Analysis Plan

Project Name: Evaluation of San Diego Small Business Relief Fund
Project Code: 2003-SD
Date Finalized: 02/12/2021

Contents

In March 2020, the City of San Diego created the Small Business Relief Fund (SBRF) to help small businesses affected by the economic fallout from the COVID-19 pandemic. By December 2020, the City had disbursed nearly USD 17M in grants and loans. Yet demand for funds far exceeded supply: of the roughly 10,500 applications submitted between March 27 and April 14, funds could only be extended to 2,327 businesses. This project aims to estimate the effect of funding on business resilience using quasi-experimental methods to compare the outcomes of those applicants who were and were not funded. The current document lays out the methods and data we plan to use to conduct that analysis.

1 Project Description

Following the closing of the submission portal on April 14, SBRF loans and grants were disbursed in three phases, each deriving from different funding sources. At each phase, different eligibility criteria were applied, and eligible businesses' applications were processed by prioritizing those applications that were either submitted first or came from businesses located in low and moderate-income areas (LMA).[1] A rough outline of the process for grant and loans disbursement is represented on Figure 1 below. The broad timeline of the three phases is as follows.

Phase 1. The first phase of available funding totaled USD 6.2M comprising three sources: 1) USD 550K in City funding for grants of USD 10K (Former Enterprise Zone funding); 2) USD 2.2M in federal funds for low- or zero-interest loans up to USD 20K (Revolving Loan Fund); and 3) USD 3.4M in federal funds for forgivable loans (Community Development Block Grant funds). To be eligible, businesses needed to employ (and retain) between 1 and 100 Full-Time Equivalent (FTE) employee(s) (inclusive),[2] be located in the City of San Diego, have a City of San Diego business tax certificate that was valid as of April 15 (the day after the application period closed, have been operational for at least six months as of March 1, and not have any outstanding bankruptcies, tax liens or legal judgments. Lending institutions, insurance firms, nonprofits, chain stores, and home-based businesses were not eligible.

Phase 2. On April 13, the second phase of funding was made available in the form of USD 700K in private donations for grants up to USD 11K. These funds were distributed by a third party, San Diego Grantmakers. Eligibility was expanded to include businesses with zero FTEs (i.e., self-employed persons) as well as home-based businesses, though it was restricted to those firms with an annual revenue of USD 100K or less. Funding priority went to businesses located in underserved areas targeted for economic development (Opportunity Zone, Promise Zone, Qualified Census Tract, or Difficult Development Area). Business owner income and the financial hardship experienced by the business were also taken into account.

Phase 3. On June 9, the CARES act made available USD 13M in funding for grants up to USD 10.5K. These new funds not only increased the size of the overall funding to over USD 19M, but also expanded the eligibility criteria to include businesses with zero employees, those who had received other forms of funding (PPP, EIFL, SBA loans), and those with owners that live outside San Diego. The third phase introduced a tiered structure, whereby grants equaling USD 2,500–10,000 were provided based on the FTE and 2019 annual gross revenue of the business.[3]

1. Following U.S. Department of Housing and Urban Development (HUD) guidelines, low- and moderate-income areas, or LMAs, are census tract block groups where 51 percent or more of the residents are low- and moderate-income.

2. An eligible business needed to be retaining at least one FTE during the pandemic restrictions. An owner was not counted among FTEs.

3. Businesses with: FTE = 0 and 2019 gross revenue (GR19) < 200K received USD 2,500; FTE = 1, GR19 < 500K received USD 5K; FTE = 5 and GR19 < 1M received USD 7.5K; FTE = 6 and GR19 < 3M received USD 10,000. All

We are interested in the causal effect of receiving a grant or a loan on business-level outcomes such as the ability to generate revenue and remain open. To estimate those effects, we plan to take two approaches. The first, "Selection on Observables" (SOO), design defines treatment assignment as a business receiving an invitation from program examiners to submit documents for further review and treatment itself as receiving funding. We reweight the data by the estimated probability a business was invited to submit documentation and use instrumental variables regression in order to estimate the average treatment effect of being awarded funding among those who would be funded if invited. In this design, we exploit the fact that we can use the same observable information as program examiners in order to predict which businesses, among those determined eligible prior to document review, would be invited to submit documentation for funding. We restrict attention to phase three of the program, for which we have detailed data on the timing and decisions made during review.

The second, "Geographic Regression Discontinuity" (GRD), design seeks to estimate the average effect of being awarded funding local to the City of San Diego boundary, among firms that were either eligible or that would have been eligible if it were not for their location outside of the City boundaries. In this design, we exploit the fact that many businesses and people living in the County of San Diego or the San Diego-Carlsbad Metropolitan Statistical Area do not understand exactly where the City of San Diego boundaries lie, and applied to the SBRF despite being geographically ineligible (see Figure 2).[4] We use the entire population of applicant businesses across all phases for this analysis.

businesses in a Promise Zone, Opportunity Zone, and/or LMA received an additional USD 500 .

4. It should be noted that business owners faced such pandemic-related economic hardship and, in desparate need of funding, may have applied for any relief program, regardless of stated criteria. City of San Diego staff noted applicants for businesses headquartered outside of the City of San Diego, including those that were home-based, virtual/e-commerce or mobile, justified applying for SBRF as a "City of San Diego business" because their customer base and target market included the City of San Diego consumers.

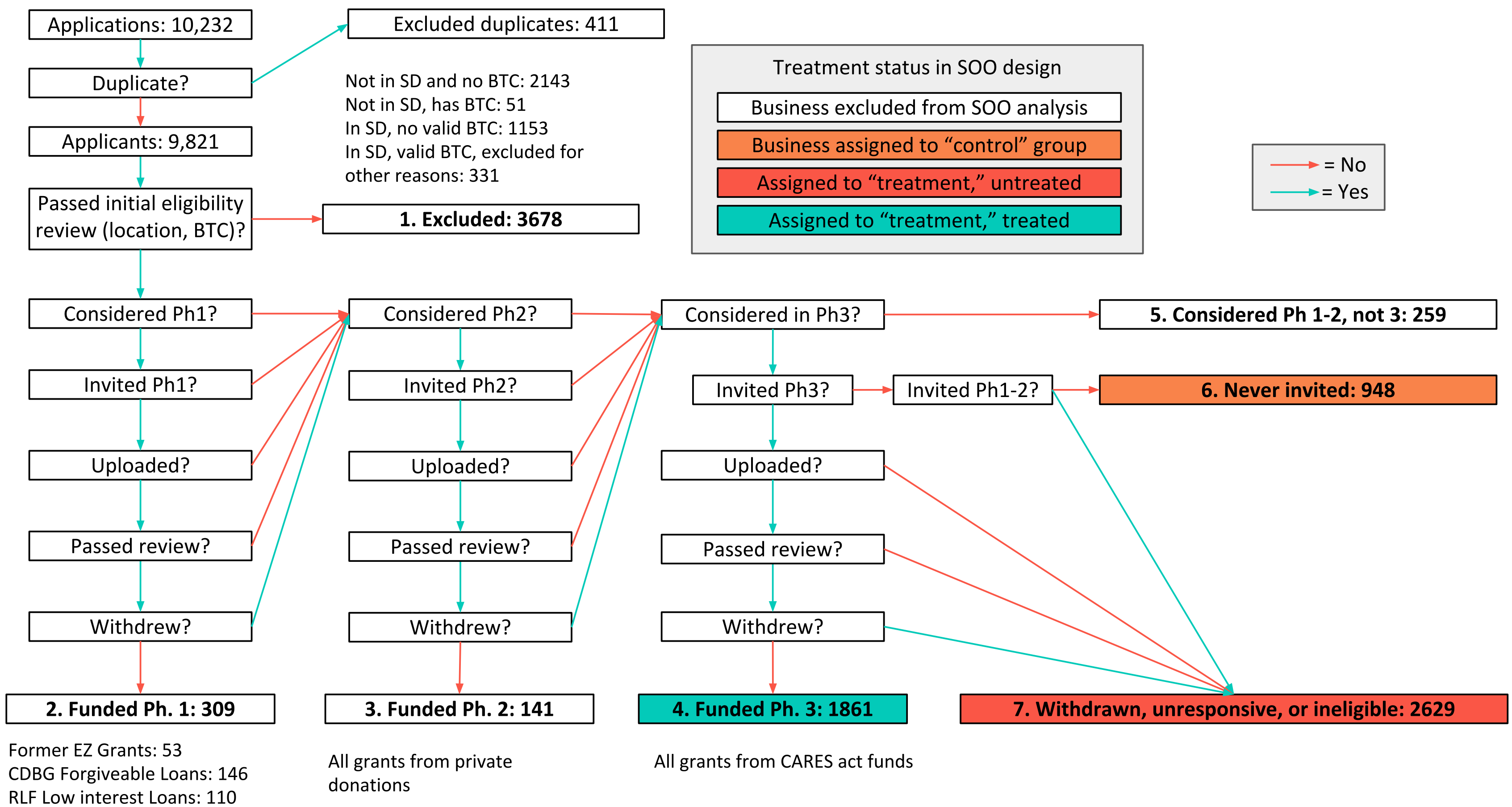

Figure 1: Process for reviewing and approving applications to the SBRF.

