# Direct and indirect effects of vaccines: Evidence from COVID-19 in schools[*]

Seth Freedman[†] Daniel W. Sacks[‡] Kosali Simon[§] Coady Wing[¶]

September 11, 2022

**Abstract**

Externalities are a key justification for vaccine subsidies and mandates. However measuring the existence and size of vaccine externalities is challenging because it requires exogenous variation in peer vaccination status, and thus work to date has not addressed the infection-reducing externalities of the COVID-19 vaccine. We overcome this challenge using unique data from Indiana that allows us to exploit exogenous variation in peer COVID-19 vaccination rates among sixth graders in Fall 2021. Sixth graders in middle schools had many vaccine eligible peers, but sixth graders in elementary schools had few, because seventh grade and older students were vaccine eligible by the start of Fall 2021, but younger students were not. Consequently the Fall 2021 vaccination rate in middle schools was 20 percent, whereas in elementary schools it was three percent.

Using population registries of COVID-19 tests and vaccinations in Indiana along with fine-grained address and school catchment maps and difference-in-differences designs, we find two key results. First, vaccines appear effective at preventing own infection, with no adverse side effects detectable in electronic health records. Second, however, vaccines appear to have essentially no effect on peer infections, at least at the low vaccination rate in our sample of adolescents. This evidence from real-world settings matches clinical evidence for COVID-19 vaccines' benefits for the vaccinated, and provides new evidence that clinical trials were unable to examine, on spillover effects. Prior work on the influenza vaccine has found substantial externalities, thus our findings suggest that prior evidence on one disease and its vaccine need not generalize to others.

JEL codes: D62, I12, I18,I21,J13
Key words: vaccines, spillovers, externalities, COVID-19

[*]We acknowledge the Indiana University Pervasive Technology Institute for providing HPC resources that have contributed to the research results reported in this paper (`https://pti.iu.edu`). This project was funded (or partially funded) by the Vice Provost for Research through the Research Equipment Fund. We thank Dr. Sarah Caddy and Dr. Buddy Creech for helpful background on immunology. We are grateful to Scott Imberman, Emily Lawler, and Emily Oster for helpful comments.

[†]O'Neill School of Public and Environmental Affairs, Indiana University (freedmas@indiana.edu).

[‡]Risk & Insurance, Wisconsin School of Business, University of Wisconsin - Madison (dansacks@wisc.edu).

[§]O'Neill School of Public and Environmental Affairs, Indiana University and NBER. (simonkos@indiana.edu).

[¶]O'Neill School of Public and Environmental Affairs, Indiana University (cwing@indiana.edu).

# 1 Introduction

The ability of vaccines to control pandemics depends on both their direct and indirect effect: how they reduce illness among the vaccinated, and how they mitigate transmission to the unvaccinated. Indirect effects are a textbook example of an externality (Gruber, 2005), and may justify interventions such as mandates and subsidies. Yet the magnitude of vaccines' direct and indirect effects are difficult to ascertain. Clinical trials can establish the safety and efficacy of the vaccines for the vaccinated—their direct effects—but the direct effects in clinical trials may differ from their effects in the field: clinical trial populations are not necessarily representative of the population, in terms of demographics (e.g. Hall (1999)), expected benefits (Malani, 2008), or study site (Allcott, 2015); behavioral responses such as increased risk taking may offset some of the health benefits of an imperfect vaccine (Chan et al., 2016); and post-clinical trial viral variants may evade the immune response generated by a vaccine. Clinical trials also are not designed to detect indirect effect—reductions in transmission to the unvaccinated—and evidence on direct effects is not necessarily informative about indirect effects, as vaccines can prevent illness or hospitalization without preventing viral replication.

In this paper, we provide novel quasi-experimental evidence from the field on the direct and indirect effects of the COVID-19 vaccines, using rich data from Indiana. We find large direct effects of the vaccines—reduced COVID-19 incidence— during the height of the Delta wave, comparable to the results observed in the clinical trials, with no detectable adverse effects. We find negligible indirect effects—no effects on infection among the unvaccinated—in our study population, which has low vaccination rates.

Our research design takes advantage of age-based roll out of vaccine eligibility, along with variation in peer vaccination arising from school district idiosyncrasies. Children aged 12-16 became vaccine eligible in May 2021 in Indiana and nationally. But children 5-11 were not eligible until November 2021. To identify direct effects of the vaccine, we compare students who became eligible in May to slightly younger students who were ineligible until November. To identify spillover effects—which requires variation in peer vaccination—we compare sixth graders in middle schools to those in elementary schools. During our study, sixth graders were themselves ineligible for the vaccine in both types of schools, as were other ele-

mentary school students. However, in middle schools, older students were vaccine eligible. Sixth graders in middle schools therefore are exposed to a much higher peer vaccination rate than the elementary school sixth graders. While the middle school / elementary school comparison identifies cross-grade spillovers, we identify within-grade spillovers in a complementary design that focuses on seventh graders (whose peers are vaccinated) who are themselves unlikely to be vaccinated. In all cases we estimate vaccine effects in a difference-in-differences framework, using 2020 as the pre-period.

We use data containing the near-universe of COVID-19 vaccination records and PCR-based COVID-19 tests in Indiana, as well as linked medical records for a large segment of the state population. We observe date of birth and detailed location (zip code and census tract) of the residential address, from which we infer school type. We estimate direct effects among the 25,000 students born just early enough to become eligible in May 2021, or just late enough to become eligible in November 2021. We estimate spillover effects using data on the 54,000 elementary and middle school sixth graders, for whom we can accurately determine school assignment.

We begin by confirming clinical trial-established vaccine efficacy and safety in our context. Early vaccine eligibility increases vaccination rates by 26 percentage points and reduces COVID-19 incidence by 0.4 percentage points, for an implied vaccine effectiveness of about 80 percent. Since our post-treatment period includes August-December 2021, this shows the vaccines remained effective against the Delta wave. Vaccine eligibility has no effect on all-cause emergency room visits, suggesting that adverse side effects of the vaccine are very rare.

Despite clear direct effects of vaccines among older students, we find no infection-reducing spillovers of going to school with vaccine-eligible peers. In the Fall of 2020, before vaccines were available, we find that elementary and middle school sixth graders have nearly identical COVID-19 incidence, suggesting little confounding. A year later, as COVID-19 cases rose during the Delta wave, middle school sixth graders attended schools with an overall vaccination rate of 25 percent, versus less than 5 percent for elementary school sixth graders. Yet we continue to see nearly identical COVID-19 rates among sixth graders in the middle and elementary schools. Overall we estimate that going to school with vaccine-eligible peers

induces a statistically insignificant 0.2 percentage point *increase* in COVID incidence among Indiana sixth graders. While our primary analysis focuses on cross-grade spillovers, we also examine within-grade spillovers using the sub-population of seventh-grade students with unvaccinated parents, for whom the vaccine take-up effect is small. Although these largely unvaccinated students attended school in classrooms with vaccinated peers, they did not experience lower COVID-19 incidence suggesting that within-grade spillover effects were negligible.

Our analysis suggests there were essentially no infection-reducing spillovers from the COVID-19 vaccine in our quasi-experimental setting. These results are consistent with the possibility that the vaccine does little to reduce the overall spread of the virus, despite have an important protective effect on people who actually take up the vaccine. However there are at least three other explanations for our results. First, it is possible that the degree of contact across peers in our quasi-experiment is not great enough for us to detect spillovers. We think this is unlikely, as other studies find spillovers from adolescents to toddlers, or from non-elderly to elderly (Ward, 2014; Carpenter and Lawler, 2019; White, 2021), whereas we investigate spillovers between grades 6 and 7 and within grade 7. Second, it is possible that infection-reducing vaccine spillovers are offset by heightened risk-taking among the unvaccinated, a Peltzman (1975) effect. However, we find that when they become eligible for the vaccine, sixth graders in middle schools are slightly more likely to take up the vaccine than sixth graders in elementary schools. This happens despite the greater number of vaccinated peers at middle schools, a pattern that is inconsistent with risk-compensating behaviors. Third, it is possible that the vaccine take-up is not large enough to generate meaningful externalities. The 20 percent take-up induced by our quasi-experiment is similar to that that in prior work finding spillovers of other vaccines Ward (2014); Carpenter and Lawler (2019); White (2021)—albeiet the base rate is much lower—and we show in a simple SIR model that if vaccines were perfectly effective in reducing transmission, and infectiousness (represented by the $R_0$ parameter, sometimes called the basic reproductive number) is not too high, then a 20 percent increase in peer take-up would produce large externalities. Nevertheless, it is possible that with imperfectly effective vaccines or very high infectiousness, a greater degree of take-up is necessary to produce meaningful external reductions in infection. Taken to-

gether, our results suggest that if the COVID-19 vaccines really can help protect the unvaccinated through spillovers, such benefits would only materialize with higher levels of take-up than is currently observed among adolescents.

Thus our primary finding is that a large increase in COVID-19 vaccination does not reduce COVID-19 incidence in a largely unvaccinated peer group. This finding contributes to two literatures. First is a small literature investigating health externalities of vaccines. Ward (2014) and White (2021) find that influenza vaccinations among the non-elderly and among healthcare workers generate large health effects among the elderly. Carpenter and Lawler (2019) find that TDap booster mandates for middle school students reduce pertusis incidence among 0-4 year-olds. We complement these papers by examining the COVID-19 vaccines in particular, a major, contentious vaccine at the center of public health efforts.[1]

Our finding on health spillovers from the COVID-19 vaccine also contribute to the large and rapidly growing literature on pandemic mitigation policy. Much of this literature has focused on the mobility (examples include Gupta et al. (2021); Cronin and Evans (2021)) and economic consequences of non-pharmaceutical interventions (such as Chetty et al. (2020); Kong and Prinz (2020); Goolsbee and Syverson (2021); Alexander and Karger (2021)). Our work is closest to the strand of literature investigating health consequences of policy interventions, such as masking (e.g. Abaluck et al. (2021); Ginther and Zambrana (2021)) and shelter-in-place orders (for example Dave et al. (2021); Berry et al. (2021); Friedson et al. (2021). Especially relevant is recent work by Acton et al. (2022) showing that college vaccine mandates reduce local COVID incidence and mortality, even among people too old to be college students. We complement their work. They show that vaccine mandates are an effective public health tool in a high take-up environment. We show that direct effects are large, but indirect effects small, in an environment with relatively low vaccine take-up.

