NBER WORKING PAPER SERIES

VACCINATION RATES AND COVID OUTCOMES ACROSS U.S. STATES

Robert J. Barro

Working Paper 29884
http://www.nber.org/papers/w29884

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2022

The research reported in this working paper was supported by the National Institute on Aging of the National Institutes of Health under Award Number P30AG012810. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. I have benefited from comments by Stephen Barro, David Cutler, Ed Glaeser, Niels-Jakob Hansen, Rui Mano, Bruce Meyer, Michael Strain, Stan Veuger, and Mark Warshawsky. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Vaccination Rates and COVID Outcomes across U.S. States
Robert J. Barro
NBER Working Paper No. 29884
March 2022
JEL No. I1

ABSTRACT

Rates of COVID deaths and cases differ markedly across U.S. states, as do rates of vaccination. This study uses cross-state regressions to assess impacts of vaccinations on COVID outcomes. A number of familiar issues concerning cross-sectional regressions arise, including omitted variables, behavioral responses to vaccination, and reverse causation. The benefits from a field context and from the broad range of observed variations suggest the value from dealing with these issues. Results along these lines reveal sizable negative effects of vaccination on deaths and cases up to early December 2021, although vaccine efficacy seems to wane over time. The estimates imply that 250 additional doses, with a marginal cost around $5000, leads to one expected life saved. This $5000 is far below typical estimates of the value of a statistical life. Results since December 2021 suggest smaller effects of vaccinations on deaths and, especially, cases. These findings may reflect diminishing effectiveness of vaccines against new forms of the virus, especially the omicron variant. A further possibility is that confidence engendered by vaccinations may have motivated individuals and governments to lessen non-pharmaceutical interventions, such as masking and social distancing.

Robert J. Barro
Department of Economics
Littauer Center 218
Harvard University
Cambridge, MA 02138
and NBER
rbarro@harvard.edu

Vaccination rates against COVID-19 differ markedly across U.S. states. For example, based on CDC data and as shown in Table 1, the rate of "full" vaccination over a recent period of roughly three months, 11/17/21-2/11/22, averaged 61% with a standard deviation of 8%. These rates varied from 48% in Alabama to 77% in Vermont. If vaccinations are effective at reducing infections and deaths, these differences should map into differences in COVID-related cases and deaths.

Table 1 shows that CDC data on reported COVID-related cases and deaths also varied substantially across the states. For example, for 12/1/21-2/25/22 (14 days after the vaccination period), the change in cumulative cases per person—corresponding to cumulations of new cases over the period—averaged 0.39 with a standard deviation of 0.069.[1] The range was from 0.27 for Idaho to 0.66 for Rhode Island. Over the same period, the change in cumulative deaths per person averaged 0.0021 with a standard deviation of 0.0007 and a range from 0.0007 for the District of Columbia to 0.0037 for Michigan.

Table 1 shows comparable statistics for earlier periods. For vaccinations, the data start at 3/5/21, corresponding to the beginning of CDC information on full vaccinations.[2] Each of the four periods shown covers roughly three months (86 days). Note that the mean of full vaccination rates rose from 0.24 in 3/5/21-5/30/21 to 0.45 in 5/30/21-8/23/21, 0.55 in 8/23/21-11/17/21, and 0.61 in 11/17/21-2/11/22. (The table also shows national averages, which differ to a minor extent from the means of values across the states.)

For cases and deaths, the corresponding periods in Table 1 are each 14 days subsequent to the periods for vaccinations. Note that the mean of Covid cases per person rose from 0.05 in

---

[1]These and subsequent numbers are expressed at annual rates; that is, the changes over 86 days were multiplied by 365/86.

[2]The national fraction reported as fully vaccinated on 3/5/21 was already positive, 0.086.

3/19/21-6/13/21 to 0.08 in 6/13/21-9/6/21, 0.13 in 9/6/21-12/1/21, and 0.39 in 12/1/21-2/25/22 (all measured at annual rates). For COVID deaths per person, the mean was 0.0007 in 3/19/21-6/13/21 and 0.0006 in 6/13/21-9/6/21, then rose to 0.0018 in 9/6/21-12/1/21 and 0.0021 in 12/1/21-2/25/22. Note that COVID cases per person rose proportionately much more than COVID deaths per person in the most recent period.

The objective of this study is to use cross-sectional regressions for the U.S. states to attempt to assess the effects of vaccinations on COVID-related outcomes. The regression framework takes as dependent variables the outcomes (cases and deaths) over the four periods shown in Table 1. That is, each dependent variable is the number of cases or deaths per person cumulated over periods of roughly three months. The corresponding explanatory variables related to levels of vaccinations are averages over periods lagged 14 days compared to the dependent variables.[3] The idea is that, at any point in time, the probabilities of infection and death depend (with some lag) on the fractions of the population vaccinated.

