Analysis Plan

Project name: Decreasing SNAP denial rates with a text campaign and document uploader

Project code: 2310

Date finalized: 10/19/2023

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.

For this project, the Office of Evaluation Sciences (OES) is collaborating with Code for America (CfA) and a state partner to evaluate the effectiveness of a text message campaign encouraging use of a new online document uploader to submit required documents for Supplemental Nutrition Assistance Program (SNAP) applications. The campaign is designed to make it easier to submit verification documents, reduce procedural denials (denials due to missing application components, rather than due to ineligibility), and increase access to benefits for eligible applicants. These findings will help identify strategies to decrease SNAP procedural denial rates, which can be scaled across other US states and possibly other public benefits programs. Conditional on data availability, this project will also build evidence on equitable outcomes by answering to what extent the intervention increased access to SNAP resources among underserved groups.

The proposed intervention has two components: a text message reminder and a link to a new document uploader. First, Code for America has developed a new web-based document uploader, which eliminates the need to log in and has a simplified interface that improves the document submission process. Second, Code for America will send a series of text messages to applicants. We are not evaluating the benefits of the text component of the intervention, as applicants in both the treatment and control groups will be sent text message reminders as part of the evaluation. However, texts to the treatment group will include a link to the new Code for America upload platform, providing direct access to and encouragement to use the Code for America document uploader.

Treatment assignment

SNAP applicants are assigned to either a treatment condition or a control condition based on the date and timestamp of the submission (the rules are outlined below). This method of treatment assignment is intended to be as good as random (i.e., ensuring statistical independence between treatment assignment and potential outcomes), as true random assignment was determined to be infeasible.[1] Assignment occurs at the level of the SNAP application.

To conduct treatment assignment, a SQL query is first run by the state's technology office to pull all applications for SNAP that were submitted online in the past day (i.e., if today is July 20, then the query pulls all applications with a submission date of July 19). The query output is then automatically sent to CfA's Salesforce, which performs a GetRecords query which takes the applicant data sent by the state, imports it as a contact object in Salesforce, and runs the randomization flow described below.

Applications are assigned to either a treatment or a control condition, based on the date of randomization (`randomization_date`) and timestamp of the original application submission (`submission_date_time`), using the following algorithm:

1. Look first at the date of randomization (`randomization_date`). If the day field of `randomization_date` is an odd number (e.g. July 17, 19, 21, etc.), then proceed to step 2. If the day field of `randomization_date` is an even number (e.g. July 18, 20, 22, etc.), then proceed to step 3.
2. Look at the last digit of the second value of the timestamp, (`submission_date_time`) i.e. the bolded digit in a timestamp of the form HH:MM:S**S**. If this digit is an odd number, then assign the application to Treatment. If this digit is an even number, then assign the application to Control.
3. Look at the last digit of the second value of the timestamp, `submission_date_time` i.e. the bolded digit in a timestamp of the form HH:MM:S**S**. If this digit is an odd number, then assign the application to Control. If this digit is an even number, then assign the application to Treatment.

The table below summarizes the results of this algorithm with four examples. We believe that this assignment algorithm is as good as random, as we have no reason to believe that the seconds digit of the date/timestamp will be correlated with any potential outcomes. Once data are collected, we will investigate whether the distribution of treatment is uniform across digits.[2]

For the full sample, we will also conduct a Kologomorov-Smirnov test to adjudicate whether the digits X even/odd day combinations are uniformly distributed. This will yield 20 combinations (10

[1] We show evidence this is the case in the Appendix on Distribution of Treatment and Control for an investigation of the treatment assignment protocol with historical data.
[2] We assessed the distribution of dates and times using data from applications received prior to the start of the intervention, and we expect applicants to be balanced between treatment and control using this assignment strategy.

seconds digits on 2 day types – odd vs. even), and we will compare the observed distribution to a uniform distribution.

	`randomization_date`: Odd	`randomization_date`: Even
Seconds field of timestamp in `submission_date_time`: Odd	Treatment *Example: 07/17/2023, 05:16:31*	Control *Example: 07/18/2023, 05:16:31*
Seconds field of timestamp in `submission_date_time`: Even	Control *Example: 07/17/2023, 05:16:32*	Treatment *Example: 07/18/2023, 05:16:32*

Based on the current volume of applications (~500 / day), we are anticipating approximately 31,000 applications.

Preregistration details

This Analysis Plan will be posted on the OES website at oes.gsa.gov before outcome data are analyzed. In addition, this project will be pre-registered in the AEA RCT Registry at https://www.socialscienceregistry.org/.

Hypotheses

This project's objective is to learn whether the intervention decreases procedural denials and increases SNAP enrollment. To evaluate this, we have three primary hypotheses, as well as several secondary hypotheses that are of interest to CfA and the implementing state partner. In addition to these outcomes, we are interested in heterogeneity in impacts across subgroups, and also intend to look at the impacts of takeup of the document uploader (as opposed to the intent-to-treat effects of encouragement to use the uploader). These analyses are described in more detail in the Exploratory Analysis section below.

In our analysis, we distinguish between procedural denials and ineligibility. Procedural denials are SNAP denials due to failure to complete part of the application process or errors in the application process. Denials due to ineligibility are when applicants are found to not be eligible for the program (e.g., above the income threshold).

Consistent with our theory of change, some of our hypotheses are unidirectional, but the hypothesis tests we will conduct are two-sided tests.

