Conditional Cash Lotteries Increase COVID-19 Vaccination Rates

Andrew Barber[a] Jeremy West[a]*

[a]University of California, Santa Cruz

Manuscript accepted December 2021 at the

Journal of Health Economics

Abstract

Conditional cash lotteries (CCLs) provide people with opportunities to win monetary prizes only if they make specific behavioral changes. We conduct a case study of Ohio's Vax-A-Million initiative, the first CCL targeting COVID-19 vaccinations. Forming a synthetic control from other states, we find that Ohio's incentive scheme increases the vaccinated share of state population by 1.5 percent (0.7 pp), costing sixty-eight dollars per person persuaded to vaccinate. We show this causes significant reductions in COVID-19, preventing at least one infection for every six vaccinations that the lottery had successfully encouraged. These findings are promising for similar CCL public health initiatives.

JEL: D78, H41, I18

Keywords: health policy, financial incentives, behavioral economics

*Barber: abarber1@ucsc.edu. West (corresponding author): westj@ucsc.edu. We thank Carlos Dobkin, Laura Giuliano, Ben Hansen, and David Lang for helpful comments. The conclusions and any errors in this paper are those of the authors. We have no material financial interests related to this research.

1 Introduction

Providing safe and effective COVID-19 vaccines to the public only nine months after declaring the pandemic is a remarkable feat of science and policymaking. Vaccine development is only the first hurdle, however, because community (herd) immunity requires a large share of the population to be vaccinated. Overcoming widespread reluctance to vaccinate remains a significant challenge, especially as "waning vaccine confidence has taken a toll on immunization programs across the globe" in recent years (de Figueiredo et al., 2020).

From a decision-making perspective, a person's choice to (not) be vaccinated boils down to whether their expected benefit—including altruistic benefit—outweighs their cost of vaccination. The United States and other governments have greatly reduced this cost by making COVID-19 vaccines free of charge, offering free transportation to vaccination sites, and providing easily accessible facts about the vaccines to smooth any information frictions. Despite these efforts, many people remain unpersuaded. In Figure 1, we use data from the U.S. Census Bureau's Household Pulse Survey to plot COVID-19 vaccine hesitancy rates by state.[1] Although there is considerable heterogeneity, ranging from 7.3 percent hesitancy in Washington, D.C. to 31.6 percent in Wyoming, it is clear that much of the U.S. population remains unwilling to vaccinate despite essentially all vaccination costs being eliminated.

Motivated by this hesitancy, a number of states have attempted to nudge people towards vaccination by also boosting the expected benefits of being vaccinated. The most prominent form of these incentive schemes, which we refer to as a conditional cash lottery (CCL), provides people with an exclusive opportunity to win large monetary prizes only if they have received a COVID-19 vaccine.[2] A CCL is similar to a conditional cash transfer in that both incentives require people to make specific behavioral changes; however, the prize-based nature of a CCL is an important distinction. Drawing on insights from behavioral economics, CCLs capitalize on "probability neglect," a cognitive bias wherein low-probability events are either neglected entirely or hugely overrated (Sunstein, 2002). Appealing to this behavioral bias can be particularly useful for public health objectives like vaccinations because CCL incentives should predominantly encourage people who both under-estimate communicable disease risks and over-estimate their likelihood of winning a lottery prize.

In this paper, we study the first CCL targeting COVID-19 vaccinations, which Ohio

[1] Vaccination hesitancy includes responses of "definitely not" and "probably not" as survey respondents' stated willingness to be vaccinated. Figure 1 uses data from February 17 to May 10, 2021, including all available vaccine hesitancy data provided by the survey prior to the intervention in Ohio that we study.

[2] We provide information about each of the state COVID-19 lottery initiatives in Appendix Table A1. In total, states have committed more than $200 million in CCL prizes for vaccinated individuals to-date.

Governor Mike DeWine announced on May 12, 2021.[3] Run by the state's Department of Health, the Vax-A-Million campaign consisted of a weekly drawing each Wednesday from May 26 through June 23, with each of the five drawings awarding one adult (18+) a prize of one million dollars and one youth (12-17) a full scholarship to any public college or university in Ohio. The total program cost was 5.6 million dollars (DeWine, 2021). A free registration provided entry into all remaining prize drawings, with the entry deadline for the final drawing ending at midnight on June 20. Importantly, only state residents who had received at least one dose of a COVID-19 vaccine prior to a drawing were eligible to win.

We evaluate how Ohio's CCL treatment affects COVID-19 vaccinations and infections by comparing how these outcomes change over time in Ohio relative to a Synthetic Ohio constructed from a weighted average of other states. To obtain this counterfactual, we employ the ridge augmented synthetic control method (SCM) developed by Ben-Michael et al. (2021), which improves on the pioneering SCM work of Abadie and Gardeazabal (2003) and Abadie et al. (2010). Whereas the classic SCM forces all unit weights to be non-negative, potentially yielding a poor pre-treatment fit of the model, the ridge augmented version allows for negative weights by modifying the synthetic control estimation via a ridge regression outcome model. The ridge regularization parameter penalizes the distance from classic SCM weights, so this approach cleverly de-biases the synthetic control estimates while also minimizing extrapolation from untreated states' convex hull. It additionally allows for incorporating pre-treatment covariates to further improve the model fit.[4]

Our study uses daily state-level data primarily from the U.S. Department of Health and Human Services and the Centers for Disease Control and Prevention (CDC). Our outcomes are COVID-19 vaccinations, COVID-19 cases (positive tests), and COVID-19-related intensive care unit (ICU) patient-days. Because state populations vary and our outcomes of interest grow monotonically over time (e.g. the total count of vaccinated people), we specify each dependent variable cumulatively as a ratio to state population (e.g. the vaccinated

[3]It is almost a fluke that Ohio was the first government to implement a CCL. Of course, most policies are not implemented at random, but Ohio, which had been plagued with a high rate of vaccine hesitancy from the outset of the pandemic, also had the distinction of being the first state to close schools for in-person learning. The combination of these two details may have motivated the state to be willing to take a radical approach to boost vaccination rates, as described by Ohio Governor DeWine in a NYT op-ed in May 2021: "In mid-April, vaccine supply in Ohio started exceeding demand. There was a striking — and scary —drop in vaccinations. People were tired. They wanted their lives back. The vaccines were our pathway out, yet fewer and fewer were getting them. Every day I asked my staff, 'What else can we do?' ... My chief adviser ... hesitantly suggested the idea of a lottery. She almost didn't mention it because of its seeming absurdity. 'This is kind of a wacky idea, but ...' " (DeWine, 2021).

[4]As we discuss and show below, our empirical findings are robust to instead using the classic synthetic control model, to including no covariates, and to a large variety of alternative model specifications.

share of state population). We also incorporate several covariates that capture residential, political, behavioral, and supply-side factors related to accessibility of or preferences about the vaccines. Our study period spans from February 19, 2021, the earliest comprehensive data on vaccinations, to July 18, 2021, 28 days after Ohio's lottery entry ended.

We find an increase in COVID-19 vaccinations in Ohio that begins almost immediately after the Vax-A-Million announcement and persists past the final prize drawing. Relative to the synthetic control, the program causes a 0.7 percentage points (1.5 percent) increase in the share of state population receiving at least a first dose of a COVID-19 vaccine by the program's end date, with most of this effect occurring within two weeks of the announcement. In levels, this amounts to about 82,000 people who were persuaded to vaccinate by the CCL incentive, implying an average program cost of 68 dollars per "complier." For context, this cost-per-complier is less than the 80 dollars in direct costs that the federal government pays a healthcare provider to fully vaccinate one person (U.S. CMS, 2021).

In turn, we find that this heightened level of vaccination subsequently reduces the spread and impact of COVID-19 within the state. Using the same framework, we estimate that Ohio's program reduces case volumes by around 125 per 100,000 population (1.3 percent) and COVID-19-related ICU patient-days by around 41 per 100,000 population (2.6 percent) by the end of our study period. In aggregate, these estimates correspond to nearly 15,000 cases and 5,000 ICU patient-days prevented (approximately 335 patients). Moreover, because of the exponential nature of disease transmission, these estimates are likely to greatly understate the total longer-run reductions relative to counterfactual.[5]

Inference is almost always the most challenging aspect of using a synthetic control method, given that there is only a single treated unit. Following Abadie et al. (2010), many SCM studies resort to a form of cross-sectional permutation inference by estimating "placebo effects" for untreated units—comparing the ratio of the post-treatment root mean squared prediction error (RMSPE) and pre-treatment RMSPE of each of these estimates to that for the treated unit. We also conduct this RMSPE-based inference, finding strong support for Ohio's candidacy as an outlier among placebo-treated states. However, as Abadie (2021) discusses, this approach suffers from some "complications," most notably that it only "reduces to classical randomization inference when the intervention is randomly assigned, a rather improbable setting." In our study context, this condition is equivalent to the assumption that Ohio is a randomly-selected state to have implemented the first CCL for COVID-19

[5]We do not attempt to model the long-run effective reproduction number of infections prevented. Such an exercise is complicated because of new genetic variants of SARS-CoV-2 and because the basic reproduction number (R_0) for compliers encouraged by the incentive likely differs from that of the broader population.

vaccinations, which—as supported by Figure 1 above—is unlikely to be the case.

Fortunately, substantial progress has been made recently in the econometrics and statistics literature pertaining to inference for the SCM. These modern approaches extend the conformal prediction techniques of Vovk et al. (2005) to leverage the pre-treatment time series variation of the treated unit and synthetic control, rather than relying on cross-sectional comparisons. For this study, we primarily conduct inference for our estimates using the conformal inference method developed by Chernozhukov et al. (2021) and applied to the ridge augmented SCM by Ben-Michael et al. (2021). Conformal inference operates by extending the estimation period to include post-treatment observations, re-estimating the unit weights using this extended period, and then testing whether the estimated treatment effect using these adjusted weights "conforms" with that estimated using the original unit weights. The key assumption for this approach is that, under the null hypothesis, the distribution of differences between the treated unit and control unit is stationary over time. We provide support for this assumption below.

Our paper makes several contributions to the health economics literature. Most directly, we provide one of the only examinations of a large-scale conditional cash lottery. Although lottery-based incentives have been used conceptually for over sixty years to encourage behavior change related to public health (British Medical Journal, 1957), the limited empirical evidence is somewhat mixed and focuses primarily on smaller interventions in clinical trials or field experiments.[6] Moran et al. (1996) find that a lottery-based gift card incentive is less effective than an educational brochure at encouraging influenza vaccinations. Volpp et al. (2008a,b) test small CCL incentives for losing weight and for anticoagulant drug adherence, finding success in encouraging behavioral change. Thirumurthy et al. (2016) show that offering lottery prizes does not increase voluntary medical circumcision by men in Kenya. In two recent field experiments, Goette and Stutzer (2020) find that blood donors in Switzerland are more likely to donate again when offered a lottery ticket and Björkman Nyqvist et al. (2018) find that a CCL for safer sexual behavior in Lesotho reduces HIV incidence. Additionally, this latter study demonstrates that lottery-based incentives primarily appeal to individuals with greater risk tolerance, further supporting a mechanism of probability neglect.

Several concurrent research studies also examine aspects of Ohio's Vax-A-Million program. Walkey et al. (2021) conduct an interrupted time series study of Ohio versus the United States during the few weeks surrounding the lottery announcement, concluding that

[6] A larger related literature examines conditional cash transfers for public health objectives including vaccinations (e.g. Barham and Maluccio, 2009; Campos-Mercade et al., 2021; Chang et al., 2021). As noted above, the uncertainty in CCLs is a key distinction.

Ohio's program does not increase vaccination rates. However, Ohio's vaccination rates track poorly with national rates during the pre-treatment period—a factor motivating our synthetic control identification strategy. Lang et al. (2021) use the classic SCM to study how Ohio's program affects the share of *fully* vaccinated residents, finding no effect. This null effect could be because lottery eligibility only required a single dose rather than full vaccination. In addition, their study stops at the final lottery drawing, weeks before many lottery-eligible participants could have obtained a second dose of a vaccine series, which require 21 or 28 day gaps between doses. Finally, Brehm et al. (2021) use county-level data from Ohio, Indiana, Michigan, and Pennsylvania to conduct pooled SCM and state-border difference-in-differences estimations of how Ohio's program affects the number of of first dose vaccinations during the treatment period. The study finds an effect on vaccinations that is very similar in magnitude to that we show here.

To our knowledge, we provide the only evidence about how Ohio's CCL initiative ultimately affects COVID-19 infections and hospitalizations, through a mechanism of increased vaccinations. We believe this is a critically important dimension for evaluating this novel policy instrument because it facilitates conducting a cost-benefit analysis of what has appeared, at least initially, to be a controversial program (Buchanan, 2021). Collectively, our estimates indicate that—by nine weeks after the announcement of the program—Ohio's CCL prevents at least one COVID-19 infection for every six vaccinations that the lottery successfully encourages and prevents at least one ICU patient-day for every 17 vaccinations that it encourages. As noted above, these effects are only growing stronger over time due to the exponential nature of disease transmission. Based on Di Fusco et al.'s (2021) values for COVID-19-associated ICU expenses, our estimates imply a reduction in hospital charges of around 66 million dollars, a social benefit that is an order of magnitude larger than the 5.6 million dollar cost of the program. Thus, even without including any other short- or long-run benefits from reducing COVID-19 incidence, Ohio's CCL program passes an economic cost-benefit analysis with flying colors.

With these findings, we also contribute to the literature evaluating how COVID-19 vaccination rates affect community infections. We provide evidence specific to the subset of the population that is persuaded to vaccinate only by a lottery-based financial incentive, in contrast to evidence from vaccinations of people motivated by altruistic reasons or seeking self-protection from the virus. Against a backdrop of increasing hesitancy globally towards vaccinations, this distinction could be quite valuable for public health policymakers. Inspired by Ohio's approach, at least nineteen other state governments have followed suit with

their own "vaccination lotteries," with substantial heterogeneity in programmatic design. We leave it to future research to provide further insights about what, specifically, serves as the optimal form of conditional cash lottery to encourage COVID-19 vaccinations.