As is familiar from the extensive work by epidemiologists on contagious disease, the high-frequency behavior of infections and deaths features waves of rising and falling outcomes. The idea in the regression analysis is to consider periods of sufficient length, such as three months, so that these short-run dynamics tend to average out. The estimated coefficients may then give estimates of the effects of vaccination rates on average probabilities of infection and death.

[3]The relevant lag may differ from 14 days and would differ for cases versus deaths. However, in practice, the regression results are not sensitive to the use of different lags between 14 and 28 days. These lags are consistent with those described by Bjornskov (2021, p. 320)

I. Issues with Cross-Sectional Regressions

As is well known, inferences from cross-sectional regressions may be difficult to draw. Because of these problems, detailed below, many researchers have moved increasingly away from these types of regressions, preferring instead to rely on randomized control trials (RCTs) or natural experiments. Although RCTs are important for assessing the efficacy of vaccines, including those recently developed for COVID-19, it is more difficult to evaluate impacts on cases and deaths in the "field." As far as I know, there are no RCTs applicable to field results connecting COVID vaccinations to COVID outcomes. In some cases, natural experiments—such as regression-discontinuity designs applied to state borders—have been used successfully in the context of COVID-19. For example, this approach has been applied to facemask mandates by Goolsbee and Syverson (2021), who consider economic impacts, and Hansen and Mano (2021a), who assess health outcomes.[4]

There are also important advantages of cross-sectional regressions. In particular, they apply to the field context and can exploit the large observed cross-sectional variations in the variables of interest—such as differences in vaccination uptake across U.S. states. Because of these major benefits, it seems worthwhile to pursue the cross-sectional regression approach in the context of COVID vaccinations and outcomes.

I consider now three major issues in interpreting the results from the cross-state regressions. The first concern is that vaccination take-up may be correlated with other variables that influence COVID outcomes. If these other variables are omitted from the regressions, the

[4]Herby, Jonung, and Hanke (2022) carry out a meta-analysis of 24 studies of the effects of facemask mandates on COVID-19 mortality. Their overall conclusion is "lockdowns have had little to no effect on COVID-19 mortality." Many of the studies considered seem to lack convincing causal evidence—the cross-border approach of Hansen and Mano (2021a) and the instrumental-variable regressions of Welsch (2020) seem superior in this regard. Oddly, these two studies were not included in the Herby, Jonung, and Hanke (2022) analysis.

estimated coefficient on the vaccination rate may proxy for the influences of these other variables. For example, if older people are more susceptible to COVID infection and, especially, death, they are likely to be vaccinated more frequently (and earlier). In this case, the observed correlation between vaccine take-up and COVID cases and deaths may be positive. This issue is handled by including as explanatory variables a set of major socio-economic variables—specifically, the fraction of the state population aged 65 and over in 2020, state life expectancy at birth in 2018, the fraction of the state adult population with education of four years of high school or more in 2019, the fraction of the state population classified by the U.S. Census as black in 2020, and the urbanization rate in 2010. The analysis also includes differences in average temperature across states at different times of the year. Inclusion of some other variables, such as population share aged 75 and over in 2020, per capita personal income in 2020, and college education in 2019, do not materially affect the results.[5]

The second issue is that persons vaccinated may alter their behavior in ways that impact probabilities of COVID infection and death. For example, a vaccinated person may feel protected and react accordingly by engaging in more social interactions or other risky behaviors. An analogous mechanism for seatbelt use, analyzed in research that began with Peltzman (1975), is that a person who uses a seatbelt (perhaps because of a legal mandate) may drive faster. These kinds of mitigating actions may not arise in clinical trials (particularly if persons do not know their vaccination status) but would apply in the field. In the regression analysis, the estimated effects of vaccinations on COVID outcomes comprise direct effects on probabilities of infection and death combined with any mitigation behavior. In some contexts, these combined effects are the objects of interest—e.g. overall effects of vaccinations on deaths (or of seatbelt use on

[5]Data by U.S. state on the socio-economic variables come from the U.S. Census Bureau. The data on personal income are from the Bureau of Economic Analysis. The temperature data are from usclimatedata.com.

automobile fatalities). In other contexts, there would be more interest in the effects of vaccinations, holding fixed the behavioral variables.[6] In any event, the present regression results apply only to the combined effects.

The third issue concerns reverse causation. Higher vaccination rates likely reduce COVID infections and deaths, and these are the effects that we seek to isolate. However, in addition, higher probabilities of infection and death likely encourage people to get vaccinated (and motivate governments to mandate or subsidize vaccinations and to support the creation and distribution of vaccines). The first channel, whereby vaccination reduces probabilities of infection and death, tends to generate a negative association between vaccination rates and rates of infection and death, whereas the second channel tends to generate a positive association. If the second channel is not held constant, the observed association between vaccination rates and rates of infection and death tends to underestimate the magnitude of the (negative) effects from vaccination.

