Did Ohio's Vaccine Lottery Increase Vaccination Rates? A Pre-Registered, Synthetic Control Study

David Lang* Lief Esbenshade† Robb Willer‡

*Graduate School of Education, Stanford University. Email: dnlang86@stanford.edu.
†Graduate School of Education, Stanford University. Email: liefesbenshade@stanford.edu.
‡Stanford University,. Email: willer@stanford.edu.

Abstract

Overcoming vaccine hesitancy is critical to containing the COVID-19 pandemic in the United States. To increase vaccination rates, the State of Ohio launched a million-dollar lottery sweepstakes in May 2021. Following a pre-registered analysis, we estimate the effects of Ohio's lottery program *Vax-a-Million* on COVID-19 vaccination rates by comparing it to a "synthetic control" made up of a weighted composite of eight other states. We find a not statistically significant 1.3 percentage point decrease in the full vaccination rate in Ohio relative to the synthetic control. We investigate the robustness of our conclusion to model specifications through a multiverse analysis of 216 possible models, including longer time periods and alternative vaccination measures. The majority (88%) find small negative effects in line with the results of our pre-registered model. Our results suggest that the likely impact of the lottery was null or a small negative effect.

Keywords: COVID-19, vaccination, synthetic control, lotteries, vaccine hesitancy, vaccine confidence, multiverse analysis

1 Introduction

The COVID-19 pandemic is the largest public health crisis in recent history. With the discovery and mass production of efficacious vaccines against the virus, motivating uptake of the vaccines has emerged as a critical challenge in the United States. To this end, a number of incentives and encouragement strategies have been implemented in the U.S., ranging from free beer to saving bonds (Kmietowicz, 2021; Treisman, 2021). Because of the rapid pace of the pandemic and response efforts, however, little is yet known about the effectiveness of these programs and whether they can help contain the virus in the United States (Mandvalli, 2021).

Lotteries are a potentially powerful and low cost way to encourage vaccination that have been shown to induce behavior change in financial (Gertler et al., 2018) and health behaviors (Van Der Swaluw et al., 2018). Particularly in the context of vaccination, lotteries may be an effective incentive in that they are uniquely attractive to risk-preferring individuals. Prior work finds risk-preference is positively correlated with vaccine skepticism (Massin et al., 2015), suggesting that the vaccine-hesitant may be uniquely high in risk-preference, and thus uniquely responsive to lottery incentive programs.

Additionally, individuals who systematically overestimate low-probability events, such as adverse vaccine reactions, may also systematically overestimate the probabilities of winning a lottery (Camerer and Kunreuther, 1989). Lastly, prior work finds that low-income communities are less likely to get vaccinated but also more likely to have high participation rates in lotteries (Price and Novak, 1999; Razai et al., 2021; Soares et al., 2021). For these reasons, lotteries may effectively encourage individuals to vaccinate where other strategies have failed.

Yet despite these appealing factors, there are serious concerns about whether and how individuals should be compensated for obtaining COVID-19 vaccinations. Efforts to promote uptake of the Human Papillomavirus (HPV) vaccine illustrate the uncertainty and challenges associated with incentivizing vaccination. One randomized controlled trial in the United Kingdom found that financial compensation boosted HPV vaccination rates in participants by nearly ten percentage points (Mantzari et al., 2015). However, in a different context when community health service providers in the Netherlands offered raffles for iPods in exchange for receiving the HPV vaccine, these communities subsequently had lower vaccination rates and the vaccination initiative received negative media coverage (Rondy et al., 2010).

Vaccine promotion efforts have also faced political challenges (Intlekofer et al., 2012; Haber et al., 2007), including opposition based on the cost of compensating vaccine recipients (Buchanon, 2021). Similar sentiments have been shared with regards to COVID-19 vaccination efforts, including concerns that payment to individuals erodes intrinsic motivation, is coercive, and could reduce confidence in the vaccine's safety (Largent and Miller, 2021). While the benefits of increasing vaccination rates may outweigh potential ethical concerns (Persad and Emanuel, 2021), the optimal compensation strategy for staggered treatments - such as the Moderna and Pfizer vaccines - are not obvious (Higgins et al., 2021).

These concerns speak to a larger issue of how to evaluate the efficacy of these vaccination efforts. To increase vaccination rates, lotteries could induce individuals who would never get vaccinated otherwise, or they could cause individuals to vaccinate sooner than they would have otherwise. Alternatively, lotteries may have no impact on vaccination, or even discourage individuals from vaccinating. The relative composition of these effects may make the net impact of vaccination lotteries positive, transitory, or deleterious (Buttenheim and Asch, 2013; Cryder et al., 2010).

In this paper, we present a pre-registered analysis of a program intervention that offered vaccine recipients a chance to participate in a series of million-dollar lotteries if they had received their first vaccination prior to the weekly drawing. Despite early results suggesting Ohioans increased their first-doses in response to this lottery announcement(Ohio Department of Public Health, 2021), results of subsequent studies of Ohio's lottery program have been mixed (Mallow et al., 2021; Robertson et al., 2021; Walkey et al., 2021; Brehm et al., 2021; Barber and West, 2021; Thirumurthy et al., 2021; Sehgal, 2021). Developments in the social sciences over the last decade have made it clear that "researcher degrees of freedom" are a significant problem leading to over-identification of treatment effects when none exist (Simmons et al., 2011; Brodeur et al., 2018; Blanco-Perez and Brodeur, 2020). Our analysis is distinct in that we pre-registered our analysis plan, including code for data processing, outcome selection, and model weights, ensuring that our analysis was neither intentionally nor unintentionally biased towards finding a specific result (Gelman and Loken, 2013; Munafò et al., 2017). We couple this analysis with a multiverse analysis of 216 alternative specifications. While pre-specifying our analysis plan ensures that we did not cherry pick a particular model, other modelling choices may lead to a more efficient estimate of the true treatment effect (Ferman et al., 2020). The multiverse analysis examines the robustness of the pre-registered results to the use of alternative outcomes, covariate selection, time-frames, and modelling methods (Steegen et al., 2016) .

We hypothesized that because lotteries offer incentives that may be uniquely motivating to many unvaccinated individuals, we would see a relative increase in vaccination rates in Ohio following the opening of the "Vax-a-million" lottery compared to states that were similar before the lottery announcement. However, our pre-registered analysis found no evidence that Ohio's vaccination rate increased any faster than in its comparison states. While absence of evidence is not the same as evidence of absence, this result should caution other states that are considering using lottery incentives to increase vaccine turnout. Details of the policy intervention and our specific causal inference strategy follow.

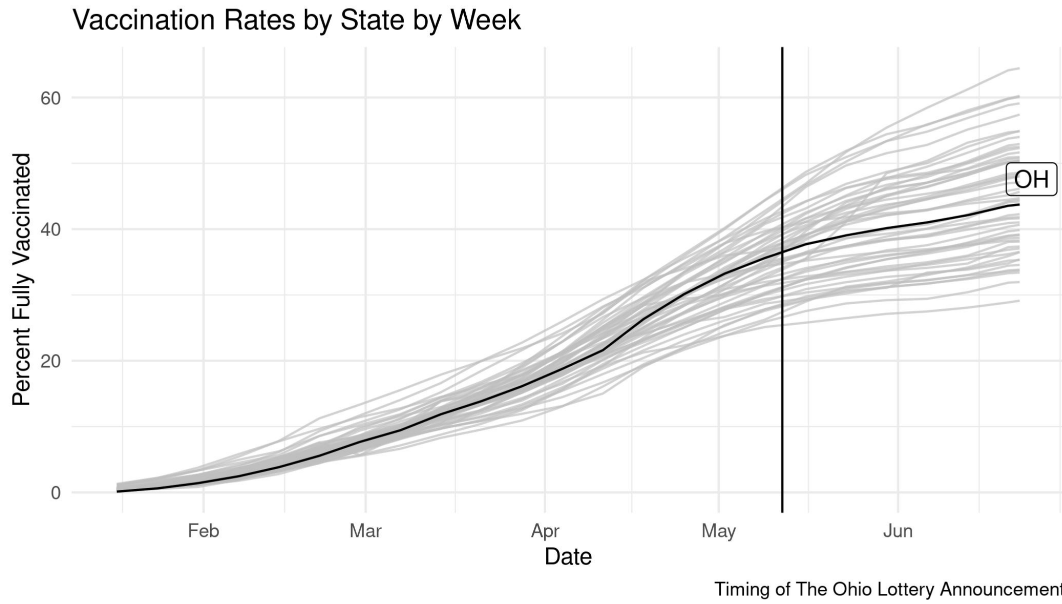

FIGURE 1: Vaccination Rates by State

2 Data

The focal program we study is called *Vax-a-Million*[1]. The intervention was announced on May 12th, 2021 by Ohio Governor Mike Dewine. Starting on May 26th, a weekly lottery drawing was conducted through June 23rd 2021. All Ohio residents who were 18 years or older and entered to participate in the lottery were eligible to receive a one-million dollar prize if they had received their first dose by the date of the drawing.

The focal outcome of our study comes from Our World in Data's COVID-19 vaccination database, which uses numbers published by the US Center for Disease Control (CDC) (Mathieu et al., 2021). The measure counts the percentage of individuals that are fully vaccinated in each US state. We chose this to be our outcome measure because it is aligned with the stated public policy goals of the State of Ohio (DeWine, 2021). The full vaccination outcome has become particularly significant with the rise of virus variants that are more virulent to single doses (Bernal et al., 2021). Notably, this measure requires that individuals who receive either the Pfizer or Moderna vaccines must receive two doses to count as fully vaccinated.

We plot fully vaccinated rates in each state in Figure 1. At the time of the lottery announcement, Ohio was in the middle of the distribution, ranking as the 25th most vaccinated state with 37.4% of the entire population being fully vaccinated. On the day after the final lottery drawing, Ohio had slipped three positions to the 28th most vaccinated state with 43.7% percent of the population fully vaccinated.

[1]https://www.ohiovaxamillion.com/index.html

All data is aggregated to the week level. All subsequent plots and analyses are recentered and denominated in weeks relative to the lottery announcement to facilitate communication. While vaccination data is updated daily, there is substantial noise and missing data in the daily numbers as records are omitted from one day and reported on the next. Thus, weekly aggregation smooths daily fluctuation in reported vaccinations.

3 Methods

We use a synthetic control methodology to create the counterfactual vaccination outcome for Ohio. This technique is useful for cases where a single-aggregated unit such as a state or country receives a treatment (Abadie and Gardeazabal, 2003; Abadie et al., 2010). One can then create a synthetic version of a state by constructing a convex combination of other states (the donor pool) using either pre-treatment outcomes and/or other relevant covariates. Researchers have recently used synthetic control methods to estimate the effectiveness of California's shelter in place orders at the beginning of the COVID-19 pandemic (Friedson et al., 2020), to estimate the impact of a super-spreading event on COVID-19 case rates (Dave et al., 2021), and to estimate the effects of lockdowns on air pollution and health in Wuhan, China (Cole et al., 2020). We are aware of three other concurrent research efforts to analyze the effectiveness of the Ohio lottery program that also use a synthetic control approach (Barber and West, 2021; Brehm et al., 2021; Sehgal, 2021). We address differences in findings and methods below in section 4.1

A particular novelty of this method is that it allows researchers to specify a counterfactual without any knowledge of post-treatment data, making it well-suited for preregistration (Cunningham, 2018). By pre-specifying the weighting of states, it provides a clearly articulated counterfactual of what would happen if no interventions occurred. In light of concerns regarding "cherry picking" with synthetic control methodologies (Ferman et al., 2020), we pre-registered the weights for the synthetic comparison group using data from January 12th to May 9th. We defined the pre-treatment period through the end of the last full week before the lottery announcement on May 12th. On June 15th, we revised our pre-registered protocol to exclude states that had announced vaccine lotteries after our original pre-registration, specifying that we would also run our analysis excluding these states from Ohio's synthetic control. We present the analysis omitting other lottery-adopting states in the main text below. In Appendix C we present findings that follow our initial plan and include all 50 states and the District of Columbia. These findings are not substantively different. We stopped data collection and calculated results after the last lottery was run on June 23rd. All code used in this paper including, but not limited to, downloading raw data, data processing, descriptive analyses, power tests, and synthetic control analysis are publicly available on Github at `https://github.com/XXX`. All the code used to generate our pre-registered synthetic controls comes from the *tidysynth* package in R (Dunford, 2021). Our initial code

TABLE 1: Weights used to construct the synthetic counterfactual to Ohio. States not listed had weights less than 0.001. These weights are based on a June 15th registration that excludes all other lottery adopting states. See Appendix C Table 7 for a comparison to our original pre-registered weights.

