

## Article

# Population-Level Effectiveness of COVID-19 Vaccination Program in the United States: Causal Analysis Based on Structural Nested Mean Model

Rui Wang <sup>1</sup>, Jiahao Wang <sup>2,3</sup>, Taojun Hu <sup>1</sup> and Xiao-Hua Zhou <sup>1,4,\*</sup>

<sup>1</sup> Department of Biostatistics, School of Public Health, Peking University, Beijing 100191, China; wangruiuishou@pku.edu.cn (R.W.); 2011110158@stu.pku.edu.cn (T.H.)

<sup>2</sup> School of Public Health, Peking University, Beijing 100191, China; jiahaowang@pku.edu.cn

<sup>3</sup> China Center for Health Development Studies, Peking University, Beijing 100191, China

<sup>4</sup> Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China

\* Correspondence: azhou@math.pku.edu.cn

**Abstract:** Though COVID-19 vaccines have shown high efficacy, real-world effectiveness at the population level remains unclear. Based on the longitudinal data on vaccination coverage and daily infection cases from fifty states in the United States from March to May 2021, causal analyses were conducted using structural nested mean models to estimate the population-level effectiveness of the COVID-19 vaccination program against infection with the original strain. We found that in the US, every 1% increase of vaccination coverage rate reduced the weekly growth rate of COVID-19 confirmed cases by 1.02% (95% CI: 0.26%, 1.69%), and the estimated population-level effectiveness of the COVID-19 program was 63.9% (95% CI: 18.0%, 87.5%). In comparison to a no-vaccination scenario, the COVID-19 vaccination campaign averted 8.05 million infections through the study period. Scenario analyses show that a vaccination program with doubled vaccination speed or with more rapid vaccination speed at the early stages of the campaign would avert more infections and increase vaccine effectiveness. The COVID-19 vaccination program demonstrated a high population-level effectiveness and significantly reduced the disease burden in the US. Accelerating vaccine rollout, especially at an early stage of the campaign, is crucial for reducing COVID-19 infections.

**Keywords:** COVID-19; vaccines; causal inference; effectiveness; structural nested mean models



**Citation:** Wang, R.; Wang, J.; Hu, T.; Zhou, X.-H. Population-Level Effectiveness of COVID-19 Vaccination Program in the United States: Causal Analysis Based on Structural Nested Mean Model. *Vaccines* **2022**, *10*, 726. <https://doi.org/10.3390/vaccines10050726>

Academic Editor: Giuseppe La Torre

Received: 10 December 2021

Accepted: 27 April 2022

Published: 5 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Since December 2019, the coronavirus disease 2019 (COVID-19) pandemic has rapidly spread and caused severe disease burden globally [1]. Countries worldwide have strived to boost COVID-19 vaccine coverage to contain the spread of the virus and reduce the burden of severe disease [2]. In the United States (US), three COVID-19 vaccines, developed by Pfizer-BioNTech, Moderna, and Johnson & Johnson, were distributed for public use, and the coverage rate of at least one dose of the COVID-19 vaccine rapidly increased from 1.3% on 1 January 2021 to 64.0% on 1 October 2021 among the general population of the United States [3].

Despite randomized clinical trials showing high efficacy of COVID-19 vaccines against infection (ranging from 66.9–95%) [4–6], the real-world effectiveness of the vaccines at the population level remains unclear. The vaccine direct effectiveness, which is a measure of the direct effect of the vaccine on vaccinated individuals, has been evaluated at the individual level using cohort studies (from the same population) [7–10] or case-control studies [11], and their results showed a lower risk of infection among vaccinated individuals compared with those unvaccinated. However, these designs are limited in capturing herd immunity (indirect effect) and population-level effectiveness (overall effect of a vaccination program in the population) [12–14]. Furthermore, these designs were unable to examine the population-level

outcomes of mass vaccination (e.g., the growth rate of infection cases [15–18], the number of cases averted [19]), which have important public health implications and provide evidence for policymaking. Some methods have been adopted to study the population-level effect (overall effect) of the COVID-19 vaccines, most of which are mathematical models. For example, compartment models and agent-based models are built to simulate how cases could evolve under different COVID-19 vaccination strategies [20–24]. However, these models rely on estimated parameters unconnected to the actual epidemic and vaccination data.

The challenges in evaluating the population-level effectiveness of the COVID-19 vaccines are the utilization and integration of real-world data on the epidemic and vaccination that evolved dramatically over time. The relationship between epidemic outcome (e.g., growth rate of new cases) and vaccinated population is complex and dynamic because various factors influenced vaccine uptake, such as perceptions of vaccination [25], severity of the epidemic, and intensity of non-pharmaceutical interventions [26,27]. The time-varying characteristics of these factors may confound the relationship between the vaccinated population at the present stage and the growth rate of infection cases in the subsequent period, and they may also be impacted by the vaccinated population (treatment) and the growth rate of new infection cases (outcome) at the previous stage. In the past, there were studies that used longitudinal data at the population level as statistical models to assess the effectiveness of other vaccines, such as influenza vaccine, meningococcal B vaccine, and BCG vaccine. However, no such studies were conducted regarding COVID-19 vaccination, to the best of our knowledge. Furthermore, in the context of COVID-19 vaccination, these methods are unable to handle the rapidly evolving relationship between time-varying treatments, outcomes, and confounders. For example, generalized estimating equations [28] or fixed-effects models [29] will generate biased estimates when there exist time-varying confounders that are impacted by previous independent variables [30–33]. The Susceptible-Infected-Recovered (SIR) model cannot handle confounders, either. As a result, these methods only give predictions based on the associations in observed data and lack causal interpretation due to the failure to adjust for confounders. Synthetic control [34] or interrupted time series [35] can only evaluate binary treatment.

In this paper, we studied the relationship between the vaccinated population and epidemic outcomes at population level. Structural nested mean models (SNMMs) were used to deal with time-varying confounders [36]. SNMMs can easily handle continuous treatments [36–38]. Moreover, as defined in the potential outcome framework [39,40], SNMMs have the advantage of providing a causal interpretation of the estimates and results. Based on the longitudinal data on COVID-19 vaccination coverage and weekly confirmed infection cases in the United States, we estimated the effect of the vaccinated population on the growth rate of confirmed new cases using SNMMs. We then calculated the population-level effectiveness of the COVID-19 vaccination program against the original COVID-19 virus strain and the disease burden averted during the study period, compared with a no-vaccination scenario. This study identified that every 1% increase in vaccine coverage would reduce the weekly growth rate of COVID-19 cases by 1.02% (95% CI: 0.26%, 1.69%), and stressed the significance of intensive vaccine rollout, especially at early stages, in averting infections and increasing vaccine effectiveness.

## 2. Materials and Methods

### 2.1. Data

We collected state-level daily COVID-19 cases and vaccine coverage data from 1 March to 30 May 2021 in the U.S. from Johns Hopkins Coronavirus Resource Center and the Centers for Disease Control and Prevention (CDC) [1,3]. We also collected state baseline characteristics, including GDP, population, health resource (number of physicians), political position (red or blue states in the 2016 election), proportion of people aged 65 or above, race composition, education level (proportion of people with advanced degrees), unemployment rate, and sex ratio. Their sources can be found in Appendix A Table A1. Overall government

response index was collected from the Oxford COVID-19 Government Response Tracker dataset [41].

Data from 49 states and Washington, DC, in the U.S. were included in the analyses, excluding New Hampshire due to errors in the vaccinated population data. The study aimed to evaluate the population-level effectiveness of COVID-19 vaccination program against the original strain. Therefore, the period after June 2021 was excluded because the proportion of SARS-CoV-2 variants changed radically since then, making the Delta variant gradually predominate as the main variant of transmission. The periods before March 2021 were also excluded due to missing data on vaccination coverage. The data were aggregated at the weekly level to avoid the collection and report patterns of the confirmed case data, resulting in the study period of 13 weeks at 50 states (including DC) for analysis (first week: 1 March to 7 March; 13th week: 24 May to 30 May).

### 2.2. Treatment and Outcome Variable

Treatment (denoted as  $A_{i,t}$ ) represents the number of people (per 10,000) who received their first dose in state  $i$  at week  $t$ . The key outcome variable, the growth rate of new COVID-19 cases for each week [15–18], is calculated as

$$Y_{i,t} = \log(C_{i,t+1}) - \log(C_{i,t}), \tag{1}$$

where  $C_{i,t}$  is the number of new cases reported in state  $i$  at week  $t$ , and  $Y_{i,t}$  represents the growth rate of new cases in state  $i$  at week  $t$ . We omitted index  $i$  when there was no ambiguity.