A common way to deal with reverse causation is to use instrumental variables that explain a substantial part of the variation in the explanatory variable, in the present context the vaccination rate, but do not enter directly as determinants of outcomes, in the present case the rates of COVID cases and deaths. That is, the instrument matters for outcomes only through the channel of affecting the frequency of vaccination. The present analysis uses as an instrument a variant of the variable proposed by Welsch (2020, Section 3.2)—the Trump (Republican) share of the 2020 Presidential vote.[7] Welsch (Table 2) used the 2016 value of this variable as an

[6]This analysis would allow for welfare benefits derived from the mitigating actions; for example, people getting pleasure from greater social interactions or from driving faster while wearing seatbelts.
[7]The voting data are from Federal Election Commission (fec.gov).

instrument for facemask usage, measured in July 2020 in a survey conducted by *The New York Times*.

Perhaps surprisingly, the Trump variable has a great deal of explanatory power for vaccination rates across states, even after holding constant key socio-economic variables, such as those mentioned before—old-age share, life expectancy, education, fraction black, and urbanization. That is, the Trump variable does not matter for vaccine take-up because it proxies for these kinds of socio-economic factors. Therefore, from the standpoint of having a lot of independent explanatory power for vaccination rates, the Trump variable is a good candidate as an instrument. In effect, the 2020 Presidential voting pattern sorts people (and states) into bins for vaccine attitudes in a manner that is largely orthogonal to socio-economic characteristics.

A reasonable concern is that the Trump variable would matter for COVID cases and deaths in ways that do not work entirely through vaccination status. A particular concern is that, in accordance with Welsch's (2020, Appendix Table A1) findings, the Trump vote share is inversely related to the incidence across U.S. states of facemask mandates. For the period 3/16/20-2/1/21, which precedes the advent of vaccines, the presence of a facemask mandate at the state level is significantly negatively related to the Trump vote share.[8] However, a combination of the estimated negative effect of the Trump vote variable on facemask mandates with the Hansen and Mano (2021a) estimated negative effect of facemask mandates on COVID deaths yields a very small implied positive effect of the Trump vote on COVID deaths, compared with the effects estimated below that work through vaccinations. Therefore, from a quantitative standpoint, the Trump variable may be a satisfactory instrument for vaccination rates even

[8]The facemask mandate is measured from information given in Raifman, et al. (2022) as the fraction of days between March 16, 2020 and February 1, 2021 in which a statewide facemask mandate was in effect.

though this variable has influences on COVID outcomes that work through facemask mandates and usage.

II. Data and Empirical Setup

Data on COVID-related deaths and cases, measured relative to population, are reported by the CDC and provided by Opportunity Insights, *Economic Tracker* (see Chetty, et al. [2022]). The data used in this study are for the 50 U.S. states plus the District of Columbia.

The two measures of COVID outcomes enter as dependent variables in the regressions and are examined over the four periods noted before. The starting date, March 19, 2021, is 14 days after the beginning of data on vaccination rates (fully vaccinated persons relative to state population), also coming from the CDC and Opportunity Insights.[9] The first three periods, shown in Table 1, are 3/9/21-6/13/21, 6/13/21-9/6/21, and 9/6/21-12/1/21. These periods are of equal length (86 days) and extend to the rough date of onset of the omicron variant in the United States. The most recent period, 12/1/21-2/25/22, is the same length as the previous three. For each period, COVID-related deaths or cases are the changes in the cumulative per capita numbers (corresponding to the cumulations of new cases), expressed at annual rates.

[9]The CDC data reported by Opportunity Insights have occasional large jumps in cumulative COVID deaths and vaccinations. (The death and case data are reported by the CDC as 7-day moving averages of daily data, whereas the vaccination data are reported daily.) My interpretation, consistent with feedback obtained from the CDC, is that the jumps do not represent real changes but rather reflect shifts in procedures or assessments of data already processed. This view accords with the observation that some of the jumps are negative. As one example of a jump, the reported cumulative COVID deaths per 100,000 persons in Oklahoma shifts from 125 on 4/6/21 to 169 on 4/13/21. In the most egregious case, for the full vaccination rate in West Virginia, the variable jumps from .415 to .489 on 12/2/21, from .492 to .690 on 12/8/21, from .690 to .710 on 12/10/21, and from .716 to .548 on 12/23/21. I modified the data to smooth out these jumps. The main inferences from the results, notably from Table 2, do not change when the original data are used. However, the overall fit of the regression system is much poorer with the original data.

III. Regression Results

Tables 2-4 show results from regressions of COVID-related deaths and cases per capita on vaccination rates. I begin with results on deaths because these data are likely much more reliable than those on cases (which are sensitive to amounts of testing and likely leave many cases uncounted or misreported).