Unit	Weights
CT	0.029
GA	0.168
HI	0.061
IA	0.066
KS	0.256
PA	0.056
VA	0.173
WI	0.192

and analysis was posted to the Open Science Foundation (OSF) repository on May 24th [2]

We construct our synthetic control using the the following expression:

$$\sum_{m=1}^{k} v_m \left(X_{1m} - \sum_{j=2}^{J+1} w_j X_{jm} \right)^2 \quad (1)$$

X_1 corresponds to our vector of pretreatment outcomes, vaccination rates before the lottery, for the state of Ohio. X_j corresponds to the pretreatment outcomes and the associated indices of other states in the donor pool. w_j corresponds to the unit weights, the associated weighting of each state in our synthetic construction. v_m corresponds to a variable importance weight of the pretreatment outcomes that we match on. We minimize this expression subject to the constraints that both our unit weights and variable weights are non-negative and sum to unity.

We trained our synthetic control model on the 17 weeks preceding the vaccination announcement. We used data from 31 non-lottery states and the District of Columbia in the donor pool. After optimizing expression 1 based on the past 17 weeks of vaccination data, we generated the synthetic control version of Ohio. Exact weights are shown in Table 1. Synthetic Ohio is a composite of Kansas, Wisconsin, Virginia, Georgia, Iowa, Hawaii, Pennsylvania and Connecticut.

While best practice on the role of covariates in synthetic control is still evolving, using outcome data for each pre-treatment period obviates the need for covariates and shrinks their variable importance weights to zero (Kaul et al., 2015). We believe using the full path of our pre-treatment outcome is a parsimonious specification. As part of the multiverse analysis we test the effect of including additional covariate adjustments.

[2]The url to our OSF page is `https://osf.io/XXX/`. Updated code can be found at our github. `https://github.com/XXX`

Synthetic Ohio and Actual Ohio match very well in their cumulative vaccination rate during the pre-treatment period. In Table 2 below, we show the value of our pre-treatment outcomes for Actual Ohio, Synthetic Ohio, and the average across our donor pool, in the weeks leading up to the vaccination announcement. In all cases, the error between Actual Ohio and Synthetic Ohio was at most 0.6 percentage points. This result suggests that the difference between Synthetic Ohio and Actual Ohio in the pre-treatment period is relatively small.

TABLE 2: Balance Table.

pretreatment outcome	Ohio	Synthetic Ohio	Difference	Donor Pool
lagged_vaccinations_week17	0.120	0.362	−0.242	0.554
lagged_vaccinations_week16	0.610	0.775	−0.165	1.088
lagged_vaccinations_week15	1.400	1.435	−0.035	1.885
lagged_vaccinations_week14	2.440	2.411	0.029	2.958
lagged_vaccinations_week13	3.830	3.757	0.073	4.333
lagged_vaccinations_week12	5.560	5.709	−0.149	6.186
lagged_vaccinations_week11	7.670	7.692	−0.022	7.914
lagged_vaccinations_week10	9.440	9.498	−0.058	9.749
lagged_vaccinations_week09	11.870	11.795	0.075	12.036
lagged_vaccinations_week08	13.860	13.837	0.023	14.027
lagged_vaccinations_week07	16.130	16.056	0.074	16.292
lagged_vaccinations_week06	18.780	18.804	−0.024	19.250
lagged_vaccinations_week05	21.610	22.218	−0.608	22.595
lagged_vaccinations_week04	26.320	26.083	0.237	25.941
lagged_vaccinations_week03	30.110	29.780	0.330	28.863
lagged_vaccinations_week02	33.230	33.021	0.209	31.600
lagged_vaccinations_week01	35.620	35.827	−0.207	34.113

4 Results

We present results for Synthetic Ohio and Actual Ohio in Figure 2. At the time of the final lottery drawing, the vaccination rate for Actual Ohio was 43.7% and the vaccination rate for Synthetic Ohio was 45.0%. This represents a decrease in full vaccinations of approximately 1.3% percentage points relative to Synthetic Ohio.

The pre-registered inference strategy we use to compute statistical significance is a permutation test of the mean squared predicted error ratios. This is calculated by repeating the minimization procedure from expression 1 to create a unique synthetic counterfactual for each of the donor states. We compute the ratio of the Mean Squared Predictive Error (MSPE) between the pre-treatment and post treatment periods using the synthetic counterfactuals for each state. We then sort these MSPE ratios in descending order and use the associated

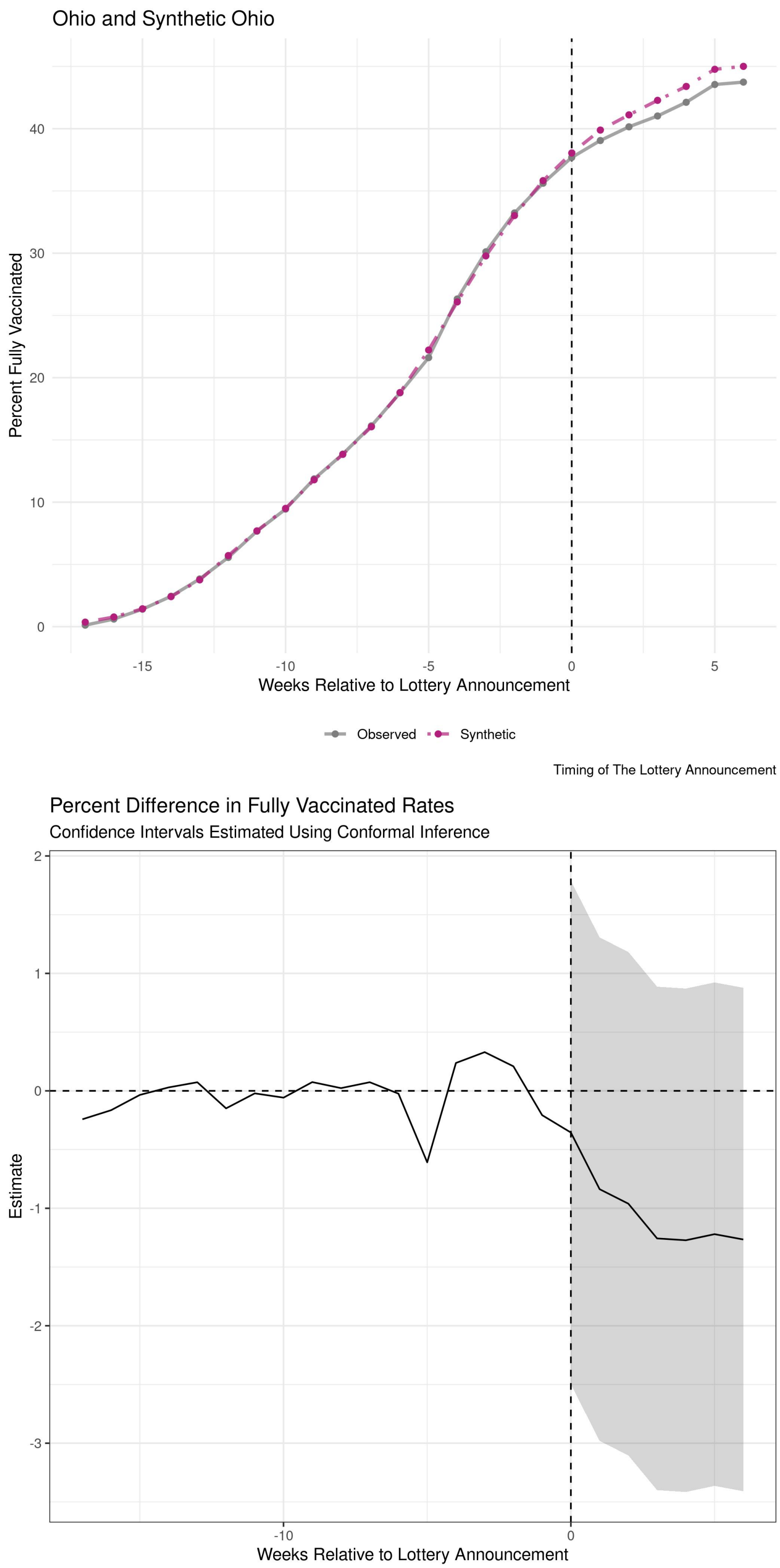

FIGURE 2: Trends in Vaccination Rates (Top). Difference in Vaccination rates between Actual Ohio and Synthetic Ohio (Bottom). Negative values show that Ohio has a lower total vaccination rate than the synthetic comparison.

rank for each state as it's associated p-value (See Appendix A for a placebo analysis and Appendix B for a power analysis). Intuitively, non-treated states should continue to match their synthetic counterfactual relatively well in the post-treatment period. If a policy exhibits a substantial change in the focal outcome, the synthetic control for the true treated state will have relatively poor out of sample fit in the post treatment period. Our pre-specified threshold for statistical significance was a p-value of 0.10. Excluding lottery states, this would correspond to our treatment state having an associated rank of 3 or higher out of the remaining 33 states.

We present the result for the MSPE ratio test in Table 3. This MSPE ratio is 27.0, suggesting the treatment period error is substantially larger than the pre-treatment period error. However, when we compute our p-value using the state's MSPE ranking in descending order, we see that the associated p-value is $12/33$, yielding an approximate p-value of 0.36. In other words, 11 other non-lottery adopting states had larger divergences from their synthetic counterfactuals than Ohio did. Thus, we cannot reject the hypothesis that Ohio's state lottery had no impact on statewide cumulative vaccination rates. This null finding, however, should not be construed as proof that the lottery had no effect.

While best practices for estimating confidence intervals for synthetic control analyses are still being established, conformal inference techniques suggest that the associated point estimates at the end of period are between -3.4% and 0.9%, see bottom panel of Figure 2 (Chernozhukov et al., 2021; Ben-Michael et al., 2021a). We emphasize that this confidence interval is descriptive and note the method is not specified in our pre-registration.

TABLE 3: Outcome Table

Measure	MSPE-Ratio	Average Difference	Last Period Difference
Value	27.0	-1.14	-1.27
Rank	12	23	24
p-value	0.36	0.70	0.73

4.1 Multiverse Analysis

We are aware of three other, concurrent research efforts that use synthetic control methods (Brehm et al., 2021; Barber and West, 2021; Sehgal, 2021). None of these studies used preregistered analysis plans. These papers have found small positive effects of the Ohio lottery on first vaccination doses. In order to fully understand why similar methods would find differing conclusions, we conducted a multiverse analysis to explore the impact of several different data processing and modelling choices, based on those used in the aforementioned papers, on the final outcome estimates through the creation of 216 distinct synthetic control models (Steegen et al., 2016; Silberzahn et al., 2018). In Appendix D we use the weights published in (Barber and West, 2021; Sehgal, 2021) and present a direct replication and comparison of their first dose effects over an extended time frame. We find that the positive effects on first doses observed at the end of the lottery period rapidly decay and turn negative.

