COVID-19 Vaccinations With a Sweepstakes

NCT04951310

6/7/2021

COVID-19 Lottery Study Pre-Registration

1) Have any data been collected for this study already?
No.

2) What's the main question being asked or hypothesis being tested in this study?
We will measure the effectiveness of a COVID-19 vaccination sweepstakes by analyzing COVID-19 vaccination data from the City of Philadelphia and neighboring regions. The sweepstakes will award multiple prizes every two weeks for at least six weeks. Half of the prizes will be awarded to Philadelphia residents in "selected" zip codes that are randomly selected from a list of 20 under-vaccinated zip codes in the city. selected zip codes will be announced two weeks prior to each drawing. The other half of prizes will be awarded to Philadelphia residents in the remaining zip codes in Philadelphia. We will examine (1) the effect of the lotteries (and accompanying attention that they bring to the vaccination process) on Philadelphia resident vaccination rates relative to people living in neighboring counties and cities and (2) the effect on vaccination rates of selected zip code residents compared to residents from other Philadelphia zip codes that were eligible for this designation ("potentially selected").

3) Describe the key dependent variable(s) specifying how they will be measured.
The primary outcome is the weekly vaccination rate for residents receiving a first dose of a COVID-19 vaccination in each zip/county/city as measured by local health departments and, for residency, as measured by the Census or other official sources.

(Rate = Total residents receiving a first dose that week / Total residents in that zip/county/city code as measured by the Census or other official sources, at the most recent data available for all geographies of interest. This source as of pre-registering is the American Community Survey 2019.)

4) How many and which conditions will participants be assigned to?
We will partner with the City of Philadelphia to launch a COVID-19 vaccination sweepstakes that will be conducted for at least 6 weeks. The city will be included in the "treatment" condition. Within the city, we will include two pools: a "potentially selected" zip code pool (including the 20 zip codes with the lowest per capita vaccination rates as of May 27, 2021) and a pool with the remaining zip codes. At the start of each two-week period, we will randomly draw one "selected" zip code from our list of 20 under-vaccinated zip codes in the city. Half the prizes in each category will go to those in the "selected" zip code, and the other half of prizes will go to those in all zip codes that are not the "selected" zip code.

Our differences-in-differences analysis will compare weekly first-dose COVID-19 vaccination rates in each of the three "selected" zip codes (drawn from 20 "priority" zip codes) against the "potentially selected" zip codes that were not selected. We will also compare COVID-19 vaccination rates in all of Philadelphia (Philadelphia county) with vaccination rates in neighboring Pennsylvania counties (Bucks County, Delaware County, and Montgomery County, as per Roberto et al. 2019) and in one comparison city (Baltimore, as a comparable and nearby city and as per Roberto et al. 2019) over the same time period.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will conduct 3 main analyses:

   1. **Comparing across cities**
To analyze the effect of living in Philadelphia as compared with neighboring counties, we will run a standard difference-in-differences model for weekly COVID-19 vaccination rates by county in the weeks leading up to, during, and after the intervention period. We will include data from May 2, 2021 through August 14, 2021 in this analysis (5 weeks of pre-intervention, 6 weeks of intervention, and 4 weeks of

post-intervention). The analysis will begin by showing a graphical event study representation of the results as follows (graph uses hypothetical data):

Hypothetical Figure 1: (linked http://bit.ly/BCFGPhilVaxPreReg)
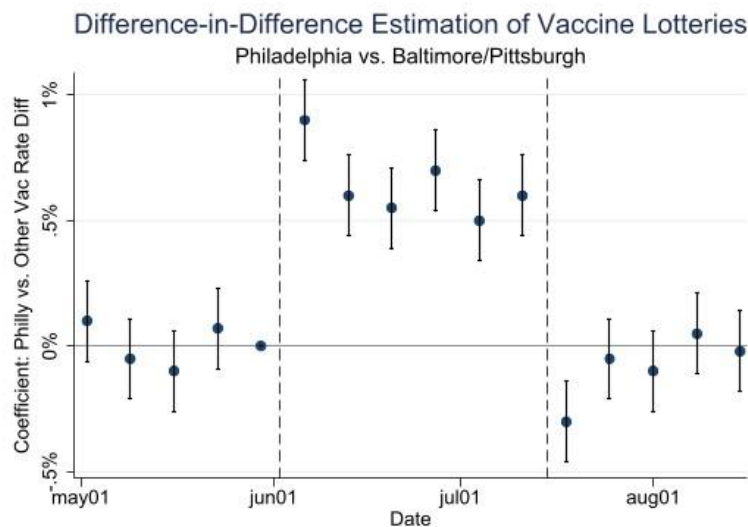
The graph above is created by regressing the weekly first-dose COVID-19 vaccination rate for each county on county fixed effects, week fixed effects, and indicators for pre-treatment and post-treatment effects (as described, for example, in Cunningham 2021) with the week just before the treatment period serving as the comparison group. This plot allows us to see the timing/dynamics of the treatment effect and also provides evidence regarding the parallel-trends assumption that is always made in a diff-in-diff analysis by checking for pre-treatment differences in trends.