Regression results in Table 2 are for COVID-related deaths per capita, observed over the four periods of 86 days: 12/1/21-2/25/22, 9/6/21-12/1/21, 6/13/21-9/6/21, and 3/19/21-6/13/21.[10] The first two columns are for seemingly-unrelated regressions, which use a least-squares procedure but compute standard errors of estimated coefficients when allowing for correlation of the error terms across the periods. The first column has on the right-hand side the average of the full vaccination rate over periods lagged 14 days relative to the dependent variable.[11] Note that, whereas the dependent variable is the change in cumulative deaths per person over the periods shown, the independent variable is the cumulative level of full vaccinations per person (with a 14-day lag compared to the dependent variable).

To allow for a possible waning effectiveness of the vaccine, the specification in column 2 of Table 2 includes two measures of vaccination rates—one for vaccinations that occurred roughly within the last six months and the other covering vaccinations from six or more months in the past. In this specification, booster shots, for which CDC information starts on 10/20/21,[12] are viewed as converting an old vaccination into a recent one. That is, when combined with the remaining efficacy of a full vaccination from six months ago, a booster is viewed as generating effectiveness equal to that of a recent full vaccination. The inclusion of booster shots applies

[10]Results are broadly similar when the data are broken down into eight periods of 43 days between 3/19/21 and 2/25/22.
[11]The results are similar with a lag of 28 days.
[12]The national fraction of reported booster shots on 10/21/21 was already positive, 0.034.

only to the two most recent periods in Table 2; that is, no boosters existed and none of the full vaccinations were "old" up to roughly September 2021.

Aside from the vaccination variables, the regressions also include on the right-hand sides the socio-economic variables mentioned before (old-age fraction, life expectancy, high school education, fraction black, and urbanization). Also included is the historical average maximum temperature over the relevant period (computed from monthly data for the largest city in each state). In the estimation, separate coefficients are estimated for each period for each independent variable, including the constant term. In this specification, the constant terms absorb variations over time in aggregate COVID outcomes.

In column 1, the estimated coefficients on the (roughly) contemporaneous vaccination rate are all negative, significant at the 1% level for 9/6/21-12/1/21 and 6/13/21-9/6/21, and significant at the 5% level for 12/1/21-2/25/22. To assess the magnitudes of the estimated responses, consider the period 9/6/21-12/1/21, for which the estimated coefficient is the largest in magnitude, -0.0091. Over this period, the mean of the vaccination rate variable is 0.548 with a standard deviation of 0.078. Therefore, a one-standard-deviation increase in the vaccination rate, which is a rise by 14.2%, is estimated to lower the death rate by 0.00071, compared to the mean death rate of 0.00175. That is, the death rate falls by 40.6%. The implied elasticity of response is the ratio of -40.6 to 14.2, which equals -2.9. The estimated elasticities are smaller in magnitude for the other periods, corresponding to the smaller sizes of the estimated coefficients in Table 2, column 1.[13]

[13]For the other explanatory variables, the fraction over age 65 is significantly positive in each period, and high school education is negative and at least marginally significant in each period. Life expectancy and fraction black are each significantly negative in two periods, and urbanization rate is significantly positive in two periods. The temperature variable is significantly negative in two periods, with a particularly strong effect in the most recent period, 12/1/21-2/25/22. This last result suggests a tendency for colder places to have more COVID deaths during

When the two measures of vaccination rates are included in column 2, the results for the period 9/6/21-12/1/21 suggest that recent vaccinations are roughly twice as effective against deaths as older vaccinations; estimated coefficients are -.0097 and -.0064, respectively. Each of these estimated coefficients is statistically significant at least at the 5% level. However, the two estimated coefficients differ from each other only with the high p-value of 0.31. For the most recent period, 12/1/21-2/25/22, there is essentially no information about a possible waning influence of vaccinations.

The small size of the estimated coefficient for the earliest period, 3/19/21-6/13/21, may reflect reverse causation from COVID deaths to vaccination propensity. This effect is likely to be powerful during the early stages of vaccination rollout, when the places most adversely impacted are especially likely to have large rollouts of vaccinations. This channel could also be operating in the most recent period, 12/1/21-2/25/22, which features the introduction of booster shots.

Another way to interpret the estimated effects of vaccinations on COVID deaths comes from the literature on the value of a statistical life (surveyed in Viscusi and Aldy [2003]). The point estimates for 9/6/21-12/1/21 from Table 2 imply that the coefficient -0.0097 applies to vaccination rates over the first six months and the coefficient -0.0064 applies over the next six months. If vaccinations are ineffective after 12 months and boosters are ignored, a quantity V of vaccinations would be expected to reduce deaths over one year by the amount $V\cdot(.5\cdot.0097 + .5\cdot.0064) = .00805\cdot V$. Therefore, to expect to save one life over the year, one needs $1/.000805 =$ 124 full vaccinations, which correspond roughly to 248 shots. If shots cost $20 each at the margin, then it costs about $5000 to expect to save one life. Since usual estimates of the value of

the winter. However, the temperature variable is not statistically significant when considered for an earlier winter period, 12/23/20-3/19/21, which precedes the advent of full vaccinations.

a statistical life are much larger than $5000 (see Viscusi and Aldy [2003]), this result indicates that vaccinations against COVID-19 are a great bargain. The results are less powerful with the smaller magnitudes of coefficients estimated for other periods. For example, with the coefficients estimated for 12/1/21-2/25/22 in Table 2, it requires 455 full vaccinations or around 900 shots or about $18000 to expect to save one life. Even this higher magnitude suggests that vaccinations are a great deal.