Multiverse analyses have been used to understand differences in empirical research findings in several research contexts including racial disparities in police use of force (Cesario et al., 2019), brain structure and depressive disorders (Kołodziej et al., 2021), smartphone use and parenting (Modecki et al., 2020) and power pose effects (Credé and Phillips, 2017). The multiverse approach can highlight when findings are robust to data processing and modelling decisions, or when statistically significant findings disappear under slightly different assumptions.

In Table 4 we outline six decision criteria that we considered. The choices indicated with a * are our pre-registered modelling decisions. Pre-registering those choices before outcome data was available ensured that those decisions were not influenced by a motivated search of the modelling space for a significant effect.

Our first consideration in our multiverse is the associated donor pool of our synthetic Ohio. We initially pre-registered the decision to use all 50 states, prior to many other states subsequently adopting lotteries. On June 15th, we amended the pre-registered plan to exclude any states that subsequently adopted vaccination lottery sweepstakes.This leaves a donor pool of 50 and 32 units respectively. We note that though other work (Brehm et al., 2021) uses county level data we restrict our multiverse analysis to state units.

Second, we extend our results with additional outcomes, testing whether first doses or total vaccines administered captured effects that did not show up in the share fully vaccinated.

Third, we consider different starting time windows. Our analysis used vaccination data compiled by Our World in Data (Mathieu et al., 2021) that were available from January 12, 2021. However others have reported that comprehensive data from the CDC was not available until February 19, 2021 and that access to vaccines may have been limited to health care workers and other at-risk individuals prior to that (Barber and West, 2021). Other work has used an even shorter 30-day pre-treatment window, which we omit here for the sake of parsimony (Sehgal, 2021).

Fourth, we vary the end date of the post-treatment period. Our original ending date – the day of the last lottery drawing – was potentially too early to find effects on full vaccinations, due to the necessity of waiting three to four weeks between doses to complete the Pfizer and Moderna vaccines (Barber and West, 2021; Brehm et al., 2021). We therefore test whether effects change four and eight weeks after the final lottery drawing to allow for the administration of second doses. This latter selection was chosen as it was the day prior to full approval of the Pfizer vaccine by the Food and Drug Administration (FDA) [3].

Fifth, we assess the inclusion of covariates to improve the construction of the synthetic counterfactual. Following the lead of other papers investigating the effect of the Ohio

[3] https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine

TABLE 4: Data Processing and Modelling Choices. Choices indicated with a * represent the pre-registered modelling decisions.

Modelling Decisions
1. States to Include in Comparison
(a*) All states + District of Columbia
(b* June 15th modification) Only states that did not adopt a lottery See table 6 for excluded states
2. Outcome
(a*) Percent fully vaccinated
(b) Percent with at least one dose
(c) Total number of vaccines given per hundred
3. Vaccination Data Start Time
(a*) OWID data availability (01/12/2021)
(b) CDC data availability (02/19/2021)
4. Vaccination End of Comparison
(a*) Lottery end date (06/24/2021)
(b) 4 weeks after lottery end to allow for full vaccination (07/23/2021)
(c) 8 weeks after lottery end (08/22/2021)
5. Covariates
(a*) None
(b) State demographics, 2020 Republican Presidential Election vote share, 2019 Influenza vaccination rate.
(c) b + Google mobility trends
6. Synthetic Control Model
(a*) Traditional Synthetic Control
(b) Augmented synthetic control with unit fixed effects and ridge regressions

lottery, we test the use of a series of state demographic variables as well as 2019 estimates of flu vaccination rates and the 2020 Presidential election republican vote share (Brehm et al., 2021), and additionally including daily estimates of individual mobility provided by Google trends (Barber and West, 2021).

Lastly, we varied the modelling technique used to compare the traditional synthetic control approach to an augmented synthetic control model (Ben-Michael et al., 2021a), exploring a technique that relaxes some of the assumptions of the traditional synthetic control model by allowing for state fixed effects and augmenting the counterfactual comparison with ridge regressions. Most notably, these approaches allow the associated donor weights to be negative and can facilitate better pre-treatment fit.

We fully interact all of the modelling decisions and data processing choices to create 216 possible models.

We analyze the results of this multiverse of modelling decisions in two stages. First we examine the variability in donor states' weights in the synthetic control for Ohio in Figure 3. Here we see that the construction of the counterfactual is sensitive to the modelling decisions made. While some states consistently have near-zero weights, the states that are used in the counterfactual have widely varying weights ranging from comprising half the weight of the counterfactual down to having no weighting at all. While the weights used in synthetic controls provide a high level of transparency in the calculation of the counterfactual, these results show that the determination of those weights is quite sensitive to researcher decisions. Prior research has found that synthetic control weights are subject to cherry picking concerns (Ferman et al., 2020).

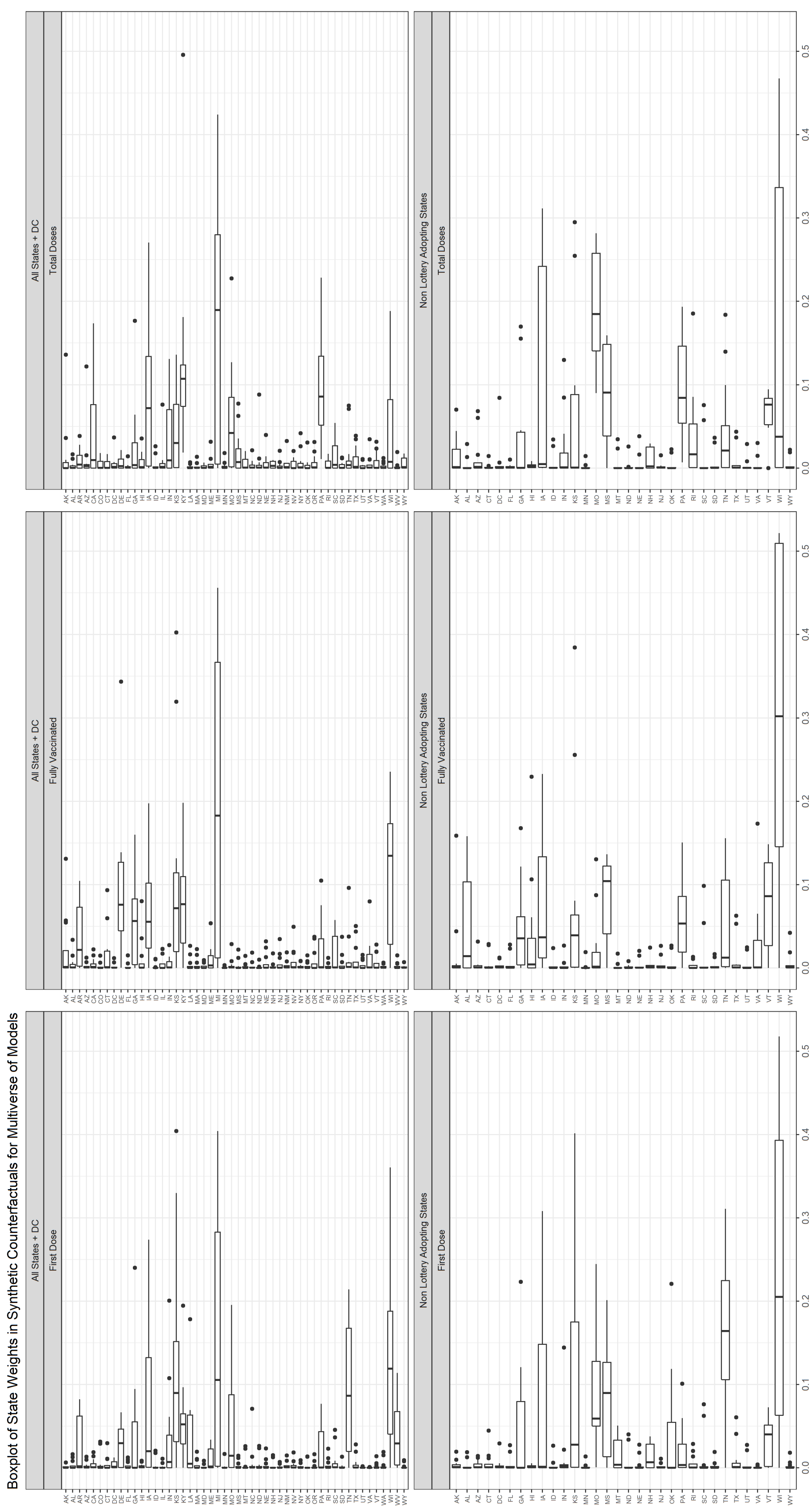

FIGURE 3: Distribution of Donor State Weights in Synthetic Counterfactuals Across Multiverse of Models

In Figure 4 we present the end of period differences between Actual Ohio and Synthetic Ohio across modelling decisions. The figure presents the estimated final period treatment effect on the y-axis while the x-axis represents different time windows (Options 3 and 4 from table 4) while color is used to denote the covariate choice used (option 5). A total of twelve figures are presented with the three rows corresponding to measuring effects on the percentage of the population receiving a first dose, percentage fully vaccinated, and the total number of doses per capita (option 5). While the four columns correspond to the counterfactual states considered and modelling choice (options 1 and 5). Across specifications we find a maximum positive effect of +3.6 percentage points on full vaccinations and a maximum negative effect of -5.9 percentage points on total doses administered. We see estimates of negative effects in 59 out of 72 models that use first doses as the dependent variable, 60 out of 72 of the models that measure effects on full vaccination rates, and all 72 of the models that estimate total vaccines administered. Our pre-registered model is indicated with a * and is near the middle of the model estimates.