### 2.3. Causal Parameters of Interest

To define our causal parameters of interest, we formalized our ideas in a potential outcome framework, which was proposed by Donald Rubin [40] and extended to longitudinal settings by James Robins [42]. For each state, our data are of the form  $\{L_t, Y_t, A_1, \dots, L_{13}, Y_{13}, A_{13}, Y_{14}\}$ , where  $L_t$  denotes covariates, including time-varying covariates and baseline covariates (we mention which confounders we included in analysis in the next section). Assume at week  $t$ , first  $L_t$  was observed, then  $Y_t$  was observed, and  $A_t$  was the last to be observed.

Let  $X_t$  be an arbitrary variable. We use  $\bar{X}_t$  to denote the history of  $X_t$ , that is to say,  $\bar{X}_t = (X_1, \dots, X_t)$ . Then,  $\bar{A}_t = (A_1, \dots, A_t)$  represents the vaccinated population (per 10,000 people) each week prior to  $t$ ; we refer  $\bar{a}_t$  as vaccination speed because it represents the first-order difference in vaccination coverage each week. The potential outcome can be written as  $Y_{t+1}(\bar{a}_t)$ , which is the outcome that would have been observed if the vaccination population had been set to  $\bar{A}_t = \bar{a}_t$ . Moreover, we use  $Y_{t+1}(\bar{a}_s, 0)$  for abbreviation of  $Y_{t+1}(\bar{a}_s, 0, \dots, 0)$ , which is the potential outcome under intervention which takes the value  $\bar{a}_s$  at the first  $s$  periods and takes value 0 for the remaining periods. The potential outcomes and observed outcomes are linked by the consistency assumption [33], which means that if a region's observed vaccination speed at stage  $t + 1$  is  $\bar{a}_t$ , then  $Y_{t+1} = Y_{t+1}(\bar{a}_t)$ .

Our first causal parameter of interest is the average treatment effect [31]:

$$E\left[Y_{t+1}(\bar{a}_t^{(1)}) - Y_{t+1}(\bar{a}_t^{(2)})\right]. \tag{2}$$

This is the contrast of the mean difference of outcome under different vaccination speeds  $\bar{a}_t^{(1)}$  and  $\bar{a}_t^{(2)}$ , representing the effect of vaccination speed (vaccine population per 10,000 people) before  $t + 1$  on the growth rate.

Moreover, how many new cases we would have observed each week in the United States if a certain vaccination speed  $\bar{a}_{i,t}^*$  had been accomplished is also of interest to us, and is

$$\sum_{i=1}^{50} C_{i,t+1}(\bar{a}_{i,t}^*). \tag{3}$$

Here  $C_{i,t+1}(\bar{a}_{i,t}^*)$  is the potential outcome of new cases at state  $i$  and week  $t$ , and  $\bar{a}_{i,t}^*$  is the segment of first  $t$  components of  $\bar{a}_{i,13}^*$ .

Lastly, define the population-level vaccine effectiveness (PVE) of vaccination speed  $\bar{a}_{i,13}^*$  as

$$1 - \frac{\sum_{t=1}^{13} \sum_{i=1}^{50} C_{i,t+1}(\bar{a}_{i,t}^*)}{\sum_{t=1}^{13} \sum_{i=1}^{50} C_{i,t+1}(0)}. \tag{4}$$

This population-level vaccine effectiveness is one minus the ratio of new cases under different vaccination speeds, similar to the definition of the usual vaccine effectiveness estimand [43]. This value can be interpreted as the proportion of new cases prevented by the vaccination speed  $\bar{a}_{i,13}^*$  compared to the no-vaccination scenario.

#### 2.4. Confounders

A sufficient set of confounders was adjusted to yield causal interpretation of our results. We adjusted the baseline cumulative number of cases and vaccination coverage. Based on the vaccine hesitancy framework and impact factors on vaccine uptake [27], we adjusted three types of confounders. The first type of confounder was related to people’s perception of risk and severity of COVID-19 pandemic; thus, we adjusted the lagged growth rate of new cases as an indicator of severity of the COVID-19 pandemic. It is a time-varying confounder. The second type of confounder was related to social influence from government and surrounding people. We used the overall government response index as an indicator of government behavior. We used vaccination coverage and vaccinated population (per 10,000) in the last week as indicators of surrounding people’s behavior. The confounders of the second type are also time-varying ones. For the third type of confounder, demographic and socio-economic factors at the state level were taken into consideration, including GDP, population, health resources (number of physicians), political position (red or blue states in the 2016 election), proportion of people aged 65 or above, racial composition, education level (proportion of people with advanced degrees), unemployment rate, and sex ratio.

#### 2.5. Data Analysis

##### 2.5.1. Structural Nested Mean Model

Our analysis was mainly based on the structural nested mean model. The structural nested mean model parameterized the conditional average treatment effect:

$$E\left[Y_{t+1}(\bar{a}_s, 0) - Y_{t+1}(\bar{a}_{s-1}, 0) \mid \bar{L}_s = \bar{l}_s, \bar{Y}_s = \bar{y}_s, \bar{A}_s = \bar{a}_s\right], \tag{5}$$

where  $t = 1, \dots, 13$ , and  $s \leq t$ . Following previous literature [15–18], we specified a linear model for the growth rate as

$$E\left[Y_{t+1}(\bar{a}_s, 0) - Y_{t+1}(\bar{a}_{s-1}, 0) \mid \bar{L}_s = \bar{l}_s, \bar{Y}_s = \bar{y}_s, \bar{A}_s = \bar{a}_s\right] = \psi a_s. \tag{6}$$

This SNMM specification was commonly used in past epidemiology literature [44,45]. Our SNMM could imply a marginal structural model (MSM) [46]:

$$E[Y_{t+1}(\bar{a}_t)] = E[Y_{t+1}(0)] + \psi \sum_{k=1}^t a_k. \tag{7}$$

Then, our causal parameter of interest could be represented as

$$E[Y_{t+1}(\bar{a}_t) - Y_{t+1}(\bar{a}_t^*)] = \psi \sum_{k=1}^t (a_k - a_k^*). \tag{8}$$

In our dataset,  $a_k$  represents the number of people out of every 10,000 people who were vaccinated for the first dose; then,  $\sum_{k=1}^t a_k$  is the number of people who were vaccinated for the first dose since week 1. The parameter  $100 * \psi$  can be interpreted as the percentage of growth rate that decreases for every percent increase in the vaccinated population, starting from the first week.

To estimate  $\psi$ , we also specified a propensity score model (i.e., models for vaccinated population per 10,000 people) and an outcome model in each week to adjust for confounders. We utilized Poisson regression models for propensity scores and linear regression models for outcomes (Appendix B).  $\psi$  can be estimated using g-estimation (Appendix B) [36–38]. The standard error and 95% confidence interval were bootstrapped [47]. The estimation is doubly robust in the sense that the estimation for  $\psi$  is unbiased if either the propensity score models or the outcome models are correctly specified [38].

### 2.5.2. Generalized Estimating Equation Models and Fixed Effects Models

We also fitted two generalized estimating equation (GEE) models [48,49] and two fixed effects models for comparison.

Two GEE models were fitted. The first GEE model regressed growth rate on the vaccinated population since week 1, adjusting baseline covariates. The second one regressed growth rate on the vaccinated population since week 1, adjusting baseline covariates, time-varying covariates, and time trend. We utilized the independent working correlation structure in our GEE analysis.

Two fixed effects models were also fitted. The first fixed effects model included state fixed effects, week fixed effects and vaccinated population since week 1. The second one included state fixed effects, week fixed effects, time-varying covariates, and vaccinated population since week 1. A detailed model specification of the GEE models and fixed effects models can be found in Appendix C.

### 2.5.3. Scenarios Analysis

We also used the estimate from SNMM and g-estimation to predict the number of new cases each week in the United States under different vaccination speeds. For each state  $i$ , we utilized the following relationship to predict the counterfactual growth rate under vaccination speed  $\bar{a}_{i,13}^*$ :

$$Y_{i,t+1}(\bar{a}_{i,t}^*) = Y_{i,t+1} + \psi \sum_{k=1}^t (a_{i,k}^* - a_{i,k}^\dagger). \quad (9)$$

Here,  $a_{i,k}^*$  is the  $k$ -th component of  $\bar{a}_{i,13}^*$ , and  $a_{i,k}^\dagger$  refers to the observed vaccinated population in state  $i$ , week  $k$ . We predicted the counterfactual growth rate under the following hypothetical scenarios. In each scenario, we set a different vaccination speed.

