Analysis Plan

Project Name: Reducing no-shows for in-person passport appointments
Project Code: 2313
Date finalized: 8/7/2024

Project description

This project evaluates a policy change designed to reduce occurrences of uncancelled no-shows for in-person appointments for passport processing. We will be using a regression discontinuity design by time to test the impact of a robotext intervention on no-show rate. The intervention is designed to both remind appointment holders about their appointments and provide a way to cancel the appointment via text.

Prior to May 22, 2024, passport applicants who scheduled an in-person appointment for passport services received an automated phone call 48 hours before their appointment. This call serves as an appointment reminder and can be used to connect to a customer service center to cancel or reschedule an appointment. Appointment holders also received a confirmation email when booking their appointment that includes instructions on how to cancel or reschedule.

On May 22, the U.S. Department of State, Bureau of Consular Affairs (CA) began asking customers when they book appointments if they agree to receive text message reminders ("robotexts"). For customers who book their appointments online, they are presented with a box to check to opt in to receive text messages. For customers who book their appointments over the phone, the customer service representative asks the customer for their permission to sign them up for text messages.

If the appointment holder has opted-in to receive robotexts they will receive a text reminder 48 hours before their appointment in addition to the automated phone call reminder. This text message includes the appointment details and the option to reply with "1" to confirm the appointment, "2" to change the appointment, or "3" to cancel the appointment. If the appointment holder does nothing or replies with "1", nothing will happen and their appointment remains unchanged. If they reply with "2", they are provided with information on how to call the National Passport Information Center (NPIC) to reschedule their appointment (currently, it is only possible to reschedule by calling NPIC). If they reply with "3", a second text is sent asking, "Are you sure you want to cancel?", and they click on the word "cancel" in order to move forward with canceling the appointment.

If an appointment holder who has opted-in to receive robotexts does not indicate any response to the 48-hour text, they will be sent a duplicate reminder 24-hours before their appointment time.

Preregistration details

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

Hypotheses

The overarching research question is: What is the impact of sending a robotext to passport appointment holders on appointment no-shows, appointment cancellations, and utilization of in-person passport appointments?

The primary hypothesis is that sending robotexts to customers with scheduled appointments will increase the likelihood that customers who no longer need their appointments will cancel their appointments, thereby reducing no-shows.

We hypothesize that the proportion of appointments that are no-shows will be lower among participants who schedule appointments after robotexting has gone into effect than among participants who schedule appointments before robotexting began.

Data and data structure

Data source(s):

The primary data source for this evaluation is data obtained from the U.S. Department of State's Consolidated Appointment System (CAS). This contains data elements that are logged when a customer makes an in-person appointment and records when customers change, cancel, attend, or do not show up for an appointment.

Outcomes to be analyzed:

Outcome	Description
Cancellation	Whether or not an appointment was canceled prior to the appointment time.
Appointment no-shows	Whether or not an appointment resulted in a no-show, where the appointment was unattended and not canceled or canceled after the appointment time.
Successfully completed appointment	Whether or not a scheduled appointment resulted in the appointment holder attending and completing the appointment.
Rescheduled (pending data availability)	Whether or not a scheduled appointment was rescheduled to a new time.

Transformations of variables:

We will create a new variable "robotext" with a value of 1 if customers created an appointment after the robotext rollout date and a value of 0 if they created an appointment before the robotext rollout date (this is an intention-to-treat variable; we do not know if all of these appointment holders actually received a text).

There are 12 data elements in the CAS data, which we will transform in order to conduct the analyses. The variables we will use include details about how and when and for whom the appointment was made, when and where the appointment was made for, type of appointment, and the outcome of the appointment. Where appropriate we will transform these variables into categorical or binary indicators for analysis. In addition to these 12 variables, we are exploring the possibility of getting access to individual-level data on: 1. opt-in indicator (have actively opted in to receive text messages), 2. robotexts sent, 3. whether the customer responded to confirm the appointment, 4. whether the customer has responded to reschedule their appointment, 5. whether the customer responded to cancel their appointment. If we do not receive individual-level data, we plan to analyze aggregate data.

Transformations of data structure:

Not applicable.

Data exclusion:

Appointments may be made up to 14 days prior to a customer's travel date. We will exclude appointments whose date is more than 14 days after the appointment-making date, as they likely reflect data-recording errors.

Treatment of missing data:

Based on the 2023 appointment data, the only field that contains missing data is "uniqueid," the identifier for individual appointment-makers. The variable is missing for about 0.52% of the observations. It is also an imperfect appointment-maker identifier because it is based on an individual's last name and different people could share the same last name. Nevertheless, we do not expect this missing data issue to impact our analysis, as we are not planning on using the *uniqueid* variable in our analysis.