Primary hypotheses

- Intent to treat (**ITT**): Encouragement to use the document uploader, in the form of links included in text messages, will:
    - increase the likelihood that an applicant submits any verification documents for SNAP applications

- decrease procedural denial rates for SNAP
- increase the likelihood that an applicant will be approved for SNAP

Secondary hypotheses

- **ITT**: Encouragement to use the document uploader, in the form of links included in text messages, will:
  - decrease the likelihood that an applicant is procedurally denied for missing verification documents
  - impact the likelihood that an applicant is procedurally denied for missing an interview
  - increase the dollar value of SNAP benefits received
  - decrease the time between initial application and a SNAP decision
  - impact the likelihood that an applicant is found ineligible for SNAP benefits
- **ITT:** Including a link in the text message to the CfA document uploader will impact whether applicants opt out of receiving text messages (pooled across texts)
- **ITT: (Subgroup effects)** Encouragement to use the document uploader, in the form of links included in text messages will do the following within each of the priority subgroups:
  - increase the likelihood that an applicant submits any verification documents for SNAP applications
  - decrease procedural denial (any) rates for SNAP
  - increase the likelihood that an applicant will be approved for SNAP
- **ITT: (Heterogeneous effects)** The effect of encouragement to use the document uploader, in the form of links included in text messages will vary across each of the priority subgroups for the following outcomes:
  - increase the likelihood that an applicant submits any verification documents for SNAP applications
  - decrease procedural denial (any) rates for SNAP
  - increase the likelihood that an applicant will be approved for SNAP
- **Local average treatment effect (LATE): (Received Text)** Receiving a message with a link to use the document uploader will:
  - increase the likelihood that an applicant submits any verification documents for SNAP applications
  - decrease procedural denial (any) rates for SNAP
  - increase the likelihood that an applicant will be approved for SNAP
- **LATE: (Clicked on Link)** Clicking on a link to the document uploader will:
  - increase the likelihood that an applicant submits any verification documents for SNAP applications
  - decrease procedural denial (any) rates for SNAP
  - increase the likelihood that an applicant will be approved for SNAP
- **LATE: (Used uploader)** Beginning to fill out the document uploader will:
  - increase the likelihood that an applicant submits any verification documents for SNAP applications
  - decrease procedural denial (any) rates for SNAP

- increase the likelihood that an applicant will be approved for SNAP

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):
Data for this project comes from the state's SNAP benefits office, the state's technology office, and CfA. Data management and processing will be coordinated by the technology office, CfA, and OES. A description of each data source and the relevant variables it will generate follows.

The data come from four sources:

(1) the state system data warehouse, which records applicant and application information, applicant actions (e.g. uploading documents, attending an interview), and application outcomes (application decision, benefits amounts).

(2) Salesforce, managed by CfA, which records treatment assignment.

(3) Twilio, managed by CfA, which is the platform they use to send text messages and which records whether text messages were successfully sent or bounced, or if recipients opted out of receiving messages.

(4) Mixpanel, managed by CfA, which records if applicants interacted with the CfA document uploader, time spent on each screen, and how many documents they uploaded in the CfA uploader.

CfA will match these four data sources via a crosswalk procedure that they will share with OES, and provide OES with the merged data.

Data will be collected at three points:

1. **Daily applicant data for new applicants (from Salesforce and Twilio):** The state technology office will push daily application data to CfA, which will then assign applicants to receive treatment and control (using the approach described above), and send the initial text messages. At this time, OES will have access to Salesforce data on treatment assignment, application ID, and date and time of initial application submission in the state application system. These variables will be used to monitor that treatment assignment in CfA's Salesforce environment works as planned and to track study intake and sample size.

2. **Daily applicant data for existing applicants (from Salesforce and Twilio):** After the first day of the study intake, the state technology office will push a second extract each day. This extract will include the updated status on all pending applications, including a record of whether they have uploaded any documents or been approved or denied for SNAP. CfA will

subset this list to applicants in the intervention, and send follow-up text messages to applicants (in treatment and control conditions) who are due for reminders to upload documents.

3. **SNAP outcome data**
   CfA will monitor the applications and collate data from the state system data warehouse. For clients who proceed through the entire application process, outcome data will be available after an application has concluded through the full application process (including the interview, document submission, and final eligibility determination). For clients who fail to submit verification documents or complete an interview, outcome data will be available 60 days after their initial application submission date, as that is when their case will be closed by the state system and procedural denials will be determined.

CfA will be responsible for merging data across the initial data extract and randomization, application upload mode, and final outcome data from the state system. Once they have provided the final data extract to OES, we will be able to observe the variables discussed below.

Outcomes to be analyzed:

Variable	Description	Definition
Primary outcomes		
Submitted verification documents	**Indicator for whether the applicant submitted one or more verification documents to the state system**	**Binary:** **0 - Applicant submitted 0 verification documents** **1 - Applicant submitted more than 0 verification documents. This will include irrelevant or incorrect verification documents (e.g. uploading an incorrect income statement or something irrelevant in lieu of an income statement)**
No Procedural Denial	**Indicator for whether the application was denied due to procedural reasons (including missing documentation) or not**	**Binary:** **0 - Application denied due to procedural denials** **1 - Application not denied due to procedural denials, which includes being accepted to SNAP *and* being denied for non-procedural reasons**
Enrolled in SNAP	**Indicator for whether the applicant enrolled in SNAP**	**Binary:** **0 - Applicant received $0 in SNAP benefits**[3] **1 - Applicant received > $0 in SNAP benefits**
Secondary outcomes		

[3] We expect a determination decision for all applicants within 60 days of submitting their application.

Monthly dollar value of SNAP benefits	Continuous variable that records the amounts of benefits reported given to each applicant, including "no benefits" for applicants that were not approved	Continuous: 0 - Receives no benefits (was not enrolled in SNAP) >0 - Dollar amount of benefit received per month. This is the total amount awarded during the eligibility dates, divided by the number of months of eligibility.
Procedural denial – missing verification documents	Indicator for whether an individual was denied for missing verification documents	Binary: 0 - Procedural denial reason is not "missing or incomplete verification documents" (includes other procedural denial reasons, ineligibility, approval, missing value) 1 - Procedural denial reason is "missing or incomplete verification documents"
Procedural denial – missing interview	Indicator for whether an individual was denied for missing an interview	Binary: 0 - Procedural denial reason is not "missing interview" (includes other procedural denial reasons, ineligibility, approval, missing value) 1 - Procedural denial reason is "missing interview"
Ineligibility	Indicator for whether an individual was found ineligible	Binary: 0 - Individual is not found ineligible (this includes all procedural denials, where eligibility is unknown) 1 - Individual is found ineligible
Time to decision	Continuous indicator that records the number of days between completing an initial application through the state system (Day 0) and receiving a decision	Continuous, transformed by taking the difference in the SNAP decision date and the date of the initial application submission) Upper values should be truncated at 60 days due to policy to close an application after 60 days
Opt out	Indicator for whether the client opted out of receiving text messages	Binary: 0 - Applicant did not opt out of receiving text messages on any text sent as part of the evaluation 1 - Applicant opted out of receiving text messages on any text sent as part of the evaluation

All variables will be calculated at the application level, and will be transformed so they are measured once per application. For example, "submitted verification documents" is calculated by summing across rows in a table of possible documents for which submission is not null. This transformation is done by our partners, so data that OES receives will already be at the application level. The outcome variables related to document submission will be for *any* document submission to the state system (e.g., via the state's online system, fax, mail, email, or in-person submission), rather than document submission strictly through the CfA document uploader.