- No-vaccination scenario: No people are vaccinated since week 1, which means  $\bar{a}_{i,13}^* = \bar{0}_{13}$ .
- Twice speed scenario: The number of people vaccinated for the first time each week is twice the actual number in each state, which means  $\bar{a}_{i,13}^* = (2a_{i,1}^\dagger, 2a_{i,2}^\dagger, \dots, 2a_{i,13}^\dagger)$ .
- Half speed scenario: The number of people vaccinated for the first time each week is half of the actual number in each state, which means  $\bar{a}_{i,13}^* = \left(\frac{a_{i,1}^\dagger}{2}, \frac{a_{i,2}^\dagger}{2}, \dots, \frac{a_{i,13}^\dagger}{2}\right)$ .
- 1% constant speed scenario: 1% of the population receive their first dose in each week in each state, which means  $\bar{a}_{i,13}^* = (100, 100, \dots, 100)$ .
- 4% constant speed scenario: 4% of the population receive their first dose each week in each state, which means  $\bar{a}_{i,13}^* = (400, 400, \dots, 400)$ .
- Speed up scenario: For the first six weeks, 1% of the population receive their first dose in each week in each state, while for the remaining seven weeks, 4% of the population

receive their first dose in each week in each state, which means  $\bar{a}_{i,13}^* = (100, 100, \dots, 100, 400, \dots, 400)$ .

- Speed down scenario: For the first seven weeks, 4% of the population receive their first dose in each week in each state, while for the remaining six weeks, 1% of the population receive their first dose in each week in each state, which means  $\bar{a}_{i,13}^* = (400, 400, \dots, 400, 100, \dots, 100)$ .

Under each counterfactual scenario, we utilized the relationship between growth rate and new cases to predict the number of new cases in each week, then calculated the population-level vaccine effectiveness.

### 2.5.4. Additional Analysis and Extension

To demonstrate the validity of the method adopted, we conducted additional analyses and extended our statistical models to two new datasets. Firstly, we predicted the new cases in each scenario during the same study period with a modified Suspected–Infected–Recovered (SIR) model. Secondly, we used the SNMM, GEE models and fixed effects models described above to estimate the effect of vaccination on weekly growth rate based on the dataset from September 2021 to December 2021 in the United States when the delta variant dominated the epidemic. See Supplementary Material for more details.

## 3. Results

### 3.1. Baseline Characteristics

From 1 March to 30 May 2021, 4.55 million confirmed cases were reported in the US. The coverage rate of at least one dose among the general population reached 50.0% from 15.2% (Figure 1). The average weekly vaccination speed in the United States was 2.7%. Table 1 shows the descriptive statistics of the baseline covariates of 50 states.

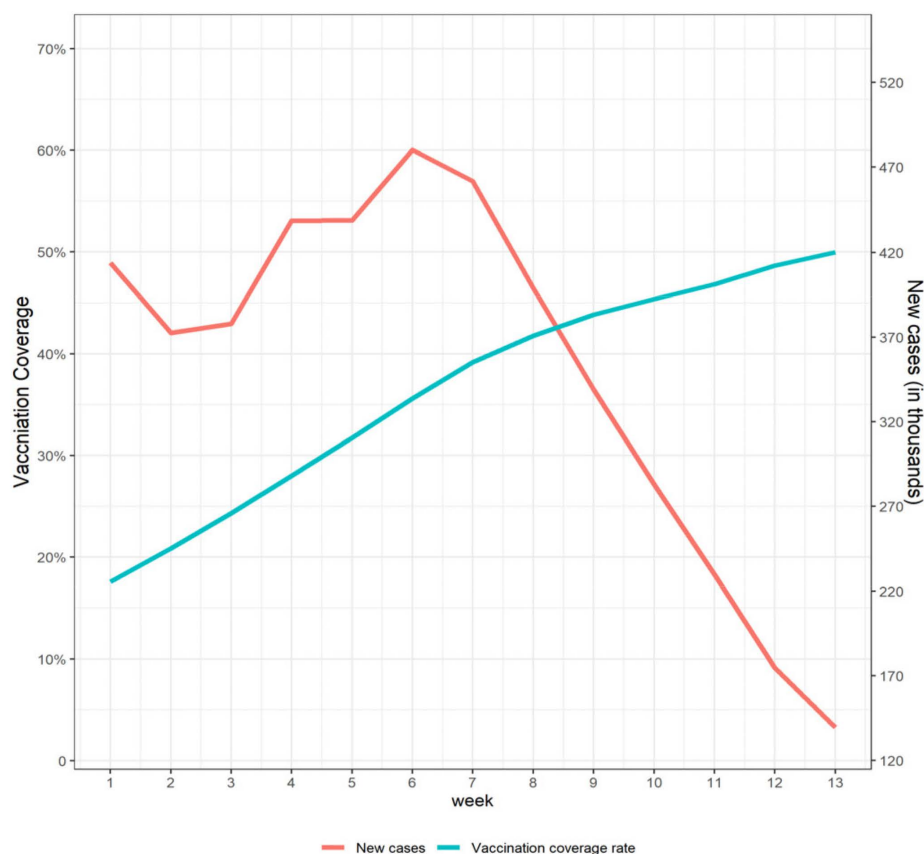


Figure 1. Vaccination coverage and new cases (in thousands) in the United States.



**Table 1.** Descriptive statistics of baseline covariates.

Variables	Mean	SD
Number of physicians (per capita)	475.2	192.6
GDP (millions of chained 2012 dollars)	375,880	490,590
Population	6,551,748	7,415,328
Race composition (proportion of black people)	0.132	0.109
Proportion of old people (aged 65 or above)	0.164	0.020
Red or Blue state in 2016 election (Red = 1)	0.6	0.495
Unemployment Rate (in March, 2021)	5.56	1.72
Proportion of people with advanced degrees	0.126	0.042
Sex ratio	0.977	0.033
Cumulative cases at baseline	569,301	664,218
Vaccination coverage at baseline (per 10,000 people)	1572	238

### 3.2. Impact of COVID-19 Vaccine Program on Weekly Growth Rate of COVID-19 New Cases

Table 2 shows the impact of the COVID-19 vaccination program on the weekly growth rate of new COVID-19 cases at the state level in the US. The result, based on the structural nested mean model (SNMM) with g-estimation, showed that for every 1% of the population becoming vaccinated, the growth rate of new COVID-19 infection cases reduced by 1.02% (95% CI: 1.69%, 0.26%). In contrast, the first GEE analysis with baseline covariates adjusted showed that every 1% increase in the COVID-19 vaccination coverage rate reduced the growth rate of new cases by 0.754% (95% CI: 0.974%, 0.533%), which was smaller than that of the g-estimation by approximately one standard error. The second GEE analysis, which adjusted the time trend and all covariates, estimated that every 1% increase in vaccination rate reduced the growth rate of new cases by 1.74% (95% CI: 2.42%, 1.05%), which was bigger than that of the g-estimation by approximately two standard errors. Results from the fixed-effects model showed similar patterns. The two-way fixed effects model showed that every 1% increase in the COVID-19 vaccination coverage rate reduced the growth rate of new cases by 1.52% (95% CI: 2.09%, 0.96%), which was bigger than that of the g-estimation by approximately 1.5 standard error. After adjusting for time-varying confounders, the two-way fixed effects model estimated that every 1% increase in the vaccination rate reduced the growth rate of new cases by 1.87% (95% CI: 2.43%, 1.30%) which was bigger than that of the g-estimation by approximately two standard errors.

**Table 2.** Impact of the COVID-19 vaccine program on the weekly growth rate of new COVID-19 cases.

	Decline of Growth Rate		
	Estimate	SE	95% CI
Main analysis			
SNMM with g-estimation	1.02%	0.0037	(1.69%, 0.26%)
GEE analysis			
GEE (adjust baseline covariates)	0.754%	0.00076	(0.974%, 0.533%)
GEE (adjust baseline and time-varying covariates)	1.74%	0.0035	(2.42%, 1.05%)
Fixed effects model			
Two-way fixed effects model	1.52%	0.0029	(2.09%, 0.96%)
Two-way fixed effects model (adjust time-varying covariates)	1.87%	0.0029	(2.43%, 1.30%)

### 3.3. Population-Level Effectiveness of COVID-19 Vaccination and Averted Disease Burden

Table 3 shows the population-level effectiveness of COVID-19 vaccination and averted disease burden in total. Based on the causal analyses, if the COVID-19 vaccination had been suspended after the first week, the total infection cases during the study period would have reached 12.60 million (95% CI: 5.55, 36.51) in the US over the 13 weeks. By contrast, there were 4.55 million new cases under the status quo. At the present vaccination speed, the population-level effectiveness of COVID-19 vaccination was estimated as 63.9%

(95% CI: 18.0%, 87.5%), and the COVID-19 vaccination program averted about 8 million cases. Figure 2a presents the comparison of cumulative cases over time between the status quo and the 'no-vaccination' scenario.