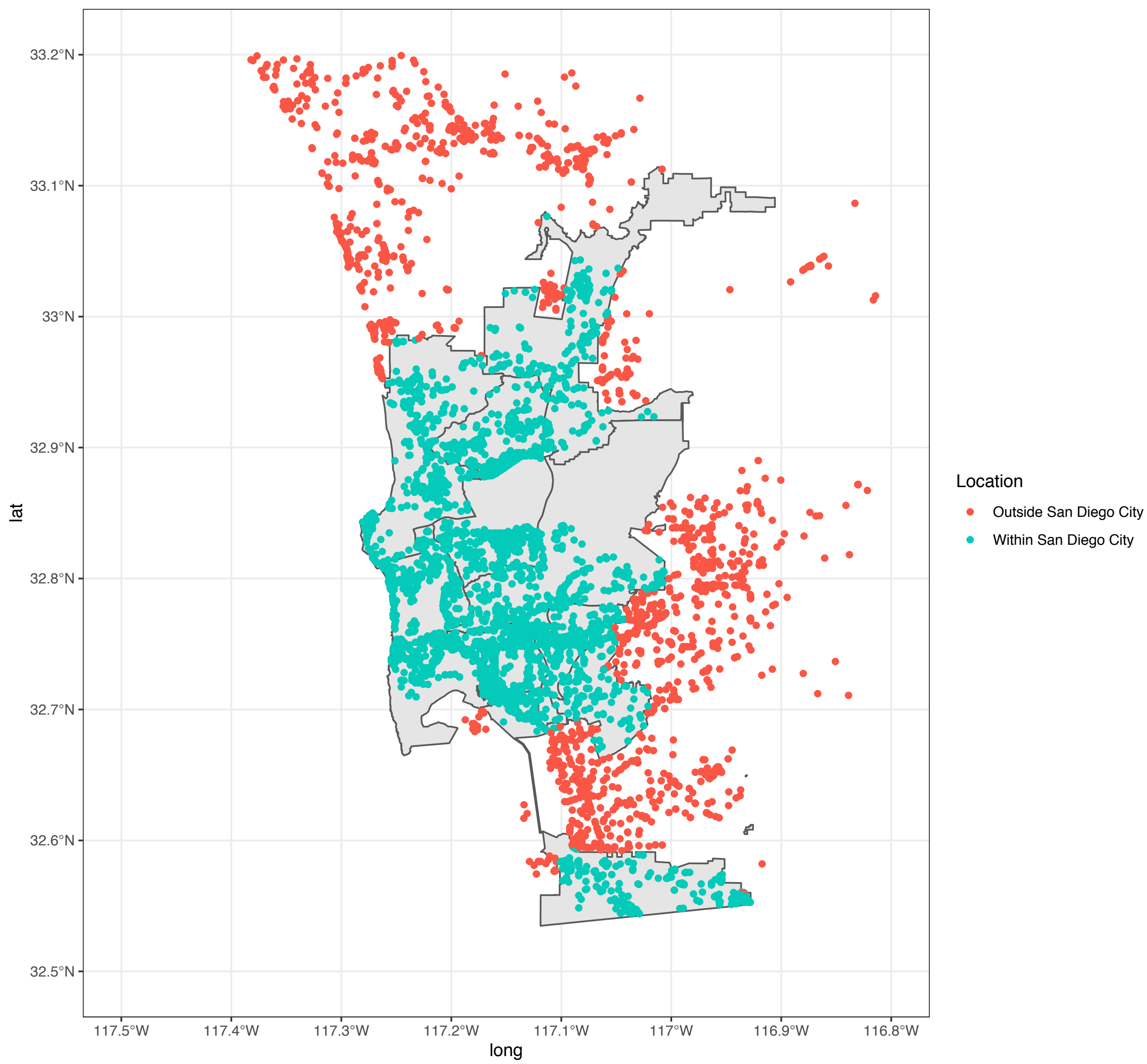

Figure 2: Map of businesses who applied for the SBRF. Green businesses are located within the City of San Diego boundaries, while red businesses are located just outside them. A similar geocoding exercise was conducted by program staff to determine business eligibility.

2 Data and Data Structure

In this section, we describe the different datasets used in the analysis. There are three broad categories: raw data on business-level outcomes from data providers such as Yelp, Facteus, and PACER; internal data on grants and loans applications from the Economic Development Department of the City of San Diego; and the panel and cross-sectional datasets that result from merging and restructuring the outcome and application datasets.

2.1 Datasets on business outcomes

Self-reported opening status or pivots to remote services

Yelp data on whether the business reports either a permanent or temporary closure helps us to measure businesses' attempts to stay open and generate revenue. The Yelp data contains daily information on whether the business has permanently or temporarily closed. The data also indicates whether the business has put up a "virtual services" banner, whether the business has partnered with Grubhub to enable delivery, and (possibly) opening hours. At time of writing, it is unclear whether we will succeed in obtaining the opening hours data due to technical difficulties in the way it is logged/overwritten.

Bankruptcy

Data on bankruptcy are downloaded from the Southern District of California U.S. Bankruptcy Court using PACER. The case report files contain a number of fields that describe the disposition of the case as well as identifiers for the parties. In total, the data contain 199 bankruptcies due to business debt and 4774 bankruptcies due to consumer debt from January 1 to November 30, 2020. These will include Chapter 7, Chapter 11, and Chapter 13 consumer and business filings. A recent working paper shows these follow different trends (Wang et al. 2020).

Credit risk

Cortera holds a credit database that spans businesses of all types across the United States. We are considering acquiring this data but it has not been obtained at date of writing and it is possible we may not include it in the analysis given the limited temporal variation. Cortera matches data based on business name and address, and was able to achieve an 87% match rate on a sample dataset of 200 businesses. Data are generated at the month level. The dataset includes a credit score, rating of credit health, and predicted likelihood of severe delinquency over the next 12 months. In the sample data, credit score information was matched for 76% of businesses.

Additional variables needed for analysis

We will merge in an indicator for whether the business received a grant or loan from the county over the period under analysis using public datasets of business names who received county funding. This data will primarily be used to assess whether there is differential receipt of county-level grants by treatment status.

Credit card transactions data

From FACTEUS we received anonymized data on individual consumer credit card transactions within San Diego and surrounding suburban ZIP codes. This data came from FACTEUS's U.S. Consumer Card Payments and Gamma Payments databases. Transactions cover February 1, 2020 to July 31, 2020. The U.S. Consumer Card Payments data consists of 6,593,636 individual credit card transactions, while the Gamma Payments data consists of 1,041,097. The data lists the transaction amount and timestamp, as well as business identifiers and the consumer's date of birth and zip code.

We had originally intended to use this data in order to track the transaction performance of businesses across time. Unfortunately, however, because the data focuses on the behavior of consumers, most businesses that do show up in the dataset show up only once. We describe an exploratory analysis using this data below instead.

2.2 Internal application and review data

We have several sets of data reflecting the material businesses used to apply to the SBRF, as well as decisions made by program staff. We have cleaned and transformed these into an "application dataset" that we use to construct treatment variables and to estimate propensities for the SOO design. The specific list of variables used in that analysis is described in section B of the internal (non-public) appendix. These include the number of employees, location in an underserved area, self-reported revenue losses due to COVID, and the eligibility criteria advertised to applicants described in the introduction.

2.3 Merging application and outcome data

To construct the analytic datasets described in the next section, it is necessary to merge the application dataset with the outcome datasets. The merge process is not complete at time of writing but should work roughly as follows:

1. **Yelp data:** we use the Yelp business search API along with the business's name (cleaned and truncated to correspond with Yelp naming conventions) and its address. The API, for businesses that match, returns a unique business identifier. The Yelp data includes businesses matched through this process and businesses Yelp is able to fuzzy match (0 additional businesses in the current analysis).

