The Virus, Vaccination, and Voting

Jeffrey Frankel and Randy Kotti, Harvard Kennedy School

2nd draft, August 12, 2021[1]

Abstract

Across US counties, vaccination rates have a statistically significant downward effect on the Covid-19 death rate, as of August 7, 2021. Controlling for poverty rates, age, and temperature lowers the magnitude of the estimate a little. Using the Biden-Trump vote in the 2020 election as an instrument for vaccination rates raises the magnitude of the estimate. Presumably it corrects for a positive effect of observed local covid deaths on the decision to get vaccinated. Overall, the beneficial effect holds up.

If we can accomplish the scientific miracle of developing vaccines capable of ending the Covid-19 pandemic, why can't we convince enough people to get vaccinated? In lower-income countries, vaccination is often limited by the availability of the vaccines.[2] But this is not the case in countries as fortunate as the United States, where the problem is primarily vaccine hesitancy, or even outright vaccine hostility.

1. Introduction: Two Americas of perceptions

To many, it is crystal clear that the advantages of getting vaccinated far outweigh the disadvantages – not just for society as a whole, but also for the individual. What explains widespread vaccine hesitancy? In the words of Dr. Anthony Fauci, there are two Americas. Their perceptions regarding vaccination are separated by a wall -- epistemologically, socially, and to an extent geographically.

The Food and Drug Administration gave emergency use authorization to three vaccines, after appropriate trials. (Moderna and Pfizer were authorized in December 2020 and Johnson & Johnson in February 2021.) On July 4th, 2021, Dr. Fauci suggested that 99.2% of covid deaths in June were among those who had still not gotten vaccinated.[3]

---

[1] A July 17, 2021, draft of this paper found similar results with data on July covid-19 rates. The findings were briefly summarized in a blog post, July 25.

[2] Çakmaklı, et al. (2020, 2021).

[3] *NBC News* (2021)

Some Americans, however, are not persuaded by appeals to the expertise of remote authorities or by the logic of scientists' methods. The skeptics need evidence that is more tangible, closer to home.

2. The negative correlation between vaccination and virus victims

Recent data across US counties show a strong negative correlation between vaccination rates, on the one hand, and rates of infection, hospitalization or death from coivd-19, on the other hand. In the week ending June 22, 2021, counties where 30 percent or fewer of the residents had been vaccinated suffered 5.6 covid deaths per 100,000, while counties in which more than 60 percent of residents had been vaccinated experienced less than half the deaths, only 2.1 per 100,000.[4] This seems like evidence that is perhaps tangible and closer to home than FDA trials.

It should be noted that the criterion for cause of death in all these studies is whether the doctor or coroner enters covid-19 on the death certificate. This probably understates the true number of deaths caused by covid-19, as international studies of excess mortality rates strongly suggest.

Table 2, below, is based on updated county-level data. A Data Appendix to this paper explains the definitions and sources of the numbers used.[5]

As shown in Column (1) of Table 2, a 1 percentage-point increase in a county's percentage of residents (12 years old and older) who were fully vaccinated as of July 10th was associated with a covid-19 death rate over the subsequent 28 days (to August 7th) that was lower by a highly significant .092 per 100,000 inhabitants. That represents 2.6% of the total monthly deaths related to covid. Extrapolating, the apparent statistical effect of going from the current nationwide vaccination rate to 100% vaccination would be to bring covid-related deaths near to 0.

But, as we are frequently reminded, correlation need not prove causality.

Perhaps the apparent beneficial effect of vaccination is really the illusory result of an omitted variable, some third factor such as the county's poverty rate. That is, perhaps low-income people are more likely to live in crowded conditions and for that reason to become covid victims, while at the same time they are less likely to get vaccinated. In

---

[4] Leonhardt (2021).

[5] Key data come from a *New York Times* site. Appendix 1 to this paper checks the robustness of the results with respect to some decisions regarding how the county-wide covid-related death rate is counted.

that case, the correlation between vaccination and covid mortality could be spurious. One can control in the regression equation for third factors such as the poverty rate or age, to isolate the effect of vaccination rates.

