The Longer-Term Effect of COVID-19 Vaccination: The Relationship between COVID Vaccination Rate and Vaccination Effectiveness in the US in 2022

ECO481 Special Topics in Economics: Health and Economic Inequalities

University of Toronto

December 8, 2023

Word Count: 1880

Introduction

The COVID-19 virus has a strong negative effect on people's health outcomes. According to the Centers for Disease Control and Prevention (CDC), people experience negative health outcomes such as infection, symptomatic illness, hospitalization, and even death. Scientists and healthcare providers have invented multiple types of vaccinations to fight back the negative consequences of COVID-19 infection and to alleviate COVID-related symptoms. The most commonly used type of vaccine worldwide is the messenger RNA (mRNA) vaccine. In the U.S., mRNA vaccines such as Pfizer-BioNTech and Moderna are mostly implemented to create an antibody in the physical immune system and to slow down the rapid transmission of COVID among individuals. According to Ssentongo et al. (2022), worldwide COVID-19 vaccines in the first month after vaccination are 94% likely to effectively reduce symptoms, and by the fourth month are 64% likely to reduce symptoms. As for those who are fully vaccinated, Pilishvili et al. (2021) show that vaccine effectiveness is 88.8% for Pfizer and 96.3% for Moderna receivers, and the result is consistent among age, race, and ethnic groups with different presence of underlying condition and level of contact.

Although the COVID-induced symptoms significantly decline after receiving vaccination, we have limited knowledge of the effectiveness of COVID-19 vaccines in the longer term. As shown in Wright et al.'s study (2022), COVID vaccine effectiveness has a notable decline after 250 days of getting the vaccine. People still need to receive a booster shot after a couple of months or one year after they receive their first vaccine. In addition, since receiving a COVID vaccine is not mandatory, the different vaccination rates among jurisdictions may also influence the overall vaccine effectiveness in terms of impeding the transmission. In this paper, I analyze a panel data created by merging two datasets to understand the longer-term relationship between the COVID vaccination rate and vaccination effectiveness (i.e., new hospitalization and new confirmed cases) in the US in 2022. I hypothesize that the confirmed case rate and hospitalization rate will decrease by increasing the vaccination rate across regions, suggesting fewer people are affected by COVID in terms of health outcomes. This study controls for demographic characteristics such as age, sex, and geographic location. However, socioeconomic status and race/ ethnicity data of the population were not collected and thus may lead to omitted variable bias in the analysis.

Data and Methodology

In the paper, I used two datasets provided by the CDC to conduct empirical analysis on COVID-19 vaccination rate and vaccination effectiveness: I merged the COVID-19 Vaccination Age and Sex Trends in the United States National and Jurisdictional data, as well as the United States COVID-19 Community Levels by County data to create a panel data with measures of vaccination rate, new confirmed case, new hospitalization, and age and sex variables at jurisdiction level. The vaccination trends data was administered from January 2021 to December 2022 to reveal the trends of vaccination in the US. The community-level data was collected from February 2022 to May 2023 to collect information on the new COVID-19 hospital admissions and the new confirmed cases per 100,000 population every week. Both datasets provide quantitative, panel data targeting the entire US population and were merged by jurisdiction names. Variables such as sex and age groups were transformed into dummies in the new dataset to perform control analysis.

Linear regressions were implemented to investigate the relationship between COVID-19 vaccination rate and its effectiveness in the year 2022. There are 2,773,714 observations in the regression, from which 30.5% are middle-aged adults (25-49 years old), 30.5% are older elderly (age 65 and above), and the rest are from newborns to younger elders (See Table 1 in Appendix). The independent variables are the vaccination rates of the first dose, series completion, and booster shot within each jurisdiction. The dependent variables are the percentage of new hospitalizations and new confirmed cases of each state's population (integrated from county level), as a proxy of measuring vaccination effectiveness. The controlled variables are age, sex, time, and jurisdiction, with each characteristic transformed into a group of dummies. Other demographic characteristics such as race/ ethnicity, socio-economic factors such as level of education, and individual differences such as medical consideration and vaccine hesitancy were not measured in the dataset and were excluded from the analysis.