Additional variables and transformations:

In addition to the outcome variables listed above, we also will collect data on:

- The randomization process, generated by Salesforce
- Treatment take-up, generated by Twilio and Mixpanel
- Application data, drawn from the state data warehouse

The table below describes these variables and the variable definitions we will use. Where relevant, variable transformations are also described.

Variable	Description	Definition
Randomization variables		
Treatment	An indicator for whether the respondent was assigned to the treatment group	Binary: 0 - Control group (Treated = FALSE) 1 - Treatment group (Treated = TRUE)
TreatmentSet	An indicator for whether the treatment has already been set. This will be used to ensure treatment assignment is retained for the follow-up text messages.	Binary: We expect this value to be 1 for our full sample. 0 - Treatment has not been assigned (TreatmentSet = FALSE) 1- Treatment has been assigned (TreatmentSet == TRUE)
Time of application submission (submission_date_time)	The timestamp for the date and time of application submission. If applications are submitted over the weekend, the date of submission is assigned to the following weekday.	Date and time field: MM/DD/YYYY HH:MM:SS
Treatment assignment date	The date that treatment was assigned in Salesforce	Date field: MM/DD/YYYY
EvenDate	An indicator for whether the date of treatment assignment is even or odd	Binary: 0 - Odd (Transformed from MM/DD/YYYY of treatment assignment date) 1 - Even (Transformed from MM/DD/YYYY of treatment assignment date)
EvenDigit	An indicator for the last seconds digit of application submission time is even or odd	Binary: 0 - Odd (Transformed from MM/DD/YYYY HH:MM:SS of submission_date_time) 1 - Even (Transformed from MM/DD/YYYY HH:MM:SS of submission_date_time)

Weekend submission	Indicator for whether the applicant submitted their application on the weekend	Binary: 0 - Applicant submitted an application on a weekday (Transformed from **MM/DD**/YYYY HH:MM:SS of application submission date, Weekend day = FALSE) 1 - Applicant submitted an application on a weekend (Transformed from **MM/DD**/YYYY HH:MM:SS of application submission date, Weekend day = TRUE)
Week of submission	Variable that measures the week of application submission relative to the start of the pilot	Integer: Transformed by taking the difference in days between the application time of application submission and the start of the evaluation period, dividing by 7, and rounding down to the nearest integer
Treatment take-up and uploading variables		
Text delivered	Indicator for whether any text message was delivered	Binary: Transformed by summing over Delivery Status = Delivered for each text message an individual applicant is sent 0 - No text was delivered 1 - Any text was delivered (Transformed from delivery status == Delivered > 0)
Treatment text delivered[4]	Indicator for whether any treatment text message (message with link to CfA Document Uploader) was delivered	Binary: Transformed by summing over Delivery Status = Delivered for each text message an individual applicant is sent, only for applicants assigned to treatment 0 - No text was delivered OR applicant is in the control group 1 - Any text was delivered (Transformed from delivery status == Delivered > 0) AND applicant is in the treatment group
Link clicked	Indicator for whether a treatment applicant's unique link was clicked on at least once	Binary: 0 - The applicant's unique link was clicked on 0 times or the applicant was not sent a link 1 - The applicant's unique link was clicked on >0 times *Note: We expect all control applicants to have a value of 0 for this measure, since they should not be sent texts with links.*
First page completed	Indicator for whether an applicant filled out the first page of the document uploader (with first and last name, DOB, and other identifying information) and hit "next"	Binary: 0 - The applicant is not matched to anyone who filled out the first page of the document uploader and hit next 1 - The applicant is matched to anyone who filled out the

[4] This variable is intended to be used in our LATE analysis (described below). We do not expect to be able to observe the actual messages sent (hence we expect to have only one-sided non-compliance); however, if we can, then we would adjust this variable to be a 1 if a control group applicant was sent a text with the link to the CfA Document Uploader (i.e., there could be two-sided non-compliance) and it was delivered. Note that we have no reason to believe that any control group applicants will be sent a text with a link to the CfA Document Uploader and it was delivered. This clarification is solely to make explicit that we are assuming – but cannot verify – that no control applicants are sent a text containing a link to the CfA Document Uploader.

		first page of the document uploader and hit next
Documents uploaded through CfA	Indicator for whether an applicant uploaded verification documents through CfA	Binary: 0 - The applicant matches 0 documents uploaded to the CfA uploader (includes applicants who uploaded documents through methods other than CfA's uploader) 1 - Applicant matches >0 documents uploaded to the Cfa uploader
Application and applicant data		
Application number	An applicant's unique application number that will be used to link applicants across the data sources (Salesforce, Twilio, the state data warehouse)	String
Household zip code	Zip code provided in the application	String NA - Information is missing
Applicant race: POC	Indicator for whether an applicant self-identifies as a person of color or as Caucasian, non-Hispanic *Note: Applicant refers to the person submitting the application; however other members of the household could be associated with the application.*	Binary: 0 - Applicant self-reports being white and non-Hispanic (Transformed from race = WHITE & ethnicity != Hispanic) 1 - Applicant self-reports being non-white or being Hispanic NA - Information is missing
Applicant language: Non-English	Indicator for whether the language that the applicant chose to complete the SNAP application is was non-English	Binary: 0 - Applicant chose to complete the language in English (Transformed from applicant language = English) 1- Applicant chose to complete the language in a language other than English (Transformed from applicant language = Vietnamese | Spanish) NA - Information is missing
Household community type: Rural	Indicator for whether the applicant reports living in a zip code that we code as rural	Binary: 0 - Zip code is urban or mixed 1 - Zip code is rural NA - Information is missing **See Imported Variables section for more.**
Caseworker ID	An identifier for the caseworker that the application was assigned to	String NA - Information is missing