Descriptive statistics, tables, and graphs

- Table of descriptive statistics: We will summarize observable covariates for appointment holders in the robotext and no-robotext groups. Covariates include:
    - Length of time from when the appointment was booked and the appointment time
    - Number of appointment attendees
    - Method of appointment booking (CSR, Agency, individual)
    - Type of appointment (regular, standby, special case)

- No-show appointments by day: We will create a line chart summarizing for each day in the study period the percentage of appointments made that day that resulted in a no-show, the percentage that were canceled, and the percentage resulting in a successfully completed appointment. The chart will include an indicator of which day robotexts were introduced.
- Table of statistical results: This table will report the results of the statistical tests conducted using the models and inference criteria described below.

Statistical models and hypothesis tests

This section describes the statistical models and hypothesis tests that will be used to conduct 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 will estimate a regression by comparing the no-show rates for appointments created before and after the implementation of robotexts. The unit of analysis is passport appointments. Specifically, we will estimate the following for appointment i for date/time t:

$$Y_{it} = \alpha + \beta T_t + \theta X_{it} + \lambda_t + e_{it} \qquad (1)$$

where Y_{it} is the outcome variable for each appointment. The indicator variable T_t takes the value of 1 when an appointment is eligible for receiving the robotext reminder, that is, if the appointment was booked on or after May 22, 2024 for an appointment time on or after May 24, 2024 (to allow for a 48-hour pre-appointment period). The vector X_{it} include covariates that could impact the outcome of an appointment: number of attendees, method of making the appointment, type of appointment, length between when the appointment was made and the appointment time, appointment location fixed effects, and days of the week fixed effects. Lastly, λ_t is a time trend as discussed later and e_{it} is the error term. We will follow a blinded analysis protocol as detailed in Appendix A.

Identifying assumptions and proposed checks

The regression continuity in time design rests on a few important assumptions. First, the timing of the change being studied should be exogenous and cause no sorting behavior. The implementation of robotexts is unexpected and exogenous to customer behavior, making this a likely case of a "local randomized experiment." People who are scheduling appointments will not have known the robotext launching date in advance. They also have limited discretion in picking an appointment time because travelers are only eligible to schedule appointments in a two week window pre-travel and are often given few availability.

Second, there should be a substantial mass of observations on either side of the time cutoff (i.e., robotext rollout date). Because our unit of observation is individual appointments and there are numerous appointments on each day across passport agencies, we likely have asymptotics in the

number of observations as opposed to asymptotics in time that plague some interrupted times series design.

Third and perhaps most importantly, there should be no confounding factors changing at the same time as the robotext rollout or the potential time-varying confounders are assumed to change smoothly across the date of the change (May 22 for appointment creation). While we are not aware of any other policy or program changes relevant to no-shows or appointment scheduling being rolled out during this time frame, other factors impacting passport customer behavior could change during the analysis period. For example, increasing daylight time may affect travel, affecting no-show behavior over time. If no-show rates increase abruptly at some point during the months of May and June due to end-of-school-year activities, our design would pick up such an increase as evidence of the robotext increasing no-shows, even though no such effects exist. We discuss below options to address this issue.

The assumption underlying the RD design is that potential time-varying confounders change smoothly over time, and thus can be controlled for through adding a parametric time trend λ_t to our analysis. This is known as controlling for the global polynomials of the running variable in an RD design, although it is worth pointing out that our "global" period is still very short. We note that the running variable is appointment creation date, as opposed to appointment date, because whether an appointment is created before or after May 22 determines the status of treatment, i.e. robotext eligibility. That means the appointment outcome will not be observed on the same date but rather sometime after the appointment creation when the appointment is canceled, attended or missed.

Specifically, we will allow λ_t to differ for appointments before and after May 24. The correct specification for the time trend is not knowable, but given the short analysis period, we will test the smoothness in covariate assumption by regressing each covariate on an indicator of post-robotext period and a polynomial time trend, that is, we will estimate equation 1 above but with each covariate as the outcome variable. A statistically insignificant coefficient estimate for the indicator variable will provide support for the smoothness in covariate assumption. For the time trend, we will test linear, quadratic or cubic functions of the appointment date; among the functions that generate a statistically insignificant estimate for the indicator variable, we prefer the more parsimonious option (i.e. linear over quadratic and quadratic over cubic). We will plot a parallel RD estimated on control variables to demonstrate continuity. We will control for the same time trend when analyzing the no-show outcome.

If significant covariate imbalance remains after controlling for polynomial time trends, the assumption for a "global" RD over the analysis period is unmet. If this occurs, we will explore the following. First, we will examine a shorter time window around May 22 where observed covariates and unobserved potential confounders are more likely to be similar. Second, we will test if the covariate balance is achieved by dropping one or more of the passport agency locations. Subsetting the sample by location enables us to focus on only locations that pass the covariate balance test. Both narrowing the analysis window and subsetting the sample by location reduce

the sample size and thus the probability of statistically detecting a covariate imbalance. Therefore, we will exclude as few days/locations as possible in meeting covariate balance.