Results

To start with our analysis, we will look at the overall trend in the percentage of people who confirmed COVID and the percentage who received hospitalizations in 2022. From Figures 1 and 2, we observe the new confirmed case and new hospitalization rates were at a peak at the beginning of 2022, a few months after the outbreak of the Omicorn variant. Each data point in Figures 1 and 2 represents the ratio of targeted group in a state population at a specific time. Since COVID vaccination was taking place in 2022, we observe the average percentage of new cases within state population has dropped from 0.25% to 0.10%. The decreasing numbers of confirmed cases over time suggest the vaccine is effective in terms of reducing people’s likelihood of getting infection and thus impeding the transmission. On the other hand, the trend in the COVID-induced hospitalization rate is more fluctuating with a second peak around August 2022, probably because COVID had triggered longer-term deterioration in patients’ health outcomes and caused hospitalization a couple of months after the immediate infection. However, the probability of getting hospitalization is small across time: On average, only 0.012% and 0.008% of state population are newly admitted to hospital due to COVID at the beginning and the end of 2022, respectively.

In Figures 3-8, I explore the relationship between different vaccination measurements (first dose, vaccine series completion, and booster rates) and new case and hospitalization rates by state. A negative correlation is found between each independent and dependent variable, implying a general pattern that a higher vaccination rate is associated with less COVID infection and hospital admission, leading to greater effectiveness. Among all 6 regressions, we found the most negative association between first-dose vaccination rate and hospitalization. This finding contradicts the time trend results in Figure 2 where the hospitalization rate is more fluctuating and less relevant to vaccine rates, implying it is crucial to separate the time effect from the initial observations.

To examine the association between vaccination rate and its effectiveness, multiple linear regressions were run to measure the time effect as well as age and sex effects across US jurisdictions. First, I measured the estimated correlation without time, place, age, and sex fixed effects. Since there are two dependent variables in my analysis, each dependent variable is combined with all three independent variables to perform a regression, as shown in Models (1) and (4) in Table 2. After controlling for other independent variables, a one percentage point increase in the proportion of people receiving their first dose is associated with a 0.0009 percentage point decrease in the ratio of newly confirmed cases and a 0.0001 percentage point decrease in the new hospitalization rate among population. Similarly, having a one percentage point increase in the proportion of people receiving a booster shot correlates with a 0.0005 percentage point decrease in new case rates across regions. However, the vaccine series completion rate is positively correlated to the confirmed case rate, as shown in Model (1). A possible reason for this unexpected result might be people are getting COVID a few months after completing the vaccine series, suggesting the ineffectiveness of vaccines for a longer time.

To further examine the patterns revealed in the regression model, I introduced four sets of control variables (i.e. age, sex, time, and place) into Models (1) and (4). Models (2) and (5) show the results of introducing the age and sex variables into the model. The results show that older elders benefit the most from vaccination: a one percentage point more vaccination is associated with a 0.02 percentage point decrease in the chance of newly getting COVID and a 0.0046 percentage point decrease in the chance of newly getting hospitalization. Additionally, Models (3) and (6) show the results of including all four control variables in the model. Although the R-squared values increased after introducing more controls in the regressions, we observe that most of the estimated coefficients turn out to be smaller in absolute values, possibly because the regression variables do not accurately capture the underlying patterns in the data. Considering all estimated coefficients are extremely small but statistically significant at the 1% significance level, we found an extremely weak negative relationship between vaccination rates and new case and hospitalization rates, providing limited information to test for my hypothesis.