Finally, our finding of large direct effects contributes to a literature investigating the performance of the COVID-19 vaccines in the field. That literature typically relies on simple comparison of COVID inci-

[1]Also relevant is the larger literature investigating the causes of vaccination: mandates (Abrevaya and Mulligan, 2011; Bugenske et al., 2012), recommendations (Lawler, 2017, 2020), financial incentives (Bronchetti et al., 2015), and disease outbreak (Oster, 2018; Schaller et al., 2019).

dence among vaccinated and unvaccinated groups, controlling for observables but not directly addressing selection on unobservables (e.g. Bernal et al. (2021); Andrews et al. (2022); Goldberg et al. (2021)). Our quasi-experimental design provides further evidence of the continued effectiveness of the vaccines among the vaccinated.

# 2 Background and research designs

## 2.1 Vaccines can prevent illness without preventing infection

Clinical trials establish the in-sample effect of vaccines on illness, hospitalization, and death, but they do not necessarily show that the vaccines reduce transmission. Preventing transmission requires "sterilizing immunity," meaning that the vaccine prevents the virus from entering cells and replicating. Most vaccines do not produce complete sterilizing immunity (Caddy, 2021). For example, the rotavirus vaccine and the Hepatitis B vaccine protect against illness without conferring sterilizing immunity (Baker et al., 2019; Werner et al., 2013).[2] Sterilizing immunity likely occurs along a spectrum; people more recently vaccinated or boosted are likely less transmissive, and vaccines are likely less effective in preventing transmission of variants than of the original virus for which they were designed. Even vaccines highly effective against disease need not prevent circulation. The reasons for this appear complicated. For example, viruses can circulate by colonizing nasal passages without body-wide infection, and vaccines may be less effective at preventing such local colonization (Bleier et al., 2021). Some partially vaccinated patients in the Moderna trial appear to have experienced such localized infections (Creech, 2022).

## 2.2 Vaccine eligibility and school mitigation

Twelve to fifteen year olds became eligible for the Pfizer-BioNTech COVID-19 vaccine on May 12, 2021, just at the end of the 2020-2021 school year and just prior to the Delta wave. Indiana school entry rules imply that when the 2021-2022 school year began in August of 2021, the vast majority of sixth graders

[2] The smallpox vaccine does produce sterilizing immunity, important for its eradication, and so does the measles vaccine, important for the control of an extremely infectious disease.

were 11 years old, and therefore not yet vaccine eligible. Students entering seventh grade were predominantly 12 years old, and therefore vaccine eligible. In Indiana, most sixth graders attend schools either with exclusively younger students, or with exclusively older students. We refer to "six-and-down" schools as elementary school and "six-and-up" schools as middle schools, though there is some variation in grades served (Appendix Table A.1).

At the start of Fall 2021, sixth graders attending elementary schools were surrounded by younger peers who were not vaccine eligible, while those attending middle schools were surrounded by older peers who were eligible. As the the Fall progressed, some sixth graders turned 12 and became eligible themselves. Then on November 3, 2021, vaccine eligibility was extended to 5-11 year olds, making all sixth graders and theirt younger peers eligible. Our externalities research design compares differences in COVID-19 test rates between sixth graders in elementary and middle schools, in Fall 2021 and Fall 2020.

Peer vaccination likely matters most if schooling occurs in person with relatively little mitigation. Indiana as a state had a high level of in person schooling even in Fall of 2021. Mobility around schools was 68% of Fall 2019 levels during the Fall of 2020; 70% of schools were fully open, 6% hybrid, and 16% were fully online (COVID-19 School Data Hub, 2022; Halloran et al., 2021). By Fall of 2021, virtually all schools were open in person with much more limited hybrid options. Indiana maintained a state-wide school mask mandate through the 2020/2021 school year. During Fall 2021, school districts had discretion over implementing mask mandates, though schools with mandates were not required to quarantine asymptomatic close contacts. Neither mask mandates nor in-person learning is a meaningful confound in our context (Appendix E.)

## 2.3 Three research designs

**Own effects:** We identify the effects of vaccines on the vaccinated using date-of-birth based eligibility criteria. We select as the treatment group people born in the six month period ending on May 12, 2009, all of whom became eligible on May 12, 2021. The control group consists of people born in the sixth month period *beginning* November 3, 2009, roughly six months after the last treatment group birthday.

Everyone in the control group became eligible on November 3, 2021. We chose these ranges to obtain treatment and control groups that differ sharply in eligibility date but remain relatively similar in age at a given time; results are robust to other ranges. We adjust for cohort-specific differences (including age effects) using 2020 data, when neither group was eligible for the vaccine. We measure own effects using a difference-in-difference design where the post period begins in June, 2021, with the following regression:

$$y_{it} = \beta_0 + \beta_1 treat_i + \beta_4 post_t + \beta_3 treat_i \cdot post_t + \epsilon_{it}. \quad (1)$$

The outcome $y_{it}$ is an indicator for vaccination status, positive COVID status, or emergency department visit (a measure of adverse event) for person $i$ in month $t$. We implement the models using a balanced panel of person-month level data, and we cluster standard errors on the individual. We also estimate instrumental variables models for the effect of vaccination, instrumenting for vaccination using the interaction $treat_i \cdot post_t$. The second stage equation there is

$$y_{it} = \gamma_0 + \gamma_1 treat_i + \gamma_2 post_t + \gamma_3 \hat{vaccinated}_{it} + \nu_{it}. \quad (2)$$

$\gamma_3$ gives the percentage point effect of vaccination on outcome $y_{it}$ for compliers. In Appendix B we show how to translate this to an estimate of the more familiar vaccine effectiveness.

**Within-grade spillovers:** Our "own effects" design essentially compares seventh and sixth graders, who differ not only in their own vaccine eligibility but also in their grademates' eligibility. Thus in principle $\beta_3$ in Equation 1 reflects both own effects and grade-level spillovers. Put differently, the exclusion restriction in our IV analysis says that the only reason for a differential trend in infection among seventh graders, relative to sixth graders, is own vaccination status, but peer vaccination status could also affect own infection rates.

To address this potential violation of the exclusion restriction, and provide evidence on within-grade spillovers, we re-estimate Equation 1 and 2, but limiting to subsamples with low first stages (as in Angrist et al. (2010)). Specifically, we limit the sample to sixth and seventh graders whose *parents* are unvaccinated

(as of April 30, prior to child eligibility). Because the take-up rate is low for this group, any treatment effect must be due to spillovers.[3] Parental vaccination status is likely correlated with observable determinants of infection, such as risk taking. Our DID strategy attempts to control for this unobserved confounding by using sixth graders with unvaccinated parents as a control group.

**Cross-grade spillover effects:** Our design to estimate within-grade spillovers conditions on parental vaccination status, and hence identifies spillovers for a select group, who may differ in their overall cautiousness. We therefore also implement a complementary strategy that avoids this problem and more cleanly isolates exogenous variation in peer vaccination, conditional on own vaccination. Specifically, we focus on sixth graders, for whom we can isolate exogenous variation in the share of peers who are eligible for the vaccine while they themselves are ineligible for the vaccine. Specifically, we define treatment as belonging to an elementary school and control as belonging to a middle school. We exclude for example K-12 schools which serve older and younger grades. Our pre-period runs through March, 2021 (before children were vaccine eligible), and our post period begins in July, 2021. Because the treatment is school-wide, we cluster standard errors at the school level.

We estimate the reduced form effect of exposure to vaccine-eligible peers with the canonical difference-in-differences regression, Equation 1, using student-month level data. The outcomes are schoolmate vaccine rate, own (i.e. sixth grade) COVID-19 incidence, and own (i.e. sixth grade) vaccination status. The difference-in-difference model adjusts for time-invariant differences between middle and elementary schools, as well as for common trends in incidence and vaccination, arising, for example, because of waves in the pandemic and time-varying eligibility.

The regression estimates the effect of having vaccine-eligible school mates on COVID-19 incidence, a kind of intent-to-treat effect. We do not translate this into a treatment-on-the-treated "effect of vaccinated peers" because such effects are inherently non-linear in the peer vaccination rate; the marginal benefit of peer vaccination rises then falls (see Appendix D and Goodkin-Gold et al. (2020)). However, the reduced form provides a test of the null hypothesis of no effect of vaccinated peers, because if there are

[3] In practice the first stage is low but not zero, yet the reduced form is small and insignificant. This evidence is consistent with the exclusion restriction, because if it were violated, we would expect large infection effects even absent a large first stage.

infection-reducing spillovers from vaccinated peers, then we should estimate a negative effect of having more vaccine-eligible peers.

# 3 Data

## 3.1 Regenstrief Institute Databases

We obtained data from two databases maintained by Regenstrief Data Service (Regenstrief Institute, 2022). First, the Indiana Network for Patient Care (INPC) database consists of encounter and other medical records from over 100 health care entities, including hospitals, health care networks, and insurance providers. The INPC was established to improve health care at participating institutions. Second, the registry records nearly all COVID-19 lab (polymerase chain reaction) tests and vaccinations conducted in the state of Indiana. Regenstrief Institute obtained these data through a partnership with the state to develop a COVID-19 dashboard. Analysts at Regenstrief have generated patient identifiers to link these databases, and provided demographic and location (zip code and census tract) information for 75% of positive tests and 83% of patients in the encounter data.[4] We extract recrods on all patients in the database.