The instrumental estimation treats the vaccination rates as endogenous. The instrument list includes the 2020 Republican vote share for President, along with the other explanatory variables mentioned before. That is, the Trump vote share is the one excluded instrument.[14] Table 3 shows first-stage regressions, with the vaccination rate over the various periods as the dependent variable. The remarkable aspect of these results is the strong explanatory power of the Republican vote share in the 2020 election (Trump vote), especially for the three most recent periods. The important point is that a higher Trump vote share strongly associates with a lower vaccination rate even when the other explanatory variables are held fixed. An increase by 0.12 in this vote share (which has a mean of 0.49 and a standard deviation of 0.12) associates in the most recent period, 11/17/21-2/11/22, with a decline by 0.065 in the vaccination rate (which has a mean in this period of 0.61). The results are similar for the two preceding periods but are much weaker for the earliest period, 3/5/21-5/3/21. In this case a rise by 0.12 in the Trump vote share associates with a fall in the vaccination rate by only 0.011, compared to the mean of 0.24.

[14]When two vaccination variables are included, an additional instrument is required. The results in column 4 of Table 2 include the 6-month lag of the full vaccination rate on the instrument list. Hansen and Mano's (2021b) county-level analysis used as an instrument the state-level vaccine allocation interacted with the county density of pharmacies. Possibly a variable along these lines could be used for the state-level analysis.

The results from instrumental estimation are in columns 3 and 4 of Table 2. For the two periods where the estimated effects from vaccinations on COVID deaths were strongest, 9/6/21-12/1/21 and 6/13/21-9/6/21, the estimated coefficients from instrumental estimation are still highly significant and now slightly larger in size. These changes go in the direction expected—if there is positive reverse causation from COVID deaths to vaccinations—but the magnitudes of change are minor.

For the earliest period, 3/19/21-6/13/21, the extent of the change in the point estimate of the coefficient is much larger under instrumental estimation, and this estimated value is now in the ballpark of those found for other periods. However, the standard error of the coefficient estimate blows up, likely because the excluded instrument—the Trump vote variable—is only marginally significant for explaining the vaccination rate in this period (see Table 3). That is, the instrument is weak.

For the most recent period, 12/1/21-2/25/22, the instrumental estimate in column 3 of Table 2, which includes only one vaccine variable, is close to that found before. In column 4, the results do not clearly distinguish the effect from recent vaccinations (including boosters) to that from older vaccinations. In any event, the main inference is that vaccinations had less effect overall against COVID deaths, compared to that in periods that preceded the rise of the omicron variant in early December 2021.

Table 4 has regression results with COVID cases per capita as the dependent variable. This setting parallels that in Table 2 for COVID deaths. Results in Table 4 for the second and third periods, 9/6/21-12/1/21 and 6/13/21-9/6/21, roughly parallel those for COVID deaths. To evaluate the magnitudes of the estimated responses for cases, consider the period 9/6/21-12/1/21, for which the estimated coefficient on the vaccination rate in column 1 is -0.42. Over this period, the mean of the vaccination rate is 0.548 with a standard deviation of 0.078, so that a one-standard-deviation increase in the vaccination rate, which is a rise by 14.2%, is estimated to lower the case rate by 0.033, compared to the mean of 0.134. That is, the case rate falls by 24.6%. The implied elasticity of response is the ratio of -24.6 to 14.2, which equals -1.7.

Results in Table 4 for the earliest period, 3/19/21-6/13/21, are also parallel to those for deaths in the sense that the vaccination rate does not have a statistically significant effect on cases. These results may again reflect reverse causation in this period—the point estimate of the coefficient on the vaccination rate is negative and much larger in magnitude in the instrumental estimation, but the standard error also blows up.

The hardest results to interpret for COVID cases are for the most recent period, 12/1/21-2/25/22, which covers the rise of the omicron variant of the virus. In particular, there is no indication in this period that vaccinations, recent or old, reduce COVID cases.