**Table 3.** Results of the base case and scenario analyses.

Scenarios	Cumulated New Cases (Million)		Vaccination Effectiveness	
	Estimate (95% CI)	Difference (%) <sup>a</sup>	Estimate (95% CI)	Difference (%) <sup>b</sup>
Base case				
Status quo	4.55 <sup>c</sup>	/	63.9% (18.0%, 87.5%)	/
Scenario analysis				
No-Vaccination	12.60 (5.55, 36.51)	8.05 (177%)	0%	−63.9%
vaccination speed: two times the status-quo speed	2.84 (2.34, 3.93)	−1.71 (−37.6%)	77.5% (29.2%, 93.6%)	13.6%
vaccination speed: half of the status-quo speed	7.10 (5.03, 10.88)	2.55 (56%)	43.7% (9.34%, 70.2%)	−20.2%
vaccination speed: 4% population per week	3.99 (3.71, 4.40)	−0.56 (−12.3%)	68.4% (20.7%, 89.8%)	4.5%
vaccination speed: 1% population per week	8.66 (5.21, 16.06)	4.11 (90.3%)	31.3% (6.07%, 56.02%)	−32.6%
Speed-down: 4% for first 7 weeks and 1% for last 6 weeks	4.13 (3.91, 4.44)	−0.42 (−9.2%)	67.3% (19.9%, 89.3%)	3.4%
Speed-up: 1% for first 6 weeks and 4% for last 7 weeks	7.52 (5.10, 11.47)	2.97 (65.3%)	40.3% (8.10%, 68.6%)	−23.6%

<sup>a</sup>. The difference is the estimate of cases under each scenario minus the cases under the status quo. The percentage in the bracket is the ratio of difference over cases under the status quo. <sup>b</sup>. The difference refers to estimates of vaccine effectiveness in each scenario minus the effectiveness under the status quo. <sup>c</sup>. The observed cases under the status quo.

### 3.4. Scenarios Analysis

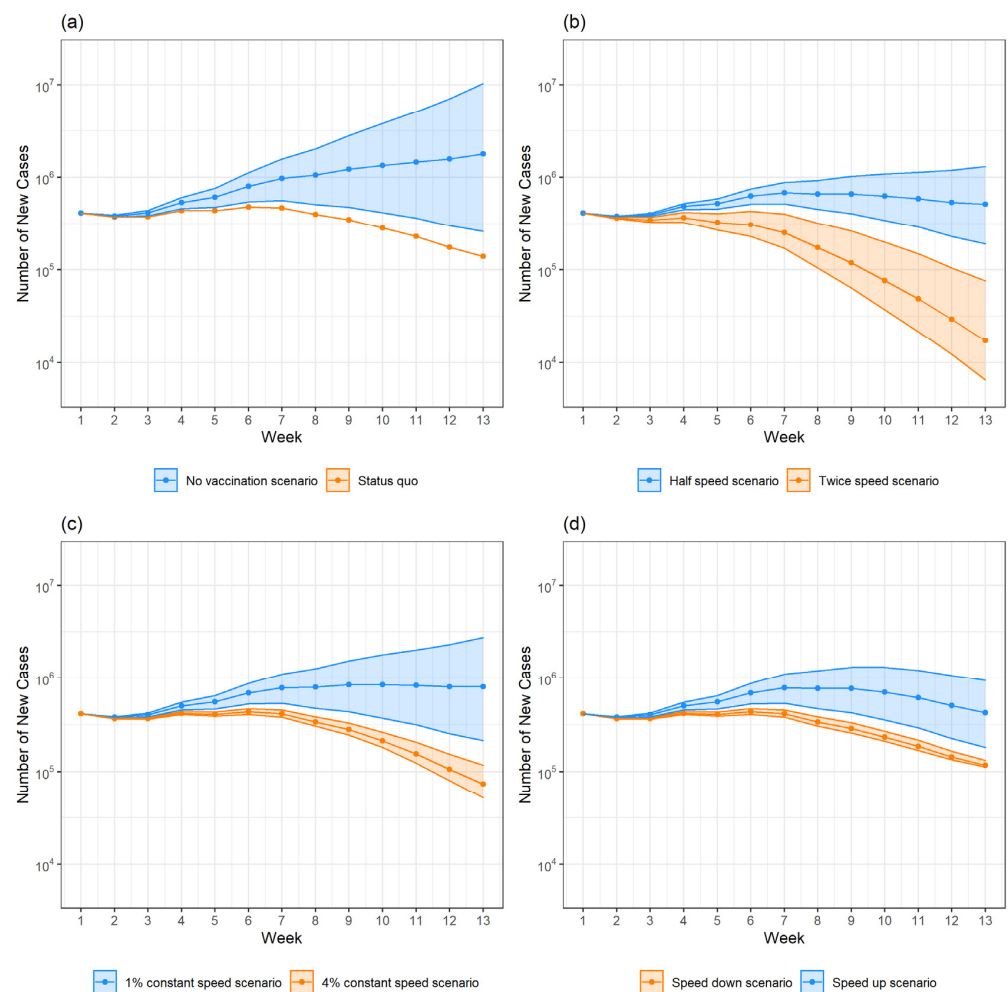
Table 3 and Figure 2b–d show the scenario analyses of the population-level effectiveness of COVID-19 vaccination and predicted disease burden of infection cases by different vaccination speeds.

In comparison to the status quo, doubling vaccination speed would have increased the vaccine effectiveness to 77.5% (95% CI: 29.2%, 93.6%), averting an additional 1.71 million cases. By contrast, if the vaccination program had been implemented at half the current speed, vaccination effectiveness would have declined to 43.7% (95% CI: 9.34%, 70.2%) with 2.55 million more infection cases. (Figure 2b).

As shown in Figure 2c, we predicted the results of the two scenarios where the vaccination programs were conducted at a constant speed of 1% or 4% of the general population per week. In the scenario with the speed of 1% of the population vaccinated per week, there would be 8.66 million cases (95% CI: 5.21, 16.06) with an estimated effectiveness of 31.3% (95% CI: 6.07%, 56.02%). In the scenario with the speed of 4% of the population vaccinated per week, there would be 3.99 million cases (95% CI: 3.71, 4.40) with an effectiveness of 68.4% (95% CI: 20.7%, 89.8%).

In the last set of comparisons (Figure 2d), we compared the results of 'speed-up' and 'speed-down' scenarios, keeping the total amount of vaccination during the 13 weeks basically the same as that of the status quo (34.8%). In the 'speed-down' scenario, where the vaccination speed was 4% of the population per week for the first seven weeks and 1% for the last six weeks, vaccination speed changed from fast to slow. There would be 4.13 million cases (95% CI: 3.91, 4.44) with the vaccine effectiveness being approximately 67.3% (95% CI: 19.9%, 89.3%). However, if the vaccination speed had been 1% of the population per week for the first six weeks and 4% for the last seven weeks, vaccination speed changed from slow to fast. The total cases would have reached 7.52 million cases (95% CI: 5.10, 11.47), and the estimated vaccine effectiveness would have declined to 40.3% (95% CI: 8.10%, 68.6%).





**Figure 2.** Predicted number of new cases in each week under different scenarios. (a) shows the comparison of predicted number of new cases under no vaccination scenario and status quo. (b) shows the comparison of predicted number of new cases under half speed scenario and twice speed scenario. (c) shows the comparison of predicted number of new cases under 1% constant speed scenario and 4% constant speed scenario. (d) shows the comparison of predicted number of new cases under speed up scenario and speed down scenario.

#### 4. Discussion

The structural nested mean model examined the causal relationship between COVID-19 vaccination coverage and the growth rate of new infection cases at the population level. From 1 March to 30 May 2021, for every 1% of the population becoming vaccinated, the growth rate of new COVID-19 infection cases was reduced by 1.02% (95% CI: 0.26%, 1.69%) in the US. In comparison to a no-vaccination scenario, the current COVID-19 vaccination program averted 8 million infection cases, with an estimated vaccine effectiveness of 63.9% (95% CI: 18.0%, 87.5%). Our results show that intensive vaccination rollout, especially at the early stages of the vaccination campaign, is crucial for reducing infections.

The real-world vaccine effectiveness at the individual level has been studied, which typically focuses on the vaccine's direct effectiveness [14]. For example, the vaccine effectiveness of BNT162b2 or mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection was 90% for full immunization and 80% for partial immunization in the US [8]. The case-control study in England estimated the effectiveness of BNT162b2 to be 70% in 10 to 13 days after the first dose and 89% in 14 days after the second dose [11]. The cohort study in Sweden showed that the estimated effectiveness of BNT162b2 in preventing infection >7 days after the second dose was 86%, but only 42% for >14 days after a single dose [7].