2 Methods

Our primary empirical strategy to estimate the effects of Ohio's lottery incentive treatment is the ridge augmented synthetic control method (Ben-Michael et al., 2021). At its core, this approach compares outcomes in Ohio to outcomes in other states over time. As we show below, Ohio's vaccination rates do not track closely with overall rates in the United States even in the weeks before the Vax-A-Million program announcement, such that a simple average across other states serves as a poor counterfactual. By using the synthetic control method (SCM) to form a weighted average of the untreated states, we obtain a much better counterfactual for Ohio. Here, we provide only a basic illustration of the method, referring interested readers to Ben-Michael et al. (2021) and Abadie (2021) for additional details.

For panel data on states i across time periods t, denote the outcome variable as $y_{i,t}$. We are interested in the treatment effect, τ^{CCL}, of a conditional cash lottery on this outcome. Suppose for simplicity that only Ohio is ever treated and that there is only a single post-treatment period when $t = T$. In a potential outcomes framework, we can express Ohio's post-treatment outcome as $y_{Ohio,T} = y_{CF,T} + \tau^{CCL}$, where $y_{CF,T}$ is the counterfactual at time T. The essence of the SCM is to form this counterfactual for post-treatment Ohio as a weighted-average of the outcome for untreated states:

$$y_{CF,T} = \sum_{i \neq Ohio} \hat{\gamma}_i^{scm} y_{i,T} \qquad \text{where} \qquad \sum_{i \neq Ohio} \hat{\gamma}_i^{scm} = 1 \qquad \text{and} \qquad \hat{\gamma}_i^{scm} \geq 0 \ \forall i \quad (1)$$

The classic SCM determines these $\hat{\gamma}_i^{scm}$ weights by minimizing the differences between Ohio and the counterfactual for the outcome in pre-treatment time periods, $t < T$, as well as optionally minimizing differences in covariates between Ohio and the counterfactual. Denote the vector of pre-treatment outcomes and covariates for a state as X_i. In an ideal setting, the SCM weights would yield a near-perfect counterfactual for Ohio, i.e. $X_{Ohio} \approx \sum_{i \neq Ohio} \hat{\gamma}_i^{scm} X_i$. In practice, it may not be feasible to determine a set of non-negative weights such that $X_{Ohio} \approx \sum_{i \neq Ohio} \hat{\gamma}_i^{scm} X_i$, and the synthetic control will yield a poor $y_{CF,T}$ counterfactual.

To improve the quality of the counterfactual, the ridge augmented SCM layers a ridge

regularized linear model onto the classic SCM:

$$y_{CF,T} = \sum_{i \neq Ohio} \hat{\gamma}_i^{scm} y_{i,T} + \left(X_{Ohio} - \sum_{i \neq Ohio} \hat{\gamma}_i^{scm} X_i \right) \cdot \hat{\eta}^{ridge} \tag{2}$$

where $\hat{\eta}^{ridge}$ are coefficients from a ridge regression of the untreated states' post-treatment outcomes $y_{i,T}$ on centered pre-treatment outcomes X_i, with a tuning parameter that limits the degree of extrapolation from the untreated states' convex hull. If the quality of the classic SCM counterfactual is very good, then $X_{Ohio} - \sum_{i \neq Ohio} \hat{\gamma}_i^{scm} X_i$ is close to zero and the ridge augmented SCM is virtually equivalent to the classic SCM. For nontrivial cases, Ben-Michael et al. (2021) demonstrate how Equation (2) can be expressed as:

$$y_{CF,T} = \sum_{i \neq Ohio} \hat{\gamma}_i^{aug} y_{i,T} \qquad \text{where} \qquad \sum_{i \neq Ohio} \hat{\gamma}_i^{aug} = 1 \tag{3}$$

Although $\hat{\gamma}_i^{aug}$ can take negative values, unlike $\hat{\gamma}_i^{scm}$, the method directly penalizes the distance between the ridge augmented SCM weights and the classic SCM weights using the tuning parameter.[7] Thus, the ridge augmentation de-biases the classic synthetic control estimates to improve the quality of the counterfactual for Ohio while also minimizing extrapolation. Empirically, we find that the classic SCM improves the quality of the counterfactual for Ohio relative to a simple average of untreated states, and the ridge augmented SCM further improves the quality of the Synthetic Ohio to support causal inference.

3 Data

3.1 Data sources

Our study compiles data from a variety of public sources. Data on COVID-19 vaccinations are provided by the U.S. Centers for Disease Control and Prevention (CDC), which aggregates information from state and local health departments. Specifically, this dataset includes the total number of vaccines that have been administered as of each date for each

[7] We determine this tuning parameter using the cross validation techniques provided in the software package accompanying Ben-Michael et al. (2021), which draws on insights from Abadie et al. (2015). Specifically, the tuning parameter λ is selected by computing the leave-one-out cross validation MSE over pre-treatment time periods: $CV(\lambda) = \sum_{t<T} (y_{Ohio,t} - y_{CF,t}^{(-t)})^2$, where $y_{CF,t}^{(-k)} = \sum_{i \neq Ohio} \hat{\gamma}_{i(-k)}^{aug} y_{i,t}$ is the estimate of $y_{CF,t}$ where time period k is excluded in estimating $\hat{\gamma}_i^{aug}$. The algorithm selects the maximal value of λ with MSE within one standard deviation of the minimal MSE.

state. These vaccination counts are separated by manufacturer: Janssen (Johnson & Johnson), Moderna, and Pfizer. The daily counts are also separated into first and/or final dose vaccinations. Although some vaccination data is sparsely available for earlier time periods, February 19, 2021 is the first date on which all states report comprehensive data, and we use this date to start the panel used in our analysis. Inspecting the data, there are some clear inaccuracies in the daily counts of vaccinations—such as a negative amount of vaccines being administered—anomalies which are also discussed in the CDC's data documentation.[8] Most of these errors are simply misattribution of some vaccinations to a date the day before or after the vaccines were actually administered. To correct for these inaccuracies, we smooth vaccination counts for a small number of state-weeks containing these "outliers" using an approach that preserves the cumulative vaccination counts for each state in each week but reduces artificial noise from erroneous data classification.[9]

To assess COVID-19 outcomes, we use data from the CDC for the total cases recorded in each state by date and we use data from the U.S. Department of Health and Human Services for the total volume of hospital intensive care unit (ICU) patients with COVID-19 by state-date. We use an outcome measure of ICU patient-days rather than patient counts because the data report the daily number of ICU-hospitalized patients, who typically stay multiple days (an average of 14.7 days per Di Fusco et al. (2021), with substantial heterogeneity). Because state populations vary and our outcomes of interest grow monotonically over time (e.g. the total COVID-19 cases recorded), we specify each dependent variable cumulatively as a ratio to state population (e.g. the total cumulative COVID-19 cases per 100,000 population). We do so using state population data for 2020 from the U.S. Census Bureau. We do not evaluate deaths from COVID-19 because this outcome is statistically under-powered—there were a total of six deaths per 100,000 population in Ohio during the post-treatment period.

Although we also present results from models without covariates, to improve the model fit we incorporate some pre-treatment state-level covariates related to accessibility of or preferences about the vaccines. We include population density (Census Bureau) and gross domestic product per capita (Bureau of Economic Analysis) as rough proxies for variation across states in the living circumstances and economic activity that could influence vaccine

[8]Discussion of data anomalies and other data reporting considerations is provided in the CDC's data documentation available at www.cdc.gov/coronavirus/2019-ncov/vaccines/distributing/about-vaccine-data.html.

[9]Specifically, we tag outlier observations using a criterion of daily vaccination volumes being greater than twice the state's seven-day moving average, adjusted for state-specific day of the week. For each state-week containing an outlier, we reallocate the total weekly volume across the days of that week using state-specific day-of-week weights. This approach leaves total vaccination counts in each state-week unaffected, reallocating only within state-week. Less than three percent of observations are outliers necessitating these corrections.

hesitancy, either directly or through the heterogeneous impact of COVID-19 across states during the pandemic. We include 2020 Republican presidential vote share because political leaning has been linked to vaccination hesitancy (Ivory et al., 2021). States' pre-pandemic influenza vaccination rates for 2019 from the Centers for Medicare and Medicaid Services are included to capture variation in more general propensities towards vaccination. We use Google's Community Mobility Reports indices to capture variation in behavior as reflected in visits to different types of places during the pre-treatment period.[10] Finally, we compute distance measures of state population to vaccination sites using Census Block Group population centers and the locations of all COVID-19 vaccination sites in the U.S. from www.vaccinatethestates.com. We use the median distance and 95th percentile distance of population to vaccination sites to proxy for differences in vaccine accessibility.

3.2 Synthetic Ohio

We use the ridge augmented synthetic control method to determine state unit weights for Synthetic Ohio. The donor pool includes all states that did not initiate their own lottery schemes for COVID-19 vaccinations before the end of Ohio's Vax-A-Million program (i.e. we exclude states listed in Appendix Table A1, sans Michigan and Missouri). In some of our robustness checks, we relax this requirement to include all 50 states and Washington, D.C.

Table 1 shows the largest five unit weights for Synthetic Ohio, using an outcome of the share of state population with any COVID-19 vaccination. Appendix Table A2 shows the full set of unit weights for Synthetic Ohio, which includes some negative weights as discussed in the methodology section above. Other than Idaho, the five states that contribute the most towards Synthetic Ohio are all also located in the Midwest Census Region: Wisconsin, Kansas, Michigan, and North Dakota. Like Ohio, these five states generally have high levels of surveyed pre-treatment vaccine hesitancy, shown in Figure 1 discussed above (Michigan, the apparent exception, has a hesitancy rate of 14.7 percent, just below the bin cutoff).

Table 2 presents summary statistics for the United States, Ohio, and Synthetic Ohio. Panel [A] shows shows values for the dependent variables during the pre-treatment time period(s) indicated. Panel [B] shows values for state covariates during 2019 or 2020, as indicated, or over the full pre-treatment period of February 19 through May 11, 2021. In the first row, Ohio's vaccination rate closely matches that of the U.S. overall as of April

[10]Google's Community Mobility Reports provide proxies for movement over time across six categories of places: retail and recreation, groceries and pharmacies, workplaces, residential, parks, and public transit. Each index can range from −100 to 100. Google defines these using movement of people's cell phones to different places, with the baseline zero-value for each index set during January 3 through February 6, 2020.

2, 2021. However, the U.S. vaccination rate greatly outpaces Ohio's during the subsequent pre-treatment weeks such that, by the May 12 lottery announcement, Ohio's vaccinated population share lags the state average by almost four percentage points, a pattern that is shown even more clearly in the time series graphs presented in the next section. In contrast, the weighted average vaccination rate of the states that comprise Synthetic Ohio remains much closer to the rate in Ohio. The remaining rows of Table 2 also show a clear improvement of the counterfactual by using a synthetic control rather than a simple average of untreated states. The three outcomes we evaluate and almost all state covariates are (often much) closer between Ohio and Synthetic Ohio than between Ohio and the United States' average.

4 Results

We begin our analysis by examining vaccinations. We focus on first dose vaccination rates rather than fully-vaccinated rates because Ohio's Vax-A-Million program only required a single dose of any COVID-19 vaccine for eligibility.[11] Although completing a vaccine series provides more protection against the virus, even a single dose of the Moderna or Pfizer vaccine has been found to provide substantial immunity (Dagan et al., 2021). Regardless, as we discuss and show below, there is no difference in the vaccination series follow-up rates in Ohio from before compared to after the Vax-A-Million program.

Figure 2 plots first dose vaccination rates over time. As shown in Panel (a), there are virtually no differences between Ohio, Synthetic Ohio, and the entire United States in the share of vaccinated population until early April. Then, there is a structural break between Ohio and the rest of the country, with this gap growing over time. In contrast, the vaccination rate continues to be nearly identical between Ohio and the synthetic control until Ohio's conditional cash lottery treatment begins on May 12. This figure highlights the importance of using the synthetic control method because it shows how poorly a simple average of other states would serve as a counterfactual to Ohio, despite having had similar vaccination behavior in earlier months prior to the start of the treatment.[12]

The magnitude of the difference across time between Ohio's first-dose vaccination rate

[11] A single vaccine dose is also the requirement for all other state initiatives shown in Appendix Table A1 other than Massachusetts, which requires full vaccination for eligibility.

[12] While states had slightly different vaccination eligibility timelines, every adult in the United States was eligible for a COVID-19 vaccine by no later than April 19, 2021 (Biden, 2021). Another potential concern is that 12-15 year olds became vaccine-eligible at roughly the same time (May 10) as Ohio's lottery announcement (May 12). We address this potential confounding variation in a robustness check in Appendix Figure A1, where we restrict analysis to adults of age 18 and older. Our results are robust to this restriction.

and its counterfactual is shown in Panel (b) of Figure 2. The introduction of the lottery incentive causes an almost immediate increase in vaccination rates in Ohio compared to the synthetic control—the difference is larger only three days after treatment begins than on any date over the nearly three months pre-treatment. Following the announcement of the lottery, the estimated treatment effect increases sharply over the first two weeks before leveling off, which also coincides with the timing of the first prize drawing. Of more general interest, the alacrity with which compliers respond suggests that a long treatment window may not be required to maximize the efficacy of a CCL to change behavior.

Although increasing vaccination uptake is the most direct effect of Ohio's program, ultimately the objective is to reduce COVID-19 infections, hospitalizations, and deaths. We show the first of these downstream effects of Ohio's increased vaccination rate in Figure 3. Given that COVID-19 vaccines take approximately 14 days to demonstrate partial efficacy, one would not anticipate to see declining infection rates until at least the end of May (Dagan et al., 2021). Furthermore, due to the exponential nature of viral transmission, any effect that is observed should grow over time as each infection prevented then also prevents additional cases. This expected pattern matches the evidence shown in Figure 3. There is little effect of Ohio's program on COVID-19 cases until early June, but then the cumulative difference between Ohio and the control widens monotonically, becoming statistically significant about a month after the lottery announcement.