Discussion and Conclusion

In this paper, I explore the relationship between vaccine rates and vaccination effectiveness in the US in the year 2022 and find a small negative relationship between the two variables. However, the study suffers from some limitations. First, the dataset does not include some socio-economic factors such as race, education, and income levels, possibly leading to an omitted variable bias. For example, people with lower education may have less health knowledge, making them less likely to get vaccinated. Individual differences in physical health outcomes and vaccine hesitancy are also not controlled in the analysis. For instance, some individuals may have underlying special medical considerations, including having chronic diseases (such as cancer, diabetes, etc.) and/ or a history of prior infection. Future researchers can introduce more socioeconomic variables and individual controls into the data to generate more insightful results, inspiring scientists and healthcare providers to develop more specific treatments for patients with special needs.

Second, using newly confirmed cases and hospitalization data as a proxy for COVID vaccine effectiveness may not accurately measure the effects of vaccination on people’s health outcomes. Only people with severe COVID-induced symptoms are likely to go to a hospital, either for a doctor visit or to visit the emergency department, while others experiencing light symptoms may choose to stay at home until they recover. Nevertheless, individuals with hospital visits may suffer from chronic diseases or weak immune systems, predisposing them to a greater risk of getting COVID infection, confounding the variables in the study and skewing the results. Based on these concerns, future scholars are encouraged to discover more accurate measures of vaccine effectiveness.

Finally, the correlational design of this study does not provide causal inference on the effect of increasing COVID vaccination rates on boosting vaccination effectiveness, and the underlying causal pattern remains unknown. Future researchers are encouraged to conduct studies adopting a difference-in-deference design or a regression discontinuity design to further investigate the effect of COVID vaccines on people's health outcomes, as well as the influences of sociodemographic variables on people's access to vaccines.

In conclusion, this paper examines the relationship between COVID-19 vaccination rates and vaccine effectiveness in 2022 in the US. Despite limitations, the study provides implications for future researchers and scientists on how they can refine methodologies and incorporate additional variables to enhance the value of findings. As for policymakers, given the effect of economic inequality on access to vaccinations and healthcare services, future policies can target healthcare strategies and programs, ultimately ensuring that no one is excluded from accessing vital healthcare. In general, future scholars can discuss more about how economic disparity hinders successful vaccine coverage across geographic regions and propose targeted interventions to address these disparities.

References

Centers for Disease Control and Prevention. (2023, May 12). COVID-19 Vaccination Age and Sex Trends in the United States, National and Jurisdictional. [Dataset]. https://data.cdc.gov/Vaccinations/COVID-19-Vaccination-Age-and-Sex-Trends-in-the-Uni/5i5k-6cmh

Centers for Disease Control and Prevention. (2023, May 12). United States COVID-19 Community Levels by County. [Dataset]. https://data.cdc.gov/Public-Health-Surveillance/United-States-COVID-19-Community-Levels-by-County/3nnm-4jni

Pilishvili, T., Gierke, R., Fleming-Dutra, K. E., Farrar, J. L., Mohr, N. M., Talan, D. A., Krishnadasan, A., Harland, K. K., Smithline, H. A., Hou, P. C., Lee, L. C., Lim, S. C., Moran, G. J., Krebs, E., Steele, M. T., Beiser, D. G., Faine, B., Haran, J. P., Nandi, U., … Schrag, S. J. (2021). Effectiveness of mRNA Covid-19 Vaccine among U.S. Health Care Personnel. *The New England Journal of Medicine*, *385*(25), e90–e90. https://doi.org/10.1056/NEJMoa2106599

Ssentongo, P., Ssentongo, A. E., Voleti, N., Groff, D., Sun, A., Ba, D. M., Nunez, J., Parent, L. J., Chinchilli, V. M., & Paules, C. I. (2022). SARS-CoV-2 vaccine effectiveness against infection, symptomatic and severe COVID-19: a systematic review and meta-analysis. *BMC Infectious Diseases*, *22*(1), 439–439. https://doi.org/10.1186/s12879-022-07418-y

Wright, B. J., Tideman, S., Diaz, G. A., French, T., Parsons, G. T., & Robicsek, A. (2022). Comparative vaccine effectiveness against severe COVID-19 over time in US hospital administrative data: a case-control study. *The Lancet Respiratory Medicine*, *10*(6), 557–565. https://doi.org/10.1016/S2213-2600(22)00042-X

Appendix

Table 1. Summary Statistics.