Date of final case decision	Date that the final eligibility determination was made	Date field: MM/DD/YYYY
Held SNAP interview	Indicator for whether the applicant held their SNAP interview	Binary: Transformed from the date field indicating the date of the SNAP interview 0 - Applicant did not hold a SNAP interview (Interview date is NULL) 1 - Applicant held a SNAP interview (interview date is not NULL)
Applicant age	Age in years of the applicant	Continuous: Transformed by taking the difference between the date of data cleaning and the applicant's birthday, rounded down to the nearest year
Household income	Continuous variable representing the amount of household income the SNAP applicant reports on their initial SNAP application	Continuous variable – the dollar amount of applicant's income reported in the initial SNAP application
Household predicted benefits amounts	Continuous variable representing the applicant's predicted benefits amount calculated via information provided in the initial application	Continuous variable – the dollar amount of reported income, reported assets, and reported deductions.[5]
Submitted applications via state only	Indicator for whether an applicant submitted verification documents directly via the state only, and not through CfA	Binary: Transformed by taking the difference between whether an applicant uploaded documents at all (0 or 1) and whether an applicant uploaded documents through CfA (0 or 1) 0 - Applicant uploaded documents via CfA OR did not upload any documents 1 - Applicant uploaded documents via the state only

Our data will be at the application level, but there may be cases where the same applicant appears in several applications – for example, if they forget they have submitted an application and reapply. In that case, we will rely on matching applicants by unique identifier (first and last name and SSN) to other applicants in the dataset. We will transform the data so that each applicant appears only once by taking the maximum value for each outcome variable across rows – e.g., if one version of the applicant's application has a "0" for uploading documents, but another has a "1", then that applicant will get a "1" for that variable.

We also expect that some applicants will share a phone number. For these applicants, we will transform their data so that all applicants using the same phone number get the maximum treatment assignment – in other words, if one applicant with a phone number is treated, then we will consider all applicants with that phone number treated. We have made this decision because we expect people who are close enough to share a phone number to be in the same social circles, and can share the link amongst each other once they are given access to it through at least one of the text messages.

[5] The exact calculation is still to be determined in consultation with the project partners.

Imported variables:
We will merge in one variable from non-administrative data: an indicator for whether an applicant's zip code is in an urban or rural area. For this we will use the 2020 Census Urban Area definition. To qualify as an urban area, the area must encompass at least 5,000 people or 2,000 housing units. Areas are typically Census tracts.

We will download data from the US Census showing the percent of the population in each ZIP code[6] that Census designates as living in an Urban or Rural area. Around 36% of ZIP codes in this state contain both areas that are considered Urban and areas that are considered Rural. For the purposes of this study, if any part of the ZIP code is considered Urban (that is, if >0% of the population is classified as living in an Urban area), we will consider the full ZIP code to be Urban. This is because these mixed Urban/Rural ZIP codes are primary suburbs of larger urban areas, and our main goal in this analysis is to identify effects for more rural communities, rather than borderline ones.

Using this information, we will assign each SNAP applicant to be Rural or Urban, based on the ZIP code they list on their initial application.

Transformations of data structure:
Our data will be collected via the state system, Salesforce, Twilio, and Mixpanel. Code for America will develop the crosswalk linking applications across data sources via their application number and phone number. We will have one dataset at the application level when Code for America provides us with the data.

Data exclusion:
We expect that the only exclusions will take place pre-randomization and will be dropped from the study. This includes applicants who applied for benefits under the expedited process or applied for benefits in-person or via a paper application (not online) are not eligible for this study, so they are automatically excluded from the sample.[7] These applicants will be excluded from the study at the initial stage prior to sending out text messages. CfA will also exclude people who do not submit phone numbers in the cell phone field in the initial state system application.

We also exclude from our primary analysis applicants during the ramp up period, which ended September 21, 2023.

Treatment of missing data:
For individuals not excluded from the dataset, missingness will primarily arise from two sources:

[6] Census provides ZIP Code Tabulation Areas (ZCTAs), which closely represent ZIP codes as defined by the USPS. For more information, see https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html
[7] These applicants are excluded from eligibility for the document uploader pilot. However, it is possible that some individuals may somehow obtain the uploader link. If an individual who is excluded from the study uploads documents via the CfA document uploader, we will drop these observations.

1. Engineering: Technical error where anticipated paradata is not collected
2. Applicant generated: Not filling optional fields in the original application

For engineering-generated missing values, it is possible that we will have to omit some outcome variables or the uploader LATE analysis if we do not have the necessary data due to technical errors across the whole sample (e.g., if opt-out data is not available at the individual level). In that case, we will not run the secondary analyses that rely on this data.

For applicant-generated missingness, we could be missing race, as SNAP cannot require that information. We will only conduct the planned subgroup analysis for race/ethnicity for those who have reported this information and if this information is available for sufficient sample size and reporting rate across race/ethnicity. For the models with covariates that include race, for individuals who are missing race/ethnicity, we will impute the value to the mean, and then include an indicator variable for missingness.

We could also be missing our indicator for urban/rural if the ZIP code given by the applicant is inaccurate and cannot be matched with Census data. We expect that this will be uncommon. As with race, we will only conduct subgroup analyses for applicants for whom this indicator is not missing. When running models that include urban/rural as a covariate, we will impute missing values of this variable to the mean, and include an indicator for missing.

Descriptive statistics, tables, & graphs

Primary analysis
We will create a bar chart showing the probability of uploading any verification documents, receiving a procedural denial and being accepted to SNAP for the treatment and control conditions, with 95% confidence intervals. We will have a separate bar chart for each of the primary outcomes.

Understanding compliance
We will also produce the following summary statistics, tables, and graphs to understand intervention implementation. The evaluation relies on an encouragement design where individuals assigned to the treatment group are sent text messages that include an easy-to-use link to the CfA uploader and individuals assigned to the control group are sent text messages without these links. However, spillover could occur if the CfA link was shared between people in the treatment group and people in the control group, for example if a person in the treatment group or caseworker shared the link. Thus, part of our analysis will describe take-up / use of the CfA uploader among applicants in the control group..

To conduct this analysis we plan to plot the proportion of applicants that used the CfA uploader by treatment assignment over the course of the whole study period and over time (to better conceptualize if spillover increased over time as more people had been assigned to the treatment group, and thus could share the link). We will also plot the proportion of applicants that uploaded

documents via means other than the CfA uploader by treatment assignment over the course of the whole study period and overtime.