A relatively small subset of COVID-19 cases require intensive hospital care and, reflecting the overall decline in Ohio's infection rates, we find that the vaccine incentive causes a decrease in ICU utilization as well (Figure 4). Panel (a) shows the cumulative total COVID-19-induced patient-days spent in hospital ICUs per 100,000 population by region and date. Panel (b) plots the difference between Ohio and its synthetic counterfactual. As discussed just above, because of the delay in immunity from the vaccine, any potential effect should not be expected until early June. Moreover, we should also expect some additional delay of a few days between disease onset and admission to a hospital ICU (Wang et al., 2020). Given this mechanical lag, it is unsurprising to see no effect on ICU patient-days in Ohio relative to counterfactual until the second week of June. However, ICU patient-days then also begin to decrease monotonically over time compared to the control, reaching statistical significance soon after.[13]

Figure 5 shows the robustness of our estimated vaccination effects to different specifica-

[13]To facilitate a more direct comparison of the timing and magnitude of these respective treatment effects, Appendix Figure A2 plots the estimates over time for all three outcomes during the post-treatment period.

tions. Similar plots for cases and ICU patient-days are provided in Appendix Figures A3 and A4. Each row of the figure(s) plots point estimates and 95 percent confidence intervals from a different model that varies either the specification or donor pool. The first row reproduces the estimates shown above from our baseline model. In the second row, we show results from the classic synthetic control model without covariates. As discussed in Section 2, the classic SCM produces a somewhat worse pre-treatment fit of the Synthetic Ohio—supporting our use of the ridge augmentation—but the estimates remain relatively close to the baseline values. The next row shows results using residuals of the outcomes to the state covariates, finding very similar results as the baseline model, although the estimates become somewhat less precise. Finally, we vary the state donor pool to assess sensitivity, both by including all 50 states and (separately) by iteratively leaving out each potential donor state.[14] On the whole, this extensive set of robustness exercises provides compelling support for the causal inference of our analysis.

The magnitudes of the estimated effects are shown in Table 3. In Panel [A], we detail the evolution of the lottery's effect on cumulative vaccination rates throughout the treatment period. One week into the intervention, we estimate that there is a 0.31 percentage points increase in vaccination rates in Ohio relative to the counterfactual. By the second week into the program, the estimated rate of increased uptake is 0.6 percentage points. The estimated effect then remains relatively stable over the remaining three and a half weeks, reaching 0.7 percentage points by the end of the incentive program. Given that the counterfactual vaccination rate is 46.5 percent of population by the end of the treatment period, this end-line effect size corresponds to a 1.5 percent increase in the vaccinated share of Ohio's population. We find a similar effect for adults age 18 and older. These effects shown in the table are all statistically significant, with the lower bound of the 95 percent confidence interval at 0.15 percentage points after the first week and at 0.13 percentage points by the end of treatment.

Panel [B] presents estimates for the lottery's effect on COVID-19 cases and ICU patient-

[14]Appendix Figure A5 presents inference using the cross-sectional permutation approach of Abadie et al. (2010) by estimating "placebo effects" for untreated units—comparing the ratio of the post-treatment root mean squared prediction error (RMSPE) and pre-treatment RMSPE of each of these estimates to that for the treated unit. We find strong support for Ohio's candidacy as an outlier among placebo-treated states. Including placebo estimates for all states in the donor pool, the permutation based p-value for Ohio is 0.059 (2/34). Restricting the pool of placebo estimates to states with a pre-treatment root mean square prediction error that is no larger than five times that of Ohio's, the permutation based p-value for Ohio is 0.042 (1/24). Appendix Figure A6 shows the distribution of synthetic control placebo estimates for the share of state population with any COVID-19 vaccination as of two weeks following Ohio's CCL announcement, along with 95 percent confidence intervals calculated using conformal inference. Again, the exercise supports causal inference about Ohio's estimated treatment effect.

days, both at the end of treatment (June 20) and at the end of the data four weeks later (July 18). We find statistically significant estimates for both of these outcomes by the time the treatment period ends, with the effect sizes increasing in magnitude until (at least) the end of the analysis period. By end of sample, Ohio has 125.4 fewer total recorded COVID-19 cases per 100,000 population (a 1.3 percent reduction relative to the counterfactual) and 41.4 fewer COVID-19-induced ICU patient-days per 100,000 population (2.6 percent of the counterfactual). At a glance, it may seem implausible that a 1.5 percent increase in vaccination rates could reduce COVID-19 outcomes by these magnitudes; however, this underscores the importance of accounting for the exponential nature of communicable disease transmission. Ohio's incentive program essentially serves as a shock, producing a surge in vaccinations in the state over a fairly short window of time. In turn, this reduction in the infection-vulnerable population abates transmission to unvaccinated people to yield further reductions in infections, effectively "bending the curve" for exponential growth of COVID-19 within the state. In addition, the type of person who is persuaded to vaccinate by a financial lottery-based incentive might be especially valuable in curtailing the spread of the disease.

Using these estimates from Table 3 along with data for Ohio, Table 4 shows calculated aggregate effect sizes and characteristics by time period. By aggregating the per-capita effects that we observe over the post-treatment period using state population, we are able to compute the total number of compliers, COVID-19 cases prevented, and ICU patient-days averted due to Ohio's intervention. The first column of the table shows calculated values for the latest 40 days of the pre-treatment period (April 2 through May 11, 2021), and the remaining three columns show various time windows during the (post-) treatment period. In the third column, we observe that about 690,000 Ohioans in total received their first/only vaccine dose during the 40 day span of the Vax-A-Million program. Using the estimates from Table 3, we calculate that 82,000 (12%) of these people did so only because of the treatment, i.e. they are treatment compliers.

Among the compliers, 86 percent chose to get vaccinated within the first two weeks of the lottery, thus making them eligible for all of the five prize drawings. Comparing first dose vaccinations across vaccine manufacturers, Janssen (Johnson & Johnson) retained about 12-13 percent market share both before and during the lottery initiative, while Moderna's market share in Ohio shrunk from 33 to 26 percent and Pfizer's grew from 55 to 62 percent. Appendix Figure A7 shows daily time series of manufacturers' market shares in Ohio and the United States during our study period. On the whole, there is little evidence that Ohio's treatment compliers systematically selected a different vaccine mix than the vaccinated population

at large. The table also explores vaccination series follow-up. Although only first dose vaccinations were required for lottery eligibility, the evidence supports that compliers exhibit typical second dose follow-up rates. In the 40 days leading into treatment, 87 percent of people who start a Pfizer or Moderna vaccine series also obtain a second dose (using 21 and 28 day lagged windows, respectively, for counts of Pfizer and Moderna second dose vaccinations). During Ohio's program, this rate is 88 percent.[15]

Turning to the final column of Table 4, we find that Ohio's program substantially affects total COVID-19 cases and ICU utilization. By four weeks after the lottery's completion, we estimate that the program led to nearly 15,000 fewer cases and almost 5,000 fewer days spent in the ICU than would have occurred absent the lottery. Based on Di Fusco et al.'s (2021) values for COVID-19 ICU hospitalization, this amounts to about 335 fewer patients in the ICU for COVID-19-related complications. To reiterate, these aggregate effects only include reductions during our sample window, and the exponential nature of disease transmission implies that prevention of additional cases and ICU patients is likely as well.

5 Conclusions

At the time of this writing, projections show that more than thirteen million people will have died from COVID-19 worldwide by March 2022 (IHME, 2021). Of course, the realized extent of this death toll greatly depends on how many people are vaccinated. Safe and effective COVID-19 vaccines are freely available in many countries, including the United States, but this widespread vaccine availability is inadequate if a large portion of the population remains unwilling to vaccinate. A vaccine mandate could be used to increase vaccination rates (Abrevaya and Mulligan, 2011; Lawler, 2017), but making COVID-19 vaccinations mandatory is both publicly unpopular and politically tangled (Largent et al., 2020; Ivory et al., 2021). Ultimately, vaccination is a choice that depends on a person's beliefs about the benefits of being vaccinated (Auld, 2003).

To increase the perceived benefits of vaccination, a growing number of governments have implemented conditional cash lotteries (CCLs) that offer opportunities to win large prizes

[15]Follow-up rate is determined by summing total second doses of Pfizer and Moderna with 21 and 28 day lags, respectively, and then dividing this value by the sum of the total first doses of each. For all dates t within each period/column of the table, this formula is:

$$\text{Follow-up rate} = \frac{\sum_t \text{Pfizer(2nd dose)}_{t+21} + \text{Moderna(2nd dose)}_{t+28}}{\sum_t \text{Pfizer(1st dose)}_t + \text{Moderna(1st dose)}_t}$$

only available to vaccinated individuals. A CCL incentive is promising in this context because of its targeted nature: people with a greater propensity to decline vaccination are also more likely to assign a higher expected value to a lottery, a behavioral phenomenon known as probability neglect. Our paper evaluates the first CCL for COVID-19 vaccinations, which Ohio implemented during May and June of 2021. We find that Ohio's initiative significantly increases vaccinations—successfully encouraging more than 82,000 Ohioans who would otherwise not be vaccinated, an increase of 1.5 percent. Furthermore, we estimate that this surge in vaccinations then decreases COVID-19 prevalence within the state, reducing infections and ICU utilization by at least 1.3 percent and 2.6 percent, respectively.

These estimates allow us to assess the cost-effectiveness of the program. In a large study of COVID-19 patients, Di Fusco et al. (2021) find that the average hospital bill per day in the ICU is around 13,500 dollars. Using our estimate of the number of ICU patient-days averted, we calculate that the total benefit from avoiding these charges is approximately 66 million dollars. Additionally, there are substantial other social benefits from the 15,000 (or more) cases prevented, such as quality of life enrichment—especially for those who avoid cases of "long-haul COVID," where symptoms persist for months or longer—and potentially lives saved. Given that the total cost of Ohio's Vax-A-Million incentive scheme is 5.6 million dollars, the benefits of the CCL unquestionably exceed the program's cost.

Hesitancy towards vaccines has been rising globally in recent years, creating a significant challenge for policymakers. In lieu of mandates, governments are increasingly turning to other instruments to improve vaccination rates. Our evidence from Ohio's program illustrates that financial incentives—and conditional cash lotteries more specifically—are an effective means to increase vaccine uptake in areas plagued by vaccine hesitancy. And while this group may be larger in a state like Ohio, it is far from an outlier in this regard. Although a CCL is certainly not a panacea, we show that it can be a cost-effective component of a broader policy mix to increase vaccine uptake, with compelling potential to support other public health objectives as well.

References

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

A. Abadie and J. Gardeazabal. The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1):113–132, 2003.

A. Abadie, A. Diamond, and J. Hainmueller. 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, 2010.

A. Abadie, A. Diamond, and J. Hainmueller. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510, 2015.

J. Abrevaya and K. Mulligan. Effectiveness of state-level vaccination mandates: Evidence from the varicella vaccine. *Journal of Health Economics*, 30(5):966–976, 2011.

M. Auld. Choices, beliefs, and infectious disease dynamics. *Journal of Health Economics*, 22 (3):361–377, 2003.

T. Barham and J. A. Maluccio. Eradicating diseases: The effect of conditional cash transfers on vaccination coverage in rural Nicaragua. *Journal of Health Economics*, 28(3):611–621, 2009.

E. Ben-Michael, A. Feller, and J. Rothstein. The augmented synthetic control method. *Journal of the American Statistical Association*, Forthcoming, 2021.

J. Biden. Remarks by President Biden marking the 150 millionth COVID-19 vaccine shot. The White House, 2021.

M. Björkman Nyqvist, L. Corno, D. de Walque, and J. Svensson. Incentivizing safer sexual behavior: Evidence from a lottery experiment on HIV prevention. *American Economic Journal: Applied Economics*, 10(3):287–314, 2018.

M. E. Brehm, P. A. Brehm, and M. Saavedra. The Ohio vaccine lottery and starting vaccination rates. Mimeo, 2021.

British Medical Journal. Vital statistics. *The British Medical Journal*, 1(5025):1013–1014, 1957.

T. Buchanan. With first drawing days away, Ohio legislator seeks to stop Vax-A-Million lottery. *The Columbus Dispatch*, 2021.

P. Campos-Mercade, A. N. Meier, F. H. Schneider, S. Meier, D. Pope, and E. Wengström. Monetary incentives increase COVID-19 vaccinations. *Science*, 374(6569):879–882, 2021.

T. Chang, M. Jacobson, M. Shah, R. Pramanik, and S. B. Shah. Financial incentives and other nudges do not increase covid-19 vaccinations among the vaccine hesitant. NBER Working Paper: w29403, 2021.

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

N. Dagan, N. Barda, E. Kepten, O. Miron, S. Perchik, M. A. Katz, M. A. Hernán, M. Lipsitch, B. Reis, and R. D. Balicer. Bnt162b2 mrna COVID-19 vaccine in a nationwide mass vaccination setting. *New England Journal of Medicine*, 384(15):1412–1423, 2021.

A. de Figueiredo, C. Simas, E. Karafillakis, P. Paterson, and H. J. Larson. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: A large-scale retrospective temporal modelling study. *The Lancet*, 396(10255):898–908, 2020.

M. DeWine. Don't roll your eyes at Ohio's vaccine lottery. *The New York Times*, 2021.

M. Di Fusco, K. M. Shea, J. Lin, J. L. Nguyen, F. J. Angulo, M. Benigno, D. Malhotra, B. Emir, A. H. Sung, J. L. Hammond, S. Stoychev, and A. Charos. Health outcomes and economic burden of hospitalized COVID-19 patients in the United States. *Journal of Medical Economics*, 24(1):308–317, 2021.

L. Goette and A. Stutzer. Blood donations and incentives: Evidence from a field experiment. *Journal of Economic Behavior & Organization*, 170:52–74, 2020.

IHME. GBD compare: COVID-19 projections. Technical report, Institute for Health Metrics and Evaluation (IHME). University of Washington, Seattle, Washington, 2021.

D. Ivory, L. Leather, and R. Gebeloff. Least vaccinated U.S. counties have something in common: Trump voters. *The New York Times*, 2021.

D. Lang, L. Esbenshade, and R. Willer. Did Ohio's vaccine lottery increase vaccination rates? A pre-registered, synthetic control study. Mimeo, 2021.