Our "full student sample" consists of roughly 990,000 students with non-missing date of birth, alive on July 1 2020, with implied ages putting students in kindergarten through 12th grade in the 2020-21 or 2021-22 school years. We impute grade levels assuming that students begin kindergarten in the year they are 5 on August 1, and advance one grade per year. Appendix Table A.3 reports the count of students by grade and school year. While the Regenstrief Institute databases are not designed to cover the full state population, the coverage here appears high: in the July 2020 Census report, there were 1,151,021 Indiana residents aged 5-17 (U.S. Census Bureau, 2022), thus our 2020-21 population corresponds to 86% of the school aged children in the state. Because most hospitals in the state participate in the INPC, most children born in the state appear in the data, regardless of insurance status or subsequent health care utilization.

[4]Some patients' location changes between 2020 and 2021. We use the earliest location, because it is more likely to reflect the school enrollment location, and to avoid conditioning on endogenous migration.

We construct separate analysis samples depending on the design. For the own effect and within-grade effects designs, we limit the sample to students born in the relevant date range, with high-quality school assignment (defined below), and non-missing encrypted address. For the spillover design, we limit the full student sample to observations with high-quality school assignment, in sixth grade, and assigned to a treatment or control school.

Appendix Table A.3 reports how our sample size changes as we impose our restrictions. The most restrictive condition (beyond the age limitation) is the requirement of high-quality school assignment, which cuts the sample by about half. Robustness tests show our results are not sensitive to the exact sample restrictions.

## 3.2 Assigning students to schools

Our spillover research design requires that we link students to their school. To do so we use the geography reported in the Regenstrief data (zip code and census tract), overlaid on school catchment area maps (National Center for Education Statistics (2022)). Because these two maps do not perfectly align, we assign observations to the sixth-grade-serving school whose catchment area covers the greatest share of the land mass of the student's geography; we refer to this school as their "modal school." Some points are covered by multiple school catchment areas, either because a given geography lies in multiple school districts, or because a school district allows students to choose among multiple schools. Appendix C provides more details on this process. We use the NCES data to classify students' modal school as "treatment" or "control," where treatment schools are six-and-up schools, and control schools are six-and down. We say an assignment is high-quality if the assigned school's catchment area covers at least 70 percent of the student's geography, accounting for double coverage.

Because geography is an imperfect predictor of school assignment, our approach introduces measurement error in the assignment of students to schools. However, our reduced form estimates are unbiased as long as assignment to *treatment* is correct—that is, correct assignment to six-and-up or six-and-down school type—even if we incorrectly assign students to particular schools. We show in Appendix C that

measurement error is likely small after limiting to high-quality school matches; errors in school assignment and school vaccination rates appear rare, and errors in treatment status appear very rare.

## 3.3 Derived measures

We define our main outcomes at the student-month level: indicators for any lab-confirmed COVID-19 case, emergency room visit, or inpatient admission, as well as *cumulative* vaccination status, i.e. an indicator for having received at least one or at least two vaccine doses by the end of a given month.[5] While our lab-confirmed case measure misses cases with rapid tests but not PCR tests, such cases are relatively rare during our sample period, when rapid tests were in short supply (Leonhardt, 2021). Our IV models focus on the effect of the second dose, but nearly everyone who receives a first dose also receives a second.

In some analyses we look at school-level or school-grade-level characteristics, such as the school level vaccination rate among all students (not just sixth graders). We construct such school-level characteristics by averaging over all students assigned to a given school or school-grade meeting the inclusion criteria described below.

Our "within grade spillovers" design requires that we condition on having unvaccinated parents. We do not observe family identifiers, but we do observe (encrypted) addresses, the address a patient has on file with a given health care provider. We treat each address as a household, and for each household we measure the share of adults (aged 26-64) at that household who are vaccinated as of April 30, 2021 (just before 12-16 year-olds became eligible). Because a patient may have multiple addresses, we define the parental vaccination rate of students in our sample as the average adult vaccination rate across all households they live in. This household imputation procedure appears to work well. The distribution of adult genders seems reasonable (Appendix Table A.2), and adult vaccination is highly predictive of child vaccination (Appendix Figure A.1).

[5] We treat the small number of students receiving the one-dose Johnson and Johnson vaccine as receiving two doses.

# 4 Results

## 4.1 Vaccines protect the vaccinated

We begin by establishing the effectiveness of the COVID-19 vaccines in our setting. We use the "own effects" research design, comparing students just old enough to become eligible for COVID-19 vaccines on May 12, 2021, to students just young enough that they are ineligible until November 3, 2021. Our key results are evident in the raw time series of vaccination rates, COVID-19 incidence, and emergency room visits for treatment and control, plotted in Figure 1.

The figure shows that vaccine eligibility increases vaccinations, reduces COVID-19 incidence, and has no discernible effect on emergency room visits. Starting from the top panel we see that vaccine take-up grows steadily for the treatment group when they become eligible, with of course no vaccination in the control group until their eligibility date six months later. The middle panel shows that, in the pre-period, the treatment and control groups had essentially equal COVID-19 incidence, suggesting little or no confounding. After becoming vaccine eligible, the treatment group diverges from the control group; the vaccine-ineligible students experienced lower COVID-19 in each of the last four months of 2021.[6] The final panel shows that vaccine eligibility has no effect on adverse events, measured here as emergency room visits. Treatment and control show nearly identical levels and trends throughout the sample period, with no divergence after vaccine eligibility. The null effect on ER visits is not driven by offsetting decreases in COVID ER visits and increases in other ER visits (Appendix Table A.4).

We report DID and IV estimates of the effect of vaccine eligibility and vaccination in Panel A of Table 1. Vaccine eligibility increases the vaccination rate by about 25 percentage points, with nearly identical impacts on first and second doses. Since most age-eligible students have age-eligible grade-mates, the impact on peer vaccination rates is quite similar. Early eligibility reduces COVID-19 incidence by 0.4 percentage points. Earlier eligibility has no discernible effect on emergency room visits; the confidence intervals

[6] This divergence does not occur until September, 2021, four months after initial vaccine eligibility. Such divergence is unsurprising: vaccine take-up grew over time, vaccines take time to generate an immune response, and COVID-19 prevalence was fairly low in May-July of 2021, but grew dramatically in August and September as the Delta variant circulated and school resumed.

rule out effects larger than about +0.1 percentage points. Our instrumental variables estimates indicate that vaccination itself reduces COVID-19 incidence by about 1.5 percentage point, for a complier vaccine effectiveness of about 80 percent. (See Appendix B for details on vaccine effectiveness.) This estimate is roughly comparable to, but somewhat smaller, than the 95% effectiveness reported in the clinical trials for the mRNA vaccines (Baden et al., 2020; Polack et al., 2020). These results are robust to alternative sample inclusion criteria (Appendix Table A.5).

Thus, relative to sixth graders, seventh graders experienced an increase in vaccinations and a decrease in COVID upon attaining vaccine eligibility. In principle our reduced form estimates reflect the combination of own and within-grade spillover effects. However we show in Panel B that spillover effects are likely small. Specifically in Panel B we limit the sample to students in households with unvaccinated adults. The first stage falls to about 10 percentage points, and the DID estimate falls to a small and statistically insignificant -0.05 percentage points. There is no treatment effect among the students with no first-stage, despite similar sized peer vaccination take-up. This result suggests limited within-grade spillovers, and validates our exclusion restriction. As further evidence for these conclusions, we see in panels C and D that both the first stage and reduced form double when we focus on students in households with partially vaccinated adults or fully vaccinated adults, while the peer vaccination rates do not change substantially across these samples. Thus the differential fall in infection among seventh graders seems driven by own vaccine take-up rather than within-grade spillovers.

## 4.2 Vaccines do not protect the unvaccinated

While early vaccine eligibility reduces COVID-19 incidence for the eligible, this reduced incidence does not spillover to the mostly unvaccinated sixth graders, as we now show with our "spillover effects" design. Our key results are again evident in the simple trends, which we plot in Figure 2, for sixth graders in six-and-up schools (the treatment group) and in six-and-down schools (the control group).

The figure shows that treatment sixth graders experience a large increase in the vaccination rate of their schoolmates, relative to control sixth graders, but no differential decrease in COVID-19 incidence.

Vaccination rates are zero for both groups until May, 2021, when they diverge sharply. By Fall 2021, treatment group sixth graders go to schools in which about one in five students are vaccinated, while for control group the number is closer to one in 100.[7] Turning to COVID-19 incidence, we see near identical levels and trends in the pre-period for treatment and control, suggesting little if any confounding. Incidence increases in fall 2021 for both groups. However, despite the large relative increase in schoolmate vaccination for the treatment group, we see no relative decrease in COVID-19 incidence during this period. By the late fall, sixth graders in treatment and control alike become vaccine eligible themselves, and we see some evidence of higher vaccine take-up among the treated sixth graders.

We report difference-in-difference estimates and standard errors, by time period, in Table 2. Going to a six-and-up school induces a 20 percentage point increase in school-wide vaccination rate; this difference persists into the late fall, when sixth graders become vaccine eligible, because take-up is fairly low, and take-up continues to grow among the older students. Despite the increase in schoolmate vaccination, we find no protective effect of six-and-up school attendance on COVID-19 incidence among sixth graders. In the early fall, when few sixth graders were vaccine eligible, we estimate a positive and significant treatment effect of 0.36 percentage points. Later in the fall the effect falls to 0.10 percentage points. The lower bounds of these confidence intervals rule out effects more negative than about -0.4 percentage points. While heightened risk taking could in principle offset the protective effects of schoolmate vaccination, we find little evidence for such Peltzman effects: we estimate small, insignificant, but positive effects on own vaccination rates, i.e. reduced risk taking in response to peer vaccination.