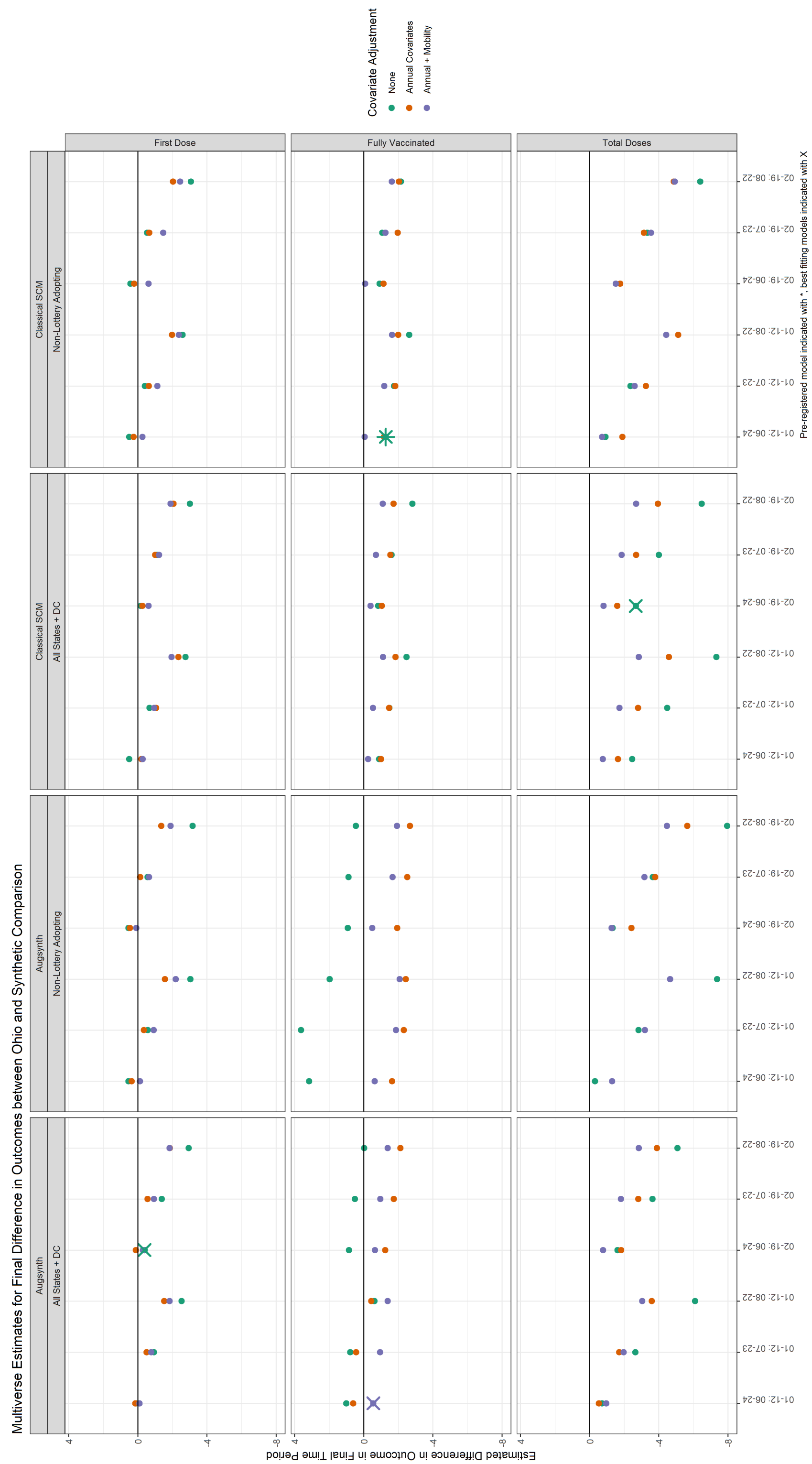

FIGURE 4: Distribution of Estimated Change in Vaccination Outcomes Across Multiverse of Models

To assess the relative validity of these models, we compute the post-treatment fit for our donor state. Put differently, we estimate how much non-treated states diverge from their synthetic counterfactuals, with smaller divergence treated as evidence of better model fit (Abadie, 2021). Intuitively, if a model has poor post-treatment fit when no intervention occurred, this model would be less preferable to one that more accurately described donor states' behavior.

We calculate the average mean squared predicted error in the post-treatment period for each model across the permutations. We find that the median model has an average post period MSPE of 11.9 in the non-treated states. Our pre-registered model has an average non-treated post period MSPE of 8.4, which is the 35th percentile of all the calculated average post period MSPE values. For each of the three outcome measures, the model with the lowest MSPE is indicated in Figure 4 with an X. We estimate a negative effect of -0.38 percentage points on first doses, -0.54 percentage points for full vaccination rates and -2.66 percentage points for total doses in the best fitting models, though none of these estimates are statistically significant according to our permutation test. Thus, we find small, negative effects for each outcome in the models that best generalize to the control states.

5 Discussion

Contrary to our pre-registered hypothesis, our work did not detect a statistically significant effect of Ohio's lottery program on state-wide vaccination rates. In our pre-registered analysis, the majority of specifications from our multiverse analysis, and in the best-fitting models from the multiverse analysis, we found convergent evidence that the effect of the lottery was generally small, slightly negative, and not statistically significant. Nonetheless, while negative point estimates predominated in these analyses, we do not have strong evidence that the lottery program had negative effects, as our estimated confidence intervals include small positive effects. These findings are particularly important as they contrast with early results suggesting that Ohio's lottery was effective at boosting vaccination rates in the short term, and those early results were used by the White House to encourage other states to adopt lotteries of their own (Ohio Department of Public Health, 2021; White House Press Briefing, 2021)

How do our results fit with other papers analyzing the effect of COVID-19 vaccine lotteries, in Ohio and beyond? Our results corroborate a pre-registered experiment that attempts to assess the efficacy of vaccination lotteries (Gandhi et al., 2021) in Pennsylvania. With respect to studies of Ohio specifically, some other work casts doubt on the efficacy of lottery sweepstakes at increasing COVID-19 vaccination rates (Walkey et al., 2021; Thirumurthy et al., 2021). County-level analyses have found some positive impact on starting-vaccination rates (Brehm et al., 2021; Robertson et al., 2021; Mallow et al., 2021). Other work that used a similar synthetic control approach found modest positive effects (Barber and West, 2021; Sehgal, 2021; Brehm et al., 2021). In the multiverse analysis we find that differences in estimates are not explained through the use of full vaccination rates versus first doses as the focal outcome, and find that the positive effects reported in these papers are replicated in only a small subset of the modelling space.

We caution however that our results are all based on state-level average vaccination rates for Ohio. A challenge faced by recent evaluations of the Ohio Lottery program is that relevant comparator states have had data revisions and corrections that are on the order of a percentage point (Cox, 2021; Kansas Vaccine Database, 2021; McDaniel, 2021). We encourage other researchers to look at this issue with more granular data and to examine heterogeneity in incentive effects for specific sub-populations, especially those with lower vaccination rates. Initial work suggests that the Ohio lottery may have been successful in increasing vaccination rates in lower income counties (Mallow et al., 2021).

As more states have adopted lottery incentives, future research should use methods that allow for multiple treated units. New methods of multi-treatment synthetic control models may be appropriate for this context (Ben-Michael et al., 2021b, 2021a). We present an exploratory analysis of the multi-state adoption in an Appendix E and find evidence for modestly positive (1.0%), non-significant effects of the lotteries on vaccination rates. One group of researchers who have studied the effect of these incentives in a similar sample of states found small negative effects on weekly vaccination rates (Thirumurthy et al., 2021), while another set found small positive effects (Robertson et al., 2021).

Although initial coverage was positive, later news reports questioned the long term effectiveness of lottery incentives(Welsh-Huggins, 2021), and our analysis suggests early effects likely did not endure. Nonetheless, 17 states followed Ohio's lead and took up their own vaccination lotteries. In Appendix D, we present analysis that suggests that research which found positive effects on initial doses during the lottery period turned negative shortly after the lottery's completion.

As the pace of vaccination continues to slow, it is important that policymakers receive rapid feedback about the effectiveness of their efforts. Our work acts as proof of concept that social science methods can be used both in prospective and policy-relevant settings in real-time. We made a pre-print of these results available on July 5th, less than 2 months after the policy was announced on May 13. We also made all code and data used publicly available. We offer the following closing thoughts on how policymakers and researchers may better facilitate such policy evaluation.

First, providing high frequency data with clearly defined policy changes can help facilitate assessment of such actions. The ease with which this analysis was conducted was due largely to the fact that researchers and public officials offered a tremendous level of data transparency. We as researchers had no privileged access. The fact that the intervention was well-defined and conducted over a short period further facilitated our analysis.

Second, given the known biases in publication processes towards positive and statistically significant results (Dwan et al., 2008; Munafò et al., 2017), multiverse analyses are a powerful tool to ensure that researchers haven't intentionally or unintentionally made apparently reasonable decisions that may bias analyses toward finding a specific result. Incorporating pre-analysis plans can help adjudicate which of these specifications are most appropriate(Ludwig et al., 2019).

Lastly, we highlight the value of synthetic control methods as a tool for prospective policy analysis for researchers. Of the nearly 80,000 registrations on the Open Science Foundation repository, only seven use synthetic controls[4]. Synthetic control methods allow researchers to generate a specific and transparent counterfactual outcome before post-treatment data is available. With these pre-defined weights, comparing treatment outcomes between a synthetic and actual state is no more complicated than computing a weighted average (as we show in Appendix D). Given the technique's alignment with pre-registration, relative simplicity, and broad utilization, we believe more researchers should consider pre-registering synthetic control analyses of timely policy matters, and coupling this approach with a multiverse analysis (Athey and Imbens, 2017; Steegen et al., 2016).

6 Acknowledgments

The authors received no specific funding for this work. The first two authors' graduate studies are supported by a grant from the Institute of Education Sciences (Award No. R305B140009). The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The authors would like to thank Ben Domingue for support in this project. They would also like to thank Klint Kanopka and Jonas Mueller for providing feedback on early drafts of this work and Joe Mernyk for help in obtaining and processing co-variate data. We also like to thank Eli Ben-Michael for guidance on the augsynth package. We also wish to thank Andrew Barber and Jeremy West for sharing data and code so that we could better understand the sources of differences in our estimates.

7 Competing interests

The authors declare no competing interests.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391–425.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, *105*(490), 493–505.

[4]https://web.archive.org/save/https://osf.io/registries/discover?provider=OSF%20Registries&q=%22synthetic%20control%22

Abadie, A., and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, *93*(1), 113–132.

Athey, S., and Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, *31*(2), 3–32.

Barber, A., and West, J. (2021). Conditional cash lotteries increase covid-19 vaccination rates. *Available at SSRN*.

Ben-Michael, E., Feller, A., and Rothstein, J. (2021a). The augmented synthetic control method. *Journal of the American Statistical Association*(just-accepted), 1–34.

Ben-Michael, E., Feller, A., and Rothstein, J. (2021b). Synthetic controls with staggered adoption. *NBER Working Papers*.

Bernal, J. L., Andrews, N., Gower, C., Gallagher, E., Simmons, R., Thelwall, S., ... others (2021). Effectiveness of covid-19 vaccines against the b. 1.617. 2 (delta) variant. *New England Journal of Medicine*.

Blanco-Perez, C., and Brodeur, A. (2020). Publication bias and editorial statement on negative findings. *The Economic Journal*, *130*(629), 1226–1247.

Brehm, M. E., Brehm, P. A., and Saavedra, M. (2021). *The Ohio Vaccine Lottery and Starting Vaccination Rates* (Tech. Rep.). Retrieved from `https://twitter.com/NewDay/status/1397539038458961920`

Brodeur, A., Cook, N., and Heyes, A. G. (2018). Methods matter: P-hacking and causal inference in economics. *IZA Discussion paper*.

Buchanon, T. (2021). With first drawing days away, ohio legislator seeks to stop vax-a-million lottery. *Ohio Capital Journal*.

Buttenheim, A. M., and Asch, D. A. (2013). Making vaccine refusal less of a free ride. *Human vaccines & immunotherapeutics*, *9*(12), 2674–2675.

Camerer, C. F., and Kunreuther, H. (1989). Decision processes for low probability events: Policy implications. *Journal of policy analysis and management*, *8*(4), 565–592.

Cesario, J., Johnson, D. J., and Terrill, W. (2019). Is there evidence of racial disparity in police use of deadly force? analyses of officer-involved fatal shootings in 2015–2016. *Social psychological and personality science*, *10*(5), 586–595.

Chernozhukov, V., Wüthrich, K., and Zhu, Y. (2021). An exact and robust conformal inference method for counterfactual and synthetic controls. *Journal of the American Statistical Association*, 1–16.

Cole, M. A., Elliott, R. J., and Liu, B. (2020). The impact of the wuhan covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach. *Environmental and Resource Economics*, *76*(4), 553–580.