IV. CDC Reports on COVID Outcomes in Relation to Vaccination Status

The regression findings can be compared with CDC reports on COVID outcomes in relation to vaccination status.[15] On October 30 2021, data from 29 jurisdictions indicate that COVID cases per capita for unvaccinated persons were 4.3 times those for fully vaccinated and 12.7 times for those with booster shots. These ratios fell to 2.0 and 3.7, respectively, on January 1, 2022, and 3.1 and 3.0, respectively, on February 19, 2022. That is, case rates for unvaccinated had become much closer to those for fully vaccinated and boosted, and the distinction between fully vaccinated and boosted no longer appeared. For COVID deaths, covering 26 jurisdictions, the ratios were 12.0 and 36.6, respectively, on October 30 2021 and

[15] See cdc.gov/covid-data-tracker/#rates-by-vaccine-status.

fell to 6.2 and 24.1, respectively, on January 1, 2022. For February 19 2022, the ratio for fully vaccinated or more was 2.8 but separate data for boosters were not yet available. In any event, death rates for unvaccinated had become much closer to those for vaccinated.

Overall, the CDC data on outcomes in relation to vaccination status seem consistent with the regression findings, which indicate weaker effects of vaccinations on COVID deaths and, especially, cases since early December 2021. It is worth keeping in mind, however, that the CDC analysis is subject to issues similar to those that apply to the cross-state regressions. For example, if less healthy people are more likely to die from COVID, for given vaccination status, and more likely to be vaccinated, then the association between vaccination and death would tend to understate the beneficial benefits from vaccination. Similar effects arise if older people are more likely to die for given vaccination status and more likely to be vaccinated, although the CDC indicates that its statistics adjust for age. However, the CDC data do not adjust for other socio-economic variables or for vintage of vaccination.

V. Speculative Thoughts and Research Plans

The results in Tables 2 and 4 reveal substantial negative effects of vaccinations on COVID deaths and cases up to roughly the emergence of the omicron variant of the virus in early December 2021. Results on deaths (Table 2) suggest that the power of vaccines wanes over time but still remains effective even after about six months. This waning influence is offset by the introduction of booster shots. In comparison to the findings from earlier periods, the results since early December 2021 indicate that vaccinations have a weaker effect in reducing COVID deaths and may no longer reduce COVID cases.

The cross-state regression results accord in a sense with the aggregate U.S. data, which do not directly enter into the regressions. That is, since early December 2021, COVID deaths and, particularly, cases surged in an upward wave, followed by a downward wave. The overall rise in cases and deaths over the roughly 3-month period after December 1, 2021 occurred despite the continuing rise in "full" vaccination rates and the spread of booster shots (see Table 1). One caveat in interpreting the aggregate data is that the rise in reported cases should be discounted because of the sharp increase in testing. Another consideration is that the rises in aggregate deaths and cases in winter 2021-2022 may to some extent reflect seasonal factors (which enter through the temperature variable in the cross-state regressions).

There are a number of possible explanations for the apparent reduction in the effectiveness of vaccinations in the cross-state analysis for the period since early December 2021 (Tables 2 and 4). One is waning efficacy of vaccinations over time, though the regression analysis attempted to take account of this channel by considering the vintages of vaccinations and allowing for the introduction of booster shots. Another factor is diminishing effectiveness of existing vaccines against new forms of the virus, notably the omicron variant. A further possibility is that the confidence engendered by vaccinations (despite the increases in overall deaths and, particularly, cases) may have motivated individuals and governments to lessen non-pharmaceutical interventions, such as masking and social distancing. These responses may have been reinforced by "COVID fatigue," which raised the perceived benefits from social interactions compared to the costs attached to health risks. Of course, this response need not be irrational; that is, the benefits from heightened social interactions may, in fact, more than offset the costs from the increases in deaths and cases.

More narrowly, in terms of research plans, the first idea is to carry out the analysis at the county level. This change will sharply raise the available number of cross-sectional observations. However, the county-level data introduce new concerns about measurement error and about the connection between location of vaccination and location of outcome.

Second, a key issue in the estimation involves the instrumental variables employed. Even if the Trump 2020 vote is viewed as an appropriate instrument, there are difficulties in extending the analysis to allow for more than one endogenous variable on the right-hand side of the regressions. This issue arises, for example, in attempting to distinguish the impact of recent from older vaccinations. Relatedly, this analysis involves the role of booster shots. At this stage, it is unclear what additional instruments are available.

References

Brornskov, C. (2021). "Did Lockdown Work? An Economist's Cross-Country Comparison," CESifo Economic Studies, 318-331.

Chetty, R., N. Hendren, J. Friedman, S. Oppenheimer, W. Van Dijk, and M. Stepner (2022), Opportunity Insights, *Economic Tracker,* available at tracktherecovery.org.

Goolsbee, A. and C. Syverson (2021). "Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020," *Journal of Public Economics* 193, January, 1-8.

Hansen, N.J. and R.C. Mano (2021a). "Mask Mandates Save Lives," IMF working paper, 21/205, August.

Hansen, N.J. and R.C. Mano (2021b). "COVID-19 Vaccines: A Shot in Arm for the Economy," IMF working paper 21/281, December.