Or perhaps the simple observed correlation between vaccination and the death rate *understates* the true effect of the former on the latter, because of the endogeneity of vaccination. In a place where the coronavirus is a greater danger (say, because it is close to a major airport or other transport hub, or simply because of chance spreading), people are more likely to see their neighbors falling victim to the virus and more likely to react by deciding to get vaccinated themselves.[6] This reverse causality could work in the direction of an apparent positive correlation between vaccination and death rates.

This might help explain why earlier studies, conducted as recently as the beginning of June 2021, did not find a clear negative correlation. Only recently, since June has the beneficial effect of vaccination been powerful enough to dominate the statistical correlation.[7] The reason for the evolution in this statistic is probably the rising challenge of the Delta variant to the health of the unvaccinated.

3. Voting as an instrumental variable for vaccination

The way to disentangle the causality is to examine the effects of variation in vaccination rates that is due not to variation in the spread of the disease, but rather to some unrelated factor, an exogenous instrument. Party affiliation or voting patterns are obvious choices and are supported by empirical evidence. Even before the vaccines were available, for example, "red-state" governors in 2020 were found less likely to fight the coronavirus by steps such as promulgation of mask mandates.[8] The idea of this paper is simply to use partisan status as an instrumental variable for the vaccination rate.

As has been extensively reported, Republicans and those who voted for Donald Trump in the 2020 presidential election are less likely to have gotten vaccinated. A PRRI-IFYC survey conducted in March 2021 found that Republicans are less likely than Democrats to accept vaccination, by 45% versus 73%. Donald Trump won 24 of the 25 states that showed the lowest vaccination rate as of August 2021.[9] An August 2021 poll

---

[6] *The Economist*, "Two roads diverged," July 24th, 2021, used international data to estimate a strong effect of death rates on the decision to get vaccinated.

[7] *Washington Post* (2021).

[8] Adeel, et al. (2000) and Neelon, et al. (2021).

[9] Hicks (2021).

found that, of those who said they had been vaccinated, only 32 percent incline to the Republican Party, versus 59 percent who incline Democratic."[10]

This paper analyzes data on US counties. A _New York Times_ article on April 17, 2021, found that the vaccination rate fell below 25% in counties where Trump had won by a margin of 50 percentage points or more.[11] The vaccination gap continued to widen in July. Figure 1 illustrates the county-level relationship between the vote in the 2020 election and the vaccination rate as of July 2021.

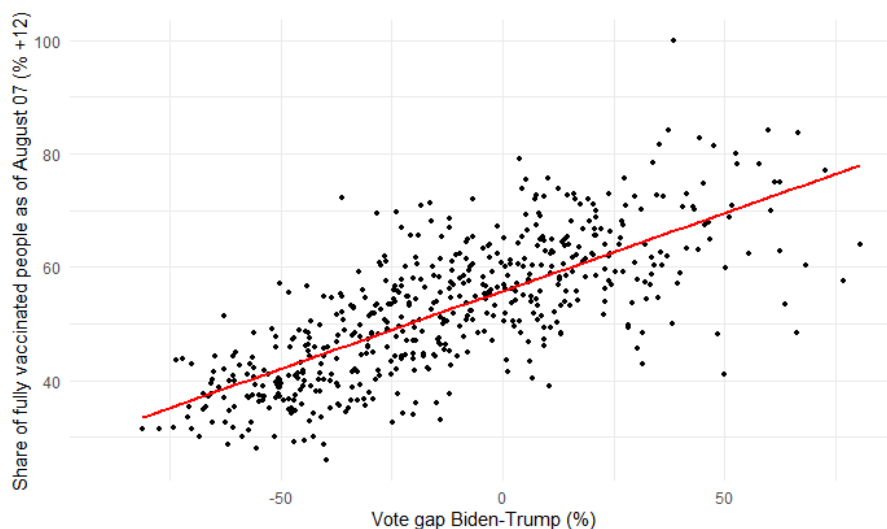**Figure 1: County-level vaccination rates in 2021 are correlated with presidential vote in 2020**