Cox, S. (2021). *Spencer cox on twitter: "welp. we screwed up. because of a reporting error we have not yet hit 70% on our adult vaccinations. i promised to admit our mistakes and hold us accountable. i hope you will forgive us—and know we have made changes to ensure it won't happen again. https://t.co/nzk8a8gabt" / twitter.* `https://twitter.com/SpencerJCox/status/1414632127182737411/photo/1`. ((Accessed on 08/22/2021))

Credé, M., and Phillips, L. A. (2017). Revisiting the power pose effect: How robust are the results reported by carney, cuddy, and yap (2010) to data analytic decisions? *Social Psychological and Personality Science*, *8*(5), 493–499.

Cryder, C. E., London, A. J., Volpp, K. G., and Loewenstein, G. (2010). Informative inducement: Study payment as a signal of risk. *Social science & medicine*, *70*(3), 455–464.

Cunningham, S. (2018). *Causal inference: The mixtape*. Yale University Press.

Dave, D., McNichols, D., and Sabia, J. J. (2021). The contagion externality of a superspreading event: The sturgis motorcycle rally and covid-19. *Southern economic journal*, *87*(3), 769–807.

DeWine, M. (2021). *Governor dewine reminds unvaccinated ohioans to continue wearing masks.* Retrieved from `https://governor.ohio.gov/wps/portal/gov/governor/media/news-and-media/dewine-reminds-unvaccinated-ohioans-to-continue-wearing-masks-06012021`

Dunford, E. (2021). *edunford/tidysynth: A tidy implementation of the synthetic control method in R.* Retrieved from `https://github.com/edunford/tidysynth`

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., ... others (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PloS one*, *3*(8), e3081.

Ferman, B., Pinto, C., and Possebom, V. (2020). Cherry picking with synthetic controls. *Journal of Policy Analysis and Management*, *39*(2), 510–532.

Friedson, A. I., McNichols, D., Sabia, J. J., and Dave, D. (2020). *Did california's shelter-in-place order work? early coronavirus-related public health effects* (Tech. Rep.). National Bureau of Economic Research.

Gandhi, L., Milkman, K. L., Ellis, S., Graci, H., Gromet, D., Mobarak, R., ... others (2021). An experiment evaluating the impact of large-scale, high-payoff vaccine regret lotteries. *High-Payoff Vaccine Regret Lotteries (August 13, 2021)*.

Gelman, A., and Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*, *348*.

Gertler, P., Higgins, S., Scott, A., and Seira, E. (2018). The long-term effects of temporary incentives to save: Evidence from a prize-linked savings field experiment. *Poverty Action Lab Working Paper*.

Haber, G., Malow, R. M., and Zimet, G. D. (2007). The hpv vaccine mandate controversy. *Journal of pediatric and adolescent gynecology*, *20*(6), 325–331.

Higgins, S. T., Klemperer, E. M., and Coleman, S. R. (2021). Looking to the empirical literature on the potential for financial incentives to enhance adherence with covid-19 vaccination. *Preventive Medicine*, *145*, 106421.

Intlekofer, K. A., Cunningham, M. J., and Caplan, A. L. (2012). The hpv vaccine controversy. *AMA Journal of Ethics*, *14*(1), 39–49.

Kansas Vaccine Database. (2021). *Data | kdhe vaccine ks.* `https://www.kansasvaccine.gov/158/Data`. ((Accessed on 08/22/2021))

Kaul, A., Klößner, S., Pfeifer, G., and Schieler, M. (2015). Synthetic control methods: Never use all pre-intervention outcomes together with covariates. *Research Papers in Economics Archive*.

Kmietowicz, Z. (2021). Sixty seconds on... vaccine bribes. *British Medical Journal*.

Kołodziej, A., Magnuski, M., Ruban, A., and Brzezicka, A. (2021). No relationship between frontal alpha asymmetry and depressive disorders in a multiverse analysis of five studies. *Elife*, *10*, e60595.

Largent, E. A., and Miller, F. G. (2021). Problems with paying people to be vaccinated against covid-19. *JAMA*, *325*(6), 534–535.

Ludwig, J., Mullainathan, S., and Spiess, J. (2019). Augmenting pre-analysis plans with machine learning. In *Aea papers and proceedings* (Vol. 109, pp. 71–76).

Mallow, P. J., Enis, A., Wackler, M., and Hooker, E. A. (2021). Covid-19 financial lottery effect on vaccine hesitant areas: Results from ohio's vax-a-million program. *The American Journal of Emergency Medicine*.

Mandvalli, A. (2021). *Reaching 'Herd Immunity' Is Unlikely in the U.S., Experts Now Believe - The New York Times.* Retrieved from `https://www.nytimes.com/2021/05/03/health/covid-herd-immunity-vaccine.html`

Mantzari, E., Vogt, F., and Marteau, T. M. (2015). Financial incentives for increasing uptake of hpv vaccinations: a randomized controlled trial. *Health Psychology*, *34*(2), 160.

Massin, S., Ventelou, B., Nebout, A., Verger, P., and Pulcini, C. (2015). Cross-sectional survey: risk-averse french general practitioners are more favorable toward influenza vaccination. *Vaccine*, *33*(5), 610–614.

Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., ... Rodés-Guirao, L. (2021). A global database of covid-19 vaccinations. *Nature Human Behaviour*, 1–7.

McDaniel, J. (2021). *Pa. vaccination doses were overcounted by 500,000 shots, state says.* `https://www.inquirer.com/health/coronavirus/covid-pa-vaccine-data-20210712.html`. ((Accessed on 08/23/2021))

Modecki, K. L., Low-Choy, S., Uink, B. N., Vernon, L., Correia, H., and Andrews, K. (2020). Tuning into the real effect of smartphone use on parenting: a multiverse analysis. *Journal of Child Psychology and Psychiatry*, *61*(8), 855–865.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Du Sert, N. P., ... Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature human behaviour*, *1*(1), 1–9.

Ohio Department of Public Health. (2021). *Ohio Vax-a-Million Details Announced | Ohio Department of Health.* Retrieved 2021-05-22, from `https://odh.ohio.gov/wps/portal/gov/odh/media-center/odh-news-releases/odh-news-release-05-17-21`

Persad, G., and Emanuel, E. J. (2021, 07). Ethical Considerations of Offering Benefits to COVID-19 Vaccine Recipients. *JAMA*, *326*(3), 221-222. Retrieved from `https://doi.org/10.1001/jama.2021.11045` doi: 10.1001/jama.2021.11045

Price, D. I., and Novak, E. S. (1999). The tax incidence of three texas lottery games: Regressivity, race, and education. *National Tax Journal*, 741–751.

Razai, M. S., Osama, T., McKechnie, D. G. J., and Majeed, A. (2021). Covid-19 vaccine hesitancy among ethnic minority groups. *BMJ*, *372*. Retrieved from `https://www.bmj.com/content/372/bmj.n513` doi: 10.1136/bmj.n513

Robertson, C. T., Schaefer, K. A., and Scheitrum, D. (2021). Are vaccine lotteries worth the money? *Available at SSRN 3897933*.

Rondy, M., Van Lier, A., Van de Kassteele, J., Rust, L., and De Melker, H. (2010). Determinants for hpv vaccine uptake in the netherlands: a multilevel study. *Vaccine*, *28*(9), 2070–2075.

Sehgal, N. K. (2021). Impact of vax-a-million lottery on covid-19 vaccination rates in ohio. *The American Journal of Medicine*.

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., ... others (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, *22*(11), 1359–1366.

Soares, P., Rocha, J. V., Moniz, M., Gama, A., Laires, P. A., Pedro, A. R., ... Nunes, C. (2021). Factors associated with covid-19 vaccine hesitancy. *Vaccines*, *9*(3), 300.

Steegen, S., Tuerlinckx, F., Gelman, A., and Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712.

Thirumurthy, H., Milkman, K. L., Volpp, K., Buttenheim, A., and Pope, D. G. (2021, aug). *Association between statewide financial incentive programs and COVID-19 vaccination rates.* SSRN. Retrieved from https://papers.ssrn.com/abstract=3912786

Treisman, R. (2021). *West Virginia Giving $100 To Young People Who Get Vaccinated : Coronavirus Updates : NPR.* Retrieved 2021-05-21, from https://www.npr.org/sections/coronavirus-live-updates/2021/04/27/991173625/west-virginia-will-pay-young-people-100-to-get-vaccinated-against-covid-19

Van Der Swaluw, K., Lambooij, M. S., Mathijssen, J. J., Schipper, M., Zeelenberg, M., Berkhout, S., ... Prast, H. M. (2018). Commitment lotteries promote physical activity among overweight adults—a cluster randomized trial. *Annals of Behavioral Medicine*, *52*(4), 342–351.

Walkey, A. J., Law, A., and Bosch, N. A. (2021, 07). Lottery-Based Incentive in Ohio and COVID-19 Vaccination Rates. *JAMA*. Retrieved from https://doi.org/10.1001/jama.2021.11048 doi: 10.1001/jama.2021.11048

Welsh-Huggins, A. (2021). Ohio ends incentive lottery with mixed vaccination results. *The Washington Post*. Retrieved from https://www.washingtonpost.com/politics/ohio-ends-incentive-lottery-with-mixed-vaccination-results/2021/06/23/396ac392-d45d-11eb-b39f-05a2d776b1f4_story.html

White House Press Briefing. (2021, May). Press briefing by white house covid-19 response team and public health officials 2021. *White House Press Briefing*. Retrieved from https://www.whitehouse.gov/briefing-room/press-briefings/2021/05/25/press-briefing-by-white-house-covid-19-response-team-and-public-health-officials-39/

TABLE 5: Vaccination Growth Rates by State, Pre Ohio Lottery Announcement

State	Mean	SD	State	Mean	SD
AK	2.141	1.004	MT	2.074	0.844
AL	1.601	0.66	NC	1.966	0.856
AR	1.672	0.737	ND	2.005	0.863
AZ	1.938	0.781	NE	2.259	1.022
CA	2.188	1.129	NH	2.058	1.221
CO	2.337	1.124	NJ	2.552	1.268
CT	2.774	1.492	NM	2.534	0.816
DC	2.176	1.58	NV	1.935	0.833
DE	2.249	1.238	NY	2.455	1.397
FL	2.063	0.849	OH	2.191	1.059
GA	1.702	0.994	OK	1.851	0.85
HI	2.503	1.252	OR	2.237	0.939
IA	2.341	1.203	PA	2.241	1.147
ID	1.772	0.745	RI	2.631	1.48
IL	2.076	1.024	SC	1.824	0.872
IN	1.855	0.672	SD	2.348	0.921
KS	2.102	1.048	TN	1.688	0.661
KY	2.097	0.974	TX	1.848	0.936
LA	1.697	0.778	UT	1.685	0.887
MA	2.678	1.376	VA	2.334	1.129
MD	2.449	1.211	VT	2.671	1.341
ME	2.825	1.517	WA	2.296	0.994
MI	2.219	0.957	WI	2.409	1.111
MN	2.385	1.036	WV	1.858	0.616
MO	1.847	0.818	WY	1.756	0.873
MS	1.5	0.746			

8 Appendix

A Placebo Analysis Plan

In the pre-registration we also presented a "placebo" test that created a false treatment period of April 5th to May 9th. In this artificial treatment period, we saw little difference between Ohio and synthetic Ohio. Once generated we plotted the difference between synthetic Ohio's vaccination rate and actual Ohio's vaccination rate (Figure 5). This placebo analysis suggested that it generated reasonable out-of-sample fit, with an error of less than a percent in any of the out-of-sample periods.