E. A. Largent, G. Persad, S. Sangenito, A. Glickman, C. Boyle, and E. J. Emanuel. U.S. public attitudes toward COVID-19 vaccine mandates. *JAMA Network Open*, 3(12):e2033324–e2033324, 2020.

E. C. Lawler. Effectiveness of vaccination recommendations versus mandates: Evidence from the hepatitis A vaccine. *Journal of Health Economics*, 52:45–62, 2017.

W. P. Moran, K. Nelson, J. L. Wofford, R. Velez, and L. D. Case. Increasing influenza immunization among high-risk patients: Education or financial incentive? *The American Journal of Medicine*, 101(6):612–620, 1996.

C. R. Sunstein. Probability neglect: Emotions, worst cases, and law. *The Yale Law Journal*, 112(1):61–107, 2002.

H. Thirumurthy, S. H. Masters, S. Rao, K. Murray, R. Prasad, J. G. Zivin, E. Omanga, and K. Agot. The effects of providing fixed compensation and lottery-based rewards on uptake of medical male circumcision in Kenya: A randomized trial. *Journal of Acquired Immune Deficiency Syndromes*, 72(Suppl 4):S299–305, 2016.

U.S. CMS. Biden-Harris administration increases Medicare payment for life-saving COVID-19 vaccine. Technical report, U.S. Centers for Medicare & Medicaid Services, 2021.

K. G. Volpp, L. K. John, A. B. Troxel, L. Norton, J. Fassbender, and G. Loewenstein. Financial incentive–based approaches for weight loss: A randomized trial. *JAMA*, 300 (22):2631–2637, 2008a.

K. G. Volpp, G. Loewenstein, A. B. Troxel, J. Doshi, M. Price, M. Laskin, and S. E. Kimmel. A test of financial incentives to improve warfarin adherence. *BMC Health Services Research*, 8(272), 2008b.

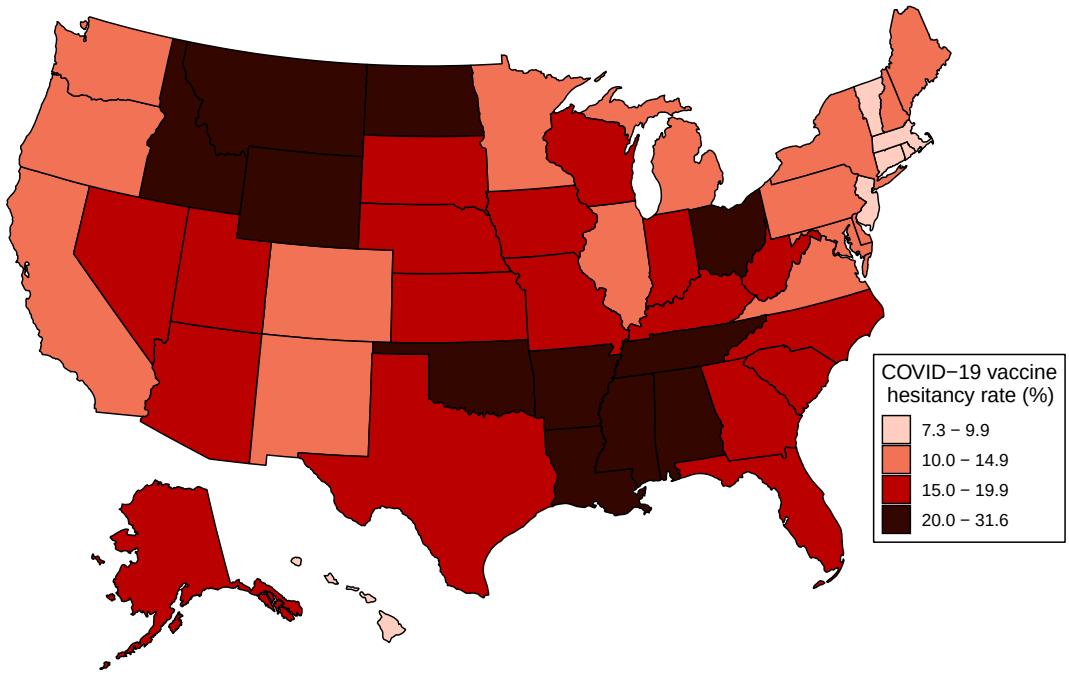
V. Vovk, A. Gammerman, and G. Shafer. *Algorithmic learning in a random world*. Springer, 2005.

A. J. Walkey, A. Law, and N. A. Bosch. Lottery-based incentive in Ohio and COVID-19 vaccination rates. *JAMA*, Forthcoming, 2021.

D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, and Z. Peng. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. *JAMA*, 323(11):1061–1069, 2020.

Figures and tables

Figure 1: Surveyed COVID-19 vaccination hesitancy by state



Notes: Data plotted in this map use an average of the Census Bureau's Household Pulse Survey responses during Weeks 25-29 (February 17 to May 10, 2021). Vaccination hesitancy includes responses of "definitely not" and "probably not" as survey respondents' stated willingness to be vaccinated for COVID-19.

Table 1: Largest five unit weights for Synthetic Ohio using ridge augmented synthetic control for an outcome of the share of state population with any COVID-19 vaccination

State	Unit weight
Wisconsin	0.321
Kansas	0.281
Michigan	0.191
Idaho	0.181
North Dakota	0.126

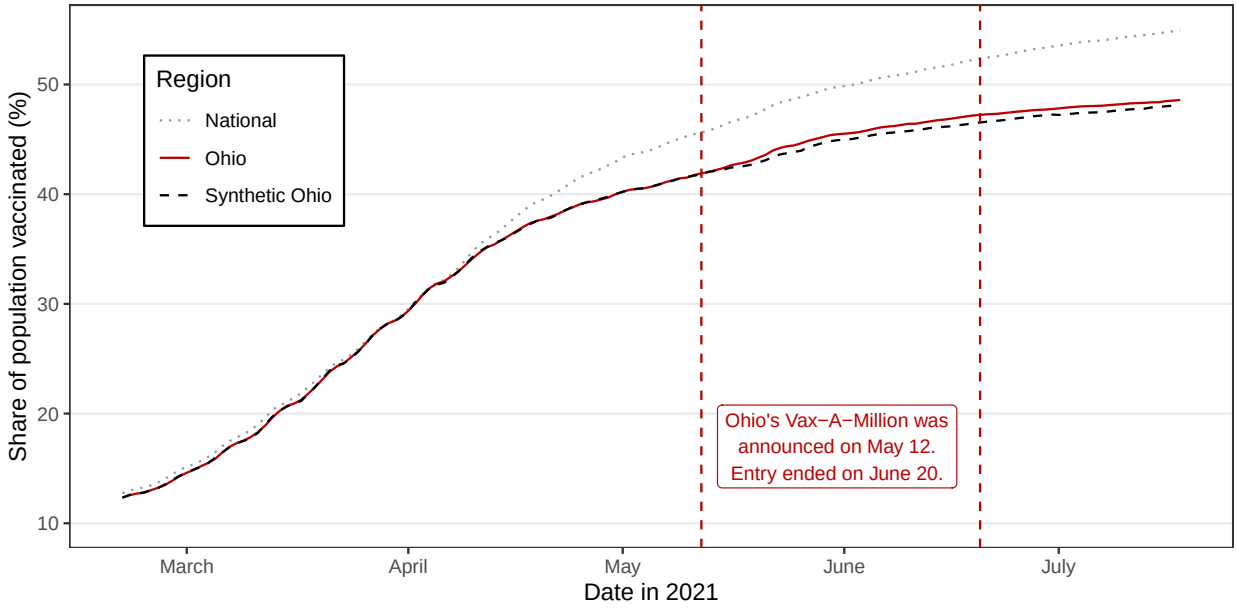
Notes: Appendix Table A2 shows the full set of unit weights for Synthetic Ohio.

Table 2: Summary statistics for the United States, Ohio, and Synthetic Ohio

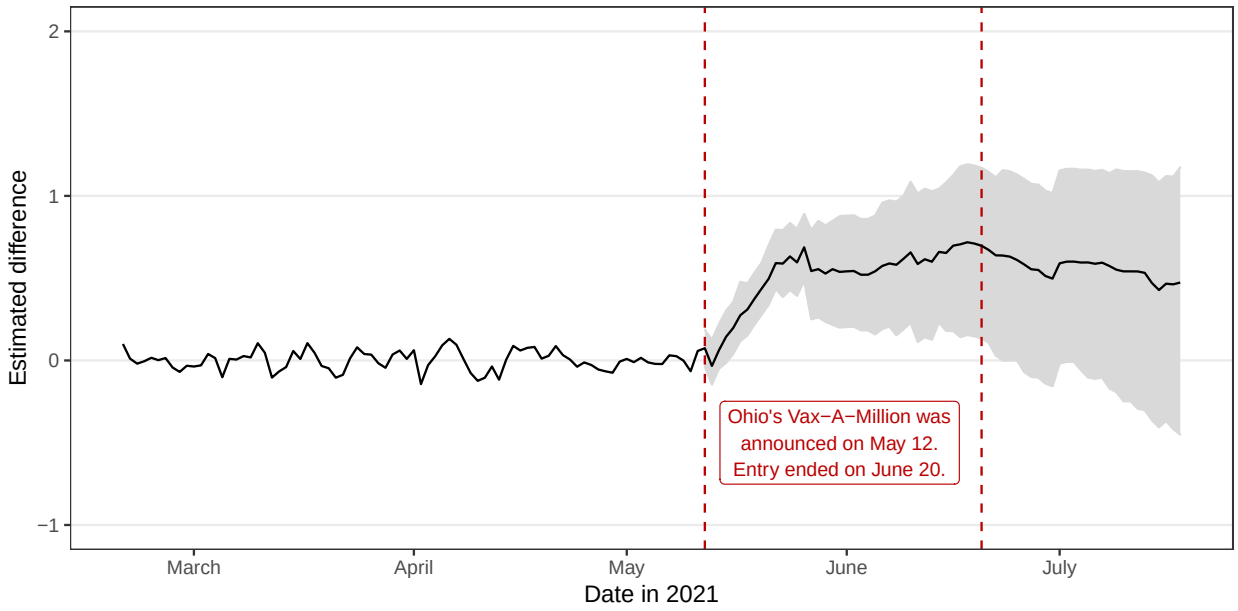
	State average	Ohio	Synthetic Ohio
Panel [A] Outcome variables in the pre-treatment period			
Share of population with any vaccination by April 2, 2021 (%)	30.21	30.04	30.18
Share of population with any vaccination by May 12, 2021 (%)	45.63	41.90	41.82
Total COVID-19 cases per 100k population by May 12, 2021	9,576	9,214	9,213
Total COVID-19 ICU patient-days per 100k pop. by May 12	1,430	1,450	1,449
Panel [B] Covariates in the pre-treatment period			
Share of population of age 12 to 17 (%)	7.47	7.54	8.16
Share of population of age 18 or older (%)	77.06	77.22	74.46
Population density in 2020 (people per square mile)	423.64	288.80	240.56
Gross domestic product per capita in 2020 ($)	61,791	57,209	57,905
Republican presidential vote share in 2020 (%)	49.12	53.27	53.53
Influenza vaccination rate in 2019 (%)	47.41	50.00	49.99
Community Mobility Report index for retail/recreation	-7.66	-5.59	-5.63
Community Mobility Report index for grocery/pharmacy	-1.22	-1.65	-1.56
Community Mobility Report index for parks	31.04	70.58	70.16
Community Mobility Report index for transit stations	-15.77	-8.59	-8.41
Community Mobility Report index for workplaces	-23.30	-20.99	-21.02
Community Mobility Report index for residences	6.13	5.30	5.32
Median distance of pop. to closest vaccination site (miles)	1.10	0.91	0.91
95th percentile distance to closest vaccination site (miles)	11.08	7.01	6.99

Notes: Table 2 presents summary statistics for the United States, Ohio, and Synthetic Ohio. Ohio's Vax-A-Million incentive program was announced on May 12, 2021 and lottery entry ended on June 20, 2021. Panel [A] shows values for the dependent variables during the pre-treatment time period(s) indicated. These outcomes are: the share of population with at least a first dose of any COVID-19 vaccination, the cumulative total COVID-19 cases per 100,000 population, and the cumulative total COVID-19 hospital ICU patient-days per 100,000 population. Panel [B] shows values for state covariates during 2019 or 2020, as indicated, or during the pre-treatment analysis period of February 19 through May 11, 2021. Each Community Mobility Report index can take values ranging from -100 to 100. Google defines these using movement of people's cell phones to different places, with the baseline zero-value for each index set during January 3 through February 6, 2020. We compute the distance measures of state population to vaccination sites using Census Block Group population centers and the locations of all vaccination sites in the United States.