Figure 2: County-level covid-related deaths are negatively correlated with vaccination rates a month earlier (showing counties that reported at least one covid-death in the period)

---

[10] Monmouth University Poll, August 2, 2021, p.2: "17% say they remain opposed to getting the vaccine at all. Among those who admit they will not get the vaccine if they can avoid it, 70% either identify with or lean toward the Republican Party while just 6% align with the Democrats. Among those who are planning to get the vaccine or are persuadable, 45% are Republicans or lean toward that party and 40% are Democrats or Democratic leaners. Among those who have already received the vaccine, 32% fall on the Republican side of the political divide and 59% are on the Democratic side."

[11] The "partisan gap holds even after accounting for income, race and age demographics, population density and a county's infection and death rate."
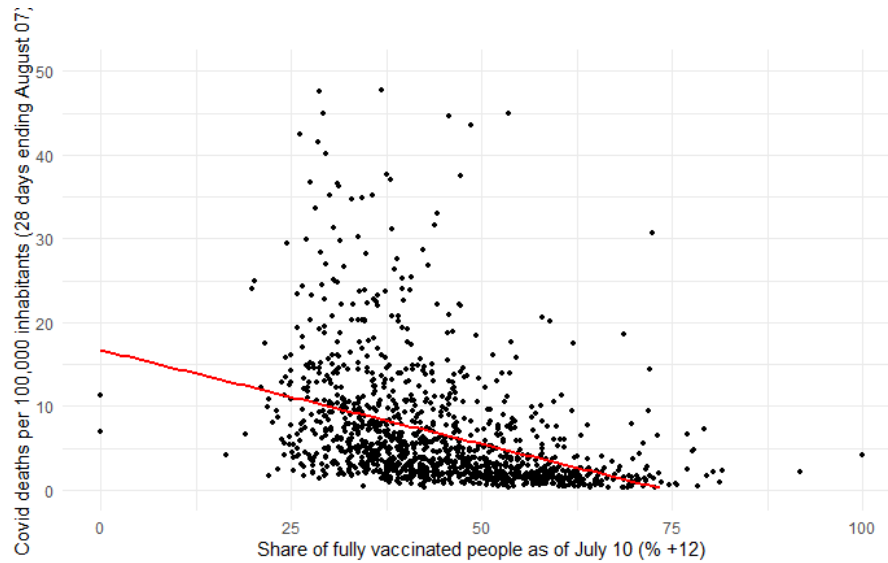
Column (2) of Table 2 controls for the poverty rate, age, and temperature, in an equation to determine the covid-19 fatality rate.[12] Poverty and age are each statistically significant, with the effect on mortality that one would expect. Low-income people are more likely to die of Covid-19, presumably because they were less healthy to begin with or do not receive as good medical care. Older people are much more vulnerable to the virus physically. Temperature has lost its beneficial effect, presumably because heat does not attract people to go outdoors in August as it does in the colder seasons.

With the controls, the results in Column (2) of Table 2 show that a 1 percentage-point increase in the number of adults (and teenagers) who were fully vaccinated in a county as of July 10th is associated with a covid-19 death rate during the subsequent 28 days (to August 7th) that was lower by an estimated .067 per 100,000 inhabitants. The effect is still highly significant statistically. Controlling for poverty and the other variables lowered the estimated coefficient slightly, but not significantly so.

But even with the controls, the estimate is biased if the vaccination decision is influenced by covid-19 prevalence, as noted. Next, we see what difference it makes to use Instrumental Variables to try to get the causality right.

Table 1 verifies that Trump voters are less likely to have gotten vaccinated than Biden voters, not just as a matter of simple correlation but also when we control for other

---

[12] The results are very similar when we add population density and race to the list of controls. These two variables do not, however, show up as statistically significant determinants of covid-19 mortality once the effects of vaccination and the poverty rate are taken into account in our regression estimates.

relevant variables: poverty, age, population density, race and temperature.[13] Use of the controls reveals an even stronger effect of partisan stance on the decision to get vaccinated: if a county supported Trump in the election, that is associated with a vaccination rate up to July 10th that increases by a highly significant 13.6 percentage points (column 3). For every percentage point in the Biden-Trump vote spread, the vaccination rate goes up by another .321 percentage points (column 4).