2. **PACER data:** we use fuzzy and exact matching to match businesses with cases in the PACER data using business EIN, name, and address.

2.4 Transformations of data structure

We will transform the data into a panel and a cross-sectional dataset. The panel data will be used in the SOO design, while the cross-sectional dataset will be used in the GRD design.

Construction of panel dataset

We will construct a business-week panel. The panel should run from January 1 2020 to December 31 2020, and cover roughly 10,000 businesses.[5]

The dataset will include the following primary outcomes:

- `bankruptcy`: A binary indicator that is 1 if the business or business owner filed for any bankruptcy chapter, commercial or consumer, on any day that week or prior, and 0 otherwise. This is recorded in the PACER data.

- `virtual_services`: A binary indicator that takes the value 1 if the business offered any "virtual services" according to Yelp any day of that week, and 0 otherwise. This includes both enabling the virtual services banner, indicating remote services, and delivery through Grubhub.

- `days_closed`: The number of days during that week for which the business had a "temporarily closed" or "permanently closed" flag on its Yelp page.

We will also include the following treatment variables and covariates:

- `invitation_wave`: Categorical variable indicating date at which that business was invited to submit funding. For those businesses never invited, date takes the value "never."

- `invited`: A binary indicator that is 1 if the business has been invited to submit documents for review by that week, 0 otherwise.

5. In practice, the exact timespan of the panel may vary between outcomes depending on the time span we are able to secure. At the time of plan posting, we do not know the full range of invitation dates. We may adjust the end date of the panel period to make sure that it encompasses a long-enough "post" period following the invitation.

- `weeks_since_invitation`: A categorical variable used for fixed effects. For businesses ever invited to submit documents, it takes negative values corresponding to the number of weeks prior to the invitation, 0 on the week of, and positive numbers following that week. For businesses never invited, it takes the value "never," which is the reference category.
- `funded`: Takes the value 1 whenever a) that business was actually funded and b) invited is equal to 1. 0 otherwise.
- `week`: A categorical variable indicating the week.
- `business`: A categorical variable serving as a unique business identifier.

To these variables, we will add additional covariates and the inverse propensity weights constructed using methods described below.

Construction of cross-sectional dataset

We create a cross-sectional dataset by collapsing the panel into pre- and post-treatment periods. For the main analyses, we define the post-treatment period as the first date at which a business was notified of funding approval. In extensions, we will illustrate how GRD estimates change when later and later subsets of the data are used for estimating outcomes, since many businesses were not funded until the early fall.

The data will contain pre- and post-treatment averages of the primary outcomes described above (`days_active`, `revenue`, `bankruptcy`, `virtual_services`, and `days_closed`).

We will also include the following treatment variables and covariates:

- `invitation_wave`: Date at which that business was invited to submit funding. For those businesses never invited, a date arbitrarily far in the future.
- `invited`: A binary indicator that is 1 if the business was invited to submit documents for review, 0 otherwise.
- `funded`: A binary indicator that is 1 if the business was funded, 0 otherwise.
- `business`: A categorical variable serving as a unique business identifier
- `latitude` and `longitude`: Geocoordinates of the business used to define distance to the boundary of the City of San Diego. Coded using street address and the `tidygeocoder` and `ggmap` packages for `R`.
- `in_SD`: A binary indicator that takes the value 1 if the business is inside the San Diego boundary and 0 otherwise. Boundaries are defined using the same council geography shapefiles used by program staff.

2.5 Data exclusion

Duplicates

Among the applications, 756 stem from businesses that appear to share the same name. A majority owner or business was limited to one SBRF award. If the majority owner already received an award in Phase 1, all other applications from that owner were deemed ineligible. The presence of multiple applications for one business poses some issues for the analysis. From a causal inference perspective, it is tempting to simply view these as taking two chances at getting funded rather than one. In this case, we could treat businesses as clusters of applications, and adjust the propensities accordingly. However, this approach is complicated by the fact that reviewers filtered out many, though not all, of these duplicates. As such, we have adopted for a more nuanced approach that filters applications down to one-per-business depending on which of the following kinds of duplicates to which the application belongs:

1. N = 396/756. The business applied from the same location multiple times and only one made it through to consideration in phase 3 (vast majority), or the same business applied from different locations and program reviewers caught it and only let one through. We retain the application that made it through in either case, since in such cases reviewers caught the duplicate and did not pass it through to phase 3.

2. N = 237/756. The business applied from the same or different locations multiple times and none made it through to consideration in phase 3. These applications will not be considered in the main analysis. We may include them in any analyses using Phase 1 and Phase 2 data however. We keep the most recent funded application or, if no application was funded, the most recent application.

3. N = 45/756. The business applied from the same or multiple locations and had two different but noncontradictory treatment assignments in phase 3. For example, they were marked as "unresponsive" and "withdrawn" for two different submittal numbers. In these cases, we retain the most recent funded application or, if no application was funded, the most recent application.

4. N = 35/756. The business applied from the same location multiple times and the statuses contradict for our purposes, in that they have the control status ("assigned but not engaged") but were also invited. We found a handful of slightly more difficult cases in which the addresses were different, the name was the same, and one was invited and the other in control. This is difficult because, in theory, these might be considered different businesses entirely. For the sake of consistency, however, we treat businesses with different locations as one business. In all such cases, the key point is that the business was assigned if it was ever invited, so we keep the invited app. This does pose the issue of heterogeneous assignment probabilities, which can lead to bias. But the number is small enough as not to provoke concern.

5. N = 41/756. Some businesses were awarded in a previous phase but were marked as "unresponsive." in phase 3. In these cases, we mark the business as assigned to treatment and keep them in the analysis. Helpfully, no business is in the control condition for phase 3 but funded in an earlier phase.

6. N = 2/756. False duplicates, in which businesses have the same name but are from different businesses: two different taxi drivers fall into this category. We give them different names and keep both.

This process results in the removal of 411 duplicate applications from the data.

Restriction to analytic sample

In approach 1 (SOO), we consider only businesses that passed the initial eligibility screening for phase 3 funding. For this phase, we have detailed information on application review and invitation timing that is required for the analysis and unavailable for other phases. See Figure 1 for more details.

In approach 2 (GRD), we consider all funding phases but will exclude any data that obviously does not meet eligibility requirements *except for* the requirement that the business be located inside the City of San Diego. The aim with this subsetting is to produce a sample whose businesses might have been eligible for funding had it not been for their spatial location, and thereby to reduce the number of “never-takers” in the sample (see below for formal definition).

2.6 Treatment of missing data

We distinguish between two types of missingness, each of which requires a slightly different approach. First, there is the issue of data that exists in principle but, in practice, we are unable to find a match. For example, whether or not a business had transactions occur on a given date in principle exists for all businesses, including those that closed. However, a business may not show up in our data. We call this “attrition.” Second, there is the issue of data that are missing for the more fundamental reason that it is observed conditional on post-treatment outcomes. For example, the average amount of revenue on a given day will be undefined for businesses that ceased to exist. We cannot define average treatment effect estimands for such outcomes for those businesses. We call this “post-treatment missingness.” Our approach to this issue is to define our outcomes such that they do not depend on post-treatment potential outcomes.

As concerns attrition, we know already that we will lack data on the outcomes of some substantial proportion of businesses: at least 30% of the businesses appear to be unmatched to the Yelp data, for example. It is conceivable that some businesses fail to match in ways that are correlated with treatment. For example, if those who applied earliest put the least time into their applications and therefore introduced more misspellings or messier address data that made it more difficult to match them based on these fields. Similarly, if brick and mortar businesses were more likely to apply earlier and also more likely to have Yelp accounts where they post hours, attrition might be correlated with treatment status.

We describe our approach to this issue in Section A.1 of the technical appendix. Briefly, for each design, we run a test for differential attrition for any outcome exhibiting missingness. If there is evidence of differential attrition, we report the results of four robustness checks.

3 Statistical Models & Hypothesis Tests

3.1 Analysis 1: Selection on observables

The selection on observables (SOO) approach leverages the fact that, for phase 3 of the SBRF, we have the same applicant data used by case reviewers to make initial determinations about whether to award businesses. In principle, the aim is to reconstruct, using our qualitative understanding of the award process gleaned from conversations with program staff and machine learning techniques, each eligible business’s probability of being invited to submit documents for funding, both for businesses that were and were not invited. These estimated probabilities can be used to construct weights that allow us to treat the comparison of invited and non-invited businesses as though it were an experimental comparison.

