluation or having an invalid phone number;<sup>1</sup>
- $rural_i$  is a binary indicator for the center located in a rural area;
- $\emptyset_i$  is block fixed effects;
- $M_t$  is month fixed effects; and
- $e_{it}$  is the idiosyncratic error term.

For all analysis we also do the following:

- We will include Lin adjusted covariates (including for month and block fixed effects);
- We will use heteroskedasticity-consistent standard errors (HC2), clustered at the provider level; and<sup>2</sup>
- We will use inverse weighting to account for differential probabilities for treatment assignment that varies by month of the intervention and the frequency in which the provider applied for CCSBG prior to the evaluation (never applied, applied 1-4 times, applied 5 times).

### Confirmatory Analyses:

Our main specification has two parameters of interest and one outcome of interest. We have three null hypotheses.

There is no effect of the text and email intervention and there is no effect of assignment to the text, email, and phone call intervention. That can be represented respectively as:

$$H_0 : \beta_1 = 0$$

$$H_0 : \beta_2 = 0$$

And, there is no differential effect of additional phone call intervention compare to the text and email intervention alone:

$$H_0 : \beta_1 = \beta_2$$

We will adjust for multiple-comparisons using a simulation-based approach.

---

<sup>1</sup> We will include this indicator only if we are able to obtain data on unsubscribing or having an invalid phone number that we can verify happened or was recorded before the start of the intervention.

<sup>2</sup> See Winston Lin. 2013. Agnostic Notes on Regression Adjustment to Experimental Data: Reexamining Freedman's Critique. *The Annals of Applied Statistics* 7(1): 295-318.



## Exploratory Analysis:

Our exploratory analysis will do the following:

- examine the effects of the intervention on exploratory outcomes;
- examine the effects of the intervention by center characteristics (i.e., subgroup analysis);
- examine whether the effects of the phone call persists over time; and
- examine treatment-on-treated effects of successful contact from the phone call and downstream effects of the intervention

### Exploratory Models for Subgroup Effects

Exploratory models that will test effects of the interventions by subgroup (i.e., family child care home where reference category is child care center or licensed provider, rural where reference category is urban, and never appliers where reference category is infrequent applier) will take the form:

$$\begin{aligned}\text{Equation 2: } y_{it} = & \beta_0 + \beta_1 \text{text\_email\_only}_{it} + \beta_2 \text{phone}_{it} \\ & + \beta_3 \text{text\_email\_only}_{it} * \text{family\_care}_i + \beta_4 \text{phone}_{it} * \text{family\_care}_i \\ & + X_1 \text{opt\_out}_i + X_2 \text{rural}_i + X_3 \text{family\_care}_i + \phi_i + M_t + e_{it}\end{aligned}$$

where:

- $\text{family\_care}_i$  is a dichotomous indicator for the center being a family child care home;
- $\text{text\_email\_only}_i * \text{family\_care}_i$  is the interaction term for being assigned to the text and email intervention and the center being a family child care home; and
- $\text{text\_email\_phone} * \text{family\_center}_i$  is the interaction term for being assigned to a phone call batch that was shared in a given month and the center being a family child care home.

Using Equation 2, we will test the hypothesis that effects of the email and text intervention and the effects of the email, text, and phone intervention does not affect the likelihood of being awarded CCSBG, for different subgroups (i.e., family child care home where the reference category is child care center or licensed provider, rural where reference category is urban, and never appliers where reference category is infrequent applier), or:

$$H_0: \beta_1 = 0; \text{ and}$$

$$H_0: \beta_2 = 0$$

Using Equation 2, we will also test the exploratory hypotheses that effects of being sent the communication bundle of emails, texts, and phone calls does not affect the likelihood of being awarded CCSBG, for different subgroups (i.e., family child care home/child care center or licensed provider, rural/urban, and never/infrequent applier), or:

$$H_0: \beta_3 = 0; \text{ and}$$

$$H_0: \beta_4 = 0$$

### Exploratory Models for Treatment on Treated Effects

We are interested in the effect of speaking to a Child Care Aware representative on awards. In general, this relationship cannot be identified because of omitted variables, most notably that whether or not an individual answers a call can be due to unobserved variables (e.g., how busy they are, trust in DHS) that also predict their likelihood of applying for CCSBG. To measure the effect of speaking to a Child Care Aware representative on awards, we will first use the following estimating equation to estimate the effect of having spoken to a Child Care Aware Representative in a given month or prior, given random assignment to the text, email, and phone call group in a given month or prior:

Equation 3:

$$Speak_{it} = \beta_0 + \beta_1 text\_email\_only_{it} + \beta_2 phone_{it} + X_1 optout_i + X_2 rural_i + \phi_i + M_t + e_{it}$$

Using exogenous variation in the predicted probability of answering a call from Equation 3, we can estimate the following model:

Equation 4:

$$Award_{it} = \beta_0 + \beta_1 text\_email\_only_{it} + \beta_2 \widehat{Speak}_{it} + X_1 optout_i + X_2 rural_i + \phi_i + M_t + e_{it}$$

We are similarly interested in the effect of awards on provider closure, which in general is not identified due to omitted variables affecting both applying and closing. We estimate the effect of the award on closures during the study period and in the three months after the study period. To measure the effect of being awarded CCSBG on closure, we will first use the following estimating equation to estimate the effect of ever being awarded CCSBG at any time in the study period given random assignment to the treatment group (either texts and emails or texts, emails, and phone calls):

$$\text{Equation 5: } award_i = \beta_0 + \beta_1 text\_email_i + \beta_2 phone_i + X_1 optout_i + X_2 rural_i + \phi_i + e_i$$

Using exogenous variation in the predicted probability of getting an award from Equation 5<sup>3</sup>, we can estimate the following model:

$$\text{Equation 6: } closure_i = \beta_0 + \beta_1 \widehat{award}_i + X_1 optout_i + X_2 rural_i + \phi_i + e_i$$

Using Equation 4, we will test the exploratory hypothesis that the effects of speaking to a Child Care Aware representative does not affect the likelihood of being awarded CCSBG, or:

$$H_0: \widehat{\beta}_2 = 0$$

---

<sup>3</sup> We pre-specify the first stage equation as ever awarded CCSBG during the study period for ease of interpretation; however, we may also consider running the first-stage equation where the outcome is the proportion of months awarded, if we find this approach improves statistical power.

Similarly, using Equation 6, we will test the exploratory hypothesis that the effects of being awarded CCSBG does not affect the probability of center closure, or:

$$H_0: \widehat{\beta}_1 = 0$$

**Inference Criteria, Including Any Adjustments for Multiple Comparisons:**

We will use a t-test to create p-values for a two-sided test with an alpha=0.05.

For our primary analysis, we will adjust for multiple hypotheses using a simulation-based approach as outlined in OES guidance for adjusting for multiple hypotheses.<sup>4</sup>

**Limitations:**

Our analysis may be limited by statistical power to detect differences between being sent emails, texts, and phone calls versus just sent emails and texts. We do not have strong priors on the expected effect size of phone calls.

Similarly, statistical power will also limit our ability to measure the differential effect of the communications bundle by provider type.

---

<sup>4</sup> <https://oes.gsa.gov/assets/files/multiple-comparison-adjustment.pdf>