Using variation in the vaccination decision attributable solely to Trump-affinity, we find in Column (3) of Table 2 that the IV-estimated coefficient on vaccination rises relative to the OLS estimate. A reasonable interpretation is that Instrumental Variables successfully addresses the reverse causality problem, that covid deaths have an effect on the decision to get vaccinated.

But this is without the controls. Column (4) adds the controls back in: poverty, age and temperature. As in the OLS estimates of Column (2), poverty and age are both statistically significant. The number of most interest is the instrumented effect of vaccination. It shows that a 1 percentage-point increase in a county's vaccination rate as of July 10th reduced the covid-19 death rate during the subsequent 28 days (to August 7th) by .070 per 100,000 inhabitants. So, controlling for the poverty rate and other variables again lowered the coefficient estimate a bit.[14] But the important finding is that it remains highly significant statistically. Furthermore, the effect of vaccination on covid mortality appears still to be growing. [15]

The reason for looking at the voting pattern was to improve the estimate of the vaccine effectiveness on anyone, regardless of political party. But perhaps some of the skeptics will take note of the higher casualty rate among their fellow skeptics and will change their minds.

Table 1
First Stage: Vaccination rates by county, determined by Biden vote and other controls

---

[13] Temperature can be interpreted as an exogenous determinant of vulnerability to Covid, which in turn influences the decision to get vaccinated.

[14] A reasonable interpretation is that lower-income people are both less likely to be vaccinated (as confirmed in the 3rd and 4th columns of Table 1) and beyond that are more likely to fall victim to Covid-19.

[15] As of July 7th, the estimate was at a highly statistically significant 0.051 (a p-value of 0.1%) when instrumenting by means of 2020 voting patterns and additional controls.

	Dependent variable:			
	Fully Vaccinated as of August 07 (% +12)			
	(1)	(2)	(3)	(4)
Trump Support	-13.505***		-13.649***	
	(0.615)		(0.640)	
Biden-Trump Vote Gap (%)		0.245***		0.321***
		(0.006)		(0.007)
Population Density			0.005***	0.0001
			(0.001)	(0.001)
Poverty Rate (%)			-0.658***	-0.455***
			(0.041)	(0.033)
Median Age			-0.005	0.169***
			(0.040)	(0.032)
Share of African American (%)			0.015	-0.279***
			(0.020)	(0.017)
August Average Temperature			-0.520***	-0.042
			(0.048)	(0.040)
Intercept	54.653***	51.159***	102.407***	58.675***
	(0.564)	(0.281)	(3.957)	(3.332)
Observations	2,410	2,410	2,410	2,410
R^2	0.167	0.381	0.364	0.600
Adjusted R^2	0.167	0.381	0.362	0.599
Residual Std. Error	10.983 (df = 2408)	9.467 (df = 2408)	9.607 (df = 2403)	7.619 (df = 2403)

*p<0.1; **p<0.05; ***p<0.01

Notes: Data at the county level. The share of fully vaccinated people above 12 years old is expressed in percentage point as reported by the CDC and the states of Texas, Colorado, and Massachusetts. Trump Support is an indicator variable equal to 1 in a given county if D. Trump received more votes than J. Biden at the 2020 Presidential Election. The Biden-Trump Vote Gap measures the difference in relative votes between Trump and Biden (positive when Biden received more votes). The population density (hab/km²), poverty rate (percentage point), median age, and share of African Americans (percentage point) are obtained from the US Census 2019 estimate. The Average Temperature is obtained from the National Centers for Environmental Information. Data exclude states of Hawaii, Georgia, West Virginia, Virginia, and Vermont due to faulty reporting. Historical vaccination data are also missing for Massachusetts, Texas, and Colorado.

Table 2

Second Stage: Covid-19 death rates by county, determined by vaccination rate and other controls.