We distinguish between an encouragement to submit further documents for funding—the business’s “assignment” or “encouragement” to be treated—and actually receiving funds—the business’s treatment “status.” Specifically, let Z denote a binary variable indicating program staff invited a business to submit documents for funding and T a binary variable indicating the business was actually awarded funds. The key distinction between these two variables is that we assume we can predict the distribution of Z using the observable variables available to program staff, whereas T depends on features of the world, U, beyond our ability to measure. We formalize the identification strategy in section A.2 of the technical appendix.

The control group is composed of those who were never invited to submit documents for review—

they fall into category 6 on Figure 1. That is, businesses in category 6 have $Z_i = 0$ and $T_i = 0$, and are a mix of compliers and never-takers. In the treatment group, we can observe compliers in category 4—they are businesses for whom $Z_i = 1$ and $T_i = 1$. Never-takers in the treatment group are businesses who were invited but never received funding. These can be found in category 7—they have $Z_i = 1$ and $T_i = 0$. Other businesses are excluded from the SOO analysis.

We suppose that there is some probability that business i is invited to submit documents for review, $\Pr(Z_i = 1) = \pi_i$, which we do not observe but can estimate using statistical models to get propensity scores, $\hat{\pi}_i$, because we have access to the same data used to make invitation decisions. If $\hat{\pi}_i$ is a good estimator of π_i, then we are able to estimate the average intent-to-treat effect by subtracting the weighted average of uninvited businesses' post-treatment outcomes from the weighted average of invited businesses' post-treatment outcomes, where `invited`[6] defines treatment and weights are proportional to $\frac{1}{\hat{\pi}_i}$ for invited businesses and $\frac{1}{1-\hat{\pi}_i}$ for uninvited businesses.

The intent-to-treat (ITT) effect is defined as the average difference between what would have happened in a world where all businesses were invited to submit documents (with some subsequently funded and unfunded) and a counterfactual world in which they were not.[7] Of course, a key question is not only what would happen if we could invite all businesses for funding, but also what average change would occur specifically for the businesses who, when invited, would actually submit documents, pass review, and be awarded funding. We refer to this in the formalization as the complier average causal effect (CACE), and describe an additional set of assumptions necessary to estimate it. Essentially, if an invitation to submit documents cannot decrease the likelihood a business is funded (monotonicity assumption), and if the only way that invitations affect outcomes like revenue is through the actual funding that follows it (conditional excludability assumption), then we can estimate the CACE by rescaling the ITT to account for the fact that many businesses invited to submit documents for funding were never funded in fact.

Estimation strategy

Our identification strategy thus depends on estimating π_i accurately using propensity scores, $\hat{\pi}_i$. Broadly speaking, there are two families of approaches to propensity score estimation for causal inference. The first seeks to reconstruct the treatment assignment process, while the second places greater emphasis on producing balance on covariates. We take the first approach.

Many early studies that used estimated propensity scores to weight observational data employed logistic regression: in a 2004 review of medical papers, for example, all 48 of the reviewed manuscripts relied on some form of logistic regression (Weitzen et al. 2004). However, more recent studies have illustrated substantial bias that can result from misspecification of the functional form: logistic regression, for example, relies on the assumption of linearity in the logit and accurate specification of any higher order terms.

More recently, machine learning techniques such as neural networks, support vector machines, and decision trees and forests have been employed to flexibly estimate propensity scores without imposing such strong assumptions about the parametric distribution or functional form of the propensity.[8]

6. See Section 2.4.

7. As described in the formalization, the definition of this estimand relies on assumptions about the number of potential outcomes businesses reveal as a result of the invitation decisions. In particular, we must assume the absence of spillovers between business outcomes and funding statuses.

8. In particular, see Diamond and Sekhon 2012 for a discussion of the relative performance of a propensity scored based on a generalized linear model and ones estimated using more flexible methods as we relax assumptions that the features predict treatment status in a linear and additive way. In the present case, for instance, we know that submission time may interact with whether the business is located in a prioritized area.

While the different machine-learning methods typically out-perform logistic regression in terms of bias, their relative performance is context-specific. Pirracchio et al. 2014 argue that meta-learning techniques that combine the predictions of different algorithms perform at least as well as the best choice among candidates. As such, they suggest using a combination of machine learning algorithms to predict propensities, and then applying a meta-learning algorithm to those predictions in order to arrive at the final prediction. We use ensemble stacking to do so.

Our propensity model uses transformations of the raw data used by program staff to make determinations about whether to invite businesses to submit documents for funding, as depicted on Figure 1. See Section 2.2 above and Section B of the internal appendix for a description of the variables employed in this analysis.

We fit the propensity score model using seven different machine-learning methods:

1. Logistic regression (from `glm` package)
2. Random forests (from `randomForest` package)
3. Support vector machines with a linear kernel (from `e1071` package)
4. Stochastic gradient boosting (from `gbm` package)
5. Neural net (from `nnet` package)
6. Decision tree (from `rpart` package)

The tuning and training of each method uses 5-fold cross validation on 80% of the data selected at random for training, with accuracy as the performance metric. We tune over five hyperparameters taken from the defaults in each software package. Finally, we take the predictions from these methods and use them to create a stacked ensemble.

We ran the prediction models on the phase three review data as of September 17, 2020. Since the review process is not yet finalized, this provides a preliminary view of how well the propensity model performs. We label businesses that either had a document folder URL or had a status that indicated they were sent an email as "invited" (including those who were found ineligible after document review) and any business that was flagged as yet to receive an email as "not invited." We excluded businesses found to be ineligible prior to any emails being sent.

Of all the models, the stacked ensemble performed the best. When trained on 80% of the data, it correctly predicted 95% of the uninvited businesses and 95% of the uninvited businesses in the 20% of the data held out for testing. Expressed in terms of accuracy, this works out to $(172+664)/(172+664+9+33) = 95\%$. We will estimate propensities using the model with the best accuracy score on the final dataset.

To estimate the ITT and CACE as defined above, we will run weighted linear regressions with robust standard errors clustered at the business-level. Weights are constructed using the estimated propensity scores as above. The p-values constructed from the standard errors from regression models will constitute our main test of the null hypothesis of no average effect. We will make no adjustments for multiple comparisons.

To estimate the ITT averaged over all periods, we run a regression of the following form using `estimatr` for `R`:

```
lm_robust(
  formula = outcome ~ invited,
```

```
  fixed_effects = ~ week + business,
  se_type = "stata",
  clusters = business,
  weights = ipw,
  data = df
  )
```

The coefficient on `invited` thus identifies the ITT by estimating the inverse propensity-weighted average two-period difference-in-difference for every business in the sample.

To estimate the CACE, we run the following weighted instrumental variables regression:

```
iv_robust(
  formula = outcome ~ funded | invited,
  fixed_effects = ~ week + business,
  se_type = "stata",
  clusters = business,
  weights = ipw,
  data = df
  )
```

The coefficient on `funded` thus identifies the CACE by estimating a two-stage least squares regression.

Extensions

Since businesses were funded at different times, averaging treatment effect estimates across all periods likely attenuates our estimate of some of the shorter-term impacts for funded businesses. To address this attenuation issue, we take two approaches to estimating dynamic ITT effects and CACEs.

The first simply estimates the ITT at each post-treatment week:

```
lm_robust(
  formula = outcome ~ weeks_since_invitation,
  fixed_effects = ~ week + business,
  se_type = "stata",
  clusters = business,
  weights = ipw,
  data = df
  )
```

The second estimates the subgroup ITT for businesses invited to submit documents in the first week of April:

```
lm_robust(
  formula = outcome ~ invited,
  fixed_effects = ~ week + business,
  subset = invitation_wave == "April-1" | invitation_wave == "never",
  se_type = "stata",
  clusters = business,
  weights = ipw,
  data = df
```

)

We will run equivalent two-stage least squares regressions for period- and subgroup-specific CACEs.

Finally, we plan to run subgroup analyses to assess any differences in impact for low-moderate income business owners.

Robustness

As a robustness check on the main model, we will run a "doubly-robust" estimator in which we include the covariates used to estimate propensities in the models listed above. We will also report a difference-in-differences model that uses T in the place of Z in the TWFE specification.

We will report how estimates change when employing different approaches other than the main one used in order to estimate propensities. Finally, we will check for robustness to different approaches to sample inclusion, such as including businesses that were reviewed in Phase 1 (this will involve imputing to which businesses invitations were sent and the timing with which this was done, as we do not have this information).