Figure 2: Share of population with any COVID-19 vaccination over time



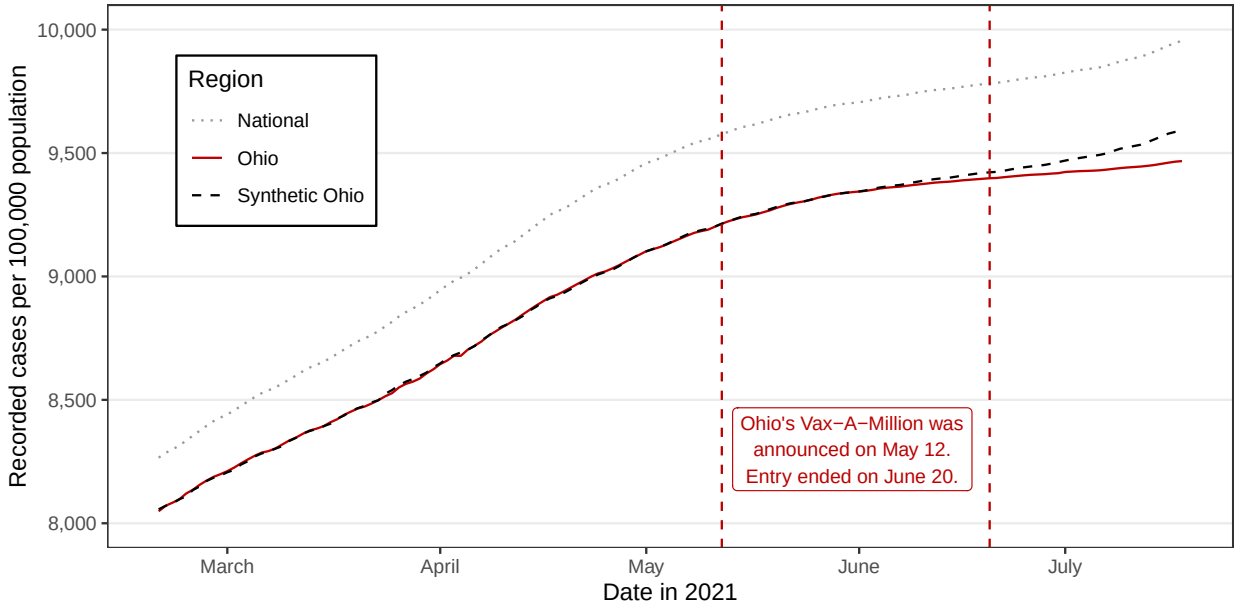
(a) Time series plotted by region



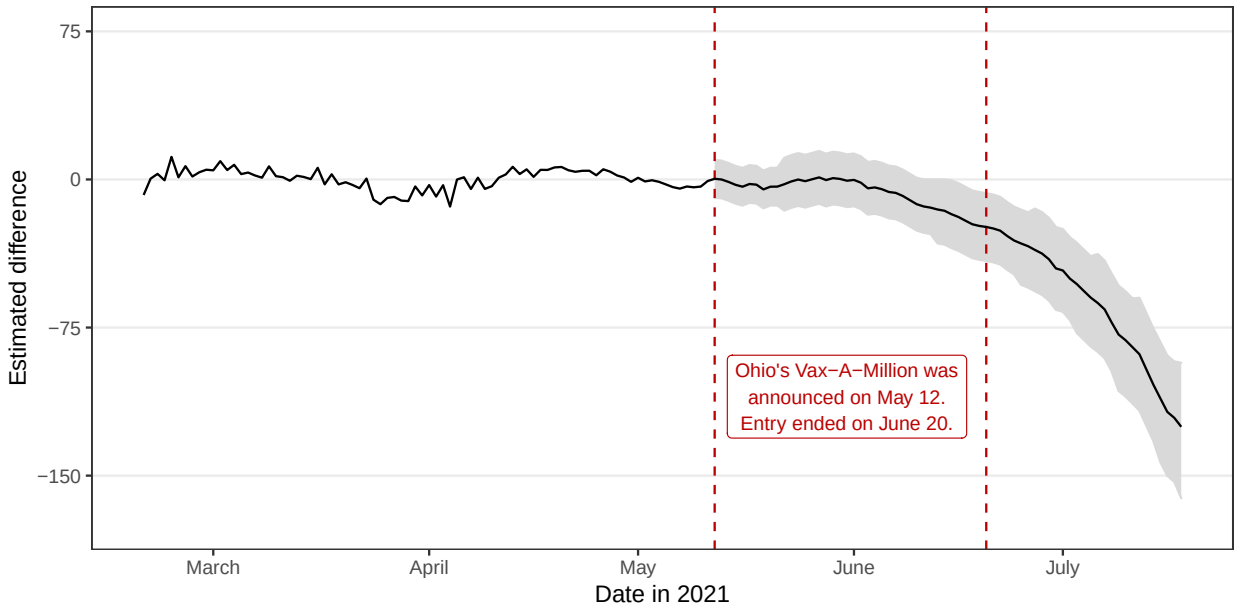
(b) Estimated difference between Ohio and Synthetic Ohio

Notes: Panel (a) of Figure 2 shows time series graphs for the share of population that had received at least a first dose of any COVID-19 vaccination by region and date. Panel (b) shows the estimated difference between Ohio and the synthetic control. The grey shading indicates 95 percent confidence intervals for each post-treatment date, calculated using conformal inference.

Figure 3: Cumulative total COVID-19 cases recorded per 100,000 population over time



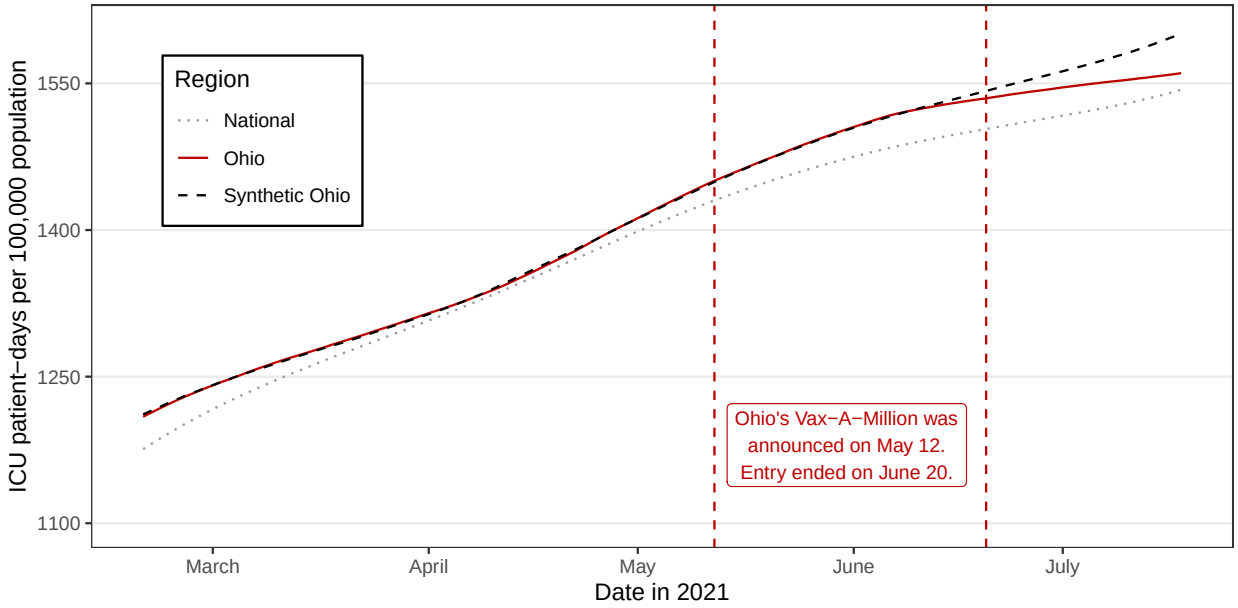
(a) Time series plotted by region



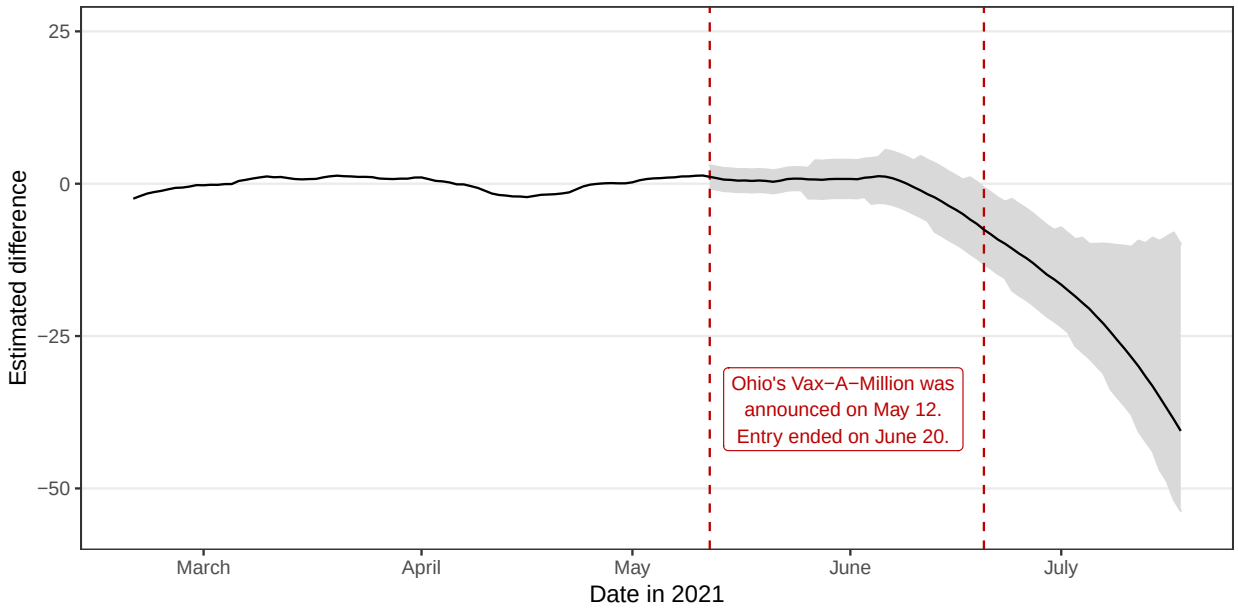
(b) Estimated difference between Ohio and Synthetic Ohio

Notes: Panel (a) of Figure 3 shows time series graphs for the cumulative total number of COVID-19 cases (positive COVID-19 tests) recorded per 100,000 population by region and date. Panel (b) shows the estimated difference between Ohio and the synthetic control. The grey shading indicates 95 percent confidence intervals for each post-treatment date, calculated using conformal inference.

Figure 4: Cumulative total COVID-19 ICU patient-days per 100,000 population over time



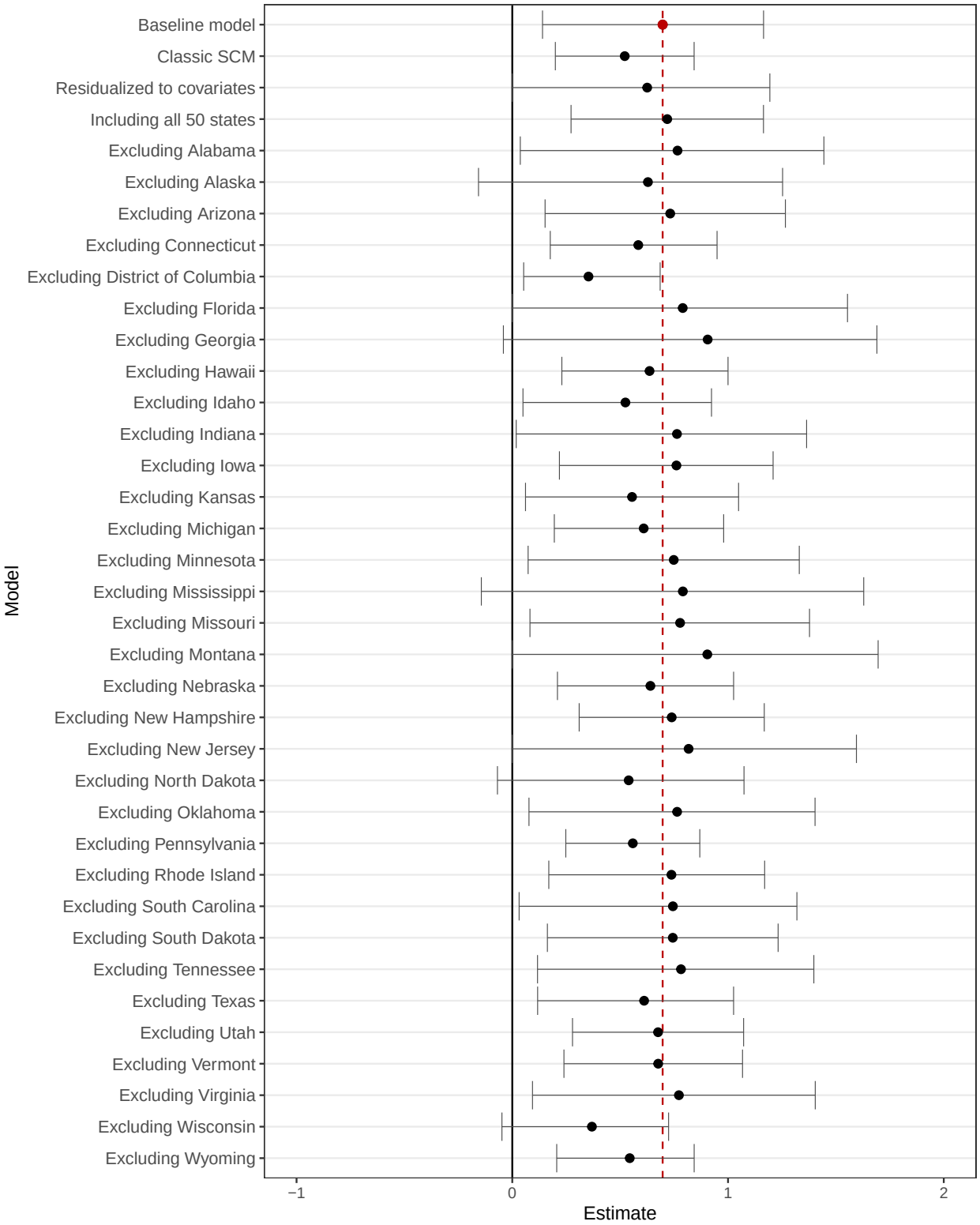
(a) Time series plotted by region



(b) Estimated difference between Ohio and Synthetic Ohio

Notes: Panel (a) of Figure 4 shows time series graphs for the cumulative total COVID-19 hospital ICU patient-days per 100,000 population by region and date. Panel (b) shows the estimated difference between Ohio and the synthetic control. The grey shading indicates 95 percent confidence intervals for each post-treatment date, calculated using conformal inference.

Figure 5: Robustness checks of the synthetic control estimates for the share of population with any COVID-19 vaccination by the end date, using different samples and specifications



Notes: Figure 5 shows estimated differences between Ohio and the synthetic control for the share of population that had received at least a first dose of any COVID-19 vaccination by June 20, 2021. Each row depicts results from a separate model using the data sample and/or specification denoted. The grey error bars indicate the respective 95 percent confidence intervals, which are calculated using conformal inference.