	Dependent variable:			
	Covid-related deaths per 100,000 inhabitants (28 days ending August 07)			
	(1)	(2)	(3)	(4)
Fully Vaccinated as of July 10 (% 12+)	-0.092***	-0.067***		
	(0.014)	(0.015)		
Poverty Rate (%)		0.125***		0.125***
		(0.032)		(0.034)
Median Age		0.067**		0.065*
		(0.032)		(0.033)
August Average Temperature		0.029		0.035
		(0.034)		(0.038)
IV:Fully Vaccinated as of July 10 (% 12)			-0.118***	-0.070***
			(0.019)	(0.022)
Intercept	7.472***	-0.304	8.608***	-0.489
	(0.619)	(3.244)	(0.843)	(3.890)
Observations	2,434	2,433	2,410	2,410
R^2	0.018	0.026	0.017	0.027
Adjusted R^2	0.018	0.024	0.017	0.025
Residual Std. Error	8.198 (df = 2432)	8.171 (df = 2428)	8.214 (df = 2408)	8.179 (df = 2405)

$^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Notes: Data at the county level. The number of Covid-related deaths are aggregated over 4 weeks and normalized per 100,000 inhabitants (CDC). The share of fully vaccinated people above 12 years old is expressed in percentage point as reported by the CDC and the states of Texas, Colorado, and Massachusetts. The poverty rate (percentage point) and median age are obtained from the US Census 2019 estimate, and the Average Temperature from the National Centers for Environmental Information. IV: Fully Vaccinated (%) denotes an instrumented version of the vaccination rate by the vote gap between Biden and Trump at the 2020 Presidential election, the poverty rate, county median age, and temperature. Data excludes states of Hawaii, Georgia, West Virginia, Virginia, and Vermont due to faulty reporting. Historical vaccination data are also missing for Massachusetts, Texas and Colorado.

<p style="text-align: center;"><u>References</u></p>

Adeel, Abdul Baseet, et al., 2020. "COVID-19 Policy Response and the Rise of the Sub-National Governments," *Canadian Public Policy* (UTP Journals Press), Vol. 46, Issue 4, Dec., <u>pp. 565-584</u>.

Çakmaklı, Cem, Selva Demiralp, Şebnem Kalemli-Özcan, Sevcan Yesiltas, and Muhammed A Yildirim. 2020. "COVID-19 and Emerging Markets: An Epidemiological Model with International Production Networks and Capital Flows." NBER Working Paper 27191.

———. 2021. "The Economic Case for Global Vaccinations: An Epidemiological Model with International Production Networks." <u>NBER Working Paper 28395</u>.

Leonhardt, David. 2021. "<u>Red America's Covid Problem</u>," *New York Times*, June 28.

Neelon, Brian, Fedelis Mutiso, Noel Mueller, John L. Pearce, and Sara Benjamin-Neelon. 2021. "Associations Between Governor Political Affiliation and COVID-19 Cases, Deaths, and Testing in the U.S.," *American Journal of Preventive Medicine*, Volume 61, Issue 1, 2021, <u>Pages 115-119</u>.

Washington Post. 2021. "Coronavirus infections dropping where people are vaccinated, rising where they are not, Post analysis finds," June 14.

NBC News. 2021. "Meet the Press". July 4.

Gaba, Charles. 2021. <u>"How much difference does population density & urban/rural status make re. vaccination levels?"</u>. *ACA Signups*, July 28.

Michael Hicks. 2021. "Hicks: Slowing economy caused by the unvaccinated". *Star Press*. August 1st.

Appendix 1: Extensions and Robustness

The number of covid-19 attributed deaths reported daily suffers approximations, errors, and omissions that the *New York Times* endeavors to correct in its calculations of weekly rolling averages. Table 3 uses these corrected averages and focuses on counties that reported at least one death in the 28 days leading to August 7th.

Using these corrections, Column (4) of Table 3 reports that a 1 percentage-point increase in a county's vaccination rate as of July 10th reduced the covid-19 death rate during the subsequent 4 weeks (to August 7th) by .116 per 100,000 inhabitants in counties that suffered at least one covid-19 death during the same period. This estimate is substantially larger than that shown in Table 2, which indicates that vaccination has a much larger effect in the counties where covid-19 still claims lives.