Our estimate at the population level was smaller than that at the individual level for the following reasons. First, it should be noted that vaccine effectiveness at the population level demonstrated the effect of the general vaccine program, which is conceptually different from that at the individual level [14]. The population-level effectiveness was estimated by comparing the effect of vaccination programs, at different vaccination speeds and in different states, on the infection. In comparison, the individual-level effectiveness was estimated by comparing the risks of vaccinated and non-vaccinated individuals. Secondly, the population-level effectiveness would be influenced by the change in population vaccination coverage at a given period of time. The increment of vaccination coverage was approximately 35% in our study period. The population-level effectiveness might be larger if a longer period were examined.

So far, a few studies have examined the population-level effectiveness of COVID-19 vaccines using an agent-based model. Moghadas et al. [20] used an agent-based model and found that compared to the no-vaccination scenario, a vaccination campaign achieving 40% coverage of the entire population within 284 days would reduce cases by 50% [20]. Their increment of vaccination coverage was close to that of our study period (approximately 35%) and their result was also close to our estimated 63.9% effectiveness. Their agent-based model depicted the whole disease dynamics, which required strong modeling assumptions about how individuals transit between different states. The simulations relied heavily on pre-estimated model parameters. Our causal approach, on the other hand, focused more on the relationship between independent and outcome variables and retrospectively utilized real data during the study period, including vaccination coverage data, epidemic data and other covariates.

When handling longitudinal data to estimate the population-level effectiveness of the COVID-19 vaccines, it is essential to address the issue of potential time-varying confounders, which may be impacted by the treatment (vaccinated population) at the prior period. The use of SNMM helped to address the issue correctly. In comparison, the GEE model adjusting baseline covariates and the two-way fixed effects models could produce biased estimates because they did not adjust the time-varying confounders. Therefore, we found that their results differed from the estimate of SNMM by at least one standard error. The GEE model adjusting baseline and time-varying confounders and the two-way fixed effects models adjusting time-varying confounders could produce biased estimates because they adjusted time-varying confounders by directly conditioning on those confounders, which would produce collider bias [30,31,33,50]. This bias may explain why their results differed from the estimate of SNMM by two standard errors. Similarly, SIR models did not adjust any confounders, which may account for the difference between the prediction by SNMM and the prediction by the SIR model. Moreover, based on the potential outcome framework, our analyses allowed us to clearly define causal estimands without ambiguity and thus enable a causal interpretation. The causal framework is advocated in evaluating vaccine performance, including efficacy and effectiveness [51]. Unlike other causal methods for assessing vaccine effectiveness, such as the synthetic control method and the fixed effects model, our model could study the causal effect of the continuous treatments under a quickly evolving and dynamic context, especially in the current COVID-19 pandemic.

One key assumption for causal interpretation is sequential ignorability [36,46]. To make this assumption hold, all confounders between treatments and outcomes in each stage must be observable and properly adjusted. The literature on vaccine hesitancy helped in determining the set of confounders that needed to be adjusted [27]. We included a sufficiently large set of variables based on the vaccine hesitancy framework to ensure that confounders were adequately adjusted in this context. Another key assumption for causal interpretation is consistency [36,46], which implicated no hidden versions of treatments [52]. A possible violation of this assumption may be due to the different types of COVID-19 vaccines used for the vaccination campaign, because their efficacy varies [4–6], but this problem is negligible because the Johnson and Johnson vaccine was only used by a small proportion of the population (approximately 6%) [3]. The two widely-used types of vaccines

(Pfizer-BioNTech and Moderna) showed similar efficacy in randomized clinical trials (94.1% vs. 95%) [4,5].

This study quantified the impact of vaccination speed on vaccine effectiveness and disease burden averted. Compared with the status quo, an additional 1.71 million cases could be averted if the vaccination program were launched at twice the actual vaccination speed; an additional 0.56 million cases could be averted if the vaccination program were launched at a constant speed of 4% of the population for each week. Simulations from mathematical models also yielded similar results. A study found that cumulative mortality would be 442, 241, and 50 (per 100,000) for daily vaccination rates of 0.1%, 0.3%, and 1% respectively [24]. Another study found that accelerating vaccine delivery would substantially reduce severe health outcomes [53].

Many countries around the world have made plans to achieve a specific vaccine coverage rate at a specific time point. For example, the U.S. government planned to achieve a goal of 70 percent vaccine coverage by 4 July 2021 [54]. The South Korean government planned to achieve a goal of 70 percent vaccine coverage by the end of the third quarter of September [55]. However, there have not been suggestions regarding targeting vaccine coverage rates at each stage over the course of the vaccination campaign. The simulated scenarios, especially the ‘speed up’ and ‘speed down’ scenarios, help to demonstrate the importance of accelerating the vaccination rollout in the early stages. First, the comparison between the ‘speed up’ scenario and ‘speed down’ scenario shows that even if the total increment of the vaccine coverage is the same over a period of the campaign, the rate of vaccination at each stage (e.g., week) of the period can still have a significant impact on the population-wide effectiveness of the vaccine. If the vaccination coverage increment over the 13 weeks is the same (approximately 35%) as the status quo, the ‘speed up’ scenario, which has a lower vaccination speed at the early stage, would cause 3.13 million additional cases compared to the ‘speed down’ scenario. Second, we found that the ‘speed up’ scenario was significantly worse than the ‘status quo’, as the 95% confidence interval of cases under ‘speed up’ (5.10 million to 11.47 million) does not cover the cases under the status quo (4.55 million). Third, we found that the ‘speed down’ scenario was significantly better than the ‘status quo’ because the 95% confidence interval of cases under ‘speed down’ (3.91 million to 4.44 million) did not cover the cases under the status quo (4.55 million) either. In addition, from a scale perspective, the ‘speed down’ scenario could reduce cases by an additional 0.42 million, which equals 10% of the status quo cases during the period. In fact, the status quo was similar to the ‘speed down’ where vaccination speed declined gradually over time, because the average speed of vaccination under the status quo was higher for the first seven weeks than for the last six weeks, which means reality also followed the pattern that the vaccination speed was faster in the early stage. This fact explains why the ‘speed down’ curve in Figure 2 is closer to status quo.

Our results have several important policy implications. First, increasing vaccination coverage is critical for reducing infections. By the end of May, the proportion of people vaccinated reached 50% in the United States, and the incidence of COVID-19 continued to decline. In many other countries and regions, however, the vaccination coverage rate is still low. For example, by the end of May, only 2 doses of COVID-19 vaccines were administered per 100 people in Africa (in comparison, the United States had administered 93 doses per 100 people), and Africa was facing a fast-surging third wave of the COVID-19 pandemic at the same time [56]. Second, accelerating the vaccination speed at the early stage is crucial for reducing infections. Therefore, interventions should be conducted from the beginning of vaccination campaigns to increase the vaccination speed, including approaches such as scaling up manufacturing, promoting effective deployment and avoiding unnecessary waste, and health education to eliminate vaccine hesitancy. These measures are also critical for the current period, given the spread of new COVID-19 variants and the need for booster shots in the US and other countries.

Our paper has several limitations. First, our structural nested mean model assumes a constant effect of vaccinated population on subsequent growth rate during the study

period for the convenience of explaining the parameter. Although this model specification has also been widely adopted in other research based on longitudinal causal inference methods [44,46], non-constant treatment effect may occur with changes in vaccination coverage. Second, due to different epidemic patterns of the delta variant, the present results only report the vaccine effectiveness against the original variants based on data from March to May. Future studies could extend the structural nested mean models (SNMMs) to examine the effect of COVID-19 vaccination programs against other variants.

## 5. Conclusions

The COVID-19 vaccination program demonstrated a high level of effectiveness at the population level and significantly reduced the disease burden in the United States. We quantified the impact of vaccination speed on the effectiveness of the COVID-19 vaccination program and the disease burden averted. Intensive vaccination rollout, especially at the early stages of the vaccination campaign, is crucial for reducing COVID-19 infections.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/vaccines10050726/s1>, Figure S1: Directed Acyclic Graph (DAG), Figure S2: Simulating the SIR model with  $\sigma = 0$ , Figure S3: Simulating the SIR model with  $\sigma = 0.05$ , Figure S4: Simulating the SIR model with  $\sigma = 0.15$ , Figure S5: Simulating the SIR model with  $\sigma = 0.25$ , Figure S6: Simulating the SIR model with  $\sigma = 0.35$ , Figure S7: Simulating the SIR model with  $\sigma = 0.45$ , Figure S8: The vaccination coverage and new cases (in thousands) in the United States (from 5 September 2021 to 4 December 2021), Table S1: Impact of COVID-19 vaccine program on weekly growth rate of COVID-19 new cases, Figure S9: Predicted Number of new cases in each week under different scenarios (September 2021–December 2021), Table S2: Results of the base case and scenarios analyses (September 2021–December 2021).