3.2 Analysis 2: Geographic regression discontinuity

In a typical, one-dimensional regression discontinuity, researchers exploit the fact that the probability of being assigned to some treatment jumps discontinuously at some point along an underlying covariate. The discontinuous point is usually called the "cutoff" and the underlying covariate is usually called the "forcing variable." In "sharp" RD designs, the probability of treatment jumps from 0 to 1, whereas in fuzzy RD designs there may be a mix of treated and untreated units on either side of the discontinuity. Very often the running variable is correlated with outcomes, so that treatment assignment and outcomes are confounded. However, provided the outcomes are a continuous function of the forcing variable, it is still possible to estimate an unbiased estimate of the average treatment effect very close to the discontinuous cutoff.

Geographic regression discontinuity (GRD) designs extend the one-dimensional RD design into a two-dimensional space. While it is tempting to simply compute the distance of each unit to the nearest point on the border, and run a single, one-dimensional RD with the distance signed positive for treatment units and negative for control units, this one-dimensional approach is prone to spatial confounding that can produce bias. Figure 3, reproduced from Rischard et al. 2018, illustrates the issue with this approach.

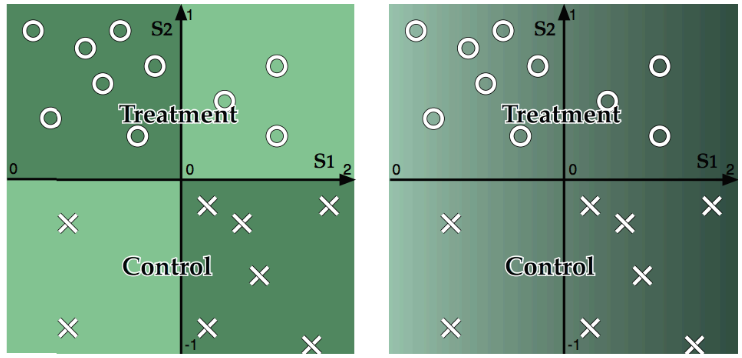

Figure 3: Spatial confounding example from Rischard et al. 2018. The boundary is located at $S_2 = 0$. The left panel background shows the density of units, while the right panel shows the spatial trend in outcomes, moving from high to low as one progresses from left to right. The average difference in outcomes at any point on the border is zero. However, if one were to run a one-dimensional RD with each point given a signed distance to the nearest point on the border as the forcing variable, one would erroneously estimate a large treatment effect. This bias occurs because treatment units happen to spatially cluster in an area with high outcomes whereas control units are clustered in an area with low outcomes. This problem can be overcome by estimating a two-dimensional geographic regression discontinuity, in which the estimand is the average difference in outcomes along the border.

A two-dimensional GRD design overcomes the spatial confounding issue explained in the caption to Figure 3 by estimating the causal effect along the length of the boundary separating treatment from control units.

There are roughly three such approaches in the literature: 1) Keele and Titiunik 2015 apply a one-dimensional RD to a grid of points along the border then average across the resultant estimates; 2) Keele et al. 2015 use matching on distance and covariates within a buffer distance to the border, then analyze matched units as though they were in an RCT; 3) Rischard et al. 2018 fit outcomes to a smooth surface, extrapolate to the border curve, then take pointwise difference between the two extrapolations to estimate the treatment effect along the border.

We will use a modified version of the first approach. We eschew the second approach as the matching often requires discarding units that cannot be matched, which makes the estimand difficult to define (we may include this as an exploratory analysis). Approach 3 is attractive but we forego using this method as the estimator encounters difficulties when covariates are included due to the large parameter space that is created (particularly with the covariance matrix). While one could conceivably residualize outcomes, this can lead to anti-conservative variance estimation (as the uncertainty in computing residuals is typically hard to propagate through into the final model). Given the rich set of covariates available, in particular pre-treatment outcome data, we are reluctant to forego the use of covariates.

We extend the first approach in two ways. First, we adopt a fuzzy approach to the discontinuity, in which all units on the treatment (control) side of the border are assigned to treatment (control), yet the proportion actually treated (untreated) is not 1. This allows for the definition of intent-to-treat and complier border causal effects. Second, we include covariates in the estimation strategy following Calonico et al. 2019.

Estimation strategy

The estimands and identification strategy used to estimate them are formalized in section A.3 of the technical appendix. Briefly stated, our estimation strategy involves extrapolating potential outcomes to a grid of points along the boundary between the inside and outside of San Diego.[9] At each point, we estimate the ITT and the CACE, then we take the population density weighted sum of the estimates to estimate a scalar summary of the average ITT and CACE.

The map of units' locations is provided on Figure 2. Any units that do not meet the eligibility criteria (with the exception of spatial eligibility) will be removed from the sample, so that the estimation is performed among units with prima facie eligibility. We will define a grid of R evenly spaced points along the border. We will set R to a large number, such as 1000. At each point, we will use the `rdrobust()` function from the `rdrobust` package to perform local polynomial regression with a triangular kernel. Specifically, we will estimate effects at each point in two different ways:

1. "Sharp" RDD of outcome on being "inside San Diego," using pre-treatment outcomes as covariates;
2. "Fuzzy" RDD of outcome on being funded, using "inside San Diego" as running variable instrument and pre-treatment outcomes as covariates;

We will calculate weighted averages of the R estimates obtained for each of the four estimation methods, and bootstrap resample observations in order to estimate standard errors (e.g., using the standard deviation of the bootstrap resampled distribution).

Extensions

A concern in both the SOO and GRD designs is that, because businesses were funded at different times, averaging over all periods may attenuate estimates of funding effects. To address this issue, we will also show how the estimates vary as the cross-sectional dataset is constructed holding the pre-treatment date fixed and using ever later post-treatment dates, excluding the data in-between.

As mentioned above, we may also try the approach suggested in Keele et al. 2015, in which we would match funded units on one side of the border to unfunded units on the other side, within a buffer distance to the border, then analyze matched units as though they were in an RCT. This may overcome issues that the main analysis faces with the density of points varying along the border.

Robustness

We will also estimate the main results without covariates.

3.3 Exploratory analyses

Analysis with FACTEUS data

As an additional exploratory analysis, we leverage consumer credit card transaction data from FACTEUS to assess the relationship between invitation status and small business outcomes. Due to limitations in the way consumer spending data is collected by FACTEUS, the credit card transaction data we obtained does not provide sufficient coverage to perform time series analysis at the level of individual businesses. We instead aggregate to broad group levels using the `invitation_wave` variable defined above, which indicates whether and when a business received an invitation.

9. Businesses located in the City of Chula Vista were eligible for RLF funds. There were 551 applications from businesses in the City of Chula Vista, and 8 were awarded. We still count these businesses as outside the City of San Diego, and treat this an issue of two-sided noncompliance.

We used the R package `fastLink` to link consumer credit card transactions in the US Consumer Spending and Gamma datasets provided by FACTEUS to small businesses that applied for funds in San Diego. Businesses may fail to appear in the dataset because they: 1) were not visited by a consumer covered by Facteus' data collection methods; 2) do not ever transact using credit cards; 3) closed. Thus, appearance in the dataset is itself an outcome. We therefore measure the number of weekly transactions by applicant businesses, including a 0 for those not in the dataset. From this business-week-level measure, we construct two weekly measures aggregated to the `invitation_wave` level using the propensity weights described above, so that we have two separate time-series measures for the businesses that were and were not ever invited for funding. The first is `n_transactions`, the count of all transactions in that week and invitation wave; and `any_transactions`, the proportion of businesses for whom we record at least one transaction in a given week in a given invitation wave.

To assess the effect of funding on these outcomes, we will conduct a difference-in-differences analysis, in which the outcome is regressed on an indicator for whether the group is treated by that week, as well as group and period specific indicators.

Heterogeneous effects of lockdown

It seems possible that receiving funding may have lessened the shock of lockdowns or enhanced recovery during the period from mid May to the end of June when businesses were allowed to reopen for in-person service. As such, we intend to assess heterogeneity in effects by the a variable recording whether a severe restriction in operations was applied during that week.

Between Feb. 1 and July 31, 2020, there are four general periods that are captured by this `lockdown` variable. These are: (1) a "pre-lockdown" period from Feb. 1 to March 18; (2) a "1st lockdown" period that runs from March 19, when an initial wave of lockdowns were imposed, to May 4; (3) an "openings" period that began on May 5 and extended until June 31; (4) a "2nd lockdown" period that began with a new wave of targeted lockdowns on specific counties, including San Diego.