Table 4 includes additional covariates (population density, race, and mask use) to the second-stage regression of covid-related deaths. Several reasons might explain how population density can be related to vaccination and contamination rates (vaccination centers accessibility, feasibility of social distancing, but also ideology), although the relation appears to be weak[16]. The mask use index was derived from a 2020 survey from the *New York Times* undertaken at the county level, asking how often people would wear a mask in public. None of these additional controls appears to have a significant effect on covid-related deaths during the period July 10th to August 7th.

However, most covariates have a significant and large effect on the rate of covid contaminations. Column (1) of Table 5 shows that a 1 percentage-point increase in the number of adults (and teenagers) who were fully vaccinated in a county as of July 10th is associated with a covid-19 contamination rate during the subsequent 28 days (to August 7th) that was lower by an estimated 13.022 cases per 100,000 inhabitants. Using an instrumented variable for the vaccination rate and controlling for poverty, mask use, and the other variables, lowered the estimated coefficient to 16.704(Column 5). This represents a relative reduction of 3.0 percent in the average infection rate over the 28 days ending August 7th.

---

[16] Gaba (2021)

Table 3
Second Stage: Covid-19 case rates by county, determined by vaccination rate and additional controls using the New York Times corrected estimates for counties reporting at least one death in the period.

	Dependent variable:			
	Covid-related deaths per 100,000 inhabitants (28 days ending August 07)			
	(1)	(2)	(3)	(4)
Fully Vaccinated as of July 10 (% 12+)	-0.203***	-0.180***		
	(0.018)	(0.020)		
Poverty Rate (%)		0.188***		0.153***
		(0.044)		(0.046)
Median Age		0.447***		0.439***
		(0.046)		(0.046)
August Average Temperature		-0.090*		-0.087*
		(0.046)		(0.049)
IV:Fully Vaccinated as of July 10 (% 12)			-0.288***	-0.218***
			(0.023)	(0.026)
Intercept	15.538***	0.441	19.258***	2.663
	(0.833)	(4.525)	(1.045)	(5.097)
Observations	1,239	1,238	1,232	1,232
R^2	0.092	0.175	0.084	0.177
Adjusted R^2	0.092	0.172	0.084	0.174
Residual Std. Error	7.791 (df = 1237)	7.439 (df = 1233)	7.793 (df = 1230)	7.399 (df = 1227)

*p<0.1; **p<0.05; ***p<0.01

Notes: Data at the county level. The number of Covid-related cases are aggregated between June 11th and July 9th over 4 weeks and normalized per 100,000 inhabitants using the NYT corrected rolling averages. Data excludes counties that did not report any covid-19 death in the period. The share of fully vaccinated people above 12 years old, as of June 11th, is expressed in percentage point as reported by the CDC and the states of Texas, Colorado, and Massachusetts. The poverty rate (percentage point) and median age are obtained from the US Census 2019 estimate, and the June Average Temperature from the National Centers for Environmental Information. The mask use indicator was obtained from a 2020 survey conducted by the *New York Times* (see data appendix for more details). IV: Fully Vaccinated as of June 9th (%) denotes an instrumented version of the vaccination rate by the vote gap between Biden and Trump at the 2020 Presidential election, the poverty rate, county median age, June temperature, log population density, share of African American, and mask use. Data excludes states of Hawaii, Georgia, West Virginia, Virginia, and Vermont due to faulty reporting. Historical vaccination data are also missing for Texas and Colorado.