Our small and insignificant estimates of infection-reducing externalities are surprisingly given the effectiveness of the vaccines and prior work on vaccine spillovers. Indeed this prior work has found that similar-sized increases in vaccination rates have had large health spillovers across more distant peer groups, relative to our context. Ward (2014) shows that a universal flu vaccination campaign in Ontario increased take-up of the non-elderly by about 11 percentage points and reduced flu hospitalizations among the elderly, with no effect on non-elderly hospitalizations; similarly, White (2021) finds that increased non-

[7] Peer vaccination rates in the control group increase and then decrease steeply in May-July 2021, because sixth graders were vaccine eligible in the spring of 2021. This pattern does not influence our estimates because we focus on fall-on-fall differences.

elderly flu vaccinations in the US reduce elderly mortality. Carpenter and Lawler (2019) find that TDap mandate for middle schoolers increased their take-up by 13.5 percentage points, and reduced pertussis incidence by similar amounts among 0-4 year-olds as among middle-school aged children. We also show theoretically in Appendix D in a calibrated SIR model that even a 20 percentage point increase in vaccination would generate substantial spillovers, if the vaccine completely prevents transmission and the infectiousness is not too high.

However, if infectiousness is very high, or the vaccine only partially prevents transmission, then it is possible that higher peer vaccination rates would indeed generate substantial externalities, even with no externality in our sample. Our results therefore imply, at minimum, that achieving external benefits of vaccination requires high vaccination rates.

Heterogeneity analysis suggests that external benefits are not detectable even at peer take-up rates of 40 percent. To show this, we re-estimate our difference-in-difference models, separately by peer vaccination rate. Specifically, we divide our treatment group into quartiles of school-wide vaccination rate (as of October, 2021). As schools with high vaccination rates may differ in potential COVID incidence, we use a DID-matching strategy to find comparable control schools. In particular we use the propensity score to match treated students to control students with similar census-tract-level vaccination rates.[8] This matching strategy addresses the main source of confounding that conditioning on high-vaccination rate schools may condition on high-COVID-prevention in general.[9]

We show the results in Table 3. Each column corresponds to a different quartile of school-wide vaccination rate. At the top quartile, treatment raises peer vaccination rates to 37 percent, from 10 percent. Even at this high take-up rate, we find a small and insignificant impact on incidence. These quartile-specific contrasts appear unconfounded (after propensity score matching), as we find precisely zero difference between treatment and control in the pre-period, consistent with little or no confounding after

[8] We model the propensity score as a logit in census-tract level vaccination rates. We start with a (logit-)linear model, and add higher degree terms until Cohen's D for the weighted control group is less than 0.05. We limit the sample to students in schools with at least 30 observations, to avoid contaminating the extreme strata with small sample sizes, and for simplicity we focus on the November-December period only.

[9] We obtain highly similar results using the alternative approach of simply stratifying on tract-level vaccination rates; see Appendix Table A.6.

matching. In the final panel we confirm our finding that the limited spillovers are not driven by risk compensation in the form of vaccine take-up; we estimate small and insignificant effects of peer vaccination on own vaccination.

Our finding of no spillovers from vaccinated schoolmates is robust to alternative specification choices and shows up in different subperiods, as we show in Appendix Tables A.7-A.8. Broadening or tightening the school coverage area (as low as 60% or as high 99%) does not substantially change our estimates, nor does limiting to students with prior medical claims (for whom coverage is better), nor does including students in "mixed type" areas, or excluding sixth graders who became vaccine eligible prior to November 3. Our results are also robust to adjusting for school mask mandates or in-person learning (Appendix E.)

# 5 Conclusion

We have shown using data on Indiana 11 and 12 year-olds that the COVID-19 vaccines produces strong protection for the vaccinated, with no detectable adverse events, but do little to protect other schoolmates from COVID-19. We emphasize three limitations of our spillover result. First, we lack data on social interactions. It is possible that sixth and seventh graders interact little, and in general that vaccinated and unvaccinated students in our sample interact less than do other groups such as family members or coworkers. Second, COVID-19 does not typically produce severe illness in children; it is possible spillovers would be different for more vulnerable populations. Third, our identifying variation increases vaccination rates to, at most, 37 percent. It is possible that spillovers would be larger at higher levels of vaccination.

Despite these limitations, our results imply that the external benefits of the COVID-19 vaccine are weaker than previously believed. At observed vaccination rates, we see no evidence of within or across grade spillovers. This finding has several implications. First, for students in our sample, relying on others' vaccination provided little protection; our results underscore that vaccination is a safe and effective means of protecting against COVID-19. Second, masking and physical distancing remain valuable even among vaccinated individuals, who still appear to carry a risk of transmission. Finally, if vaccination mandates provide infection-reducing externalities—reducing infections and illness among those not subject to the

mandate—they require high levels of compliance, above the vaccination rates observed in our data.

While our results suggest that infection-reducing externalities do not justify a COVID-19 mandate in schools, other justifications remain. One such justification is paternalism: the vaccines are safe and effective, vaccination rates are far from 100 percent, there are few if any severe side effects, and vaccine mandates increase vaccination (Abrevaya and Mulligan, 2011; Carpenter and Lawler, 2019), although so do recommendations (Lawler, 2017, 2020). Vaccine mandates (and other interventions in the vaccine market) can also be justified by externalities other than infection reduction. One key possibility is that vaccinations reduce the strain on a health care system. Another is that, by reducing individual risk, they increase economic activity, allowing the economy to function; this benefit of vaccines plays an important role in the very high value of vaccine capacity estimated by Castillo et al. (2021). Further research on these external benefits would be valuable.

# References


Abaluck, J., L. H. Kwong, A. Styczynski, A. Haque, M. A. Kabir, E. Bates-Jefferys, E. Crawford, J. Benjamin-Chung, S. Raihan, S. Rahman, et al. (2021). Impact of community masking on covid-19: A cluster-randomized trial in bangladesh. *Science 375*(6577), eabi9069.

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

Acton, R. K., W. Cao, E. E. Cook, S. A. Imberman, and M. F. Lovenheim (2022). The effect of vaccine mandates on disease spread: Evidence from college covid-19 mandates. Technical report, National Bureau of Economic Research.

Alexander, D. and E. Karger (2021). Do stay-at-home orders cause people to stay at home? effects of stay-at-home orders on consumer behavior. *The Review of Economics and Statistics*, 1–25.

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics 130*(3), 1117–1165.

Andrews, N., J. Stowe, F. Kirsebom, S. Toffa, T. Rickeard, E. Gallagher, C. Gower, M. Kall, N. Groves, A.-M. O'Connell, et al. (2022). Covid-19 vaccine effectiveness against the omicron (b. 1.1. 529) variant. *New England Journal of Medicine 386*(16), 1532–1546.

Angrist, J., V. Lavy, and A. Schlosser (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics 28*(4), 773–824.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association 91*(434), 444–455.

Baden, L. R., H. M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech, et al. (2020). Efficacy and safety of the mrna-1273 sars-cov-2 vaccine. *New England journal of medicine*.

Baker, J. M., R. M. Dahl, J. Cubilo, U. D. Parashar, and B. A. Lopman (2019). Effects of the rotavirus vaccine program across age groups in the united states: analysis of national claims data, 2001–2016. *BMC infectious diseases 19*(1), 1–11.

Bernal, J. L., N. Andrews, C. Gower, E. Gallagher, R. Simmons, S. Thelwall, J. Stowe, E. Tessier, N. Groves, G. Dabrera, et al. (2021). Effectiveness of covid-19 vaccines against the b. 1.617. 2 (delta) variant. *New England Journal of Medicine*.

Berry, C. R., A. Fowler, T. Glazer, S. Handel-Meyer, and A. MacMillen (2021). Evaluating the effects of shelter-in-place policies during the covid-19 pandemic. *Proceedings of the National Academy of Sciences 118*(15), e2019706118.

Bleier, B. S., M. Ramanathan Jr, and A. P. Lane (2021). Covid-19 vaccines may not prevent nasal sars-cov-2 infection and asymptomatic transmission. *Otolaryngology–Head and Neck Surgery 164*(2), 305–307.

Bronchetti, E. T., D. B. Huffman, and E. Magenheim (2015). Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination. *Journal of Economic Behavior & Organization 116*, 270–291.

Bugenske, E., S. Stokley, A. Kennedy, and C. Dorell (2012). Middle school vaccination requirements and adolescent vaccination coverage. *Pediatrics 129*(6), 1056–1063.

Caddy, S. L. (2021). Coronavirus: Few vaccines prevent infeciton - here's why that's not a problem. *The Conversation*. January 5, 2021. Available at `https://theconversation.com/coronavirus-few-vaccines-prevent-infection-heres-why-thats-not-a-problem-152204`.

Carpenter, C. S. and E. C. Lawler (2019). Direct and spillover effects of middle school vaccination requirements. *American Economic Journal: Economic Policy 11*(1), 95–125.

Castillo, J. C., A. Ahuja, S. Athey, A. Baker, E. Budish, T. Chipty, R. Glennerster, S. D. Kominers, M. Kremer, G. Larson, et al. (2021). Market design to accelerate covid-19 vaccine supply. *Science 371*(6534), 1107–1109.

Chan, T. Y., B. H. Hamilton, and N. W. Papageorge (2016). Health, risky behaviour and the value of medical innovation for infectious disease. *The Review of Economic Studies 83*(4), 1465–1510.

Chetty, R., J. N. Friedman, N. Hendren, M. Stepner, and T. O. I. Team (2020). *How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data*, Volume 27431. National Bureau of Economic Research Cambridge, MA.

COVID-19 School Data Hub (2022). CSDH Data. `https://www.covidschooldatahub.com`.

Creech, C. B. (2022). Phone call with Daniel W. Sacks, August 5, 2022.

Cronin, C. J. and W. N. Evans (2021). Total shutdowns, targeted restrictions, or individual responsibility: How to promote social distancing in the covid-19 era? *Journal of Health Economics 79*, 102497.

Dave, D., A. I. Friedson, K. Matsuzawa, and J. J. Sabia (2021). When do shelter-in-place orders fight covid-19 best? policy heterogeneity across states and adoption time. *Economic inquiry 59*(1), 29–52.

Friedson, A. I., D. McNichols, J. J. Sabia, and D. Dave (2021). Shelter-in-place orders and public health: evidence from california during the covid-19 pandemic. *Journal of Policy Analysis and Management 40*(1), 258–283.

Ginther, D. K. and C. Zambrana (2021). Association of mask mandates and covid-19 case rates, hospitalizations, and deaths in kansas. *JAMA Network Open 4*(6), e2114514–e2114514.