**Author Contributions:** Conceptualization, R.W., J.W. and X.-H.Z.; Data curation, R.W.; Formal analysis, R.W. and T.H.; Methodology, R.W., T.H. and J.W.; Supervision, X.-H.Z.; Writing—original draft, R.W.; Writing—review and editing, R.W., J.W., T.H. and X.-H.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Bill & Melinda Gates Foundation (INV-016826) and National Natural Science Foundation of China (grant number: 82041023).

**Institutional Review Board Statement:** All data used in this study are publicly available.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. This data can be found here: [https://github.com/wangrui24/COVID\\_Vaccine\\_paper](https://github.com/wangrui24/COVID_Vaccine_paper) (accessed on 5 April 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Data sources.

Variables	Value Ranges	Source, Period of Data Collection
Number of physicians (per capita)	Not Applicable	<a href="https://www.fsmb.org/siteassets/advocacy/publications/2018census.pdf">https://www.fsmb.org/siteassets/advocacy/publications/2018census.pdf</a> 2018 Accessed on 6 June 2021
GDP (millions of chained 2012 dollars)	Not Applicable	<a href="https://apps.bea.gov/regional/histdata/releases/0720gdpstate/index.cfm">https://apps.bea.gov/regional/histdata/releases/0720gdpstate/index.cfm</a> 2019 Accessed on 6 June 2021
Population	Not Applicable	<a href="https://apps.bea.gov/regional/histdata/releases/0320sqpi/index.cfm">https://apps.bea.gov/regional/histdata/releases/0320sqpi/index.cfm</a> Fourth Quarter 2019 Accessed on 6 June 2021

**Table A1.** *Cont.*

Variables	Value Ranges	Source, Period of Data Collection
Racial composition (proportion of black people)	0–1	<a href="http://www.census.gov/library/visualizations/interactive/race-and-ethnicity-in-the-united-state-2010-and-2020-census.html">http://www.census.gov/library/visualizations/interactive/race-and-ethnicity-in-the-united-state-2010-and-2020-census.html</a> 2019 Accessed on 6 June 2021
Proportion of old people (aged 65 or above)	0–1	<a href="https://acl.gov/sites/default/files/Aging%20and%20Disability%20in%20America/2019ProfileOlderAmericans508.pdf">https://acl.gov/sites/default/files/Aging%20and%20Disability%20in%20America/2019ProfileOlderAmericans508.pdf</a> 2019 Accessed on 6 June 2021
Red or Blue state in 2016 election (Red = 1)	0 and 1	<a href="https://www.fec.gov/resources/cms-content/documents/federalelections2016.pdf">https://www.fec.gov/resources/cms-content/documents/federalelections2016.pdf</a> 2016 Accessed on 6 June 2021
Unemployment Rate	Not Applicable	<a href="https://www.ncsl.org/research/labor-and-employment/state-unemployment-update.aspx/">https://www.ncsl.org/research/labor-and-employment/state-unemployment-update.aspx/</a> March, 2021 Accessed on 10 November 2021
Proportion of people with advanced degrees	0–1	<a href="https://acl.gov/sites/default/files/Aging%20and%20Disability%20in%20America/2019ProfileOlderAmericans508.pdf">https://acl.gov/sites/default/files/Aging%20and%20Disability%20in%20America/2019ProfileOlderAmericans508.pdf</a> 2019 Accessed on 10 November 2021
Sex ratio	Not Applicable	<a href="https://worldpopulationreview.com/state-rankings/male-to-female-ratio-by-state">https://worldpopulationreview.com/state-rankings/male-to-female-ratio-by-state</a> 2021 Accessed on 10 November 2021

**Appendix B. Details about the Estimation of SNMM**

In this paper, we adopted the estimation procedure proposed by Vansteelandt [47]. The estimation procedure is described in the following steps:

Step 1: We fit a pooled Poisson regression model for vaccination population (per 10,000 people) in each week:

$$\begin{aligned}
 E[A_{i,t} | \bar{A}_{i,t-1}, \bar{L}_{i,t}, \bar{Y}_{i,t}] &= \exp(\alpha_0 + \alpha_1 t + \alpha_1 t^2 + \alpha_3 \text{physicians}_i + \alpha_4 \text{election}_i + \alpha_5 \text{GDP}_i + \alpha_6 \text{race}_i \\
 &+ \alpha_7 \text{old}_i + \alpha_8 \text{populatiin}_i + \alpha_9 \text{cumcasei0}_i + \alpha_{10} \text{basovac}_i + \alpha_{11} \text{education}_i \\
 &+ \alpha_{12} \text{sexratio}_i + \alpha_{13} \text{unemployment}_i + \alpha_{14} Y_{it} + \alpha_{15} \sum_{k=1}^{t-1} A_{ik} + \alpha_{16} A_{i,t-1} \\
 &+ \alpha_{17} \text{government}_{it})
 \end{aligned}$$

$A_{it}$  is the independent variable, the vaccinated population (per 10,000 people) in state  $i$  at week  $t$ .

$Y_{it}$  is the outcome variable, the growth rate of new cases in state  $i$  at week  $t$ .  $L_{it}$  denotes other confounders in state  $i$  at week  $t$ .

$\text{physicians}_i$  is number of physicians (per capita) in state  $i$ .

$\text{election}_i = 1$  denotes state  $i$  is red state in 2016 president election.

$\text{GDP}_i$  is the GDP (millions of chained 2012 dollars) in state  $i$ .

$\text{race}_i$  is the racial composition (proportion of black people) in state  $i$ .

$\text{old}_i$  is the proportion of old people (aged 65 or above) in state  $i$ .

$\text{pop}_i$  is the population in state  $i$ .

$cumcase0_i$  is the cumulated confirmed cases at baseline in state  $i$ .  
 $basvac_i$  is the vaccinated population at baseline in state  $i$ .  
 $education_i$  is the education level (Proportion of people with advanced degrees) in state  $i$ .  
 $sexratio_i$  is the sex ratio in state  $i$ .  
 $unemployment_i$  is the unemployment rate in state  $i$ .  
 $government_{it}$  is the government response index at week  $t$  in state  $i$ .  
 Step 2: Predict the fitted value of  $A_{it}$ , denote it as  $p_{it}$ .  
 Step 3: Fit the following linear regression:

$$E[Y_{i,t+1} | \bar{A}_{i,t}, \bar{L}_{i,t}, \bar{Y}_{i,t}] = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 physicans_i + \beta_4 election_i + \beta_5 GDP_i + \beta_6 race_i + \beta_7 old_i + \beta_8 population_i + \beta_9 cumcase0_i + \beta_{10} basvac_i + \beta_{11} education_i + \beta_{12} sexratio_i + \beta_{13} unemployment_i + \beta_{14} Y_{is} + \beta_{15} \sum_{k=1}^{t-1} A_k + \beta_{16} A_{i,t-1} + \beta_{17} government_{it} + \beta_{18} p_{it} + \psi A_{it}$$

We can obtain a preliminary estimate of  $\psi$ , denoted  $\hat{\psi}^0$ , through this regression.  
 Step 4: Predict the potential outcome:

Let

$$H_{i,t+1,s+1} = Y_{i,t+1} - \hat{\psi}^0 \sum_{k=s+1}^t A_{ik}$$

where  $s$  ranges from 1 to  $t - 1$ , and  $H_{i,t+1,s+1}$  can be viewed as a mimic of  $Y_{i,t+1}(A_1, \dots, A_s, 0, \dots, 0)$ .

Step 5: Fit the following linear regression:

We can obtain an updated estimate of  $\psi$ , denoted  $\hat{\psi}^1$ , through this regression.

$$E[H_{i,t+1,s+1} | \bar{A}_{i,s}, \bar{L}_{i,s}, \bar{Y}_{i,s}] = \beta_0 + \beta_1 s + \beta_2 s^2 + \beta_3 physicans_i + \beta_4 election_i + \beta_5 GDP_i + \beta_6 race_i + \beta_7 old_i + \beta_8 population_i + \beta_9 cumcase0_i + \beta_{10} basvac_i + \beta_{11} education_i + \beta_{12} sexratio_i + \beta_{13} unemployment_i + \beta_{14} Y_{is} + \beta_{15} \sum_{k=1}^{s-1} A_k + \beta_{16} A_{i,s-1} + \beta_{17} government_{is} + \beta_{18} p_{is} + \psi A_{is}$$

Step 6: Using the updated estimate of  $\psi$  to repeat Step 4 to Step 5 several times [47], we can obtain a final estimate of  $\psi$ .