Table 3: Estimation results for Ohio compared to the synthetic control

Outcome	Date	Estimate	95 pct. conf. interval		Cf. value
			Low bd.	Up. bd.	
Panel [A] COVID-19 vaccinations during the Vax-A-Million treatment period					
Population vaccinated (%)	May 18	0.3098	0.1520	0.4676	42.6
Population vaccinated (%)	May 25	0.5959	0.3930	0.7988	43.9
Population vaccinated (%)	June 01	0.5415	0.2034	0.8797	45.0
Population vaccinated (%)	June 08	0.5815	0.1531	0.9647	45.6
Population vaccinated (%)	June 15	0.6531	0.1797	1.081	46.1
Population vaccinated (%)	June 20	0.6970	0.1334	1.170	46.5
Pop. vaccinated 18-older (%)	June 20	0.7761	0.2953	1.257	57.8
Panel [B] COVID-19 infections during the treatment period and post-treatment					
Cases per 100k population	June 20	-24.06	-41.19	-6.932	9,422
Cases per 100k population	July 18	-125.3	-161.4	-92.90	9,593
ICU patient-days per 100k pop.	June 20	-7.540	-13.17	-0.6617	1,542
ICU patient-days per 100k pop.	July 18	-40.56	-53.70	-9.927	1,601

Notes: Table 3 shows results from ridge augmented synthetic control estimations for Ohio's Vax-A-Million incentive program, which was announced on May 12, 2021. Lottery entry ended on June 20, 2021. The outcomes in rows are the share of population with any COVID-19 vaccination, the cumulative total number of COVID-19 cases recorded per 100,000 population, and the cumulative total COVID-19 hospital ICU patient-days per 100,000 population. The 95 percent confidence intervals are calculated using conformal inference. The final column shows the counterfactual values from Synthetic Ohio.

Table 4: Aggregate estimated effects and characteristics for Ohio by time period

		Vax-A-Million treatment period		
	Pre-treatment	First two weeks	Full period	Post-treatment
Date range included	April 2 - May 11	May 12 - May 25	May 12 - June 20	May 12 - July 18
Number of days	40	14	40	68
Ohio population	11,799,448	11,799,448	11,799,448	11,799,448
Vax-A-Million program cost ($)		5,600,000	5,600,000	5,600,000
Total first dose vaccinations	1,488,978	339,226	690,135	
First dose compliers		70,315 (21%)	82,239 (12%)	
First dose always-takers		268,911 (79%)	607,896 (88%)	
First dose of Janssen	174,651 (12%)	40,186 (12%)	87,648 (13%)	
First dose of Moderna	491,329 (33%)	92,300 (27%)	176,943 (26%)	
First dose of Pfizer	822,998 (55%)	206,740 (61%)	425,544 (62%)	
Program cost per complier ($)		80	68	
2nd dose Moderna in 28 days	403,991	80,568	142,786	
Moderna follow-up rate	82%	87%	81%	
2nd dose Pfizer in 21 days	736,942	168,624	388,493	
Pfizer follow-up rate	90%	82%	91%	
Overall follow-up rate	87%	83%	88%	
COVID-19 cases prevented				14,779
ICU patient-days prevented				4,786

Notes: Table 4 uses data for Ohio and estimates from Table 3 to calculate aggregate effect sizes and characteristics for Ohio by time period. All dates included are in 2021. The row for second doses of Moderna uses a time period shifted forward by 28 days, e.g. using second doses during June 9 - July 18 for the full treatment period of May 12 - June 20 column. Similarly, the Pfizer second dose row uses a time period shifted by 21 days. The second dose follow-up rates are calculated by dividing the total Moderna and/or Pfizer second dose values by the total Moderna and/or Pfizer first dose values.

A Appendix tables and figures

Table A1: State COVID-19 vaccination initiatives using conditional cash lottery incentives

Announcement date (in 2021)	State	Program name	Eligible vaccinations	Registration process	Largest prize ($)	Total prize value (approximate $)	Exclusive prizes	Drawing frequency
May 12	Ohio	Vax-A-Million	All	Opt-in	1,000,000	5,600,000	Yes	Weekly
May 20	Maryland	VaxCash	All	Auto	400,000	2,000,000	Yes	Daily
May 20	New York	Vax And Scratch	New	Opt-in	5,000,000	Unknown	No	Instant
May 21	Oregon	Take Your Shot	All	Auto	1,000,000	1,500,000	Yes	Once
May 24	Delaware	DEWins	New/All	Auto	302,000	5,000,000	Yes	Semi-weekly
May 25	Arkansas	Not named	New	Opt-in	1,000,000	2,000,000	No	Instant
May 25	Colorado	Comeback Cash	All	Auto	1,000,000	6,250,000	Yes	Weekly
May 27	California	Vax For The Win	New/All	Auto/Opt-in	1,500,000	116,500,000	Yes	Weekly
May 27	West Virginia	Do It For Babydog	All	Opt-in	1,588,000	10,000,000	Yes	Weekly
June 01	New Mexico	Vax 2 The Max	All	Opt-in	5,000,000	10,000,000	Yes	Weekly
June 03	Washington	Shot Of A Lifetime	All	Auto	1,000,000	2,400,000	Yes	Weekly
June 04	Kentucky	Shot At A Million	All	Opt-in	1,000,000	4,200,000	Yes	Monthly
June 10	North Carolina	Summer Cash	New/All	Auto	1,000,000	4,500,000	Yes	Bi-weekly
June 15	Massachusetts	VaxMillions	All	Opt-in	1,000,000	5,500,000	Yes	Weekly
June 17	Illinois	All In For The Win	All	Auto	1,000,000	10,000,000	Yes	Weekly
June 17	Louisiana	Shot At A Million	All	Opt-in	1,000,000	2,300,000	Yes	Weekly
June 17	Maine	Don't Miss Your Shot	All	Opt-in	896,809	896,809	Yes	Once
June 18	Nevada	Vax Nevada Days	All	Auto	1,000,000	5,000,000	Yes	Weekly
July 01	Michigan	Shot To Win	All	Opt-in	2,000,000	5,495,000	Yes	Bi-weekly
July 21	Missouri	MO VIP	All	Opt-in	10,000	9,000,000	Yes	Bi-weekly

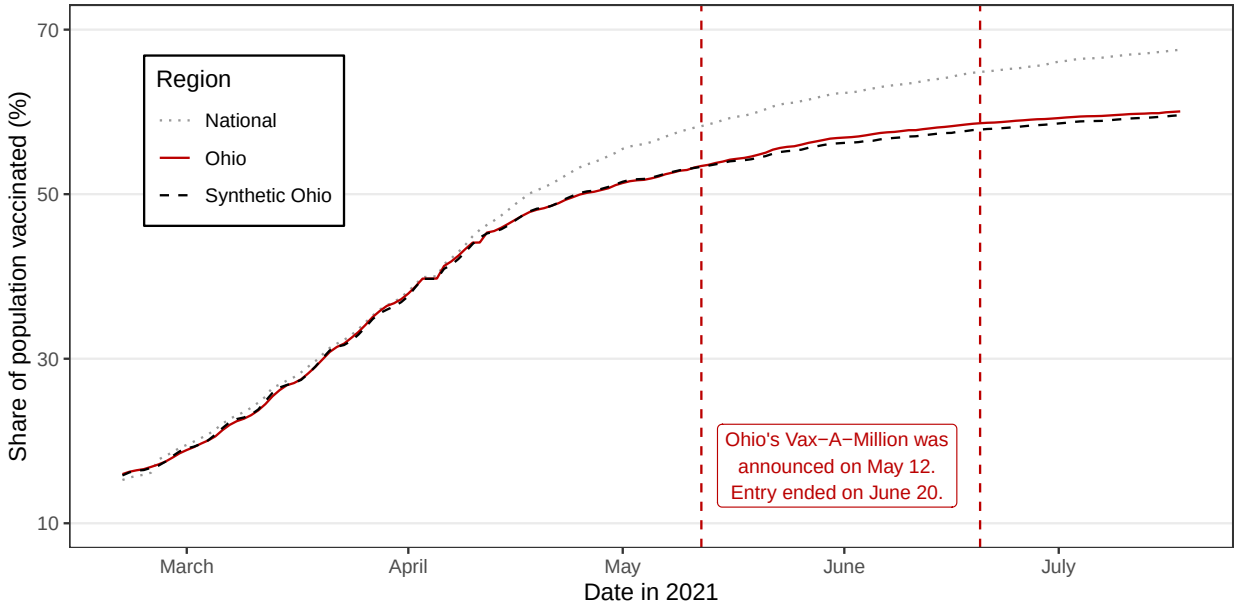
Notes: Table A1 lists all state-run conditional cash lottery incentive schemes for COVID-19 vaccinations in the United States. Ohio's Vax-A-Million incentive program was the first and was announced on May 12, 2021 and lottery entry ended on June 20, 2021. The eligible vaccinations column indicates whether people who were vaccinated prior to the program's announcement could win prizes in the CCL. The registration process column indicates whether any action is required to be considered for prizes (typically filling out a form on the state's website). The total prize value column includes monetary and nonmonetary prizes, such as college tuition scholarships. The exclusive prizes column indicates whether the CCL prizes are exclusively available to vaccinated individuals, compared to non-exclusive prizes like free scratch-off tickets for state lotteries.

Table A2: State unit weights for the ridge augmented synthetic control models

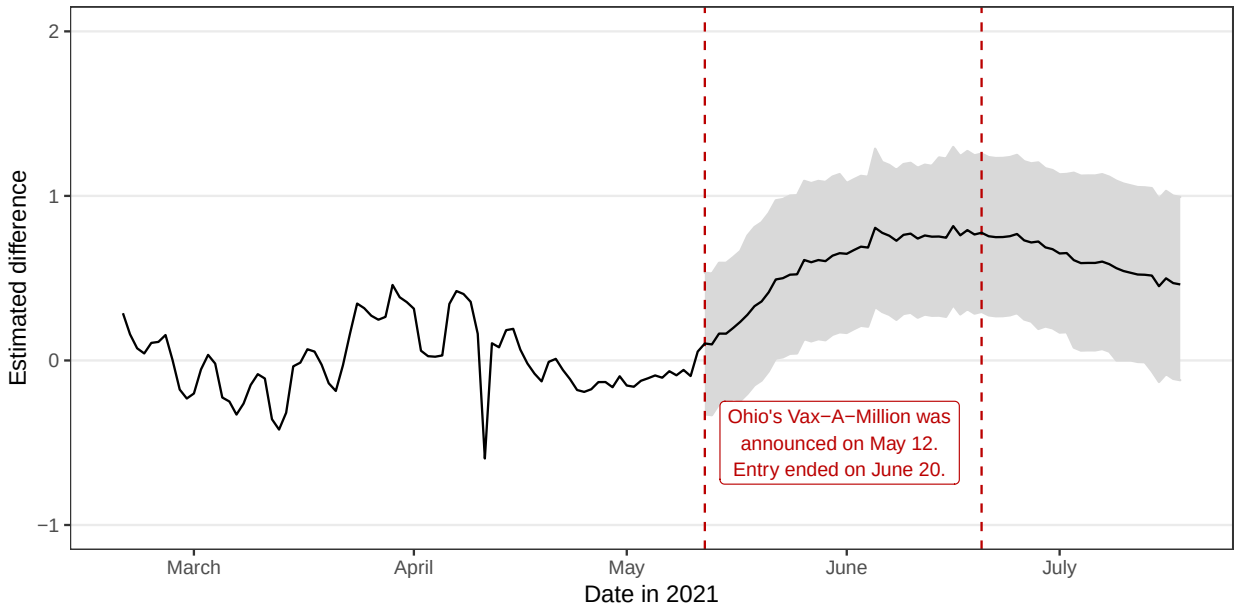
State	Unit weights for model		
	Vaccinations	Cases	ICU days
Alabama	-0.065	-0.000	-0.000
Alaska	-0.021	-0.000	-0.000
Arizona	0.008	-0.000	-0.000
Connecticut	-0.058	-0.000	0.000
District of Columbia	0.009	0.000	-0.000
Florida	-0.067	-0.000	-0.000
Georgia	0.107	-0.000	0.000
Hawaii	-0.045	-0.000	-0.001
Idaho	0.181	0.000	0.000
Indiana	0.001	0.136	0.419
Iowa	-0.066	0.000	0.000
Kansas	0.281	0.304	0.082
Michigan	0.191	0.033	0.078
Minnesota	0.015	0.000	0.000
Mississippi	0.040	-0.000	-0.000
Missouri	-0.038	0.000	0.197
Montana	-0.119	-0.000	-0.000
Nebraska	0.115	0.000	0.000
New Hampshire	-0.044	0.000	-0.000
New Jersey	0.079	-0.000	0.060
North Dakota	0.126	0.000	0.000
Oklahoma	-0.007	0.000	0.000
Pennsylvania	-0.057	0.000	0.000
Rhode Island	0.024	0.109	0.165
South Carolina	0.043	-0.000	-0.000
South Dakota	-0.024	-0.000	0.000
Tennessee	0.049	0.000	0.000
Texas	-0.056	-0.000	0.000
Utah	0.091	0.141	0.000
Vermont	0.026	-0.000	-0.000
Virginia	0.031	0.000	0.000
Wisconsin	0.321	0.277	0.000
Wyoming	-0.067	-0.000	-0.000

Notes: States not listed are not in the donor pool. For outcomes in column titles:
Vaccinations = Share of population with any COVID-19 vaccination (at least a first dose).
Cases = Cumulative total COVID-19 cases recorded per 100,000 population.
ICU days = Cumulative total COVID-19 hospital ICU patient-days per 100,000 population.