Goldberg, Y., M. Mandel, Y. M. Bar-On, O. Bodenheimer, L. S. Freedman, E. Haas, R. Milo, S. Alroy-Preis, N. Ash, and A. Huppert (2021). Waning immunity of the bnt162b2 vaccine: A nationwide study from israel. *MedRxiv*.

Goodkin-Gold, M., M. Kremer, C. M. Snyder, and H. L. Williams (2020). Optimal vaccine subsidies for endemic and epidemic diseases. Technical report, National Bureau of Economic Research.

Goolsbee, A. and C. Syverson (2021). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of public economics 193*, 104311.

Gruber, J. (2005). *Public finance and public policy*. Macmillan.

Gupta, S., T. Nguyen, S. Raman, B. Lee, F. Lozano-Rojas, A. Bento, K. Simon, and C. Wing (2021). Tracking public and private responses to the covid-19 epidemic: evidence from state and local government actions. *American Journal of Health Economics 7*(4), 361–404.

Hall, W. D. (1999). Representation of blacks, women, and the very elderly (aged $\geq$ 80) in 28 major randomized clinical trials. *Ethnicity & disease 9*(3), 333–340.

Halloran, C., R. Jack, J. C. Okun, and E. Oster (2021). Pandemic Schooling Mode and Student Test Scores: Evidence from US States. Working Paper 29497, National Bureau of Economic Research.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Indianapolis Public Schools (2022). Enrollment in ips - indianapolis public schools. `https://myips.org/enrollment-options/enrolling-in-ips/`, accessed June 1, 2022.

Kong, E. and D. Prinz (2020). Disentangling policy effects using proxy data: Which shutdown policies affected unemployment during the covid-19 pandemic? *Journal of Public Economics 189*, 104257.

Lawler, E. C. (2017). Effectiveness of vaccination recommendations versus mandates: Evidence from the hepatitis a vaccine. *Journal of health economics 52*, 45–62.

Lawler, E. C. (2020). Giving teens a boost? effects of adolescent meningococcal vaccine recommendations. *American Journal of Health Economics 6*(2), 251–287.

Leonhardt, D. (2021). Where are the tests? *New York Times*. The Morning Newsletter. September, 21, 2021.

Malani, A. (2008). Patient enrollment in medical trials: Selection bias in a randomized experiment. *Journal of Econometrics 144*(2), 341–351.

National Center for Education Statistics (2022). ELSI - elementary/secondary information system. `https://nces.ed.gov/ccd/elsi/`.

Oster, E. (2018). Does disease cause vaccination? disease outbreaks and vaccination response. *Journal of health economics 57*, 90–101.

Peltzman, S. (1975). The effects of automobile safety regulation. *Journal of political Economy 83*(4), 677–725.

Polack, F. P., S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. P. Marc, E. D. Moreira, C. Zerbini, et al. (2020). Safety and efficacy of the bnt162b2 mrna covid-19 vaccine. *New England Journal of Medicine*.

Regenstrief Institute (2022). RDS Data. https://www.regenstrief.org/rds/data/.

Sadoff, J., G. Gray, A. Vandebosch, V. Cárdenas, G. Shukarev, B. Grinsztejn, P. A. Goepfert, C. Truyers, H. Fennema, B. Spiessens, et al. (2021). Safety and efficacy of single-dose ad26. cov2. s vaccine against covid-19. *New England Journal of Medicine 384*(23), 2187–2201.

Schaller, J., L. Schulkind, and T. Shapiro (2019). Disease outbreaks, healthcare utilization, and on-time immunization in the first year of life. *Journal of health economics 67*, 102212.

U.S. Census Bureau (2022). Sc-est2020-agesex-civ: Annual estimates of the civilian population by single year of age and sex for the united states, states, and the district of columbia: April 1, 2010 to july 1, 2020. https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-state-detail.html, accessed June 1, 2022.

Waldron, D. (2022a). Email correspondence with Kosali Simon, August 1, 2022.

Waldron, D. (2022b). Covid-19 indiana schools. https://github.com/dawaldron/covid-19-indiana-schools, last accessed July 19, 2022.

Ward, C. J. (2014). Influenza vaccination campaigns: is an ounce of prevention worth a pound of cure? *American Economic Journal: Applied Economics 6*(1), 38–72.

Werner, J. M., A. Abdalla, N. Gara, M. G. Ghany, and B. Rehermann (2013). The hepatitis b vaccine protects re-exposed health care workers, but does not provide sterilizing immunity. *Gastroenterology 145*(5), 1026–1034.

White, C. (2021). Measuring social and externality benefits of influenza vaccination. *Journal of Human Resources 56*(3), 749–785.

Table 1: Effects of vaccine eligibility and take-up on illness

| Dep. var. | 1+ dose | 2+ doses | Peer 1+ dose | Peer 2+ dose | Any COVID | Any ER |
|---|---|---|---|---|---|---|
| A. All students | | | | | | |
| DID estimate | 0.2722 | 0.2645 | 0.2283 | 0.2256 | -0.0040 | -0.0008 |
| | (0.0041) | (0.0039) | (0.0012) | (0.0011) | (0.0009) | (0.0010) |
| IV estimate | | | | | -0.0153 | -0.0031 |
| | | | | | (0.0034) | (0.0039) |
| Vaccine effectiveness | | | | | 0.792 | 0.021 |
| | | | | | (0.044) | (0.181) |
| # Students | | | | | | 24,648 |
| B. Students in households with unvaccinated adults | | | | | | |
| DID estimate | 0.1212 | 0.1085 | 0.2058 | 0.2006 | -0.0005 | -0.0010 |
| | (0.0050) | (0.0045) | (0.0018) | (0.0017) | (0.0015) | (0.0016) |
| IV estimate | | | | | -0.0049 | -0.0089 |
| | | | | | (0.0141) | (0.0149) |
| Vaccine effectiveness | | | | | -0.044 | -0.846 |
| | | | | | (3.253) | (3.820) |
| # Students | | | | | | 8,519 |
| C. Students in households with partially vaccinated adults | | | | | | |
| DID estimate | 0.2899 | 0.2818 | 0.2358 | 0.2336 | -0.0042 | 0.0008 |
| | (0.0063) | (0.0060) | (0.0018) | (0.0017) | (0.0014) | (0.0017) |
| IV estimate | | | | | -0.0149 | 0.0028 |
| | | | | | (0.0048) | (0.0059) |
| Vaccine effectiveness | | | | | 0.853 | -0.008 |
| | | | | | (0.054) | (0.219) |
| # Students | | | | | | 10,941 |
| D. Students in households with fully vaccinated adults | | | | | | |
| DID estimate | 0.6228 | 0.6305 | 0.2725 | 0.2760 | -0.0117 | -0.0065 |
| | (0.0117) | (0.0115) | (0.0038) | (0.0036) | (0.0024) | (0.0025) |
| IV estimate | | | | | -0.0186 | -0.0104 |
| | | | | | (0.0039) | (0.0040) |
| Vaccine effectiveness | | | | | 0.779 | 0.296 |
| | | | | | (0.051) | (0.162) |
| # Students | | | | | | 3,345 |

Notes: Table reports DID estimates for each outcome, and DID-IV estimates and vaccine effectiveness for effect of two vaccine doses on monthly COVID incidence and ER visits. Treatment group is born in the six months before May 12, 2009; control group is born in the six months after November 3, 2009. The post period is Fall 2021, pre-period is Fall 2020. See Appendix B for details on vaccine effectiveness. In Panel A the sample is the full own-effect sample. Panels B, C, and, D are restricted to 6th and 7th graders living in households where no adults are vaccinated, some but not all adults vaccinated, and all adults are vaccinated (as of April 30, 2020).

Table 2: Vaccine-eligible schoolmates do not reduce COVID incidence among vaccine-ineligible

| Period | August-October | November-December | August-December |
|---|---|---|---|
| Effect of vaccine-eligible school mates on... | | | |
| School-wide vaccination rate | 0.1864 | 0.2055 | 0.1941 |
| | (0.0106) | (0.0120) | (0.0111) |
| Sixth grade COVID-19 incidence | 0.0035 | -0.0002 | 0.0020 |
| | (0.0016) | (0.0018) | (0.0011) |
| Sixth grade vaccination rate | 0.0038 | 0.0164 | 0.0088 |
| | (0.0037) | (0.0104) | (0.0062) |