Treatment assignment and implementation checks

We will inspect the distribution of the timestamp digit for uniformity. If there is some computer glitch or other unknown factor that might cause applications to "bunch" at certain seconds values, we wanted to be sure we were not accidentally assigning more applications to either the treatment or the control conditions. As data from fielding come in, we will:

- Reassign applicants to treatment and control following the protocol we outline above to ensure that it matches the treatment assignment that Salesforce determined. We will calculate the proportion of units where their assigned treatment and control condition matches the protocol using an R script that we write based on `randomization_date` and `submission_date_time`.
- Visually examine the distribution of treatment and control assignment to verify that the distributions appear uniform.

Balance checks

We will conduct balance checks between treatment and control for the vector of applicant-level covariates included in our primary adjusted model (see equation 2 below). We will test for balance using a single overall omnibus balance test where we will regress treatment on the above list of covariates and assess balance by looking at the test statistic of the F-test, following this approach. Additionally, we will test for balance on assignment to caseworkers using this same approach.

Funnel

We will calculate summary statistics of how many applicants made it through the intended flow, showing the following by treatment condition as a percentage of the total sample for each condition:

- % Applicants sent texts
- % Applicants who had texts delivered
- [Sent texts with links only] % unique links that were clicked on
- % applicants who uploaded documents at all
- % applicants who were *not* procedurally denied for benefits
- % applicants who were enrolled in SNAP

Cell phone numbers

There is a possibility that cell phone numbers may become more inaccurate over time as people change cell phone numbers (e.g., people who rely on prepaid phones for cell coverage). For numbers that have gone out of service, text messages may be undelivered. To better understand the extent to which changing numbers may influence the efficacy of the intervention, we will calculate summary statistics that plot the proportion of texts that were undelivered for each text in the sequence by day they were sent, relative to when the first text went out (e.g. 3 days later, 10 days later, 28 days later). We will generate these summary statistics only for people who were sent every text in the sequence, in order to keep the same composition of applicants over time.

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:

We specify two different types of models for the ITT ("intention to treat" effect), which is our primary specification. These types are: the unadjusted model (UAM) and the fully-adjusted model (FAM). We describe which models we will use for the main confirmatory test, and which are robustness checks for the ITT specification across our three primary outcomes. For these specifications, we use a linear probability model that estimates the effect of receiving a link to (encouragement to use) the CfA document uploader on our primary outcomes.

As we discuss in the Appendix List of Statistical Models and Interpretations, **we will report the ITT using the fully adjusted model as our primary confirmatory model**. For the fully-adjusted model, we include a set of covariates that existing literature suggests is predictive of whether people are approved for SNAP benefits, and covariates that capture differences in how the intervention was implemented across applicants. For the fully-adjusted model, we also include the three variables that indicate membership in the three subgroups (language preference, person of color, and urban or rural status) for which we will test for treatment effectiveness separately.

$$Y_i = \beta_0 + \beta_1 treatment_i + e_i \quad \text{(1) Unadjusted Model}$$

$$Y_{itc} = \beta_0 + \beta_1 treatment_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc} \quad \text{(2) Fully-Adjusted Model}$$

where:

- Y_{itc} is the outcome of interest for applicant i, who submitted their initial application in week t, and was assigned to caseworker c;
- β_0 is the intercept;
- $treatment_i$ is an indicator for assignment assignment to the text with the uploader link group;
- X_i is a vector of applicant-level baseline characteristics collect during the initial applicant (prior to assignment), where the vector of covariates include:
    - $shared\_phone_i$ is a binary indicator for whether the applicant shared a phone number with another applicant in the sample;

    - $weekend\_application_i$ is a binary indicator for whether an applicant applied on a Friday or Saturday;
    - $rural_i$ is a binary indicator for whether an applicant reported a zip code or address that is rural;
    - $language\_english_i$ is a binary indicator for whether an applicant reported preferring English (versus Vietnamese or Spanish, the other two languages offered for the state SNAP application);
    - poc_i is a binary indicator for whether an applicant is a person of color (versus white)
    - age_i is an integer for the applicant's age in years reported in the initial application
    - hh_size_i is an integer for the total number of people in the household / SNAP unit reported in the initial application
    - $need_i$ is a continuous variable showing total household expected benefit amount or, if we are unable to calculate predicted benefit amount,total household income;
    - $mean\_imputed_i$ is a vector of indicators for each measure for which the applicant was missing data and their value for that measure was imputed to the sample mean; and
- X_t are fixed effects for the week in which the initial application was submitted; and
- X_c are fixed effects for the casework assigned to the applicant, which we will only include in our adjusted model if caseworker data are available for 95% or more of the sample, median caseload size is greater than 100, and caseworker assignment is orthogonal to treatment assignment[8]; and
- ϵ_{itc} is the idiosyncratic error term.

We will estimate each of these statistical models with heteroskedastic-consistent standard errors (CR2) that are clustered on **applicant phone number**, per OES guidance. For the covariate-adjusted models, we will use Lin-adjusted covariates.[9]

Robustness checks

The FAM will be our primary statistical model and the basis for our confirmatory test. Our robustness checks will be as follows, and apply only to our main, confirmatory model.

- Estimate the binary outcomes using logistic regression and report the predicted probabilities calculated using the `marginal` package in R.
- Estimate the unadjusted (linear) model with no covariates

[8] Orthogonality to treatment will be evaluated through an omnibus test and a significant F-test statistic.

[9] 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.

Confirmatory analyses:

Our main specification (the fully adjusted LPM model) will be estimated three times, yielding three parameters of interest (β_1) that capture the causal effect of encouragement to use the document uploader on our three primary outcomes: uploading any documents, procedural denial, and acceptance into SNAP. We will test the null hypothesis that the effect of encouragement to use the document uploader on SNAP outcomes is statistically indistinguishable from 0. This can be expressed as the following:

$$H_{upload.docs}: \beta_1 = 0$$

$$H_{pro.denial}: \beta_1 = 0$$

$$H_{SNAP.acceptance}: \beta_1 = 0$$

We will adjust for multiple comparisons using a simulation-based approach, correcting for testing three hypotheses in the same family. We consider this family the core family testing whether the document uploader improved priority outcomes. We will not include the UAM in our adjustment for multiple hypothesis testing corrections.