In addition to the graphical analysis above, we will also produce the standard diff-in-diff 2x2 estimate by taking the average vaccination rate pre-treatment for each city and the average vaccination rate during the treatment period for each city and regressing these DVs on City Indicators, a Treatment Indicator, and a Philadelphia*Treatment indicator. The coefficient on the interaction term will be our preferred estimate of the overall effect of the lotteries' impact on Philadelphia relative to nearby cities.

2. Comparing across counties
We will do the exact same analysis as discussed above to test the lottery's impact on Philadelphia relative to nearby Pennsylvania counties (Bucks County, Delaware County, and Montgomery County). We will simply replace the weekly city vaccination rates in the above regressions with the weekly county vaccination rates and include county indicators instead of city indicators.

3. Comparing across zip codes
To analyze the impact of being chosen as a "selected" zip code, we will do a similar analysis to the one above for each of the three "selected" zip codes. Once again, we will begin by showing a graphical event study representation of the results, but this time we will be comparing a "selected" zip code to the other "potentially selected" zip codes that did not win ("unselected zip codes"). The analysis for the first "selected" zip code may look something like the following (graph uses hypothetical data):

Hypothetical Figure 2: (linked http://bit.ly/BCFGPhilVaxPreReg)
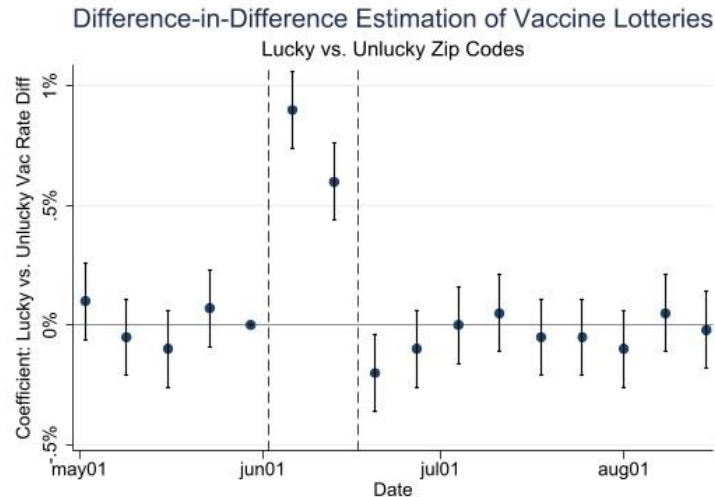
Difference-in-Difference Estimation of Vaccine Lotteries
Lucky vs. Unlucky Zip Codes

The graph above is created by regressing the weekly first-dose COVID-19 vaccination rate for each "potentially selected" zip code (20 total) on zip code fixed effects, week fixed effects, and indicators for pre-treatment and post-treatment effects (as described, for example, in Cunningham 2021). This plot allows us to see the timing/dynamics of the treatment effect and also provides evidence regarding the parallel-trends assumption that is always made in a diff-in-diff analysis by checking for pre-treatment differences in trends.

In addition to the graphical analysis above, we will also produce the standard diff-in-diff 2x2 estimate by taking the average vaccination rate pre-treatment for each zip code and the average vaccination rate during the treatment period for each zip code and regressing these DVs on Zip Code Indicators, a Treatment Indicator, and a selected Zip Code*Treatment indicator. The coefficient on the interaction term will be our preferred estimate of the overall effect of each lottery's impact on each of the "selected" zip codes relative to "unselected" zip codes (across the full treatment period, this will be 17 remaining zip codes).

There will be three different figures above for each of the 2-week lottery periods (one figure for each of the 3 "selected" zip codes). After showing the effects separately for each of the 3 "selected" zip codes, we can also do a pooled analysis where we combine/average the results across the 3 "selected" zip code analyses. To do this, we can stack the data from the 3 studies and produce just one figure like the one above where we show the same results relative to week 0 of the event window. To produce the final overall effect of the lotteries impact on the "selected" zip code, we will regress the weekly vaccination rate for each zip code on zip code fixed effects, week relative to treatment fixed effects (1 week before treatment, 2 weeks before treatment, etc.), and a "selected" zipcode*treatment indicator that equals one when the "selected" zip code was in the treatment period.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
No outliers will be excluded.

7) How many observations will be collected or what will determine sample size?
No need to justify the decision, but be precise about exactly how the number will be determined.
Weekly vaccination totals by zip code in Philadelphia will be collected from the Philadelphia department of public health. Weekly vaccination totals by county will be collected from the CDC. For each region of interest we will include weekly vaccination data from May 2nd through August 14th (for 5 weeks prior to the announcement of the sweepstakes, during the entire sweepstakes, and for 4 weeks after the conclusion of the sweepstakes.)

8) Anything else you would like to pre-register?
Assuming we find positive effects from the treatment, the increased vaccinations could be a result of persuading people who would not have otherwise been vaccinated to become vaccinated or a result of causing people who were eventually going to be vaccinated to move up their vaccination date. If the latter is true, we may see a reduction in vaccination rates in the post-treatment period. This is what in economics has been referred to as a "harvesting" effect where you are simply causing intertemporal substitution for when people perform an action like vaccination. The graphs discussed above will allow us to look for this intertemporal substitution effect by seeing if the dots are less than 0% in the post-treatment period and a similar analysis as listed above will allow us to estimate the magnitude of this substitution.

Assuming we are able to collect data on awareness of the lottery through our partner research organizations we also intend to run a mediation analysis with "awareness" as a mediator.