### Appendix C. Specification of GEE Models and Fixed Effects Models

GEE (adjust baseline covariates):

$$E[Y_{i,t+1} | \bar{A}_{i,t}, \bar{L}_{i,t}, \bar{Y}_{i,t}] = \gamma_0 + \gamma_3 physicans_i + \gamma_4 election_i + \gamma_5 GDP_i + \gamma_6 race_i + \gamma_7 old_i + \gamma_8 population_i + \gamma_9 cumcase0_i + \gamma_{10} basvac_i + \gamma_{11} education_i + \gamma_{12} sexratio_i + \gamma_{13} unemployment_i + \delta \sum_{k=1}^t A_{ik}$$

GEE (adjust baseline covariates and time-varying covariates):

$$E[Y_{i,t+1} | \bar{A}_{i,t}, \bar{L}_{i,t}, \bar{Y}_{i,t}] = \gamma_0 + \gamma_1 t + \gamma_2 t^2 + \gamma_3 physicans_i + \gamma_4 election_i + \gamma_5 GDP_i + \gamma_6 race_i + \gamma_7 old_i + \gamma_8 population_i + \gamma_9 cumcase0_i + \gamma_{10} basvac_i + \gamma_{11} education_i + \gamma_{12} sexratio_i + \gamma_{13} unemployment_i + \gamma_{14} Y_{it} + \gamma_{17} government_{it} + \delta \sum_{k=1}^t A_{ik}$$



Two-way fixed effects model:

$$Y_{it} = \mu_i + \theta_t + \delta \sum_{k=1}^t A_{ik} + \varepsilon_{it}$$

where  $\mu_i$  is the state fixed effect,  $\theta_t$  is the week fixed effect,  $\varepsilon_{it}$  is the error term.

Two-way fixed effects model (adjust time-varying covariates):

$$Y_{it} = \mu_i + \theta_t + \gamma_{12}Y_{it} + \gamma_{13}government_{it} + \delta \sum_{k=1}^t A_{ik} + \varepsilon_{it}$$

In Table 2, we reported the estimate of  $\delta$ .

## References

- Dong, E.; Du, H.; Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **2020**, *20*, 533–534. [[CrossRef](#)]
- Mathieu, E.; Ritchie, H.; Ortiz-Ospina, E.; Roser, M.; Hasell, J.; Appel, C.; Giattino, C.; Rod s-Guirao, L. A global database of COVID-19 vaccinations. *Nat. Hum. Behav.* **2021**, *5*, 956–959. [[CrossRef](#)] [[PubMed](#)]
- COVID-19 Vaccinations in the United States, Jurisdiction. Available online: <https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-Jurisdiction/unsk-b7fc> (accessed on 9 December 2021).
- Polack, F.P.; Thomas, S.J.; Kitchin, N.; Absalon, J.; Gurtman, A.; Lockhart, S.; Perez, J.L.; Marc, G.P.; Moreira, E.D.; Zerbini, C.; et al. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *N. Engl. J. Med.* **2020**, *383*, 2603–2615. [[CrossRef](#)] [[PubMed](#)]
- Baden, L.R.; El Sahly, H.M.; Essink, B.; Kotloff, K.; Frey, S.; Novak, R.; Diemert, D.; Spector, S.A.; Rouphael, N.; Creech, C.B.; et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **2021**, *384*, 403–416. [[CrossRef](#)]
- Sadoff, J.; Gray, G.; Vandebosch, A.; C rdenas, V.; Shukarev, G.; Grinsztejn, B.; Goepfert, P.A.; Truyers, C.; Fennema, H.; Spiessens, B.; et al. Safety and efficacy of single-dose Ad26. COV2. S vaccine against COVID-19. *N. Engl. J. Med.* **2021**, *384*, 2187–2201. [[CrossRef](#)]
- Bjork, J.; Inghammar, M.; Moghaddassi, M.; Rasmussen, M.; Malmqvist, U.; Kahn, F. Effectiveness of the BNT162b2 vaccine in preventing COVID-19 in the working age population—first results from a cohort study in Southern Sweden. *medRxiv* **2021**. [[CrossRef](#)]
- Thompson, M.G.; Burgess, J.L.; Naleway, A.L.; Tyner, H.L.; Yoon, S.K.; Meece, J.; Olsho, L.E.W.; Caban-Martinez, A.J.; Fowlkes, A.; Lutrick, K.; et al. Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers—Eight US locations, December 2020–March 2021. *Morb. Mortal. Wkly. Rep.* **2021**, *70*, 495–500. [[CrossRef](#)]
- Vasileiou, E.; Simpson, C.R.; Shi, T.; Kerr, S.; Agrawal, U.; Akbari, A.; Bedston, S.; Beggs, J.; Bradley, D.; Chuter, A.; et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: A national prospective cohort study. *Lancet* **2021**, *397*, 1646–1657. [[CrossRef](#)]
- Pawlowski, C.; Lenehan, P.; Puranik, A.; Agarwal, V.; Venkatakrishnan, A.J.; Niesen, M.J.M.; O’Horo, J.C.; Virk, A.; Swift, M.D.; Badley, A.D.; et al. FDA-authorized COVID-19 vaccines are effective per real-world evidence synthesized across a multi-state health system. *Med* **2021**, *2*, 979–992. [[CrossRef](#)]
- Bernal, J.L.; Andrews, N.; Gower, C.; Robertson, C.; Stowe, J.; Tessier, E.; Simmons, R.; Cottrell, S.; Roberts, R.; O’Doherty, M.; et al. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on COVID-19 related symptoms, hospital admissions, and mortality in older adults in England: Test negative case-control study. *BMJ* **2021**, *373*. [[CrossRef](#)]
- Kim, J.H.; Marks, F.; Clemens, J.D. Looking beyond COVID-19 vaccine phase 3 trials. *Nat. Med.* **2021**, *27*, 205–211. [[CrossRef](#)] [[PubMed](#)]
- Gallagher, M.E.; Sieben, A.J.; Nelson, K.N.; Kraay, A.N.M.; Orenstein, W.A.; Lopman, B.; Handel, A.; Koelle, K. Indirect benefits are a crucial consideration when evaluating SARS-CoV-2 vaccine candidates. *Nat. Med.* **2021**, *27*, 4–5. [[CrossRef](#)] [[PubMed](#)]
- Hanquet, G.; Valenciano, M.; Simondon, F.; Moren, A. Vaccine effects and impact of vaccination programmes in post-licensure studies. *Vaccine* **2013**, *31*, 5634–5642. [[CrossRef](#)] [[PubMed](#)]
- Chernozhukov, V.; Kasahara, H.; Schrimpf, P. Causal impact of masks, policies, behavior on early COVID-19 pandemic in the U.S. *J. Econom.* **2021**, *220*, 23–62. [[CrossRef](#)]
- Courtemanche, C.; Garuccio, J.; Le, A.; Pinkston, J.; Yelowitz, A. Strong Social Distancing Measures In The United States Reduced The COVID-19 Growth Rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed COVID-19 cases across the United States. *Health Aff.* **2020**, *39*, 1237–1246. [[CrossRef](#)] [[PubMed](#)]
- Hsiang, S.; Allen, D.; Annan-Phan, S.; Bell, K.; Bolliger, I.; Chong, T.; Druckenmiller, H.; Huang, L.Y.; Hultgren, A.; Krasovich, E.; et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **2020**, *584*, 262–267. [[CrossRef](#)]