Figure A1: Share of population 18-older with any COVID-19 vaccination over time



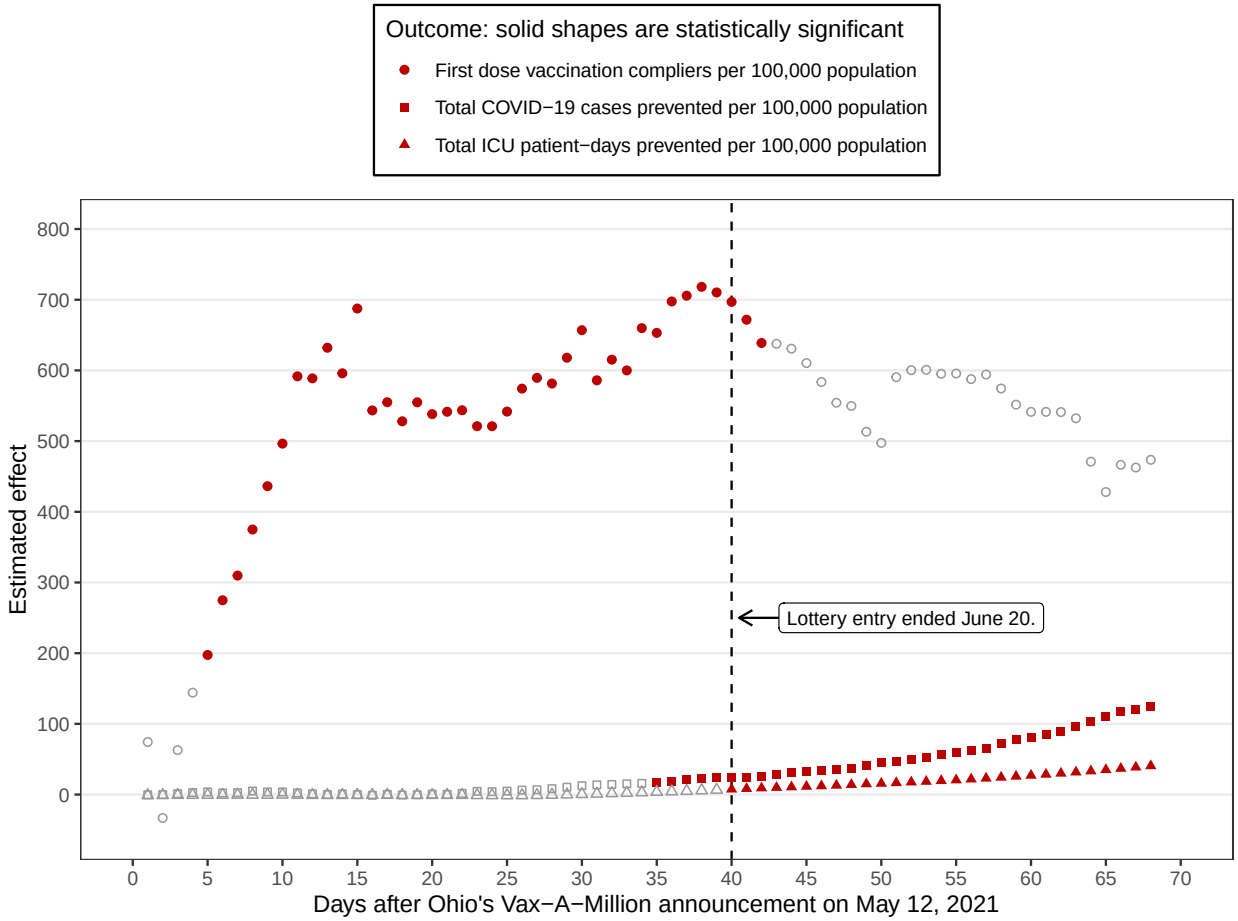
(a) Time series plotted by region



(b) Estimated difference between Ohio and Synthetic Ohio

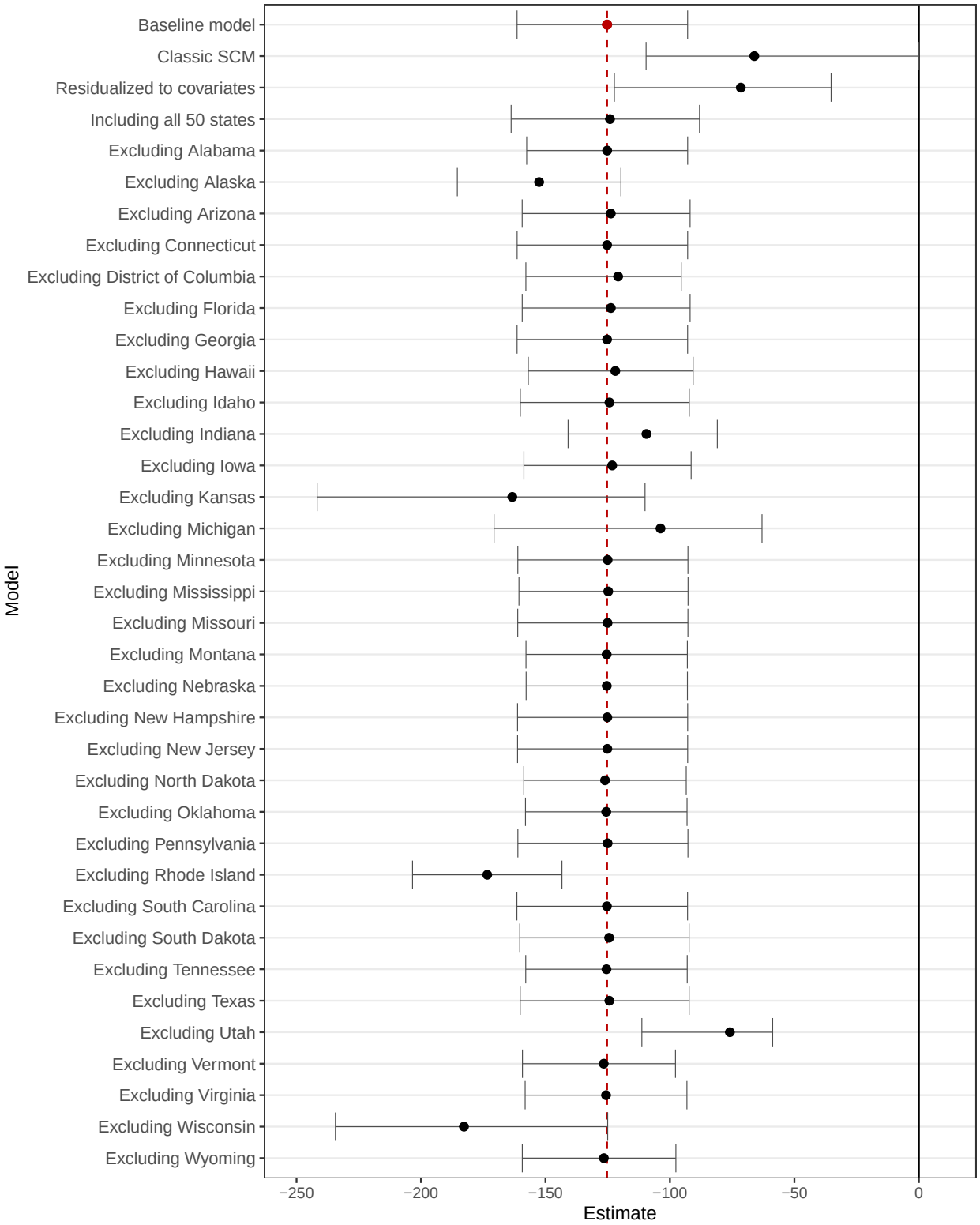
Notes: Panel (a) of Figure A1 shows time series graphs for the share of population of age 18 or older that had received at least a first dose of any COVID-19 vaccination by region and date. Panel (b) shows the estimated difference between Ohio and the synthetic control. The grey shading indicates 95 percent confidence intervals for each post-treatment date, calculated using conformal inference.

Figure A2: Estimated effects for transformations of each outcome into a common scale



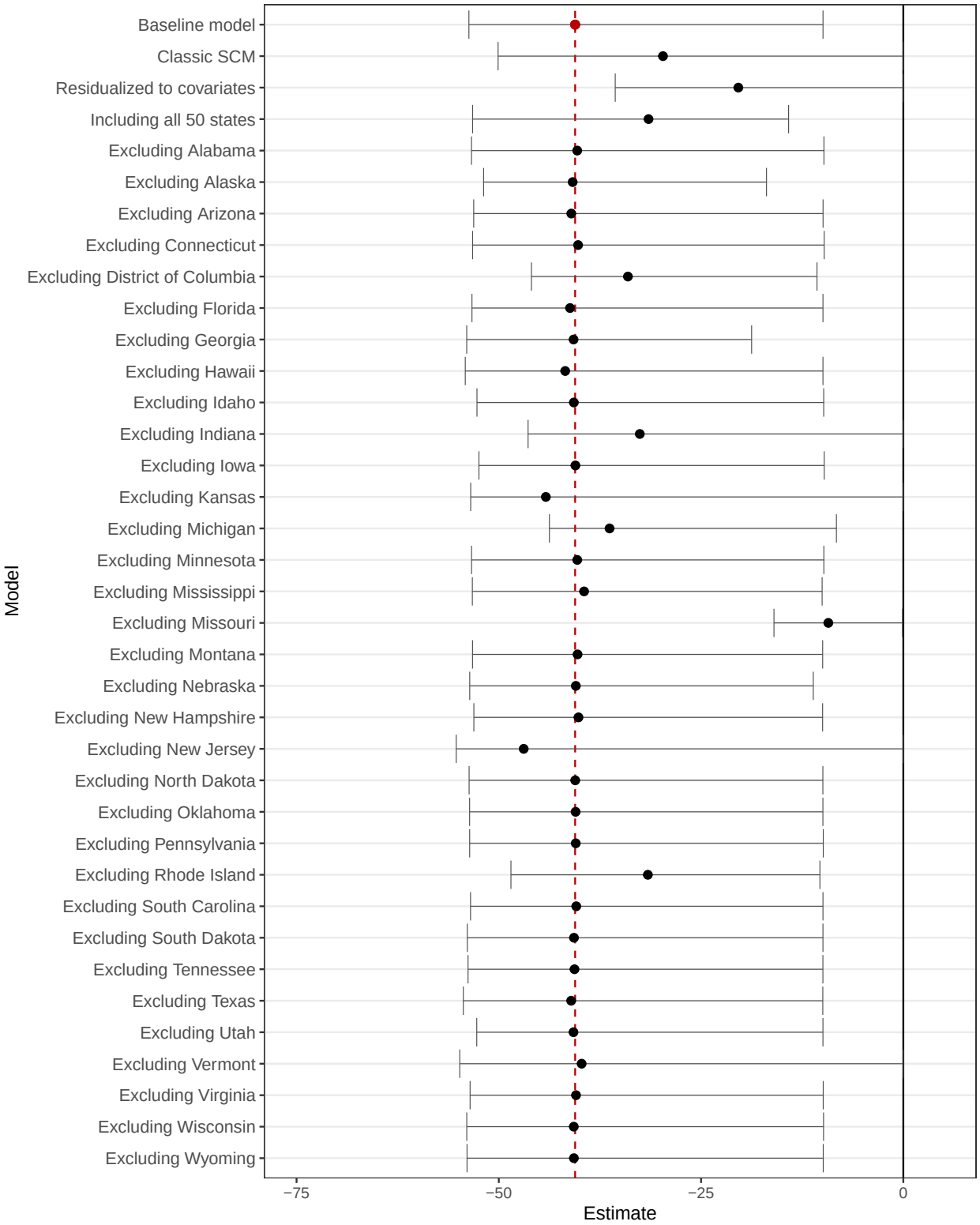
Notes: Figure A2 shows estimated differences between Ohio and the synthetic control for the three outcomes examined in this study, transformed to use a common scale. These transformations are: (1) The fraction of population with any COVID-19 vaccination – multiplied by 100,000. (2) The cumulative total COVID-19 cases recorded per 100,000 population – multiplied by negative one. (3) The cumulative total COVID-19 ICU patient-days per 100,000 population – multiplied by negative one. These effects are plotted by day following Ohio's Vax-A-Million lottery announcement. The shapes are solid if the 95 percent confidence interval does not overlap with zero, as calculated using conformal inference.

Figure A3: Robustness checks of the synthetic control estimates for the cumulative total COVID-19 cases recorded per 100,000 population, using different samples and specifications



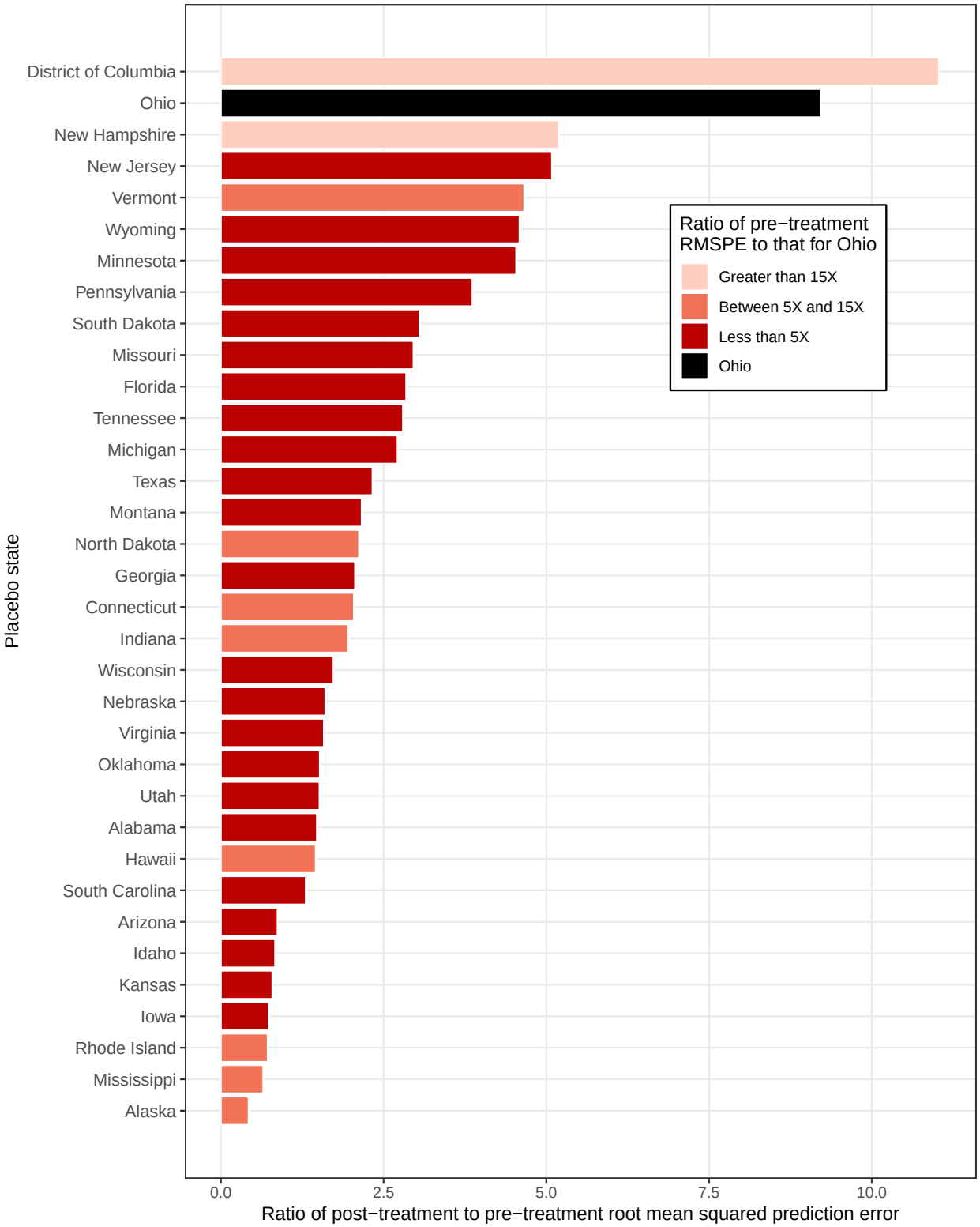
Notes: Figure A3 shows estimated differences between Ohio and the synthetic control for the cumulative total COVID-19 cases recorded per 100,000 population by July 18, 2021. Each row depicts results from a separate model using the data sample and/or specification denoted. The grey error bars indicate the respective 95 percent confidence intervals, which are calculated using conformal inference.