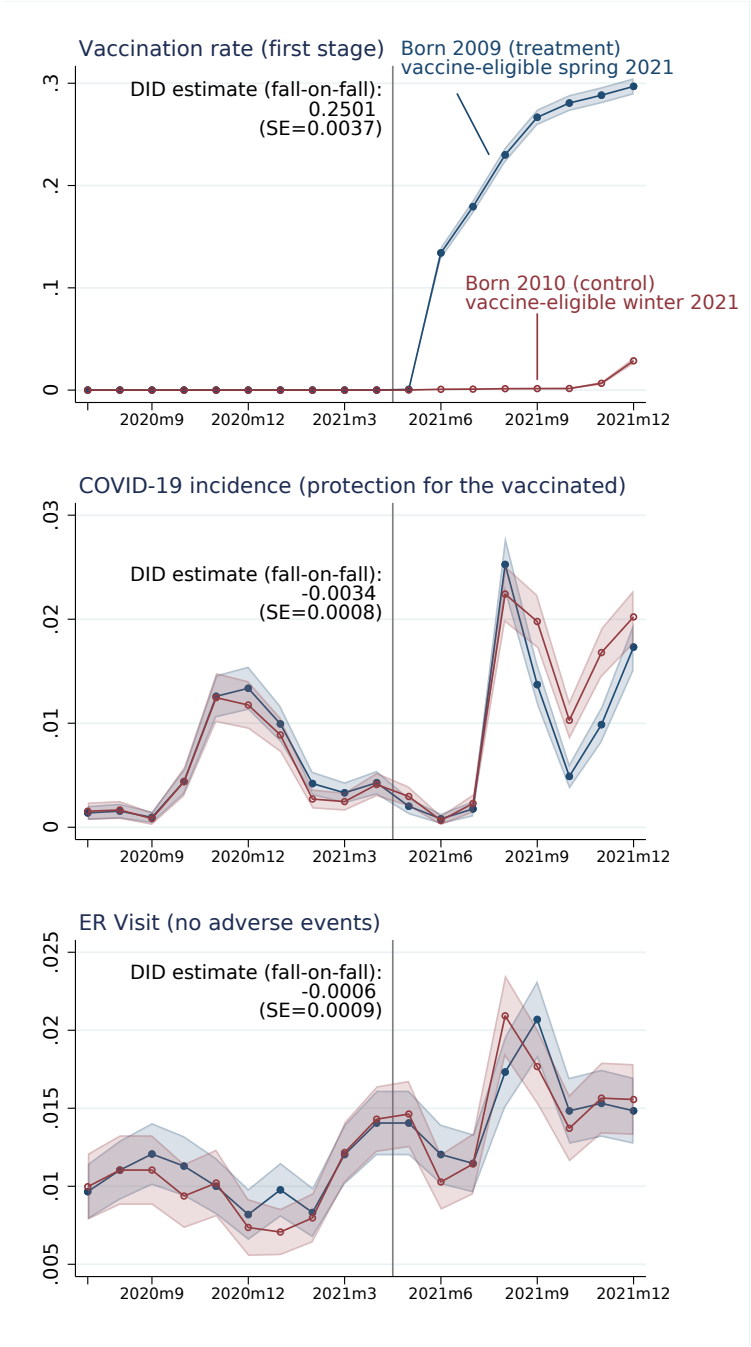
Notes: Each cell reports difference-in-differences estimate of the effect of vaccine eligible schoolmates on the indicated outcome. The sample consists of monthly observations of sixth grade students with reliable school assignment, in the indicated months. The pre-period is 2020 and the post-period 2021. The treatment group is students who go to school with older grades (who are vaccine eligible in May 2021 and later the post-period) and the control group students who go to school with younger students (who are not vaccine-eligible until November 2021). Robust standard errors, clustered on school, in parentheses.

Table 3: Effect of vaccine eligible schoolmates, by school-wide vaccination rate

| School vaccine take-up quartile | First | Second | Third | Fourth |
|---|---|---|---|---|
| Y = school vaccination rate | | | | |
| DID estimate | 0.1265 | 0.1702 | 0.2217 | 0.2632 |
| | (0.0058) | (0.0088) | (0.0124) | (0.0194) |
| Control mean, post | 0.0336 | 0.0478 | 0.0723 | 0.0942 |
| | (0.0033) | (0.0056) | (0.0090) | (0.0109) |
| Y = covid Incidence | | | | |
| DID estimate | 0.0009 | -0.0010 | -0.0016 | -0.0016 |
| | (0.0022) | (0.0020) | (0.0025) | (0.0032) |
| Pre-period difference | 0.0001 | 0.0011 | 0.0017 | 0.0011 |
| | (0.0013) | (0.0012) | (0.0014) | (0.0018) |
| Control mean, post | 0.0159 | 0.0163 | 0.0178 | 0.0184 |
| | (0.0011) | (0.0011) | (0.0015) | (0.0026) |
| Y = own vaccination rate | | | | |
| DID estimate | -0.0052 | 0.0040 | 0.0089 | 0.0007 |
| | (0.0050) | (0.0067) | (0.0106) | (0.0178) |
| Control mean, post | 0.0623 | 0.0827 | 0.1172 | 0.1557 |
| | (0.0035) | (0.0042) | (0.0089) | (0.0147) |
| Cohen's D | 0.0064 | 0.0257 | 0.0106 | 0.0299 |

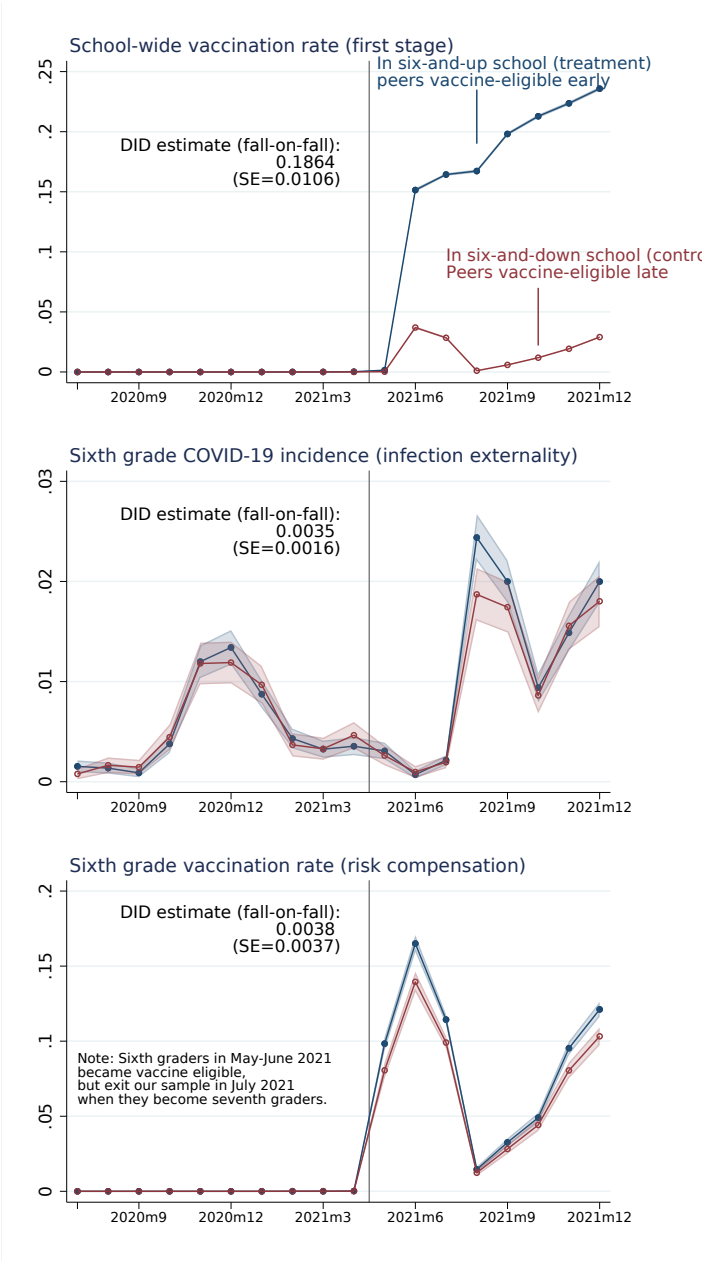
Notes: Table reports propensity-score weighted difference-in-differences estimates of the effect of vaccine eligible schoolmates on the indicated outcome, for each quartile of school-wide vaccination rates. We also report the outcome mean in the post period for the control group, and where indicated, the pre-period difference between treatment and control The sample consists of monthly observations of sixth grade students with reliable school assignment, in November-December. The pre-period is 2020 and the post-period 2021. The treatment group is sixth graders who go to school with older grades (who are vaccine eligible in May 2021 and later the post-period) and the control group is sixth graders who go to school with younger students (who are not vaccine-eligible until November 2021). Within quartile, we use the propensity score weighting to match treatment and control students, matching on on census-tract wide vaccination rates. Robust standard errors, clustered on school, in parentheses.

Figure 1: Vaccines reduce COVID-19 incidence among the vaccinated, with no adverse events




Notes: Figure plots means of the indicated variables, for Indiana residents born with the indicated birth dates, drawn from Regenstrief institute data on Indiana COVID-19 vaccinations, COVID-19 testing, and emergency room visits. Shaded area shows 95% confidence intervals, derived from robust standard errors clustered on individuals. The vertical line is the date the treatment group became vaccine-eligible. The DID estimate compares August-October 2020 and August-October 2021.

Figure 2: Peer vaccinations do not reduce own COVID-19 incidence




Notes: Figure plots means of the indicated variables, for Indiana sixth graders in the indicated school type. Sample is limited to students for whom we can reliably impute public school assignment. "Six-and-up" schools are schools where the youngest grade is six, and "six-and-down" school are where the oldest grade is six. Shaded area shows 95% confidence intervals, derived from standard errors clustered on school. The vertical line is the earliest date children were vaccine-eligible.

# A Appendix Exhibits

Table A.1: Distribution of schools in analysis sample, by grades served

| Grades served | Count of students | Share of students |
|---|---|---|
| K-6 | 10,312 | 0.191 |
| 1-6 | 500 | 0.009 |
| 2-6 | 124 | 0.002 |
| 3-6 | 1,022 | 0.019 |
| 4-6 | 863 | 0.016 |
| 5-6 | 5,452 | 0.101 |
| 6-6 | 2,270 | 0.042 |
| 6-8 | 32,862 | 0.608 |
| 6-9 | 131 | 0.002 |
| 6-12 | 540 | 0.010 |

Notes: Table report the count and share of students in our spillover analysis sample assigned to a school with the indicated grade range.

Table A.2: Joint frequency distribution of adult genders, within address

| | # female 25-64 year-olds | | | |
|---|---|---|---|---|
| | 0 | 1 | 2 | $\geq 3$ |
| # male 25-64 year-olds | | | | |
| 0 | 0.104 | 0.136 | 0.045 | 0.019 |
| 1 | 0.041 | 0.244 | 0.102 | 0.049 |
| 2 | 0.007 | 0.045 | 0.053 | 0.044 |
| 3+ | 0.001 | 0.010 | 0.016 | 0.085 |

Notes: Each cell is a fraction (0-1); they sum to 100 percent across all 16 cells. Table reports the fraction of sixth and seventh graders living with the indicated number of male and female adults, among sixth and seventh graders in 2020 or 2021, with non-missing address information. We average over all addresses at which the student lives, and round to the nearest integer. The bottom right cell, for example, means that 11 percent live at addresses where three or more adult females and three or more adult males also live.