The exact inference strategy we used to compute statistical significance is a permutation test. We treat each of the donor states in turn as though it was the treated state, and re-estimate a unique synthetic counterfactual. We computed the ratio of the Mean Squared Predictive

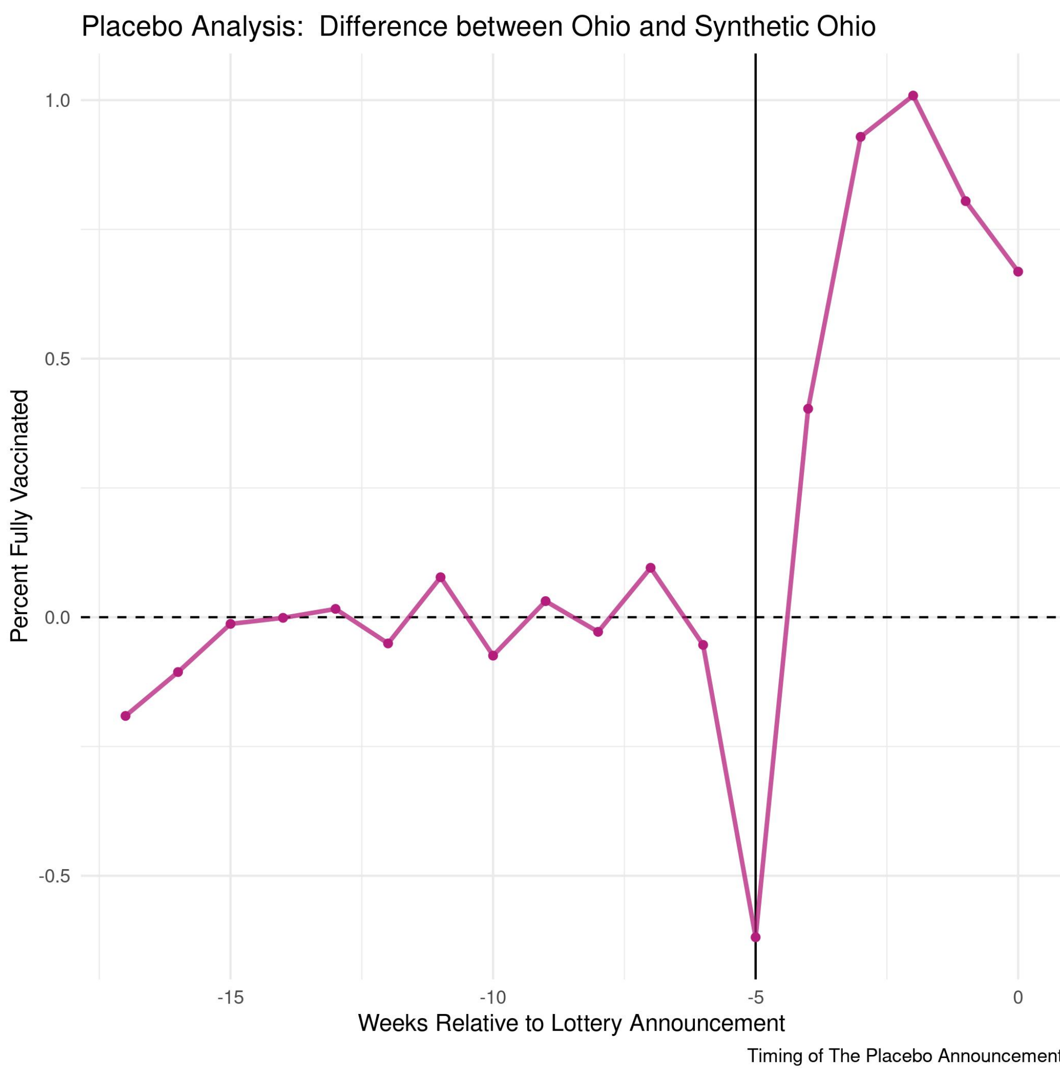

FIGURE 5: Placebo difference in vaccination rates between Ohio and Synthetic Ohio, with a false treatment date set 5 weeks prior to the true treatment date. Positive values show Ohio with a higher percentage of the population fully vaccinated than the synthetic comparison.

Error (MSPE) between our pre-treatment and post-treatment data for each state. We then sorted them in descending order based on the ratio of MSPE and used the associated rank for each state as it's associated p-value (See Figure 6). In the case of our placebo analysis, synthetic Ohio had a rank of 41 out of 51 units and an associated p-value of 0.804, indicating that we fail to reject null effects.

For our actual analysis, we repeated these exact same steps using the 17 weeks prior to the lottery announcement as our pre-treatment period and the six weeks following the lottery announcement as our post-treatment period. Failure to reject the null effect hypothesis is not interpreted as proof of null effects.

To describe the net effect of the program, we took the point estimate from the last period's difference between actual Ohio and synthetic Ohio. In the case of Figure 5, our point estimate would suggest the lottery program increased participation by 0.6% . We also computed the average difference between synthetic and actual Ohio across the treatment period. This information can be quite descriptive if, for instance, a program had no effect in the long run but encouraged some individuals to get vaccinated several weeks earlier.

B Power Analysis

We conducted power analyses with different potential effect sizes. We generated subsequent outcomes assuming that states would continue to grow at their weekly rate as sampled from historical mean and standard deviation (See Table 5). We truncate these distributions such that vaccination rates cannot decrease from week to week. We then assumed that the effect of the lottery would have an increase between zero and two percentage points per week.

We computed 200 bootstrap simulations and conducted our permutation tests. Based on Figure 7, we were reasonably powered to detect effect sizes on the order of 1.75% percentage points or larger using a p-value cutoff of 0.10, this correspond to the top 5 states. The associated power with this cutoff is 0.97. This effect is roughly on the order of the absolute effect associated with compensating individuals to receive the HPV vaccine, which saw between a 9.8% to 13.2% percentage point increase in first-time vaccination rates (Mantzari et al., 2015). To put this in context of statewide vaccination rates at the time of the lottery announcement, Ohio has a fully vaccinated rate of 37%. Such an effect would make it the second-most vaccinated state in the country just behind Maine at 48%.

C Fifty State Specification

In our original pre-registration we did not anticipate that other states would so quickly follow Ohio's lead in adopting lottery sweepstake prizes. See Table 6 for a list of state vaccine lottery announcements. We modified our OSF pre-registration on June 15, 2020 to note that we would exclude states that adopted lotteries from the synthetic comparison. This was before the final two lottery drawings in Ohio occurred. In the interest of full