Figure A4: Robustness checks of the synthetic control estimates for the total COVID-19 ICU patient-days per 100,000 population, using different samples and specifications



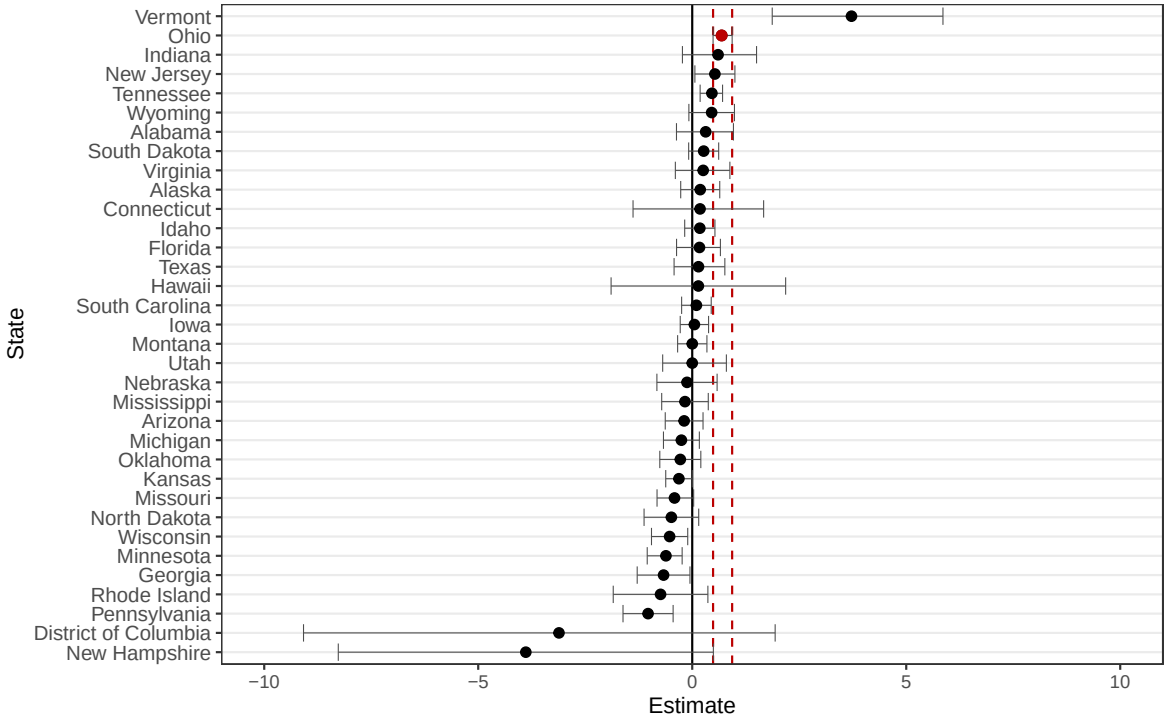
Notes: Figure A4 shows estimated differences between Ohio and the synthetic control for the cumulative total COVID-19 ICU patient-days per 100,000 population by July 18, 2021. Each row depicts results from a separate model using the data sample and/or specification denoted. The grey error bars indicate the respective 95 percent confidence intervals, which are calculated using conformal inference.

Figure A5: Synthetic control placebo effects and rankings for other states

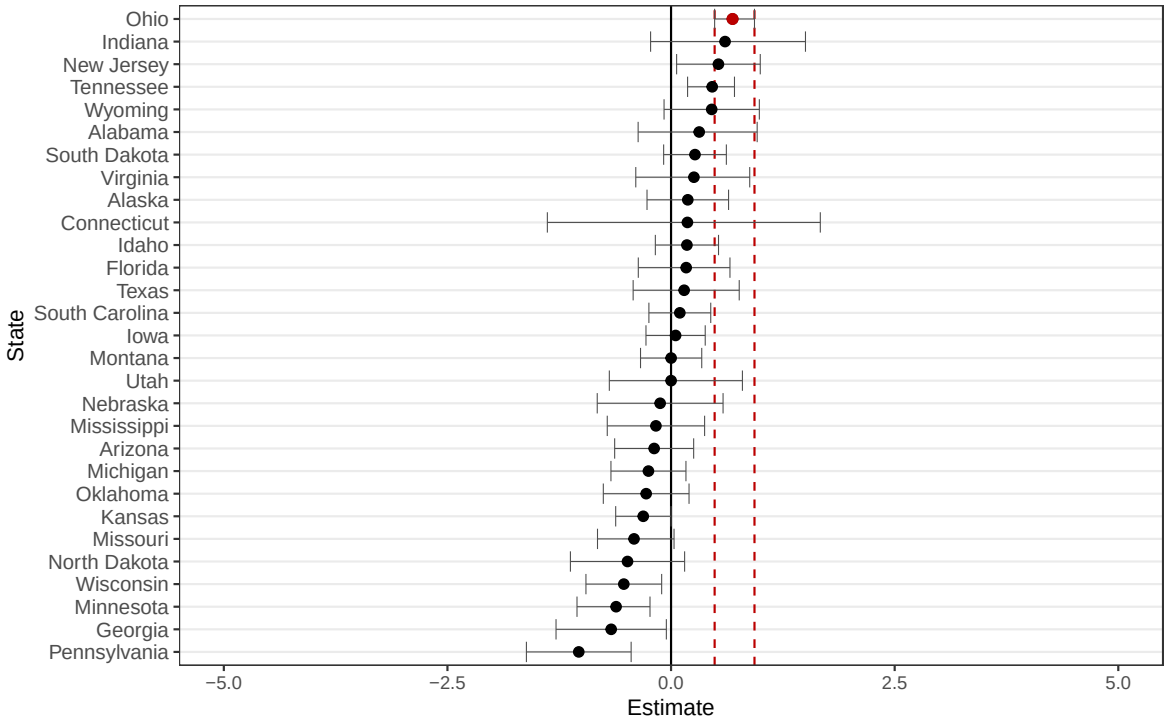


Notes: States not listed are not in the donor pool. The outcome is the share of population with any COVID-19 vaccination (at least a first dose). Post-treatment RMSPE are computed using the full treatment period, starting with Ohio's Vax-A-Million announcement on May 12, 2021 and ending with the lottery entry end-date on June 20, 2021. Pre-treatment RMSPE are computed using the full pre-treatment period in the data, starting on February 19, 2021 and ending on May 11, 2021.

Figure A6: Synthetic control placebo estimates for other states



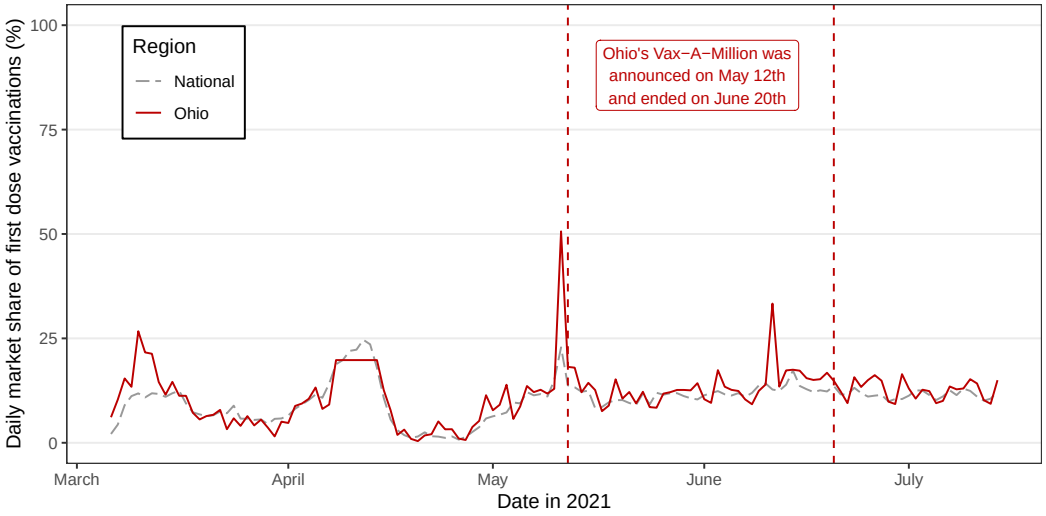
(a) Including the full donor pool



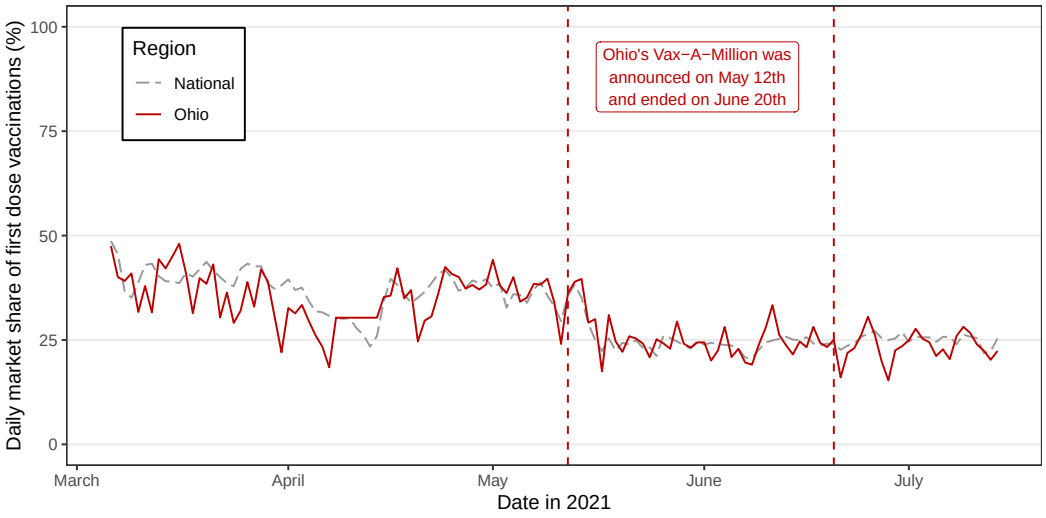
(b) Including states with pre-treatment RMSPE less than 10X that of Ohio

Notes: States not listed are not in the donor pool. The outcome is the share of population with any COVID-19 vaccination (at least a first dose) as of May 26, 2021. Pre-treatment RMSPE are computed using the full pre-treatment period in the data, starting on February 19, 2021 and ending on May 11, 2021. The grey error bars indicate the respective 95 percent confidence intervals, which are calculated using conformal inference.

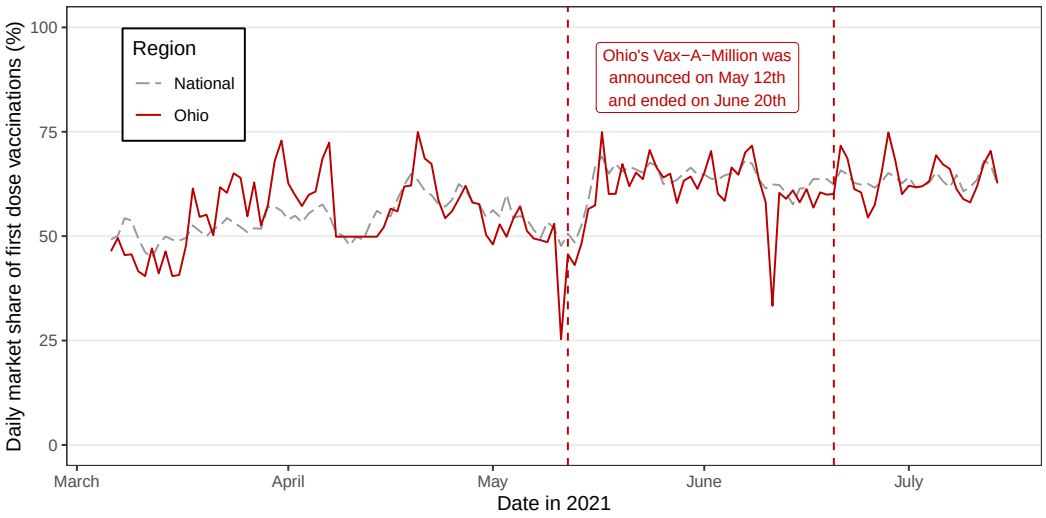
Figure A7: Manufacturers' daily market shares of first dose vaccinations over time



(a) Janssen



(b) Moderna



(c) Pfizer