Other robustness checks

We plan to use data from the previous year during the same date range for a placebo test. If no effect is detected through this placebo test, we have evidence that seasonality is unlikely to explain any changes we observe in the 2024 data. This test assumes that, holding all else equal, seasonality in appointment behavior is the same each year.

We will also conduct several robustness checks as recommended in Hausman and Rapson (2018). First, we will show the robustness of the model to the polynomial order of the running variable.

We will explore a "local donut" RD by focusing on a narrow time window around robotext implementation date to examine the robustness of the finding: May 17-21 as the pre-period (business days of May 17, May 20, May 21), and May 24-29 (business days of May 24, May 28, and May 29) as the post-period. Within a very narrow time bandwidth around the event time cutoff, factors other than the robotext setup are likely very similar on either side of the cutoff. We exclude May 22 and 23 in case there is incomplete rollout of the intervention for the post-period dates specified due to unknown operational issues. However, this approach faces two constraints. First, because the two sides of the time cutoff will be on different days of the week, if there is a day-of-the-week pattern in no-show rate, that by itself may become a confounder. Second, the narrow bandwidth will reduce the sample size and the statistical power of our test. We will apply the augmented local linear methodology to increase power of the specification. This two-step procedure uses the full sample to identify covariates that are statistically significantly correlated with no-show outcome, then estimates the conditioned second stage on a smaller sample bandwidth controlling for these covariates.

Confirmatory analyses:

The confirmatory analyses will focus on the proportion of no-shows as the main outcome. Specifically, the outcome variable is whether an appointment is a no-show (i.e. not canceled or attended). The estimate for coefficient β from the statistical model above may be interpreted as the percentage points change in no-show rate as a result of robotexts. Dividing this estimate by the baseline no-show rate (pre-treatment no-show rate) will give us the rate of change.

Exploratory analysis:

We will examine three additional outcomes for the exploratory analysis. We will test whether the probability of attending appointments and the probability of cancellations change as a result of the robotexts. If the robotext reminders make people who will otherwise skip an appointment more likely to cancel the appointment, the probability of cancellations will increase as a result of decreased no-shows, but we should not expect any changes in the proportion of attended appointments. However, if the robotexts reminders cause some people to attend an appointment they might otherwise forget, we could observe an increase in the share of attended appointments.

Additionally, we will test the number of attended appointments each day. If the reduction in no-shows frees up appointments for others such as additional walk-ins, we would see an increase in the number of attended appointments. The unit of analysis would be the number of attended appointments at a local passport agency l:

$$N_{lt} = \lambda + \rho T_{lt} + \theta X_{lt} + \epsilon_{lt} \quad (2)$$

where N_{lt} is the number of attended appointments in natural logarithm function at the location on appointment date t and X_{lt} includes appointment location-by-day of week fixed effects and time trend in appointment date (a quadratic function of the date variable centered around the robotext initiation date). The variable T_{lt} represents the share of robotext-eligible appointments at each location on a specific date.

This location-day-level analysis will have a smaller sample size and thus lower statistical power as compared to the individual-level analysis. Therefore, we will expand the analysis period to four weeks before and after May 22.

In addition, if we have data on the share of robotext opt-ins by location day, we will be able to estimate a treatment-on-the-treated (TOT) effect.

Inference criteria, including any adjustments for multiple comparisons:

We will cluster the standard error at the passport agency level, as appointment availability/outcome is likely correlated within an agency. We will make inferences based on the p-values of the coefficient estimates for T_t using two-sided tests.

Limitations:

This has the limitation of being a quasi-experimental design as opposed to a randomized-controlled experiment. We plan to conduct sensitivity analyses to consider how sensitive our findings may be to different model specifications.

It is also possible that the findings are limited to the current cohort of passport applicants. Major events and changes in passport processing policies over time may mean that people applying for passports in 2024 are different from those who applied in previous years or who will need passport services in future years. For example, periods when there are long delays for standard passport processing services could change the makeup of people scheduling in-person appointments; results from this evaluation may not be directly applicable to a different population of appointment schedulers.

We will also be limited by the administrative data available. For example, we do not yet know if we will be able to observe whether appointment holders opted in to receive robotexts or whether those who opted in to those texts. It may be that information about opt-in and response behavior is only available at the aggregate level.

Appendix

Appendix A. Blinded analysis protocol

To minimize the risk that the analysis will be influenced by observations and initial findings from the data, we will follow a blinded analysis protocol:

1. A project team member will draft analysis code using 2023 data with a placebo robotext date of May 24, 2023. The analysis codes reflect the analysis plan detailed in this document. A null effect is expected using this data, as there was not a robotext initiative in 2023.
2. Once the analysis codes are finalized, they will be used to analyze the 2024 data to estimate the effect of robotext reminders. Significant changes to the codes may require an amendment to the analysis plan, especially with regard to the confirmatory analysis.