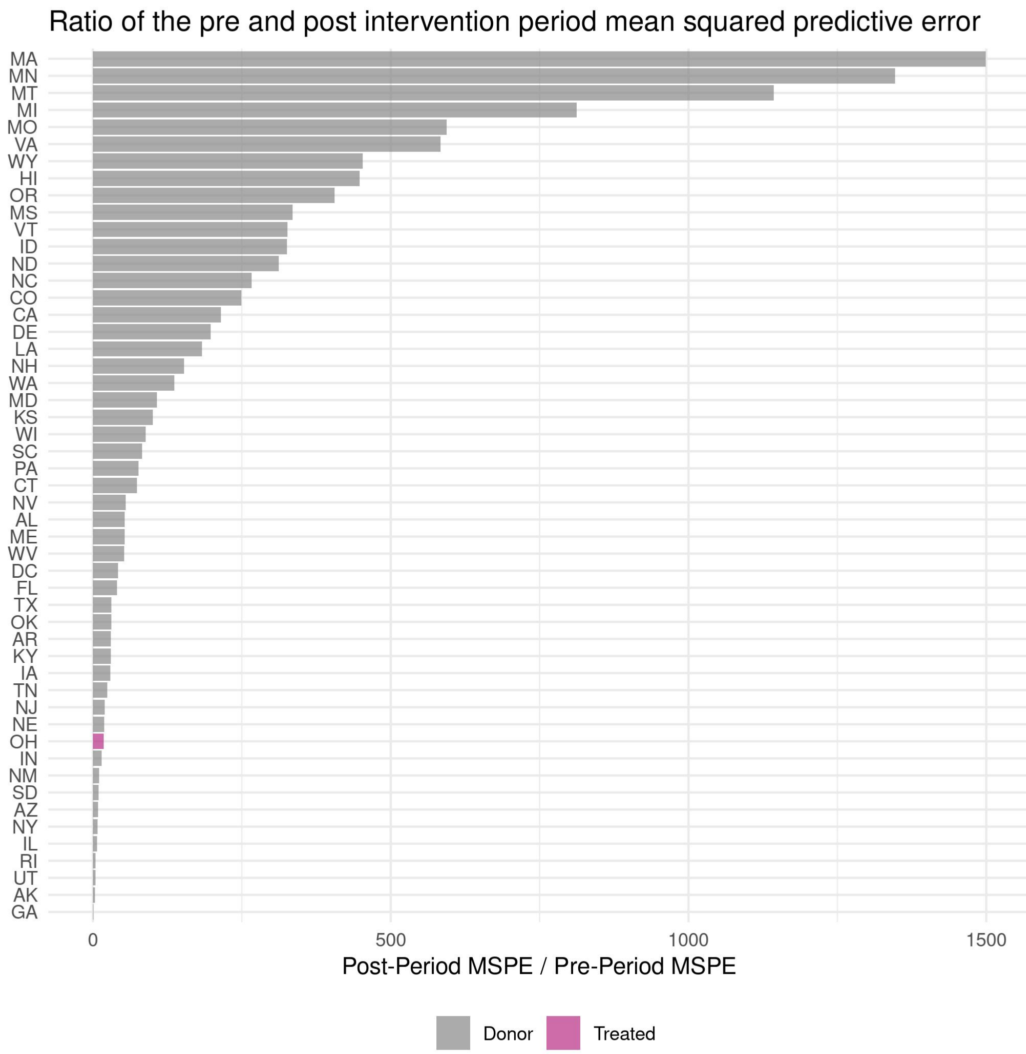

FIGURE 6: Placebo Test

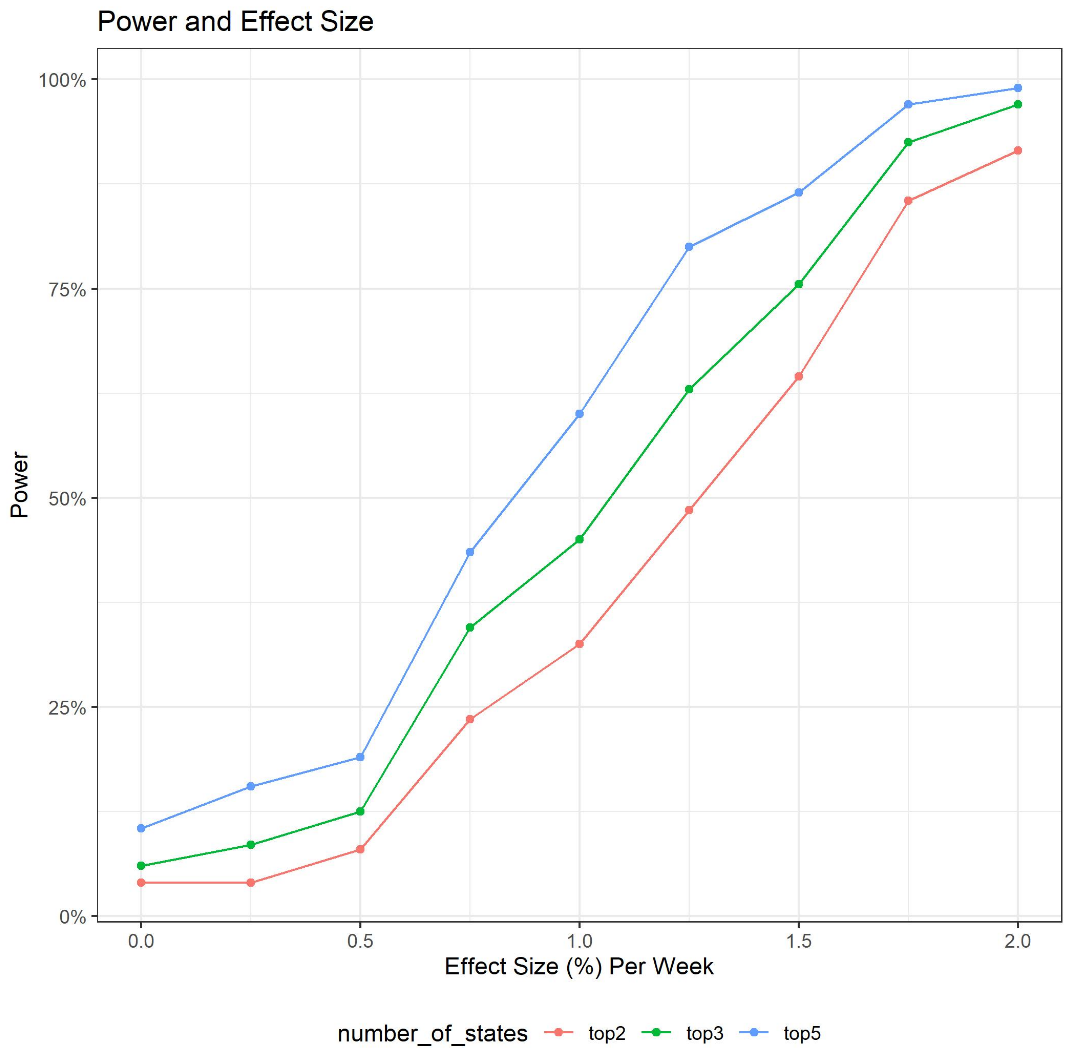

FIGURE 7: Effect Size and Power

TABLE 6: State Lottery Announcement Dates as of July 2nd, 2021

State	Lottery Announcement Date
OH	5/13/21
MD	5/20/21
NY	5/20/21
OR	5/21/21
AR	5/25/21
CO	5/25/21
DE	5/25/21
CA	5/28/21
NM	6/1/21
WV	6/1/21
WA	6/3/21
KY	6/4/21
NC	6/10/21
MA	6/15/21
ME	6/16/21
NV	6/16/21
IL	6/17/21
LA	6/17/21
MI	6/30/21

transparency, we present here results using our original pre-registration that includes all 50 states in the synthetic comparison. The pre-registered weights for the composition of the synthetic control can be seen in Table 7 and are presented alongside the weights used in the primary analysis. The most notable change is that Delaware was originally part of the synthetic control for Ohio, but due to their lottery adoption ended up being excluded. This shifted weights on other states, causing Alaska to be removed as well and adding Pennsylvania.

The quality of the match between Actual Ohio and this version of Synthetic Ohio exhibits marginally better fit in the pre-treatment period. In total the error in this period is at most 0.6% in any given week (see Table 8).

We present differences between Actual and Synthetic Ohio in Figure 8. Through the entire treatment period, vaccination rates for Ohio are below our synthetic counterfactual. At the end of the period, this difference was approximately -0.9%. The associated confidence interval associated with this point estimate is between -2.4% and 0.6%.

We present the same set of outcome measures as in our main analysis in Table 9 below. The associated p-value for our pre-registered metric is created from an MSPE ratio rank of 29/51 which is approximately 0.57. Related measures such as the average difference or end of period differences are also negative but not statistically significant. These results are not substantively different from those we present in the main body of the paper.

TABLE 7: Synthetic Ohio weights, including states that also adopted lotteries from the donor pool. For comparative purposes we include here the weights used in the main analysis, excluding other lottery adopting states.

Unit	Including Lottery State Weights	Excluding Lottery State Weights
AK	0.009	0.000
CT	0.060	0.029
DE	0.128	Excluded
GA	0.160	0.168
HI	0.035	0.061
IA	0.039	0.066
KS	0.319	0.256
PA	0.000	0.056
VA	0.080	0.173
WI	0.170	0.192

TABLE 8: Balance Table (Alternative Specification)

Pretreatment Outcome	Ohio	Synthetic Ohio	Difference	Donor Pool
lagged_vaccinations_week17	0.120	0.398	−0.278	0.556
lagged_vaccinations_week16	0.610	0.839	−0.229	1.083
lagged_vaccinations_week15	1.400	1.472	−0.072	1.871
lagged_vaccinations_week14	2.440	2.451	−0.011	2.961
lagged_vaccinations_week13	3.830	3.745	0.085	4.346
lagged_vaccinations_week12	5.560	5.673	−0.113	6.157
lagged_vaccinations_week11	7.670	7.651	0.019	7.972
lagged_vaccinations_week10	9.440	9.538	−0.098	9.809
lagged_vaccinations_week09	11.870	11.777	0.093	12.057
lagged_vaccinations_week08	13.860	13.823	0.037	14.081
lagged_vaccinations_week07	16.130	16.042	0.088	16.373
lagged_vaccinations_week06	18.780	18.748	0.032	19.368
lagged_vaccinations_week05	21.610	22.181	−0.571	22.757
lagged_vaccinations_week04	26.320	26.101	0.219	26.153
lagged_vaccinations_week03	30.110	29.802	0.308	29.197
lagged_vaccinations_week02	33.230	33.076	0.154	32.036
lagged_vaccinations_week01	35.620	35.822	−0.202	34.709

TABLE 9: Outcome Table (Alternative Specification)

Measure	MSPE-Ratio	Average Difference	Last Period Difference
Value	15.7	−0.81	−0.90
Rank	27	32	30
p-value	0.53	0.63	0.59

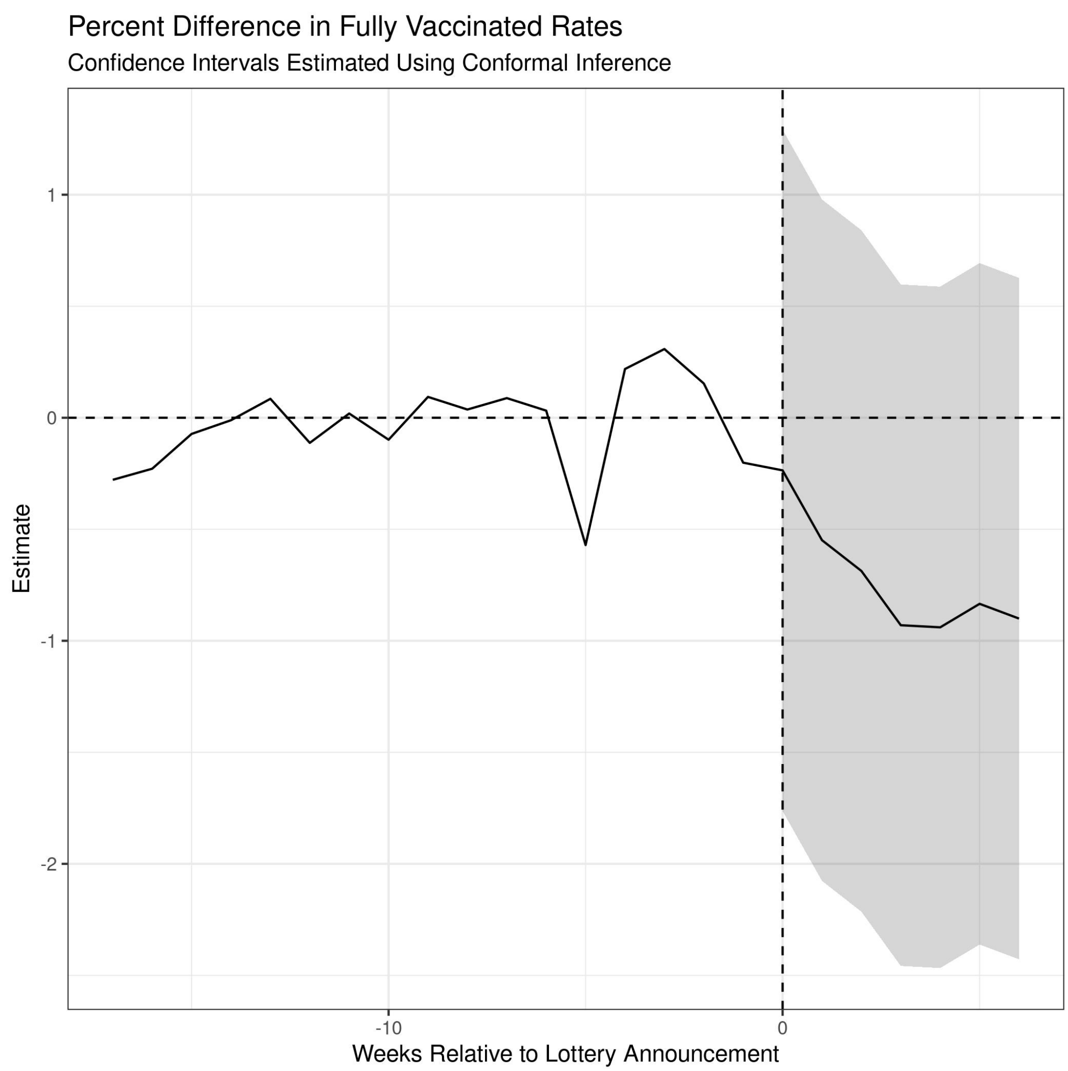

FIGURE 8: Alternate specification difference in Vaccination rates between Ohio and Synthetic Ohio, all states that adopted lotteries after Ohio have been included. Negative values show that Ohio has a lower total vaccination rate than the synthetic comparison

D Existing Point Estimates

One benefit of the synthetic control method is that only the primary outcome variable and associated weights are necessary to replicate and extend estimates from other model specifications. Here we have replicated the first dose effect estimates reported in two other synthetic control papers (Barber and West, 2021; Sehgal, 2021) and extended the estimates through August 22nd, 2021 using the weights available in those papers. We also present estimates for a first dose model that we estimate that is identical to the model we present in the main body of this paper except for the choice of outcome variable. While there may be some minor differences between these estimates and their original authors' work due to different data pre-processing decisions, these results are broadly comparable.

The results are presented in Figure 9. We note that Sehgel (2021) trained their weights on only 30 days of data before the lottery announcement and thus it is expected that their pre-fit quality would degrade for earlier dates. All three models show positive point estimates during the lottery period. However, all models show that the effect rapidly turned negative after the lottery ended. This suggests that the lottery may have shifted some individuals to get vaccinated earlier than they would have otherwise, but in the long term vaccination rates in Ohio fell below the synthetic control group [5]. These findings highlight the importance of researchers aligning the effect studied with the policy relevant outcomes that are most important to change, and checking for the persistence of effects over the long term. We note here that long term is denoted by extending the study period by a matter of weeks.

E Exploratory analysis of multiple lottery announcements

After the initial positive news coverage of the Ohio lottery and with encouragement from the White House (White House Press Briefing, 2021), seventeen states have so far followed Ohio's lead by announcing lotteries of their own (see Table 6). We estimate a synthetic control model that explicitly allow for multiple treated units with differential treatment timing. This analysis was not included in our pre-registration plan and is therefore included here as exploratory. We present the descriptive trends of total vaccination rates across states in Figure 10. In this figure we see that states with high vaccination rates, like Massachusetts and Maine, and states with low vaccination rates, like Louisiana and Arkansas, have adopted lottery incentives. Descriptively, we see that most states appear to maintain a roughly constant relative ranking after adopting the lottery incentive, which suggests that lotteries did not have large effects on vaccination rates.