Herby, J., L. Jonung, and S.H. Hanke (2022). "A Literature Review and Meta-Analysis of the Effects of Lockdowns on COVID-19 Mortality," Studies in Applied Economics no. 200, Johns Hopkins University, January.

Peltzman, S. (1975). "The Effects of Automobile Safety Regulation," *Journal of Political Economy* 83, August, 677-726.

Raifman, Julia, Kristen Nocka, David Jones, Jacob Bor, Sarah Lipson, Jonathan Jay, Megan Cole, Noa Krawczyk, Emily Benfer, Philip Chan, and Sandro Galea (2022). *COVID-19 US State Policy Database*, available at tinyurl.com/statepolicies.

Viscusi, W.K. and J.E. Aldy (2003). "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World," *Journal of Risk and* Uncertainty, 27, August, 5-76.

Welsch, D.M. (2020). "Do Masks Reduce COVID-19 Deaths? A County-Level Analysis Using IV," *COVID Economics* 57, November 13, 20-45.

Table 1 Means and Standard Deviations of Variables

Variable	Mean	Std. Dev.	Natl Avg
COVID cumulative deaths per capita (change per year)			
12/1/21-2/25/22	.00207	.00066	.00203
9/6/21-12/1/21	.00175	.00090	.00168
6/13/21-9/6/21	.00058	.00043	.00061
3/19/21-6/13/21	.00068	.00031	.00077
COVID cumulative cases per capita (change per year)			
12/1/21-2/25/22	.389	.069	.387
9/6/21-12/1/21	.134	.057	.110
6/13/21-9/6/21	.077	.044	.082
3/19/21-6/13/21	.049	.022	.049
"Full" vaccinations per capita			
11/17/21-2/11/22	.606	.083	.614
8/23/21-11/17/21	.548	.078	.553
5/30/21-8/23/21	.450	.071	.448
3/5/21-5/30/21 (data start 3/5/21)	.237	.031	.228
Booster vaccinations per capita			
11/17/21-2/11/22	.198	.053	.196
8/23/21-11/17/21 (data start 10/20/21)	.025	.008	.024
Fraction over age 25 with completed high school, 2019	.901	.027	.886
Population fraction 65 and older, 2020	.173	.020	.169
Life expectancy at birth, 2018	78.8	1.8	79.3
Population fraction black, 2020	.110	.101	.124
Urbanization rate, 2010	.741	.149	.809
Fraction of votes Republican, 2020 Presidential election	.492	.120	.469
Population fraction 75 and older, 2020	.068	.010	.067
Fraction over age 25 with completed college, 2019	.327	.065	.331
Per capita personal income ($1000s), 2020	57.7	9.4	59.6
Maximum temperature, December 1-February 25	25.2	12.7	51.0
Maximum temperature, September 6-December 1	66.2	8.9	69.7
Maximum temperature, June 13-September 6	84.7	6.5	85.0
Maximum temperature, March 19-June 13	69.6	8.1	71.6

Note: COVID-related deaths and cases are differences in cumulative values per person for dates shown (corresponding to cumulations of new deaths and cases), expressed at annual rates. Full and booster vaccinations are averages per person over periods shown. The averages apply to dates at the start, end, and middle of each period, with the middle value getting double weight. Maximum temperature is average high temperature in degrees Fahrenheit over dates shown. Underlying values are monthly for largest city in each state.

Table 2 Regressions for COVID Deaths per Capita

	(1)	(2)	(3)	(4)
Estimation method	**SUR**	**SUR**	**Instruments**	**Instruments**
12/1/21-2/25/22				
vaccination rate	-.0022** (.0011)	-.0021* (.0012))	-.0024* (.0015)	-.0047 (.0053)
vaccination rate, older	--	-.0023 (.0015)	--	.0017 (.0065)
p-value for equal coeffs		0.93		0.58
9/6/21-12/1/21				
vaccination rate	-.0091*** (.0016)	-.0097*** (.0017)	-.0098*** (.0021)	-.0101*** (.0022)
vaccination rate, older	--	-.0064** (.0031)	--	-.0064** (.0033)
p-value for equal coeffs		0.31		0.30
6/13/21-9/6/21				
vaccination rate	-.0041*** (.0008)	-.0041*** (.0008)	-.0042*** (.0012)	-.0042*** (.0012)
3/19/21-6/13/21				
vaccination rate	-.0007 (.0015)	-.0006 (.0015)	-.0052 (.0095)	-.0052 (.0095)
R-squared	.64 .66 .64 .31	.64 .67 .64 .31	.64 .66 .64 .19	.54 .67 .64 .19
s.e.	.0004 .0006 .0003 .0003	.0004 .0006 .0003 .0003	.0004 .0006 .0003 .0003	.0005 .0006 .0003 .0003