Exploratory analysis:
We have four sets of exploratory analyses: (1) ITT analyses that look at our secondary outcome measures, (2) ITT analyses that compare the effectiveness of encouragement to use the document uploader across subgroups, (3) ITT analyses that look at whether treatment encouraged individuals with greater/less economic need to enroll in SNAP, and (4) IV analyses that look at the local average treatment effect for the primary outcomes. The hypotheses for the exploratory analyses are outlined in the Hypotheses section for Secondary Hypotheses.

1) ITT Analyses for Secondary Outcomes

We will estimate the FAM for our secondary outcomes using the same equation as above, testing the following null hypotheses.

$$H_{pro.denial.documents}: \beta_1 = 0$$

$$H_{pro.denial.interview}: \beta_1 = 0$$

$$H_{SNAP.amount}: \beta_1 = 0$$

$$H_{time.to.approval}: \beta_1 = 0$$

$$H_{ineligible.SNAP}: \beta_1 = 0$$

$$H_{opt.out.messages}: \beta_1 = 0$$

2) Subgroup effects

We also include our subgroup analyses in our exploratory analyses. The subgroup models will allow us to answer two research questions:

1. What is the effect of sending text with the document uploader link on SNAP outcomes for applicants with certain demographic characteristics?
2. To what extent do these effects differ among these groups?

As with the primary model, we will use the adjusted model as our primary regression specification. We will use these models to test the following hypotheses:

- Encouragement to use the document uploader will increase the likelihood that an applicant submits any verification documents for SNAP applications within these sub-groups:
    - Primary language (English speakers vs. non-English speakers)
    - Ethnic / racial group (Minority ethnic / racial group vs. non-minority)
    - Community type (non-rural or rural areas)
- Encouragement to use the document uploader will decrease procedural denial rates for SNAP within these sub-groups:
    - Primary language (English speakers vs. non-English speakers)
    - Ethnic / racial group (Minority ethnic / racial group vs. non-minority)
    - Community type (non-rural or rural areas)
- Encouragement to use the document uploader will increase the likelihood of getting approved for SNAP within these sub-groups:
    - Primary language (English speakers vs. non-English speakers)
    - Ethnic / racial group (Minority ethnic / racial group vs. non-minority)
    - Community type (non-rural or rural areas)

Our model specification for the subgroup models is (using the example of community type):

$$Y_{itc} = \beta_0 + \beta_1 treatment_i + \beta_2 rural_i + \beta_3 treatment_i * rural_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc}$$

where:

- y_i is an outcome of interest;
- $intervention_i$ is a binary indicator for assignment to the intervention group;
- $rural_i$ is a binary indicator for living in a rural area; and
- $intervention_i * rural$ is an interaction term between assignment to the intervention group and living in an rural area; and

The other terms are the variables outlined elsewhere in the analysis plan.

The confirmatory analysis will test the null hypothesis that assignment to the intervention group has no effect on the outcome of interest among those in living in non-rural areas (the left out group):

$$H_0: \beta_1 = 0$$

and the difference in effects of the intervention between those living in the in rural and non-rural areas is zero:

$$H_0: \beta_3 = 0$$

For this analysis, we consider each category to be a separate subgroup (e.g., rural vs. non-rural), but group these into three pairs of subgroups (community type, preferred language, race). We will run this analysis six times for each primary outcome in order to measure the effect among the six subgroups of interest (non-rural, rural, English language preferred, non-English language preferred, white, and person of color) and the difference in effects among the three pairs of subgroups.

For the subgroup models, we will use Lin-adjusted covariates for all but the subgroup in question. We will use CR2 standard errors that are clustered on applicant phone number.

3) Effects by household need

We also will look at the average benefit amount (actual, not predicted) — conditional on enrollment in SNAP — to examine whether the intervention induces take-up among those with more or less need.

We also plan to plot treatment effects by quintile of household need (e.g., total household income and/or expected benefit amount).

4) Local Average Treatment Effect (LATE) / Complier Average Causal Effect (CACE))

Our main specification relies on an intent-to-treat estimand to identify the effects of the document uploader on SNAP applicant behavior. There are three limitations to this estimand that we can address with an exploratory analysis that estimates the LATE using a 2SLS ("two-stage least squares") model. First, not everyone who is randomized to be sent a text message will actually receive a text. Second, not all treatment applicants will engage with the document uploader, and we may be interested in disentangling treatment effects for those who actually use the uploader from the group who is sent the texts. Third, some applicants in the control group will get access to the document uploader (i.e. there may be "always takers" for the uploader treatment). We will estimate three different 2SLS models, each designed to address these

limitations. We will use the adjusted model with covariates for these specifications, and will only run this model for the primary outcome variables.

First, we will estimate the effect of receiving the link to the uploader among those for whom the text message with the link to the uploader (the treatment text) was successfully delivered (text status is **`delivered`**). We assume one-sided non-compliance, where some people who were sent the text message with the link will not receive it, but no one who was assigned the text message without the link will be sent a text that includes the link. We assume no defiers. We then measure the first stage by regressing whether an individual i receives a treatment text message on treatment status, and including the same set of covariates and fixed effects in both the first and second stages.

$$receives.text_{itc} = Y_{itc} = \beta_0 + \beta_1 treatment_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc}$$

We then obtain the fitted values on $receives.text$, estimated in the first stage above, and regress the outcomes of interest on the fitted values.

$$y_{itc} = \beta_0 + \beta_{2sls} \widehat{receives.text}_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc}$$

From this, we can test the exploratory null hypotheses that receipt of the text message including the document uploader link is statistically significantly different from 0. We estimate this model for each of our three outcome variables.

Second, we are also interested in understanding the effect of the uploader among those who click on the link. Note that we can only observe clicking on the link for applicants in the treatment condition, and we measure it by seeing whether an applicant's unique link was clicked on at least once. For this model, we measure the first stage, regressing whether an individual i was given a link that was clicked on at least once if they were assigned to the treatment group:

$$link.click_{itc} = \beta_0 + \beta_1 treatment_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc}$$

We then obtain the fitted values on $link.click$, estimated in the first stage above, and regress the outcomes of interest on the fitted values.

$$y_{itc} = \beta_0 + \beta_{2sls} \widehat{link.click}_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc}$$

From this, we can test the exploratory null hypotheses that clicking on the link results in applicant behavior that is statistically significantly different from 0.