18. Pozo-Martin, F.; Weishaar, H.; Cristea, F.; Hanefeld, J.; Bahr, T.; Schaade, L.; El Bcheraoui, C. The impact of non-pharmaceutical interventions on COVID-19 epidemic growth in the 37 OECD member states. *Eur. J. Epidemiol.* **2021**, *36*, 629–640. [[CrossRef](#)]
19. Kaufman, B.G.; Whitaker, R.; Mahendraratnam, N.; Hurewitz, S.; Yi, J.; Smith, V.A.; McClellan, M. State variation in effects of state social distancing policies on COVID-19 cases. *BMC Public Health* **2021**, *21*, 1239. [[CrossRef](#)]
20. Moghadas, S.M.; Vilches, T.N.; Zhang, K.; Wells, C.R.; Shoukat, A.; Singer, B.H.; Meyers, L.A.; Neuzil, K.M.; Langley, J.M.; Fitzpatrick, M.C.; et al. The impact of vaccination on coronavirus disease 2019 (COVID-19) outbreaks in the United States. *Clin. Infect. Dis.* **2021**, *73*, 2257–2264. [[CrossRef](#)]
21. Makhoul, M.; Ayoub, H.H.; Chemaitelly, H.; Seedat, S.; Mumtaz, G.R.; Al-Omari, S.; Abu-Raddad, L.J. Epidemiological impact of SARS-CoV-2 vaccination: Mathematical modeling analyses. *Vaccines* **2020**, *8*, 668. [[CrossRef](#)]
22. Makhoul, M.; Chemaitelly, H.; Ayoub, H.H.; Seedat, S.; Abu-Raddad, L.J. Epidemiological differences in the impact of COVID-19 vaccination in the United States and China. *Vaccines* **2021**, *9*, 223. [[CrossRef](#)] [[PubMed](#)]
23. Foy, B.H.; Wahl, B.; Mehta, K.; Shet, A.; Menon, G.I.; Britto, C. Comparing COVID-19 vaccine allocation strategies in India: A mathematical modelling study. *Int. J. Infect. Dis.* **2021**, *103*, 431–438. [[CrossRef](#)] [[PubMed](#)]
24. Romero-Brufau, S.; Chopra, A.; Ryu, A.J.; Gel, E.; Raskar, R.; Kremers, W.; Anderson, K.S.; Subramanian, J.; Krishnamurthy, B.; Singh, A.; et al. Public health impact of delaying second dose of BNT162b2 or mRNA-1273 COVID-19 vaccine: Simulation agent based modeling study. *BMJ* **2021**, *373*, n1087. [[CrossRef](#)] [[PubMed](#)]
25. Wheelock, A.; Thomson, A.; Sevdalis, N. Social and psychological factors underlying adult vaccination behavior: Lessons from seasonal influenza vaccination in the US and the UK. *Expert Rev. Vaccines* **2013**, *12*, 893–901. [[CrossRef](#)] [[PubMed](#)]
26. Soares, P.; Rocha, J.V.; Moniz, M.; Gama, A.; Laires, P.A.; Pedro, A.R.; Dias, S.; Leite, A.; Nunes, C. Factors associated with COVID-19 vaccine hesitancy. *Vaccines* **2021**, *9*, 300. [[CrossRef](#)] [[PubMed](#)]
27. Bish, A.; Yardley, L.; Nicoll, A.; Michie, S. Factors associated with uptake of vaccination against pandemic influenza: A systematic review. *Vaccine* **2011**, *29*, 6472–6484. [[CrossRef](#)]
28. Kelly, C.; Arnold, R.; Galloway, Y.; O'Hallahan, J. A prospective study of the effectiveness of the New Zealand meningococcal B vaccine. *Am. J. Epidemiol.* **2007**, *166*, 817–823. [[CrossRef](#)]
29. Fu, W.; Ho, P.-C.; Liu, C.-L.; Tzeng, K.-T.; Nayeem, N.; Moore, J.S.; Wang, L.-S.; Chou, S.-Y. Reconcile the debate over protective effects of BCG vaccine against COVID-19. *Sci. Rep.* **2021**, *11*, 8356. [[CrossRef](#)]
30. Imai, K.; Kim, I.S. When should we use unit fixed effects regression models for causal inference with longitudinal data? *Am. J. Political Sci.* **2019**, *63*, 467–490. [[CrossRef](#)]
31. Blackwell, M.; Glynn, A.N. How to make causal inferences with time-series cross-sectional data under selection on observables. *Am. Political Sci. Rev.* **2018**, *112*, 1067–1082. [[CrossRef](#)]
32. Rosenbaum, P.R. The consequences of adjustment for a concomitant variable that has been affected by the treatment. *J. R. Stat. Soc. Ser. A* **1984**, *147*, 656–666. [[CrossRef](#)]
33. Hernán, M.A.; Robins, J.M. *Causal Inference: What If*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2020.
34. Bruhn, C.A.W.; Hetterich, S.; Schuck-Paim, C.; Kürüm, E.; Taylor, R.J.; Lustig, R.; Shapiro, E.D.; Warren, J.L.; Simonsen, L.; Weinberger, D.M. Estimating the population-level impact of vaccines using synthetic controls. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 1524–1529. [[CrossRef](#)]
35. Lau, W.C.Y.; Murray, M.; El-Turki, A.; Saxena, S.; Ladhani, S.; Long, P.; Sharland, M.; Wong, I.C.K.; Hsia, Y. Impact of pneumococcal conjugate vaccines on childhood otitis media in the United Kingdom. *Vaccine* **2015**, *33*, 5072–5079. [[CrossRef](#)] [[PubMed](#)]
36. Robins, J.M. Causal inference from complex longitudinal data. In *Latent Variable Modeling and Applications to Causality*; Springer: New York, NY, USA, 1997; pp. 69–117.
37. Robins, J.M. Correcting for non-compliance in randomized trials using structural nested mean models. *Commun. Stat.-Theory Methods* **1994**, *23*, 2379–2412. [[CrossRef](#)]
38. Vansteelandt, S.; Joffe, M. Structural nested models and G-estimation: The partially realized promise. *Stat. Sci.* **2014**, *29*, 707–731. [[CrossRef](#)]
39. Rubin, D.B. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **1974**, *66*, 688–701. [[CrossRef](#)]
40. Rubin, D.B. Causal inference using potential outcomes: Design, modeling, decisions. *J. Am. Stat. Assoc.* **2005**, *100*, 322–331. [[CrossRef](#)]
41. Hale, T.; Angrist, N.; Goldszmidt, R.; Kira, B.; Petherick, A.; Phillips, T.; Webster, S.; Cameron-Blake, E.; Hallas, L.; Majumdar, S.; et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat. Hum. Behav.* **2021**, *5*, 529–538. [[CrossRef](#)]
42. Robins, J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Model.* **1986**, *7*, 1393–1512. [[CrossRef](#)]
43. Haber, M. Estimation of the population effectiveness of vaccination. *Stat. Med.* **1997**, *16*, 601–610. [[CrossRef](#)]
44. Tanaka, S.; Matsuyama, Y.; Shiraki, M.; Ohashi, Y. Estimating the effects of time-varying treatments: Incidence of fractures among postmenopausal Japanese women. *Epidemiology* **2007**, *18*, 529–536. [[CrossRef](#)] [[PubMed](#)]
45. Aryaie, M.; Sharifi, H.; Saber, A.; Nazempour, M.; Mansournia, M.A. Longitudinal Causal Effects of Normalized Protein Catabolic Rate on All-Cause Mortality in Patients With End-Stage Renal Disease: Adjusting for Time-Varying Confounders Using the G-Estimation Method. *Am. J. Epidemiol.* **2021**, *190*, 1133–1141. [[CrossRef](#)] [[PubMed](#)]

46. Robins, J.M.; Hernan, M.A.; Brumback, B. Marginal structural models and causal inference in epidemiology. *Epidemiology* **2000**, *11*, 550–560. [[CrossRef](#)] [[PubMed](#)]
47. Vansteelandt, S.; Sjolander, A. Revisiting g-estimation of the effect of a time-varying exposure subject to time-varying confounding. *Epidemiol. Methods* **2016**, *5*, 37–56. [[CrossRef](#)]
48. Zeger, S.L.; Liang, K.-Y.; Albert, P.S. Models for longitudinal data: A generalized estimating equation approach. *Biometrics* **1988**, *44*, 1049–1060. [[CrossRef](#)] [[PubMed](#)]
49. Liang, K.-Y.; Zeger, S.L. Longitudinal data analysis using generalized linear models. *Biometrika* **1986**, *73*, 13–22. [[CrossRef](#)]
50. Cole, S.R.; Platt, R.W.; Schisterman, E.F.; Chu, H.; Westreich, D.; Richardson, D.; Poole, C. Illustrating bias due to conditioning on a collider. *Int. J. Epidemiol.* **2010**, *39*, 417–420. [[CrossRef](#)]
51. Ainslie, K.E.C.; Haber, M.; Orenstein, W.A. Challenges in estimating influenza vaccine effectiveness. *Expert Rev. Vaccines* **2019**, *18*, 615–628. [[CrossRef](#)]
52. Imbens, G.W.; Rubin, D.B. *Causal Inference in Statistics, Social, and Biomedical Sciences*; Cambridge University Press: Cambridge, UK, 2015.
53. Sah, P.; Vilches, T.N.; Moghadas, S.M.; Fitzpatrick, M.C.; Singer, B.H.; Hotez, P.J.; Galvani, A.P. Accelerated vaccine rollout is imperative to mitigate highly transmissible COVID-19 variants. *EClinicalMedicine* **2021**, *35*, 100865. [[CrossRef](#)]
54. U.S. Reaches 70% COVID Vaccine Milestone for Adults about a Month behind Biden’s Goal. Available online: <https://www.cnbc.com/2021/08/02/covid-vaccine-us-reaches-bidens-70percent-goal-for-adults-a-month-behind.html> (accessed on 5 April 2022).
55. South Korea Coronavirus: PM Aims for ‘Herd Immunity by Autumn’. Available online: <https://www.bbc.com/news/world-asia-56156234> (accessed on 5 April 2022).
56. Africa Faces Steepest COVID-19 Surge yet. Available online: <https://afrocoms.newsweaver.com/1rz5qc4oi8/imj9z0y2kyw?lang=en> (accessed on 9 December 2021).