Table 4

Second Stage: Covid-19 death rates by county, determined by vaccination rate and additional controls

	Dependent variable:				
	Covid-related deaths per 100,000 inhabitants (28 days ending August 07)				
	(1)	(2)	(3)	(4)	(5)
Fully Vaccinated as of July 10 (% 12+)	-0.092***	-0.067***	-0.064***		
	(0.014)	(0.015)	(0.018)		
Poverty Rate (%)		0.125***	0.108***	0.125***	0.104***
		(0.032)	(0.038)	(0.034)	(0.039)
Median Age		0.067**	0.066*	0.065*	0.064*
		(0.032)	(0.034)	(0.033)	(0.035)
August Average Temperature		0.029	0.021	0.035	0.013
		(0.034)	(0.040)	(0.038)	(0.048)
Log Population Density			-0.041		-0.005
			(0.132)		(0.145)
Share of African American (%)			0.016		0.015
			(0.016)		(0.016)
Mask Use (%)			-0.010		-0.001
			(0.025)		(0.027)
IV:Fully Vaccinated as of July 10 (% 12)				-0.070***	-0.082***
				(0.022)	(0.031)
Intercept	7.472***	-0.304	1.174	-0.489	1.903
	(0.619)	(3.244)	(3.740)	(3.890)	(4.366)
Observations	2,434	2,433	2,432	2,410	2,410
R^2	0.018	0.026	0.027	0.027	0.027
Adjusted R^2	0.018	0.024	0.024	0.025	0.024
Residual Std. Error	8.198 (df = 2432)	8.171 (df = 2428)	8.175 (df = 2424)	8.179 (df = 2405)	8.184 (df = 2402)

*p<0.1; **p<0.05; ***p<0.01

Notes: Data at the county level. The number of Covid-related deaths are aggregated over 4 weeks and normalized per 100,000 inhabitants (CDC). The share of fully vaccinated people above 12 years old is expressed in percentage point as reported by the CDC and the states of Texas, Colorado, and Massachusetts. The poverty rate (percentage point), median age, population density (hab/km²), and share of African American population (percentage point) are obtained from the US Census 2019 estimate, and the Average Temperature from the National Centers for Environmental Information. The mask use indicator was obtained from a 2020 survey conducted by the New York Times (see data appendix for more details). IV: Fully Vaccinated (%) denotes an instrumented version of the vaccination rate by the vote gap between Biden and Trump at the 2020 Presidential election, the poverty rate, county median age, temperature, log population density, share of African American, and mask use. Data excludes states of Hawaii, Georgia, West Virginia, Virginia, and Vermont due to faulty reporting. Historical vaccination data are also missing for Texas and Colorado.

Table 5

Second Stage: Covid-19 case rates by county, determined by vaccination rate and additional controls

	Dependent variable:				
	Covid cases per 100,000 inhabitants (28 days ending August 07)				
	(1)	(2)	(3)	(4)	(5)
Fully Vaccinated as of July 10 (% 12+)	-13.022***	-6.555***	-9.389***		
	(0.857)	(0.862)	(1.013)		
Poverty Rate (%)		12.042***	13.057***	9.863***	11.864***
		(1.846)	(2.137)	(1.838)	(2.108)
Median Age		-4.715**	-0.841	-1.874	4.636**
		(1.864)	(1.951)	(1.809)	(1.901)
August Average Temperature		32.917***	25.425***	42.857***	31.567***
		(1.947)	(2.249)	(2.077)	(2.560)
Log Population Density			44.987***		83.897***
			(7.515)		(7.811)
Share of African American (%)			1.773**		0.052
			(0.901)		(0.868)
Mask Use (%)			-0.685		-0.003
			(1.402)		(1.472)
IV:Fully Vaccinated as of July 10 (% 12)				-7.820***	-16.704***
				(1.198)	(1.673)
Intercept	1,113.799***	-1,574.086***	-1,160.329***	-2,352.985***	-1,667.982***
	(38.596)	(186.401)	(212.708)	(211.111)	(234.747)
Observations	2,434	2,433	2,432	2,410	2,410
R^2	0.087	0.231	0.247	0.293	0.306
Adjusted R^2	0.086	0.230	0.245	0.292	0.304
Residual Std. Error	511.350 (df = 2432)	469.443 (df = 2428)	464.894 (df = 2424)	443.911 (df = 2405)	440.056 (df = 2402)