Table 1 - Summary Statistics.

	25%	50%	75%	No. Observations	Max	Mean	Min	Standard Deviation
First Dose	66.500000	82.200000	95.000000	2777816.000000	95.000000	77.004251	2.300000	19.912333
Series Completion	56.400000	71.300000	86.000000	2777816.000000	95.000000	68.604800	1.100000	20.495261
Booster Dose	28.800000	43.500000	65.200000	2777816.000000	91.200000	45.402475	0.000000	21.541246
Covid Case	0.066166	0.110478	0.172524	2777816.000000	2.351971	0.128970	0.000000	0.089659
Hospitalization	0.004759	0.007661	0.010473	2777816.000000	0.027247	0.007891	0.000745	0.003992
Male	0.000000	0.000000	1.000000	2777816.000000	1.000000	0.499942	0.000000	0.500000
Baby	0.000000	0.000000	0.000000	2777816.000000	1.000000	0.011020	0.000000	0.104397
Child	0.000000	0.000000	0.000000	2777816.000000	1.000000	0.073077	0.000000	0.260263
Teenager	0.000000	0.000000	0.000000	2777816.000000	1.000000	0.101767	0.000000	0.302342
Young Adult	0.000000	0.000000	0.000000	2777816.000000	1.000000	0.101767	0.000000	0.302342
Mid-age Adult	0.000000	0.000000	1.000000	2777816.000000	1.000000	0.305301	0.000000	0.460535
Young Elder	0.000000	0.000000	0.000000	2777816.000000	1.000000	0.101767	0.000000	0.302342
Old Elder	0.000000	0.000000	1.000000	2777816.000000	1.000000	0.305301	0.000000	0.460535
Eastern	0.000000	0.000000	1.000000	2777816.000000	1.000000	0.257002	0.000000	0.436981
Western	0.000000	0.000000	0.000000	2777816.000000	1.000000	0.142361	0.000000	0.349421
Central	0.000000	1.000000	1.000000	2777816.000000	1.000000	0.600637	0.000000	0.489768
Time	2022-05-19 00:00:00	2022-08-04 00:00:00	2022-10-20 00:00:00	2777816	2022-12-29 00:00:00	2022-08-01 21:02:21.210217216	2022-02-24 00:00:00	nan

Figure 1. COVID-19 Confirmed Case Rate Over Time.