Table A.3: Creating the analysis samples

| Restriction imposed | # Students |
|---|---|
| A. Own effects and within-grade sample | |
| Full student sample | 991,323 |
| ...& Born 2008-11-13 to 2009-5-12, or 2009-11-3 to 2010-5-2 | 64,422 |
| ...& Has high-quality school match | 31,150 |
| ...& Has encrypted address | 24,648 |
| B. Cross-grade spillover sample | |
| Full student sample | 991,323 |
| ...& Sixth grade | 132,885 |
| ...& High-quality school match | 61,323 |
| ...& Six-and-up or six-and-down school | 54,076 |

Notes: Table reports the count of students as we impose successive criteria to create our analysis samples. "High quality match" means the assigned school's catchment area catchment area covers at least 70 percent of the students geography.

Table A.4: Effect of early vaccine eligibility on ER visits with and without COVID

| Type of ER visit | All | With positive test | With negative test | With no test |
|---|---|---|---|---|
| DID estimate | -0.0008 | -0.0005 | -0.0009 | 0.0006 |
| | (0.0010) | (0.0002) | (0.0005) | (0.0009) |
| IV estimate | -0.0030 | -0.0017 | -0.0034 | 0.0021 |
| | (0.0039) | (0.0008) | (0.0019) | (0.0032) |
| Vaccine effectiveness | 0.021 | 0.763 | 0.327 | -0.343 |
| | (0.181) | (0.129) | (0.176) | (0.370) |

Notes: Table shows the effect of vaccine eligibility (DID) and vaccine take-up (IV) on all ER visits, ER visits with positive COVID test in surrounding days, and ER visits with negative (and no positive) test in surrounding days, and ER visits with no COVID test in surrounding days. Surrounding days are 5 days before to four days after the ER visit. The sample and specification are defined in the notes to Table 1.

Table A.5: Robustness of estimated effects of vaccines on the vaccinated to alternative samples

| | First stage | | | | Reduced form | |
|---|---|---|---|---|---|---|
| Dep. var. | 1+ dose | 2+ doses | Peer 1+ dose | Peer 2+ dose | COVID | ER |
| A. Baseline sample (born within 6 months of cutoff) | | | | | | |
| DID estimate | 0.2722 | 0.2645 | 0.2283 | 0.2256 | -0.0040 | -0.0008 |
| | (0.0041) | (0.0039) | (0.0012) | (0.0011) | (0.0009) | (0.0010) |
| IV estimate | | | | | -0.0153 | -0.0031 |
| | | | | | (0.0034) | (0.0039) |
| Vaccine effectiveness | | | | | 0.792 | 0.021 |
| | | | | | (0.044) | (0.181) |
| # Students | | | | | | 24,648 |
| B. Limit to born within 3 months of cutoff | | | | | | |
| DID estimate | 0.2689 | 0.2665 | 0.2307 | 0.2280 | -0.0047 | -0.0019 |
| | (0.0059) | (0.0056) | (0.0017) | (0.0016) | (0.0013) | (0.0014) |
| IV estimate | | | | | -0.0176 | -0.0070 |
| | | | | | (0.0047) | (0.0053) |
| Vaccine effectiveness | | | | | 0.864 | 0.313 |
| | | | | | (0.044) | (0.139) |
| # Students | | | | | | 12,425 |
| C. Expand to born within 12 months of cutoff | | | | | | |
| DID estimate | 0.1701 | 0.1637 | 0.1431 | 0.1402 | -0.0027 | -0.0009 |
| | (0.0044) | (0.0042) | (0.0014) | (0.0013) | (0.0007) | (0.0008) |
| IV estimate | | | | | -0.0168 | -0.0055 |
| | | | | | (0.0043) | (0.0050) |
| Vaccine effectiveness | | | | | 0.803 | 0.192 |
| | | | | | (0.059) | (0.162) |
| # Students | | | | | | 48,947 |
| D. Broadest possible sample | | | | | | |
| DID estimate | 0.2474 | 0.2380 | 0.2035 | 0.1995 | -0.0042 | 0.0003 |
| | (0.0025) | (0.0023) | (0.0009) | (0.0008) | (0.0005) | (0.0006) |
| IV estimate | | | | | -0.0176 | 0.0011 |
| | | | | | (0.0021) | (0.0025) |
| Vaccine effectiveness | | | | | 0.796 | 0.077 |
| | | | | | (0.029) | (0.123) |
| # Students | | | | | | 64,422 |

Notes: Table reports DID estimates for each outcome, and the DID-IV estimates for effect of two vaccine doses on monthly COVID incidence and ER visits. The post period is Fall 2021, pre-period is Fall 2020. Treatment group is born in six (panels A and D), three (panel B), or 12 (panel C) months before May 12, 2009; control group is born in the same number of months after November 3, 2009. The sample is limited to students with high quality school assignments, except in panel D, which drops those restrictions. See Appendix B for details on vaccine effectiveness.

Table A.6: Effect of vaccine eligible schoolmates on ineligible students, by census tract vaccination rate

| Census tract vaccination quartile | First | Second | Third | Fourth |
|---|---|---|---|---|
| Y = school vaccination rate | | | | |
| DID estimate | 0.1318 | 0.1615 | 0.2072 | 0.2652 |
| | (0.0067) | (0.0084) | (0.0126) | (0.0235) |
| Control mean, post | 0.0234 | 0.0425 | 0.0653 | 0.1068 |
| | (0.0019) | (0.0044) | (0.0108) | (0.0128) |
| Y = covid Incidence | | | | |
| DID estimate | -0.0000 | 0.0007 | 0.0018 | -0.0034 |
| | (0.0027) | (0.0030) | (0.0037) | (0.0038) |
| Pre-period difference | -0.0014 | 0.0017 | 0.0016 | 0.0003 |
| | (0.0018) | (0.0017) | (0.0024) | (0.0025) |
| Control mean, post | 0.0154 | 0.0166 | 0.0148 | 0.0215 |
| | (0.0016) | (0.0015) | (0.0020) | (0.0030) |
| Y = own vaccination rate | | | | |
| DID estimate | 0.0030 | 0.0012 | 0.0033 | -0.0014 |
| | (0.0061) | (0.0069) | (0.0078) | (0.0167) |
| Control mean, post | 0.0472 | 0.0748 | 0.1017 | 0.1814 |
| | (0.0043) | (0.0049) | (0.0063) | (0.0117) |

Notes: Table reports difference-in-differences estimates of the effect of vaccine eligible schoolmates on the indicated outcome, for each quartile of census-tract level vaccination rate. We also report the outcome mean in the post period for the control group, and where indicated, the pre-period difference between treatment and control The sample consists of monthly observations of sixth grade students with reliable school assignment, in November-December. The pre-period is 2020 and the post-period 2021. The treatment group is sixth graders who go to school with older grades (who are vaccine eligible in May 2021 and later the post-period) and the control group sixth graders who go to school with younger students (who are not vaccine-eligible until November 2021). Robust standard errors, clustered on school, in parentheses.

Table A.7: Robustness of cross-grade spillover estimates, August-December

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sixth grade vaccination rate | 0.1941 | 0.1934 | 0.1983 | 0.1941 | 0.1907 | 0.1795 | 0.1987 | 0.1950 |
| | (0.0111) | (0.0108) | (0.0118) | (0.0113) | (0.0109) | (0.0092) | (0.0115) | (0.0112) |
| COVID-19 incidence | 0.0020 | 0.0018 | 0.0026 | 0.0020 | 0.0017 | 0.0032 | 0.0013 | 0.0016 |
| | (0.0011) | (0.0011) | (0.0012) | (0.0012) | (0.0013) | (0.0014) | (0.0015) | (0.0013) |
| Sixth grade vaccination rate | 0.0045 | 0.0039 | 0.0064 | 0.0063 | 0.0061 | 0.0057 | 0.0054 | 0.0026 |
| | (0.0041) | (0.0039) | (0.0043) | (0.0040) | (0.0042) | (0.0034) | (0.0054) | (0.0012) |
| # Observations | 270,380 | 298,640 | 240,985 | 239,610 | 207,345 | 141,775 | 159,340 | 199,430 |
| # Children | 54,076 | 59,728 | 48,197 | 47,922 | 41,469 | 28,355 | 31,868 | 39,886 |
| # Schools | 336 | 345 | 325 | 322 | 311 | 271 | 328 | 334 |
| Minimum catchment area coverage | 70% | 60% | 70% | 80% | 90% | 99.99% | 70% | 70% |
| Exclude ambiguous geography | No | No | Yes | No | No | No | No | No |
| Require prior encounters | No | No | No | No | No | No | Yes | No |
| Exclude if early eligibility | No | No | No | No | No | No | No | Yes |

Notes: Table reports difference-in-differences estimates of the effect of vaccine eligible schoolmates on the indicated outcome. Each cell is a DID estimate. Column (1) contains our baseline estimates. In columns (2)-(6) we drop the requirement that students live in a geography with unambiguous treatment/control status, and vary the minimum coverage share of the modal school catchment area. In column (7) we limit to students with encounter data in the INPC in the prior two years. In column (8) we exclude students turning 12 before November 3, 2021, who are eligible for the vaccine in the early fall.