*p<0.1; **p<0.05; ***p<0.01

Notes: Data at the county level. The number of Covid-related cases are aggregated over 4 weeks and normalized per 100,000 inhabitants (CDC). The share of fully vaccinated people above 12 years old is expressed in percentage point as reported by the CDC and the states of Texas, Colorado, and Massachusetts. The poverty rate (percentage point), median age, population density (hab/km²), and share of African American population (percentage point) are obtained from the US Census 2019 estimate, and the Average Temperature from the National Centers for Environmental Information. The mask use indicator was obtained from a 2020 survey conducted by the New York Times (see data appendix for more details). IV: Fully Vaccinated (%) denotes an instrumented version of the vaccination rate by the vote gap between Biden and Trump at the 2020 Presidential election, the poverty rate, county median age, temperature, log population density, share of African American, and mask use. Data excludes states of Hawaii, Georgia, West Virginia, Virginia, and Vermont due to faulty reporting. Historical vaccination data are also missing for Texas and Colorado.

Appendix 2: Data

The analysis relies on data gathered from different sources at the county level. We used the data made available by the Centers for Disease Control and Prevention (CDC) for covid-19 cases and deaths. The vaccination rates also come from the CDC, except for the states of Texas, Colorado, and Massachusetts, who report independently the progress of their vaccination campaigns on dedicated websites [Texas Department of State Health Services, Colorado Department of Public Health & Environment, Massachusetts Department of Public Health]. The vaccination numbers reported at the CDC level are missing for more than a quarter of the counties located in the states of Hawaii, Georgia, West Virginia, Virginia, and Vermont. Where specified, we have excluded those states altogether. We complemented covid-related data with the 2020 Presidential Election county results as reported by the MIT Election Data and Science Lab[17].

We also added a set of covariates at the county level likely to explain covid mortality and vaccination decisions. On the demographic side, we relied on 2019 estimates of population, poverty rate, race, and median age, computed by the US Census Bureau based on the 2019 American Community Survey. We also used counties' land area reported by the US Census Bureau to calculate population densities. We included monthly average temperatures reported by the National Centers for Environmental Information and a survey conducted by the *New York Times* between July 2nd and July 14th, 2020 relative to mask use. The specific question was "How often do you wear a mask in public when you expect to be within six feet of another person?". Based on the answers "never, rarely, sometimes, frequently, always", we created a scale between 0 and 100, 100 meaning that the entire population in a given county reported "always" using a mask in public.

Table A1 summarizes the variables used in this paper.

---

[17] MIT Election Data and Science Lab, 2018, "County Presidential Election Returns 2000-2020", https://doi.org/10.7910/DVN/VOQCHQ, Harvard Dataverse, V9, UNF:6:qSwUYo7FKxI6vd/3Xev2Ng== [fileUNF]

Table A1

Variable	Description	Source
Biden-Trump Vote Gap (%)	Computed as the difference in relative votes obtained by J. Biden and D. Trump at the 2020 Presidential Election. In a county where D. Trump received 42% of the votes, and J. Biden 46%, the metric would be 4%.	MIT Election Data and Science Lab
Fully Vaccinated (% +12)	Share of the population above 12 years old who received a complete vaccination.	CDC and State Sources (TX, CO, MA)
Covid-related deaths per 100,000 inhabitants (28 days)	Deaths attributed to covid summed over 28 days and normalized per 100,000 inhabitants.	New York Times
Poverty rate (%)	Share of population living under the income threshold as defined by the US Census, which varies by family size and composition, but not geographically[18] (2019).	US Census Bureau
Median Age	County median age estimate (2019).	US Census Bureau
Share of African American (%)	Census estimate of the share of African American population, including people reporting multiple ethnicities (2019).	US Census Bureau
Population Density	Obtained by dividing population estimates by land areas, expressed in inhabitant per km² (2019).	US Census Bureau
Mask Use	2020 survey asking "How often do you wear a mask in public when you expect to be within six feet of another person? never, rarely, sometimes, frequently, always", rebased between 0 and 100.	New York Times
Average Temperature	Monthly average temperature by county (°F).	National Centers for Environmental Information

---

[18] https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html