Notes to Table 2

Sample is 50 U.S. states plus District of Columbia. Sample dates shown in the left-most column refer to the dependent variable. This variable is the change in cumulative reported COVID-related deaths per capita over each period (values expressed per year). Vaccination rate in columns 1 and 3 is the fraction of the population fully vaccinated against COVID-19 (not counting booster shots). This variable is lagged 14 days from the dependent variable and is entered as an average over each period, as described in Table 1. In columns 2 and 4, vaccination rate is the fraction of the population fully vaccinated over roughly the last 6 months plus the fraction fully vaccinated earlier who have received booster shots. In these columns, "vaccination rate, older" is the fraction fully vaccinated roughly 6 or more months in the past less the fraction who have received booster shots. Other explanatory variables, shown in Table 1, are fraction of population aged 65 and over in 2020, life expectancy at birth in 2018, fraction of population aged 25 and over who completed high school or more in 2019, fraction of population black in 2020, urbanization rate in 2010, and average maximum temperature over periods corresponding to the dependent variable. Coefficients on these variables, constant terms, and the vaccination rates differ across periods. Standard errors of coefficient estimates are in parentheses. SUR (seemingly-unrelated regression) allows for correlation of the error terms across periods. s.e. is the standard error of each regression. In columns 1 and 3, instrumental estimation (three-stage least-squares) uses as the excluded instrument the fraction of the population voting in 2020 that voted Republican (as shown in Table 1). In columns 2 and 4, the instrument list also includes the fraction of the population fully vaccinated roughly 6 or more months in the past.

***Significant at 1%, **significant at 5%, *significant at 10%.

Table 3 First-Stage Regressions for Vaccination Rates

	(1)	(2)	(3)	(4)
Periods for vaccination rates	**11/17/21-2/11/22**	**8/23/21-11/17/21**	**5/3/21-8/23/21**	**3/5/21-5/3/21**
Constant	.12 (.41)	.17 (.38)	.06 (.37)	.00 (.29)
Over-65	.77*** (.27)	.85*** (.25)	.79*** (.23)	.26 (.19)
Life expectancy	.0108** (.0052)	.0081* (.0048)	.0051 (.0046)	.0022 (.0036)
High School Education	-.19 (.24)	-.12 (.21)	.14 (.20)	.11 (.17)
Black	-.214*** (.070)	-.223*** (.065)	-.255*** (.062)	-.134*** (.050)
Urban	-.046 (.060)	-.029 (.047)	-.041 (.045)	-.047 (.035)
Average Maximum Temperature	.0000 (.0003)	-.0001 (.0003)	.0002 (.0004)	.0001 (.0005)
Trump vote	-.543*** (.061)	-.515*** (.057)	-.464*** (.055)	-.090** (.044)
R-squared	.83	.83	.81	.43
s.e.	.037	.035	.034	.026

Notes: Sample is 50 U.S. states plus District of Columbia. Dependent variables, over the periods shown in the top row, are the averages of full vaccination rates, as used in Table 2. Over-65 is the fraction of the population in 2020 that was aged 65 or more. Life expectancy at birth is for 2018. High School Education is fraction of the population in 2019 aged 25 or more that had completed four years of high school or more. Black is the fraction of the population in 2020 classified as black. Urban is the fraction of the population urbanized in 2010. Trump vote is the fraction of votes for President in 2020 that went Republican. Estimation is by seemingly-unrelated regression, which allows for correlation of the error terms across periods. Standard errors of estimated coefficients are in parentheses. s.e. is the standard error of each regression.

***Significant at 1%, **significant at 5%, *significant at 10%.

Table 4 Regressions for COVID Cases per Capita

	(1)	(2)	(3)	(4)
Estimation method	**SUR**	**SUR**	**Instruments**	**Instruments**
12/1/21-2/25/22				
vaccination rate	.296* (.169)	.151 (.184)	.291 (.231)	.002 (.765)
vaccination rate, older	--	.558** (.228)	--	.800 (.938)
9/6/21-12/1/21				
vaccination rate	-.419*** (.090)	-.512*** (.087)	-.339*** (.126)	-.370** (.121)
vaccination rate, older	--	.032 (.163)	--	.136 (.177)
6/13/21-9/6/21				
vaccination rate	-.477*** (.085)	-.469*** (.085)	-.528*** (.122)	-.528*** (.122)
3/19/21-6/13/21				
vaccination rate	-.037 (.094)	.024 (.096)	-.736 (.813)	-.736 (.813)
R-squared	.15 .71 .65 .43	.16 .76 .65 .44	.17 .71 .66 -.24	.13 .75 .66 -.24
s.e.	.069 .033 .028 .018	.069 .030 .028 .018	.068 .033 .028 .026	.071 .031 .028 .026

Notes: See notes to Table 2. The only difference is that the dependent variable is based on COVID-related reported cases per capita.

***Significant at 1%, **significant at 5%, *significant at 10%.