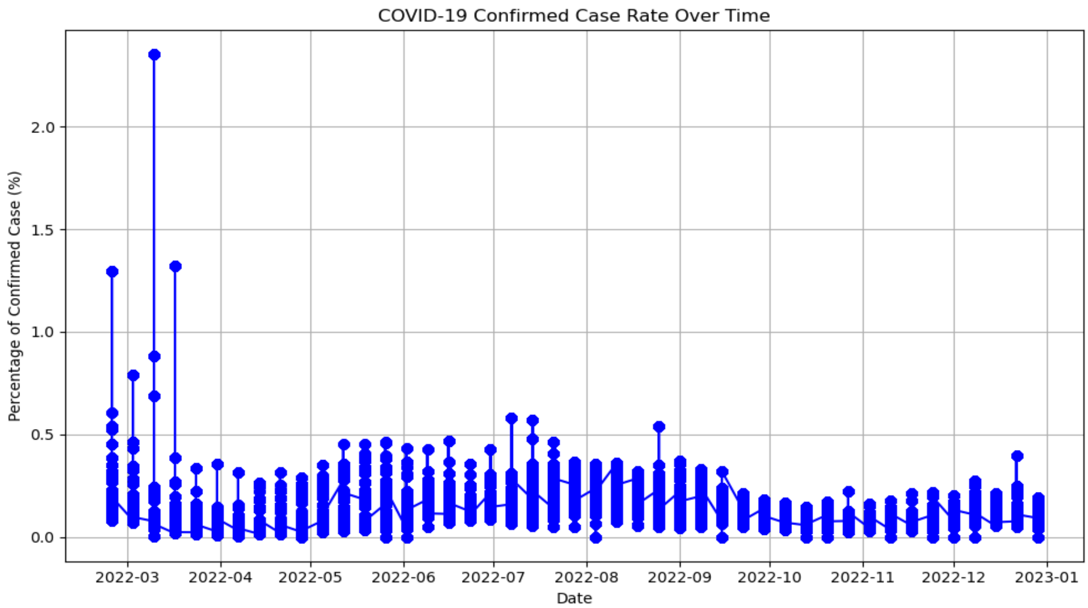


Figure 2. COVID-induced Hospitalization Rate Over Time.