Finally, we want to understand the effect of the uploader among those in both the treatment and control group who interact with it by estimating the Local Average Treatment Effect (LATE) among the compliers. Because applicants select into using the document uploader – and that selection process is a function of omitted variables – actually using the uploader is not orthogonal to potential outcomes. To address this, we estimate two equations using the 2SLS estimator. We assume two-sided non-compliance (people who were given the doc uploader link may choose not to use it, and people who are not given the link may otherwise obtain it). We assume no defiers, and that Never-Takers and Always-Takers have the same potential outcomes in treatment and control. We measure uptake of the treatment by whether an individual has filled out the first page of the document uploader and hit "next". This is the information we need to match an individual's use of the document uploader to their record in the applicant dataset.[10]

First, we measure the first stage, regressing whether an individual i uses the document uploader on if they were assigned to the treatment group:

$$upload\_use_{itc} = \beta_0 + \beta_1 treatment_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc}$$

We then obtain the fitted values on $upload.use$, estimated in the first stage above, and regress the outcomes of interest on the fitted values.

$$y_{itc} = \beta_0 + \beta_{2sls} \widehat{upload\_use}_i + \gamma_1 X_i + \gamma_2 X_t + \gamma_3 X_c + \epsilon_{itc}$$

From this, we can test the exploratory null hypotheses that use of the document uploader results in applicant behavior is statistically significantly different from 0.

We summarize the points in the treatment implementation and uptake where we will be able to observe the variable for the first stage of the 2SLS for both analyses in Figure A3 in the Appendix, Data Sources and Observed Variables for 2SLS .

Inference criteria, including any adjustments for multiple comparisons:
We will apply multiple hypothesis corrections through simulations to account for the three primary outcomes and control the family-wise error rate, per OES methodological guidance. We will reject the null hypothesis that $\beta_1 = 0$ using a cutoff of p = 0.05 to determine statistical significance using two-tailed tests and the multiple hypothesis correction.

Limitations:
Between finalization of the design document and the project launch a limitation arising from anticipated spillover or treatment noncompliance emerged: While waiting to launch the project,

[10] Ideally, we would measure whether an applicant has opened the document uploader at all, rather than whether they have gotten partially through the process. However, the first time we will be able to observe whether an individual has used the uploader is when they enter in identifying information on the first page of the uploader.

Code for America identified two community-based organizations with whom to conduct user testing on the document uploader. These organizations will have worked with CfA to do user testing with a generic document uploader link, and it is possible they will continue to use a generic link after the user testing, despite communication from CfA that they should not share the link once the pilot officially launches.

It is also possible that any effects we detect could be unique to this time period because the state may implement its own text-for-benefits campaign part way through the study period, which may impact how recipients view the text messages sent by CfA on behalf of the state

Our ability to estimate primary effects could be limited due to statistical power if SNAP applications are submitted at lower levels than in past months and years, or if we otherwise have low statistical power.

See also the Appendix, List of Statistical Models and Interpretations for a list of how different models address particular limitations.