Cost effectiveness

In an additional exploratory analysis, we may estimate the cost effectiveness of the loan and grant programs that comprise the treatment. This will require identifying the funding source for each loan/grant. Working with the administrative offices of each program, we will identify the systems holding budget and expenditure data used to deliver these awards. A data request to capture dispersed amounts as well as administrative costs of the program will be generated per agency budget holder. When possible, interviews with program managers will be held to collect qualitative information about personnel time spent to review and monitor the grant/loan programs.

Results from the cost-of-program-delivery analysis will enable another policy relevant cost analysis; estimating the overall budgetary impact these programs have on government. To understand the macro-budgetary impacts, it is necessary to measure avoided costs to government (costs not-incurred from usage of government programs for unemployed individuals) as well as the impact on future government revenues (sales other taxes) collected from these businesses.

Firstly, an analysis may measure differences between treatment and control groups in usage of government programs. This will require identification and tracking of outcomes for individual employees rather than for the business as a whole.[10] A key outcome to track will be unemployment benefits

10. This analysis is pre-registered as exploratory because we currently only have identifiers for the (1) business and (2) business owner, and do not have comprehensive identifiers for the business' employees at the time of application or award.

utilized, which we hypothesize will be less for the treatment than the control group. Pulling from administrative databases, usage (and cost) of welfare programs, including SNAP, Medicaid/CHIP, TANF, and state/local programs such as emergency utility funds, will be analyzed. Finally, if the state can provide access to tax records, this cost analysis will estimate differences in sales tax paid by businesses over a 3 year window. Thus, to capture the macro-budgetary impacts of such grants/loans, an estimate of net budgetary cost to the government will be completed several years after grant release.

Counterfactual business outcomes under different disbursement processes

In addition to the two sets of inferential analyses, which estimate the causal impact of invitations to and actual receipt of funding, we will conduct an exploratory analysis to be outlined further in a separate pre-analysis plan that investigates hypothetical outcomes for SBRF applicants if the program had been implemented differently. Here, we briefly preview.

The OES team recently released a report, "Increasing Access to Small Business Grant and Loan Programs for Historically Underserved Groups", that documents a variety of methods local agencies used to target funds to underserved groups like minority-owned businesses, women-owned businesses, and businesses located in economically-deprived areas. The SBRF program took one approach—geocoding business locations and moving businesses more quickly through the first-come first-served process if they were located in a high priority area (e.g., LMA)—while other approaches included a points system that gave higher priority for any funding to businesses with certain characteristics and lotteries with separate pools for businesses in and outside of underserved areas.

Building on that report, we will analyze *hypothetical* outcomes for the SBRF application pool under three example models:

1. **Strict first-come first-served:** if SBRF had not moved businesses up in the queue, what would funding decisions for underserved groups look like?
2. **Points system:** if SBRF had given points to businesses for certain characteristics (tenure; revenue loss; LMA area), and then chosen businesses with the highest points values, what would funding decisions for underserved groups look like?
3. **Lottery:** if SBRF had put all applicants in a lottery, what would funding decisions for underserved groups look like?

References

Calonico, Sebastian, Matias D Cattaneo, Max H Farrell, and Rocio Titiunik. 2019. "Regression discontinuity designs using covariates." *Review of Economics and Statistics* 101 (3): 442–451.

Diamond, Alexis, and Jasjeet S. Sekhon. 2012. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *The Review of Economics and Statistics* 95, no. 3 (October): 932–945. ISSN: 0034-6535, accessed December 7, 2020. https://doi.org/10.1162/REST_a_00318. https://doi.org/10.1162/REST_a_00318.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71 (4): 1161–1189.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142 (2): 615–635.

Keele, Luke, Rocio Titiunik, and José R Zubizarreta. 2015. "Enhancing a geographic regression discontinuity design through matching to estimate the effect of ballot initiatives on voter turnout." *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 223–239.

Keele, Luke J., and Rocio Titiunik. 2015. "Geographic boundaries as regression discontinuities." *Political Analysis* 23 (1): 127–155.

Lee, David S. 2009. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76 (3): 1071–1102.

Manski, Charles F. 1990. "Nonparametric Bounds on Treatment Effects." *American Economic Review: Papers and Proceedings* 80:319–323.

Pirracchio, Romain, Maya L. Petersen, and Mark van der Laan. 2014. "Improving Propensity Score Estimators' Robustness to Model Misspecification Using Super Learner." *American Journal of Epidemiology* 181, no. 2 (December): 108–119.

Rischard, Maxime, Zach Bransona, Luke Miratrix, and Luke Bornnc. 2018. "A Bayesian Nonparametric Approach to Geographic Regression Discontinuity Designs: Do School Districts Affect NYC House Prices?" Working paper. https://arxiv.org/pdf/1807.04516.pdf.

Rosenbaum, Paul R, and Donald B Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70 (1): 41–55.

Wang, Jialan, Jeyul Yang, Benjamin Charles Iverson, and Raymond Kluender. 2020. "Bankruptcy and the COVID-19 Crisis." *Available at SSRN 3690398.*

Weitzen, Sherry, Kate L Lapane, Alicia Y Toledano, Anne L Hume, and Vincent Mor. 2004. "Principles for modeling propensity scores in medical research: a systematic literature review." *Pharmacoepidemiology and drug safety* 13 (12): 841–853.

Appendix

A Technical Appendix

A.1 Approach to attrition

We will run the following tests for any outcome that exhibits missingness:

1. **Selection on Observables**: We will conduct an F-test between two linear models, both of which will be run using the same analytic sample and inverse propensity weighting scheme as the main analyses, defined below. The first will regress an indicator for missingness on the available pre-treatment covariates used in the propensity prediction model. The second will supplement the first with a treatment term, interacted with those covariates. The F-test thus tests the null hypothesis that there is no *differential* attrition between the treatment and control group.

2. **Geographic Regression Discontinuity**: We will run the same analysis using the same bootstrap approach to standard errors as that used in the main analyses for the GRD design, except that the outcome will be an indicator for missingness. The p-value from this regression tests the hypothesis that there is no discontinuity in the probability of attrition along the boundary.

When a test is statistically significant at the $\alpha \leq .05$ level, we will report the following additional robustness analyses:

1. **Poststratification**: For both SOO and GRD, we use the same machine learning procedures used to predict propensity of assignment to treatment in the SOO design to predict probability of not attriting. Observed units are then weighted by the inverse of this propensity. This results in the upweighting of units whose outcomes we *are* able to observe but whose characteristics otherwise resemble those we do not observe. In the SOO design, where units are already weighted by the inverse of propensity of the assigned condition, we will use the multiple of the two weights.

2. **Imputation**: We will use `randomForest::rfImpute()` along with the covariates used in the propensity score analysis to impute missing values and run the main analyses on the full, imputed, dataset.

3. **Trimming bounds**: For the SOO design, we will apply Lee 2009 trimming bounds. Suppose, for example, that there are more matches in the treatment than in the control group. In that case, we define a proportion to be trimmed, Q. Let R^1 denote the rate of missingness in the treatment group and R^0 that in control. Then $Q = \frac{R^1 - R^0}{R^1}$. The approach requires an assumption that the treatment exerts a monotonic effect on the missingness. In this example, that means no unit would have failed to appear in the credit card transaction data if treated and appeared if untreated. The upper bound on the treatment effect is obtained by removing the Q% of units in the treatment with the lowest outcomes and estimating the effect as usual on this subset. Ties will be broken at random. The lower bound on the treatment effect is obtained by removing the Q% of units in the treatment with the highest outcomes and estimating the effect as usual. If the imbalance in attrition runs in the opposite direction, the opposite monotonicity assumption is imposed and the trimming is applied to the control group. Importantly, this approach does not necessarily bound the sample average treatment effect. Instead, it bounds the sample average treatment effect for those whose appearance in the credit card data is unaffected by assignment to treatment. It is unclear whether this approach is feasible for the GRD analysis, so we do not plan to run it for the GRD design.

4. **Extreme-value bounds**: This approach involves imputing missing values using the extrema of their support (Manski 1990), or using their most plausible extreme values. We will use the minimum and maximum outcomes for the extreme value bound analysis.