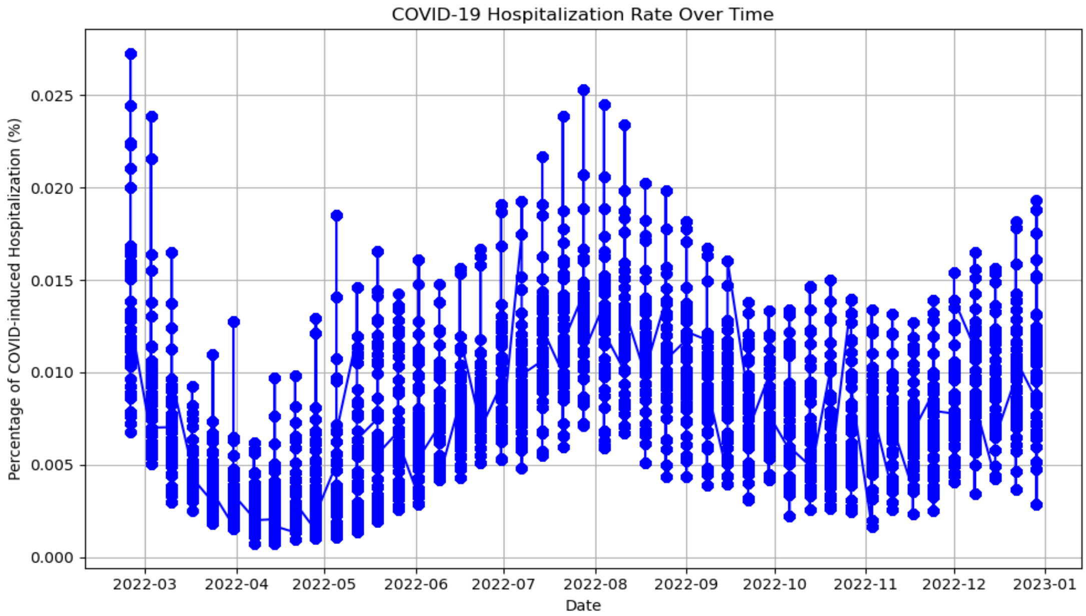
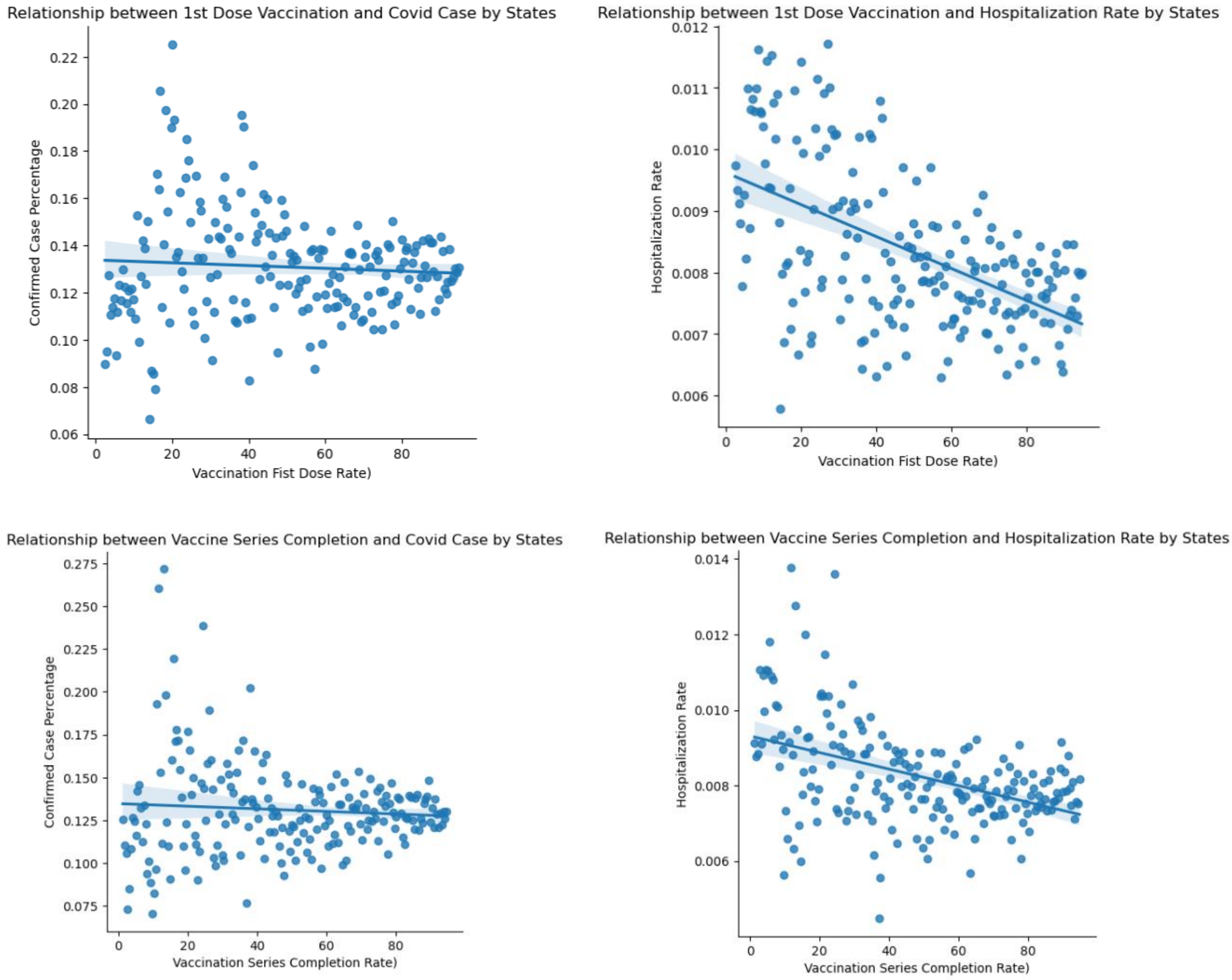
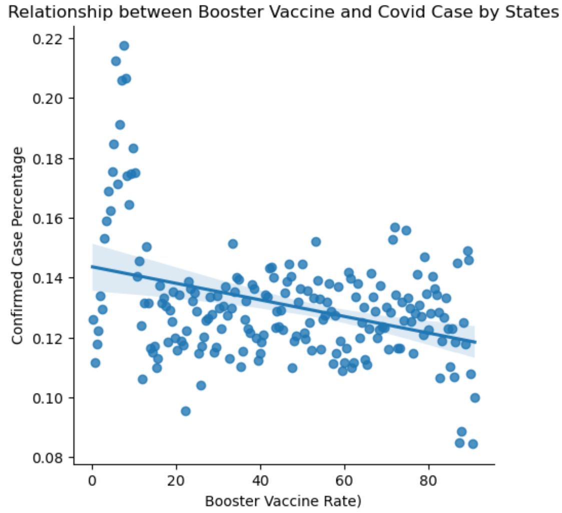


Figure 3-8. Initial Observation Results.