We use the augmented synthetic control method to estimate the average treatment effect across states that adopt lotteries (Ben-Michael et al., 2021a, 2021b). This method is a natural

[5] We note that Kansas had data revision on July 23rd `https://www.kansasvaccine.gov/158/Data` . Exclusion of Kansas from the synthetic control does not change the sign of our estimates. Seghal's estimates do not attribute any weight to Kansas and finds a similar estimate.

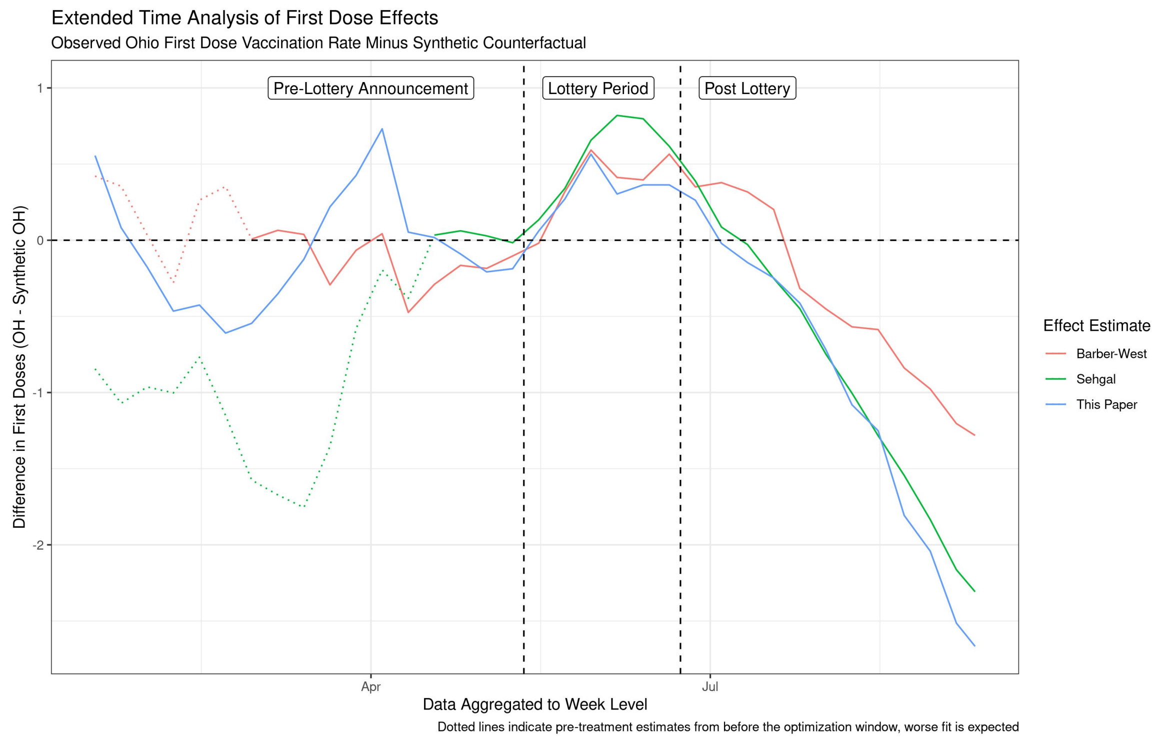

FIGURE 9: Comparison of estimates of effects of lottery on first doses(Barber and West, 2021; Sehgal, 2021)

extension of our pre-registered plan. We preserve the same outcome - the percent of the population fully vaccinated - for this analysis. We examine the change in vaccination rates up to 12 weeks after a state's initial lottery announcement. This allows us to detect whether the lottery had any lasting impact on vaccination rates.

This approach does have several key distinctions from the traditional synthetic control approach for the single-state case. First, it provides more flexibility in terms of the possible search space for generating the synthetic control, allowing weights to be negative and a unit-intercept term. Second, it adds regularization to the construction of the match, both to adjust for over-fitting and to help ensure unique solutions to the optimization. Third, it allows flexibility to balance the quality of match for an individual treated state and the composite average of all treated states.

The results of this analysis are presented in Figure 11. We now observe a small, statistically insignificant, average increase of 1.0 percentage points per week in the fully vaccinated rate of states that adopt lotteries relative to the synthetic counterfactual. This effect is consistent with another recent working paper that analyzed twelve state lotteries and also found small positive effects in ten of the lotteries they studied (Robertson et al., 2021).

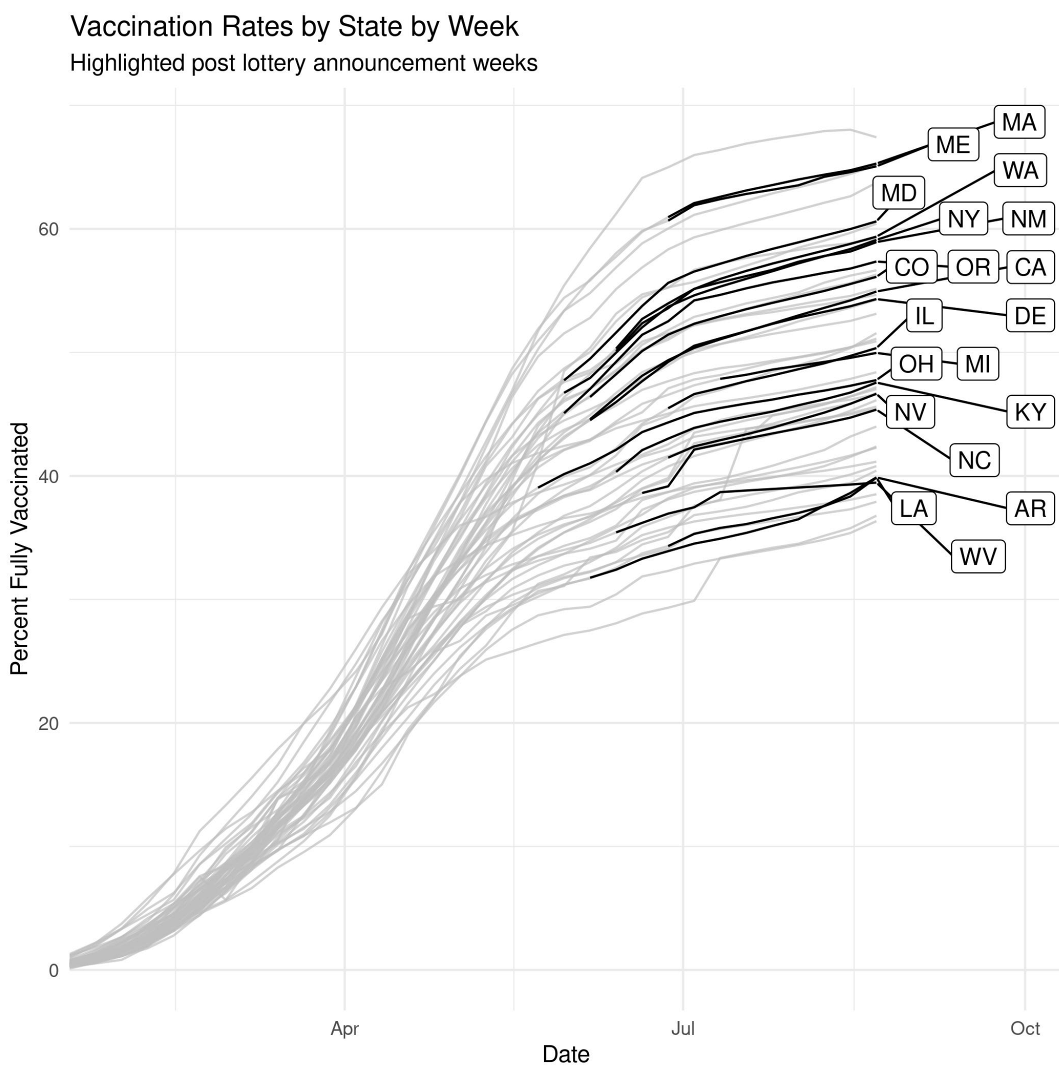

FIGURE 10: Vaccination Rates by State with Lottery Adoption Highlighted

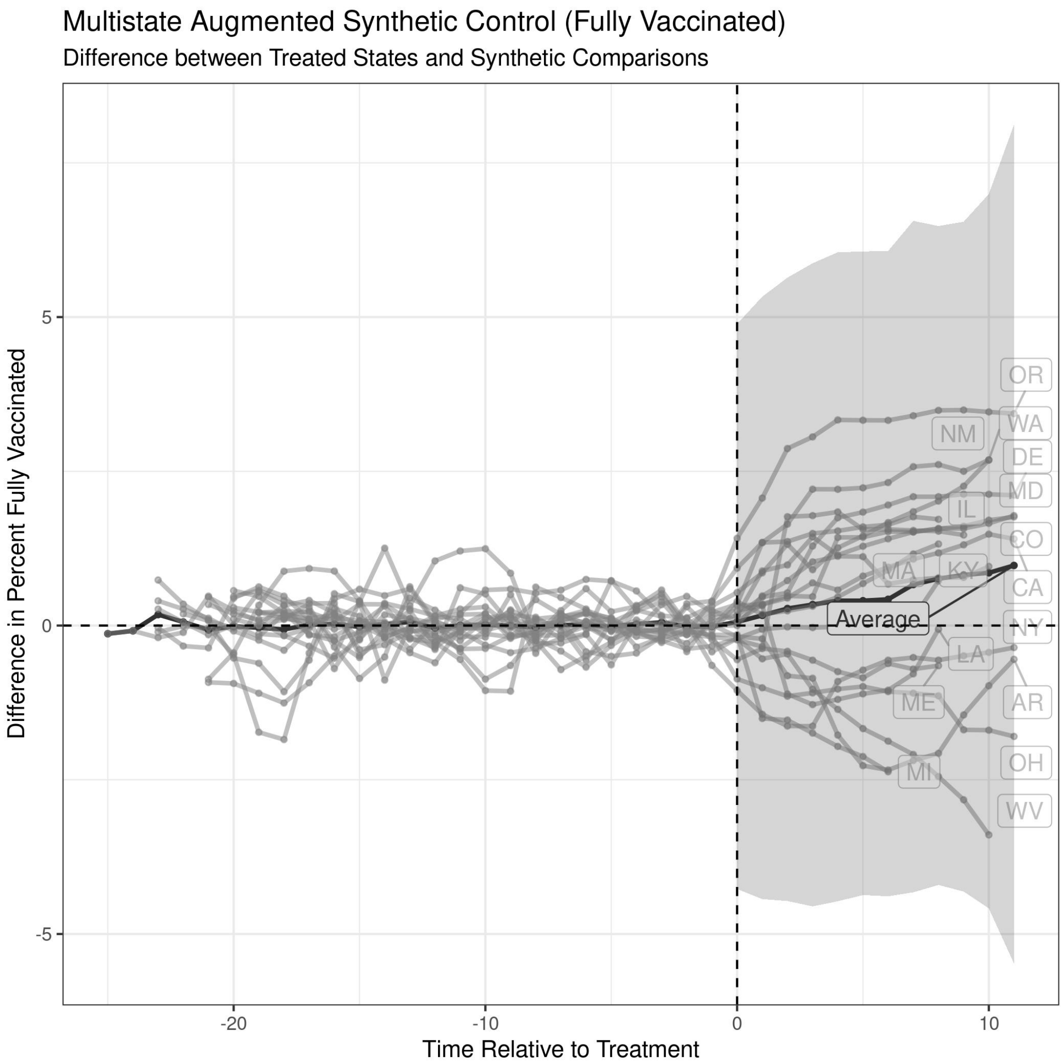

FIGURE 11: Augmented Synthetic Comparison with Multiple Adopting States