A.2 Formalization of SOO Design

Definition of treatment and identifying assumptions

We can formalize this notion using DAG notation. Let X refer to the set of variables we have available in the application datasets provided by San Diego, while U refers to features of the world we are unable to observe, such as follow-up documentation or information gleaned through phone calls to businesses. Let Y denote the outcome in which we are interested, and ϵ_k an exogenous error term for variable k. For the purposes of causal identification, our design assumes $Z = f_Z(X, \epsilon_Z)$, while $T = f_T(Z, X, U, \epsilon_T)$. In words, whether a business is selected for an invitation to submit documents is a function of exogenous noise and of the data we are able to observe—businesses are "selected on observables"—whereas their ability to actually receive funding is the outcome of a process we cannot observe and that confounds a simple estimate of T on Y. However, even if $Y = f_Y(U, X, Z, T, \epsilon_Y)$, we can estimate the effect of Z on Y simply by conditioning on the confounder X. If, in addition, we can assume $Y = f_Y(U, X, T, \epsilon_Y)$ (an exclusion restriction defined in more detail below), we can also instrument for T using Z in order to estimate the complier average effect on Y, despite unobserved confounding by U. The DAG is presented on Figure 4.

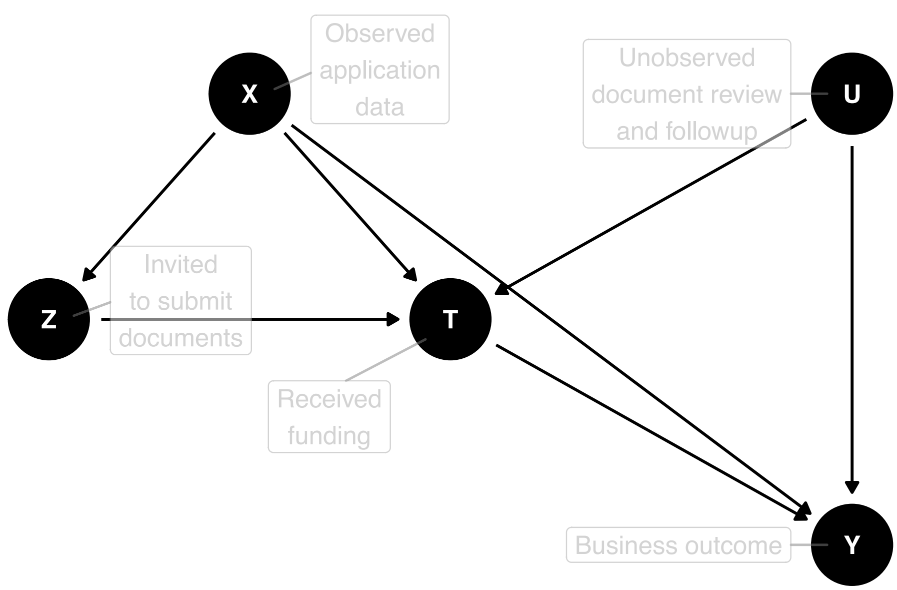

Figure 4: Theoretical assumptions behind the selection on observables design. Conditioning on X removes all paths from X to its descendants, leaving no backdoor paths from Z to Y. While T may be confounded by the unobservable U, we can nevertheless instrument for T using Z to get the complier-local effect of T on Y.

In potential outcomes notation, let Y_i denote the i'th business's observed outcome, e.g., average credit card activity on Monday of March 16 2020. And let $Y_i(T_i(Z_i = z))$ denote its *potential* outcome when T_i takes the value it would take if Z_i was set to z. For example, $Y_i(T_i(Z_i = 1))$ might represent the business's average credit card activity on Monday of March 16 2020 if they were ever invited to submit documents during the SBRF disbursement, and $Y_i(T_i(Z_i = 0))$ its average activity on the same day in a different potential state of the world in which the business was not invited to submit documentation for funding.

We invoke Stable Unit Treatment Value Assumptions (SUTVAs) according to which each business reveals at most one of those two potential outcomes: $Y_i(T_i(Z_i = 1))$ or $Y_i(T_i(Z_i = 0))$. Let **Z** denote a vector of invitations and **Z**$'$ denote a different vector of invitations. Formally, the SUTVA holds that, if $\mathbf{Z}_i = \mathbf{Z}'_i$, then $T_i(\mathbf{Z}) = T_i(\mathbf{Z}')$ and $Y_i(T_i(\mathbf{Z}) = Y_i(T_i(\mathbf{Z}'))$ for all i.[11]

As represented in the $U \rightarrow T$ relationship on Figure 4 some businesses in the sample who are invited to submit documentation will and will not get funded, for reasons we cannot fully observe. We can define four different types of outcomes to the invitation to funding "treatment." Compliers are businesses who are funded if and only if they are invited $(T_i(Z_i = 1, U_i = u) > T_i(Z_i = 0, U_i = u))$. Never-takers are businesses who do not ever receive funding, whether invited or not ($T_i(Z_i = 1, U_i = u) = T_i(Z_i = 0, U_i = u) = 0$). From the data, it is clear that there are no types who receive SBRF funding without an invitation. Thus, we rule out the existence of two other types: always-takers ($T_i(Z_i = 1, U_i = u) = T_i(Z_i = 0, U_i = u) = 1$) and defiers (($T_i(Z_i = 1, U_i = u) < T_i(Z_i = 0, U_i = u)$). This amounts to a monotonicity assumption. Finally, as the DAG suggests, we assume the exclusion restriction is satisfied. In potential outcomes, this implies: $Y_i(T_i = t, Z_i = z) = Y_i(T_i = t, Z_i = z')$. In other words, changing the value of Z_i makes no difference to outcomes except insofar as it changes T_i. This assumption is implied on Figure 4 through the absence of an edge pointing from Z to Y.

How can we identify compliers and never-takers in the data? The control group is composed of those who were never invited to submit documents for review—they fall into category 6 on Figure 1. That is, businesses in category 6 have $Z_i = 0$ and $T_i = 0$, and are a mix of compliers and never-takers. In the treatment group, we can observe compliers in category 4—they are businesses for whom $Z_i = 1$ and $T_i = 1$. Never-takers in the treatment group are businesses who were invited but never received funding. These can be found in category 7—they have $Z_i = 1$ and $T_i = 0$. Other businesses are excluded.

Estimands and identification strategy

Under the SUTVAs enumerated above, we can define the intent-to-treat estimand, which describes the average difference in potential outcomes when a business is or is not invited to submit funding:

$$\tau_{ITT} = \frac{1}{N} \sum_{i}^{N} Y_i(T_i(Z_i = 1)) - Y_i(T_i(Z_i = 0)).$$

The assumptions about Z and X encoded in the DAG above imply that the distribution of the potential outcomes is independent of Z if we condition on X:

$$\{Y_i(T_i(Z_i = 1)), Y_i(T_i(Z_i = 0))\} \perp\!\!\!\perp Z_i \mid X_i.$$

Let $\pi_i = Pr(Z_i = 1 \mid X_i = x) = E[Z_i \mid X_i = x]$. Since the potential outcomes are distributed independently of Z conditional on X, following Hirano et al. 2003:

$$\{Y_i(T_i(Z_i = 1)), Y_i(T_i(Z_i = 0))\} \perp\!\!\!\perp Z_i \mid \pi_i.$$

11. It is worth noting that this assumption may not hold, particularly with regard to the possible spillovers between funding that result from reallocations when one business refuses or accepts. This is an issue we plan to address in future.

Under these assumptions, conditioning on the true treatment assignment probability, π_i, would be sufficient to balance the full joint distribution of covariates (Rosenbaum and Rubin 1983). However, since we do not have access to π_i, we must estimate it using $\hat{\pi}_i$, commonly known as the "propensity score." We can use the propensity score to construct inverse propensity weights:

$$w_i = \begin{cases} \frac{1}{\hat{\pi}_i} & \text{if } Z_i = 1 \\ \frac{1}{1-\hat{\pi}_i} & \text{if } Z_i = 0. \end{cases}$$

If the propensity score perfectly approximates the assignment probability (i.e., $\pi_i = \hat{\pi}_i$ for all i), then in expectation across hypothetical repetitions of the random invitations to submit funding, Z:

$$E[\hat{\mu}_{Y(1)}] = E[\frac{\sum_{i:Z_i=1} Y_i w_i}{\sum_{i:Z_i=1} w_i}] = \frac{1}{N} Y_i(T_i(Z_i = 1)),$$

and

$$E[\hat{\mu}_{Y(0)}] = E[\frac{\sum_{i:Z_i=0} Y_i w_i}{\sum_{i:Z_i=0} w_i}] = \frac{1}{N} Y_i(T_i(Z_i = 0)).$$

From the additive property of expectations, it follows that

$$\tau_{ITT} = E[\hat{\mu}_{Y(1)} - \hat{\mu}_{Y(0)}].$$

Thus, we identify the ITT of Z on Y by estimating propensity scores and constructing weights.