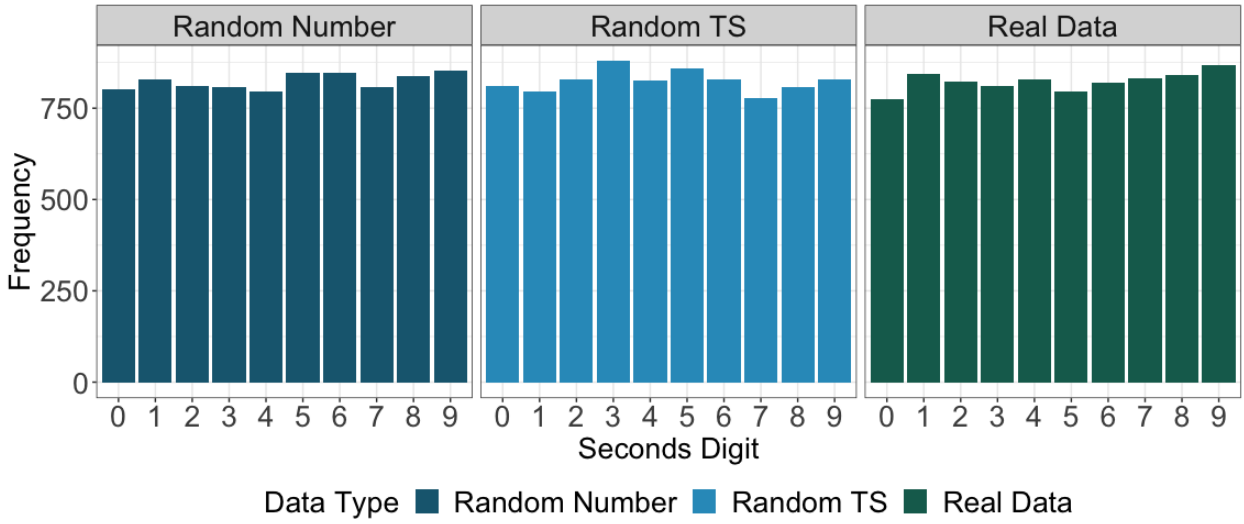
Link to an analysis code/script:
N/A

Appendix

Distribution of treatment and control under day/timestamp assignment strategy

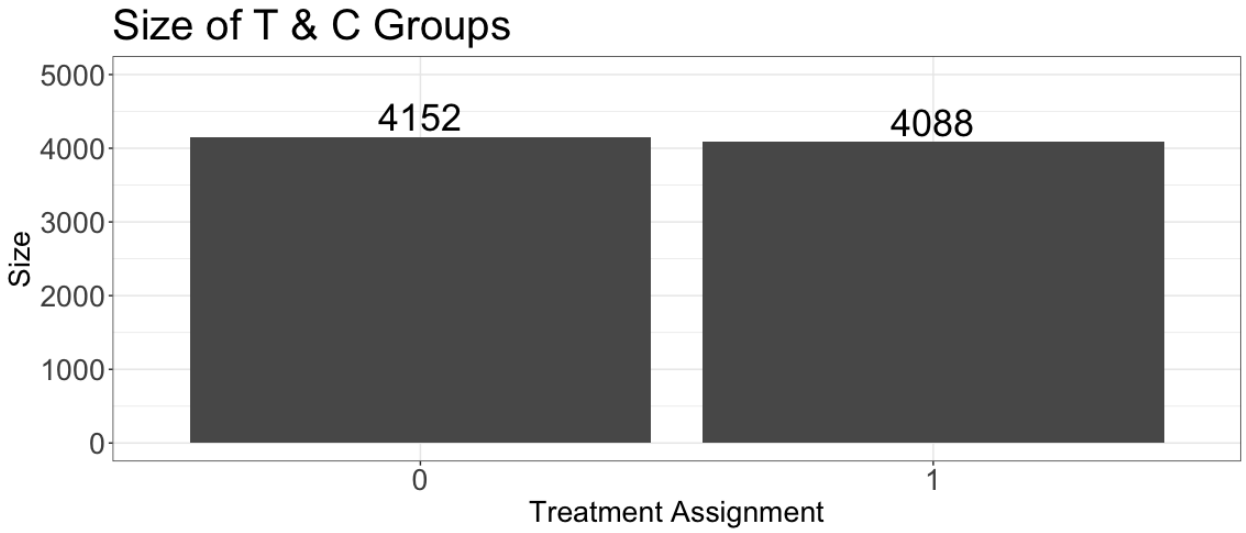
Our treatment assignment strategy is based on timestamps and the date of randomization. We have no reason to believe that the seconds digit, combined with the last digit in the date of randomization, would be correlated with potential outcomes. While we do not have the baseline data to test this assumption, we do visually compare the distribution of last digits over one month of historical real application data to the distribution of last digits in a randomly drawn number sequence and a randomly drawn timestamp. The distributions (displayed in Figure A1) are similar.

Figure A1: Visual comparison of the frequency of last digits, comparing the frequency of each digit drawn using a random number sequence, a random timestamp generator, and the real data.



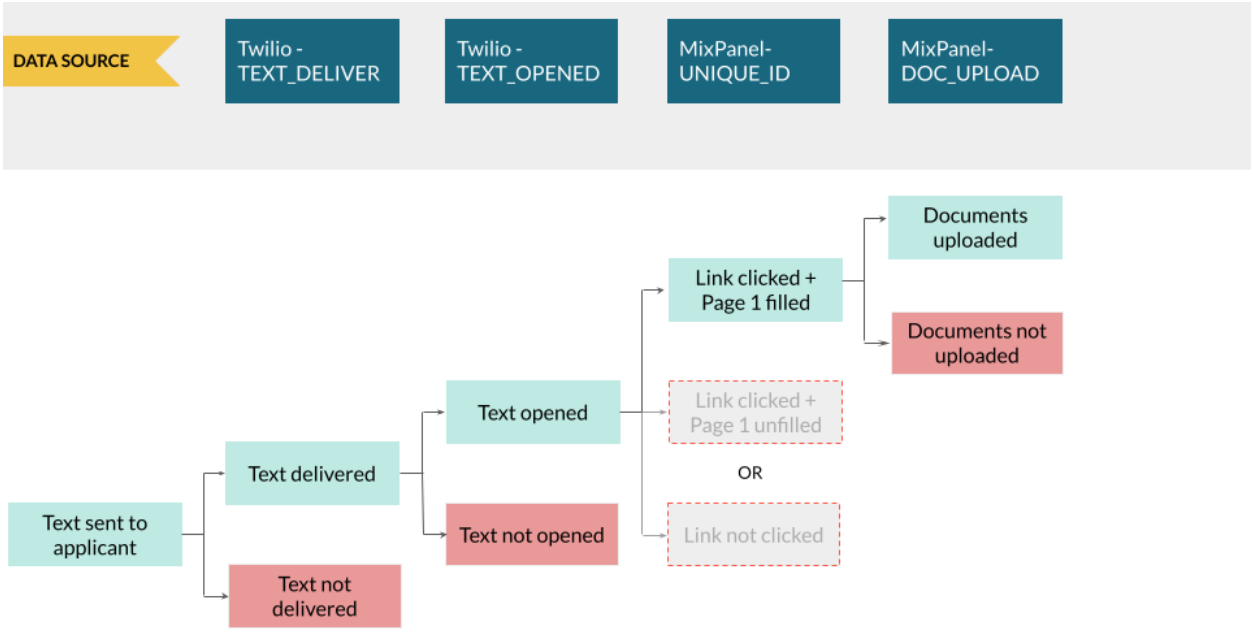
We also wanted to ensure that we would have roughly the same number of applicants in the treatment and control groups. Figure A2 shows the volume of applicants sorted into treatment and control using the procedure we outline based on thirty days of historical application data. This yields roughly equal numbers of applications into treatment and control.

Figure A2: Volume of applicants sorted into T&C, based on the `randomization_date` and `submission_date_time`



Data flow

Figure A3: Data sources and observed variables for 2SLS



List of statistical models and interpretations

Below, we list out the different models we will run, the rationale for each of the models, and how we will interpret the results.

Model	Sample	Covariates	Purpose and interpretation
ITT	Full	Full	Primary model to estimate the effectiveness of encouragement to use the document uploader
ITT	Full	None	Robustness check to show that the point estimates do not change substantially when we do not include covariates
ITT, Subgroups	Subgroup of interest	Full	Primary model to estimate the effectiveness of encouragement to use the document uploader within priority subgroups
ITT, Heterogeneous Effects	Full	Full, subgroup of interest and interaction term	Primary model to compare the effectiveness of encouragement to use the document uploader between priority subgroups
LATE, Received Text Messages	Full	Full	Robustness check to account for one-sided non-compliance in who receives a link to the document uploader. In the case that many phone numbers are unreachable, this analysis will help us account for the noise generated by people not receiving texts.
LATE, Clicks on the Link	Full	Full	Robustness check to account for the fact that not everyone will click on the link. This answers a slightly different but important question of **"what is the effect of the intervention on SNAP outcomes among those who engage (or takeup) with the intervention?"** but defines take-up minimally
LATE, Used the Document Uploader	Full	Full	Robustness check to account for contamination in treatment in who uses the document uploader. With high upload use in the control group and a positive treatment effect of the uploader, we may not see treatment effects in the ITT. This answers a slightly different but important question of **"what is the effect of the intervention on SNAP outcomes among those who engage (or takeup) with the intervention?"** but defines take-up as actually filling out part of the document uploader.