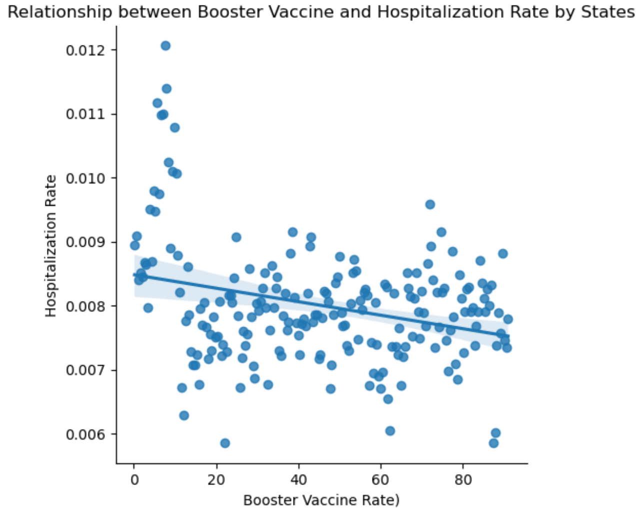


Table 2. Regression Outputs.

Table 2 - Regression Outputs.

	Case	Case w Demo	Case w Time/Plce	Hospital	Hsptl w Demo	Hsptl w Time/Plce
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.1310***	0.1173***	0.1131***	0.0092***	0.0101***	0.0096***
	(0.0002)	(0.0005)	(0.0004)	(0.0000)	(0.0000)	(0.0000)
1st Dose Received	-0.0009***	-0.0005***	-0.0003***	-0.0001***	-0.0000***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Series Complete	0.0013***	0.0012***	0.0006***	0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Booster Received	-0.0005***	-0.0004***	0.0001***	0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Sex (Male)		0.0004***	0.0011***		0.0001***	0.0000***
		(0.0001)	(0.0001)		(0.0000)	(0.0000)
Child		0.0151***	-0.0126***		-0.0015***	-0.0005***
		(0.0006)	(0.0005)		(0.0000)	(0.0000)
Teenager		-0.0117***	-0.0238***		-0.0030***	-0.0003***
		(0.0006)	(0.0005)		(0.0000)	(0.0000)
Young Adult		-0.0109***	-0.0247***		-0.0030***	-0.0002***
		(0.0007)	(0.0006)		(0.0000)	(0.0000)
Mid-age Adult		-0.0130***	-0.0285***		-0.0034***	-0.0005***
		(0.0007)	(0.0006)		(0.0000)	(0.0000)
Young Elder		-0.0164***	-0.0344***		-0.0039***	-0.0008***
		(0.0007)	(0.0006)		(0.0000)	(0.0000)
Old Elder		-0.0200***	-0.0411***		-0.0046***	-0.0014***
		(0.0008)	(0.0007)		(0.0000)	(0.0000)
Place	No	No	Yes	No	No	Yes
Time	No	No	Yes	No	No	Yes
Observations	2777816	2777816	2777816	2777816	2777816	2777816
R^2	0.0020	0.0049	0.3743	0.0070	0.0147	0.5522
Adjusted R^2	0.0020	0.0049	0.3743	0.0070	0.0147	0.5522
Residual Std. Error	0.0896	0.0894	0.0709	0.0040	0.0040	0.0027
F Statistic	1891.2892***	1380.8781***	75533.5075***	6533.3076***	4158.0904***	155726.7558***
Note:						*p<0.1; **p<0.05; ***p<0.01