Finally, we are interested in the average effect of the treatment on compliers (CACE):

$$\tau_{CACE} = \frac{1}{N_C} \sum_{i:T(1)>T(0)} Y_i(T_i(Z_i = 1)) - Y_i(T_i(Z_i = 0)),$$

where N_C is the number of compliers in the sample. It is straightforward to show that, under monotonicity, τ_{CACE} can be rewritten as the ratio between τ_{ITT} and the intent to treat effect of Z on T. As such, we can used weighted two-stage least squares as a consistent (albeit possibly biased) estimator of the average effect of the treatment on compliers.

A.3 Formalization of GRD design

Definition of treatment and identifying assumptions

We use a quasi-experimental framework to define the GRD. Let Z denote a binary variable indicating eligibility for funding under SBRF, and T actually being awarded funds. Thus, Z is analogous to an encouragement to take the treatment and T is the actual treatment status. Unlike an experiment, Z is determined nonrandomly by the border, while T is an unobservable, endogenous function of Z. We invoke a stable unit treatment value assumption (SUTVA), according to which each businesses' outcomes and treatment status depend only on their individual treatment encouragement, and not on any other units' treatment encouragements. As such, each business reveals at most one of two potential outcomes: $Y_i(T_i(Z_i = 1))$ or $Y_i(T_i(Z_i = 0))$. Let $\mathbf{Z}$ denote a vector of encouragements and $\mathbf{Z}'$ denote a different vector of encouragements. Formally, the SUTVA holds that, if $\mathbf{Z}_i = \mathbf{Z}'_i$, then $T_i(\mathbf{Z}) = T_i(\mathbf{Z}')$ and $Y_i(T_i(\mathbf{Z}) = Y_i(T_i(\mathbf{Z}'))$ for all i.[12]

12. It is worth noting that this assumption may not hold, particularly with regard to the possible spillovers between funding that result from reallocations when one business refuses or accepts. This is an issue we plan to address in future.

Following Keele and Titiunik 2015, the geographic location of any unit is given by two coordinates in space, such as latitude and longitude, $(S_{i1}, S_{i2}) = \mathbf{S}_i$ (see Figure 3). Treatment encouragement is not random, but is a function of this score. Let $\mathcal{A}^1$ and $\mathcal{A}^0$ denote the sets of all points in space in which units are encouraged to take the treatment or the control, respectively. Thus, $Z(\mathbf{s}) = 1$ for $\mathbf{s} \in \mathcal{A}^1$, $Z(\mathbf{s}) = 0$ for $\mathbf{s} \in \mathcal{A}^0$, and any given business's treatment encouragement can be written $Z_i = Z_i(\mathbf{S}_i)$ (the lowercase **s** indicates we are referring to a specific realization of the quasi random or exogenous variable, **S**).

Let $\mathcal{B}$ denote the set of all boundary points, **b**, with $\mathbf{b} = (S_1, S_2) \in \mathcal{B}$. We assume that the conditional expectation functions of the potential outcomes are continuous in **s** at all points **b** along the boundary separating $\mathcal{A}^1$ and $\mathcal{A}^0$. Formally, this assumption can be written $\lim_{\mathbf{s}\to\mathbf{b}} E[Y_i(T_i(Z_i = 1)) \mid \mathbf{S}_i = \mathbf{s}] = E[Y_i(T_i(Z_i = 1)) \mid \mathbf{S}_i = \mathbf{b}]$ and $\lim_{\mathbf{s}\to\mathbf{b}} E[Y_i(T_i(Z_i = 0)) \mid \mathbf{S}_i = \mathbf{s}] = E[Y_i(T_i(Z_i = 0)) \mid \mathbf{S}_i = \mathbf{b}]$, for all $\mathbf{b} \in \mathcal{B}$. In other words, as the outcomes of units when encouraged to take funding (not take funding) approach the border, they converge to the outcome that would be realized exactly at the boundary.

Finally, we define four types of responses to the treatment encouragement in the sample and monotonicity assumption. As above, although defined over conceptually distinct variables: compliers are those for whom $T_i(Z_i = 1) > T_i(Z_i = 0)$—they get funding when inside the San Diego boundary but don't get funding when outside of it. Never-takers never get funding ($T_i(Z_i = 1) = T_i(Z_i = 0) = 0$) and always-takers always get funding ($T_i(Z_i = 1) = T_i(Z_i = 0) = 1$), irrespective of their location inside or outside San Diego. Finally, defiers get funding but only when outside San Diego: $T_i(Z_i = 1) < T_i(Z_i = 0)$. This is very difficult to imagine, and we assume that such defiers do not exist, so that T_i cannot be decreasing in Z_i.

Estimands and identification strategy

These assumptions allow us to define and identify two main estimands. First, define the intent-to-treat effect (ITT) evaluated at a single point along the boundary as:

$$\tau_{ITT}(\mathbf{b}) = E[Y_i(T_i(Z_i = 1)) - Y_i(T_i(Z_i = 0)) \mid \mathbf{S}_i = \mathbf{b}].$$

In words, this is the expected difference, at point **b** along the border, between the outcome of all units when encouraged and not encouraged to be funded through SBRF. Under the assumptions enumerated above, $\tau_{ITT}(\mathbf{b})$ can be expressed in terms of observable data using the following equality (Keele and Titiunik 2015):

$$\tau_{ITT}(\mathbf{b}) = \lim_{\mathbf{s}^1\to\mathbf{b}} E[Y_i \mid \mathbf{S}_i = \mathbf{s}^1] - \lim_{\mathbf{s}^0\to\mathbf{b}} E[Y_i \mid \mathbf{S}_i = \mathbf{s}^0],$$

where $s^1 \in \mathcal{A}^1$ and $s^0 \in \mathcal{A}^0$. With this ability to define, conceptually and empirically, a treatment effect at any point in the border, we can define a scalar-valued average effect like the population density-weighted mean integral of the intent-to-treat effects along the border. This can be approximated using the weighted sum along a finite grid of border points indexed $r \in \{1, ..., R\}$ (Rischard et al. 2018):

$$\tau_{ITT} = \int_{\mathbf{s}\in\mathcal{B}} \tau(\mathbf{s}) f(\mathbf{s} \mid \mathbf{S} \in \mathcal{B}) d\mathbf{s} \approx \frac{\sum_{r=1}^{R} w_{\mathcal{B}}(\mathbf{b}_r)\tau_{ITT}(\mathbf{b}_r)}{\sum_{r=1}^{R} w_{\mathcal{B}}(\mathbf{b}_r)},$$

where $w_{\mathcal{B}}(\mathbf{b})$ is the local population density at point **b** and is set to approximate $f()$. In words, τ_{ITT} is the effect of the encouragement among the hypothetical population residing right on the border. Conversely, if the border passes through lakes, parks, or other areas with no businesses in the sample, differences in outcomes there are given little to no weight.

Second, we define the complier average causal effect (CACE) at a point along the border as:

$$\tau_{CACE}(\mathbf{b}) = E[Y_i(T_i = 1) - Y_i(T_i = 0) \mid \mathbf{S}_i = \mathbf{b}, T_i(1) > T_i(0)].$$

Under the monotonicity assumptions spelled out above, we can estimate $\tau_{CACE}(\mathbf{b})$ using the following observable quantities (Imbens and Lemieux 2008):

$$\tau_{CACE}(\mathbf{b}) = \frac{\lim_{\mathbf{s}^1 \to \mathbf{b}} E[Y_i \mid \mathbf{S}_i = \mathbf{s}^1] - \lim_{\mathbf{s}^0 \to \mathbf{b}} E[Y_i \mid \mathbf{S}_i = \mathbf{s}^0]}{\lim_{\mathbf{s}^1 \to \mathbf{b}} E[T_i \mid \mathbf{S}_i = \mathbf{s}^1] - \lim_{\mathbf{s}^0 \to \mathbf{b}} E[T_i \mid \mathbf{S}_i = \mathbf{s}^0]},$$

which is just the familiar instrumental variables ratio estimator applied to a specific point on the border. The population density weighted mean integral of the CACEs along the border, τ_{CACE}, can be approximated as above: by taking a population weighted sum of $\tau_{CACE}(\mathbf{b_r})$, the CACEs estimated at a grid of points along the border.