Table A.8: Robustness of cross-grade spillover estimates, August-October

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sixth grade vaccination rate | 0.1864 | 0.1854 | 0.1903 | 0.1867 | 0.1836 | 0.1721 | 0.1909 | 0.1873 |
| | (0.0106) | (0.0103) | (0.0113) | (0.0108) | (0.0104) | (0.0087) | (0.0110) | (0.0108) |
| COVID-19 incidence | 0.0035 | 0.0032 | 0.0036 | 0.0033 | 0.0029 | 0.0038 | 0.0033 | 0.0033 |
| | (0.0016) | (0.0016) | (0.0017) | (0.0017) | (0.0018) | (0.0021) | (0.0021) | (0.0018) |
| Sixth grade vaccination rate | 0.0024 | 0.0022 | 0.0033 | 0.0034 | 0.0034 | 0.0025 | 0.0029 | 0.0007 |
| | (0.0023) | (0.0023) | (0.0025) | (0.0023) | (0.0024) | (0.0021) | (0.0029) | (0.0006) |
| # Observations | 162,228 | 179,184 | 144,591 | 143,766 | 124,407 | 85,065 | 95,604 | 119,658 |
| # Children | 54,076 | 59,728 | 48,197 | 47,922 | 41,469 | 28,355 | 31,868 | 39,886 |
| # Schools | 336 | 345 | 325 | 322 | 311 | 271 | 328 | 334 |
| Minimum catchment area coverage | 70% | 60% | 70% | 80% | 90% | 99.99% | 70% | 70% |
| Exclude ambiguous geography | No | No | Yes | No | No | No | No | No |
| Require prior encounters | No | No | No | No | No | No | Yes | No |
| Exclude if early eligibility | No | No | No | No | No | No | No | Yes |

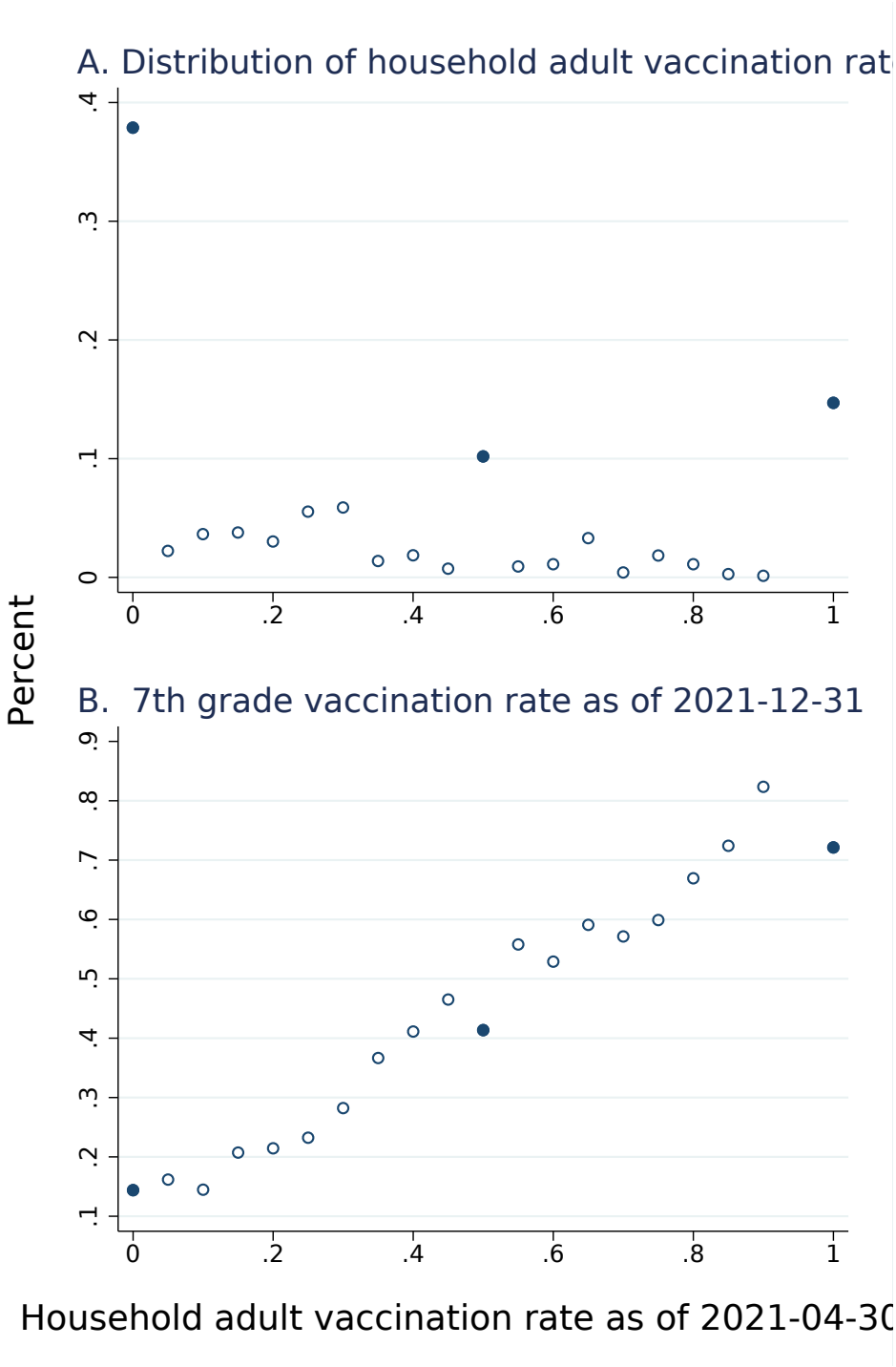
Notes: Table reports difference-in-differences estimates of the effect of vaccine eligible schoolmates on the indicated outcome. Each cell is a DID estimate. Column (1) contains our baseline estimates. In columns (2)-(6) we drop the requirement that students live in a geography with unambiguous treatment/control status, and vary the minimum coverage share of the modal school catchment area. In column (7) we limit to students with encounter data in the INPC in the prior two years. In column (8) we exclude students turning 12 before November 3, 2021, who are eligible for the vaccine in the early fall.

Table A.9: Robustness of cross-grade spillover estimates, November-December

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Sixth grade vaccination rate | 0.2055 | 0.2055 | 0.2102 | 0.2053 | 0.2014 | 0.1906 | 0.2105 | 0.2066 |
| | (0.0120) | (0.0117) | (0.0128) | (0.0123) | (0.0120) | (0.0103) | (0.0125) | (0.0122) |
| COVID-19 incidence | -0.0002 | -0.0002 | 0.0010 | 0.0001 | -0.0002 | 0.0022 | -0.0017 | -0.0010 |
| | (0.0018) | (0.0018) | (0.0019) | (0.0021) | (0.0023) | (0.0027) | (0.0022) | (0.0021) |
| Sixth grade vaccination rate | 0.0078 | 0.0065 | 0.0110 | 0.0107 | 0.0101 | 0.0104 | 0.0092 | 0.0055 |
| | (0.0069) | (0.0065) | (0.0073) | (0.0068) | (0.0071) | (0.0058) | (0.0093) | (0.0025) |
| # Observations | 108,152 | 119,456 | 96,394 | 95,844 | 82,938 | 56,710 | 63,736 | 79,772 |
| # Children | 54,076 | 59,728 | 48,197 | 47,922 | 41,469 | 28,355 | 31,868 | 39,886 |
| # Schools | 336 | 345 | 325 | 322 | 311 | 271 | 328 | 334 |
| Minimum catchment area coverage | 70% | 60% | 70% | 80% | 90% | 99.99% | 70% | 70% |
| Exclude ambiguous geography | No | No | Yes | No | No | No | No | No |
| Require prior encounters | No | No | No | No | No | No | Yes | No |
| Exclude if early eligibility | No | No | No | No | No | No | No | Yes |

Notes: Table reports difference-in-differences estimates of the effect of vaccine eligible schoolmates on the indicated outcome. Each cell is a DID estimate. Column (1) contains our baseline estimates. In columns (2)-(6) we drop the requirement that students live in a geography with unambiguous treatment/control status, and vary the minimum coverage share of the modal school catchment area. In column (7) we limit to students with encounter data in the INPC in the prior two years. In column (8) we exclude students turning 12 before November 3, 2021, who are eligible for the vaccine in the early fall.

Figure A.1: Adult vaccination and child vaccination




Notes: The top panel shows the distribution of adult vaccination rates as of April 30, 2021, among adults aged 26-64 and living in the same addresses as seventh graders in our sample. The bottom panel shows the vaccination rate (as of October 31, 2021) among seventh graders in each bin of adult vaccination rates. The solid circles represent the modal adult vaccination rates of 0, 50, and 100 percent.

# B Vaccine effectiveness in an instrumental variables framework

Studies of causal effects in empirical microeconomics typically focus on treatment effect parameters that are expressed as differences in the expected value of treated and untreated potential outcomes for specified sub-populations. Instrumental variable estimators that account for incomplete take up or non-compliance with assigned treatments are interpreted as average casual effects among members of the complier sub-population. The clinical trials used to evaluate the effects of the Covid-19 vaccines focused primarily on a somewhat different causal parameter, which is often referred to as "vaccine efficacy".

In this appendix, we define a new parameter called "complier average vaccine efficacy" (CAVE). We derive an instrumental variables estimator of the CAVE that is valid under standard instrumental variable assumptions. We use the estimator in the paper to estimate the CAVE in our own effects study design.

## B.1 Notation and Assumptions

Use $i = 1...N$ to index members of a study population. $C_i$ is a binary observed outcome variable that indicates whether the person has a confirmed positive Covid-19 test during a specified follow up window. $V_i$ is a binary treatment variable indicating whether the person was vaccinated for Covid-19 before the start of the follow up window. And $Z_i$ is a binary instrumental variable, which is supposed to affect vaccine take up but is unrelated to Covid-19 infection risk. Values of $(C_i, V_i, Z_i)$ are observed for each member of the study population.

Observed vaccine take up and Covid-19 infections are realizations of underlying potential outcomes. Specifically, let $V_i(z)$ be the vaccination status of person $i$ when her instrument is set to $z$ for $z = [0, 1]$. That means that realized vaccine take up is $V_i = V_i(0) + Z_i[V_i(1) - V_i(0)]$, where $V_i(1) - V_i(0)$ represents the causal effect of the instrument on person $i's$ vaccine take up. Similarly, let $C_i(z, v)$ be person $i's$ downstream Covid-19 infection status if person $i's$ instrument is set to $z$ and her vaccination status is set to $v$ for $v = [0, 1]$.

We work with a set of five instrumental variable assumptions, which were originally described in papers by Imbens and Angrist (1994) and Angrist et al. (1996).

**A1 SUTVA** Covid-19 infection outcomes are individualistic and do not depend on the vaccination status or instrumental variable assignments of any other members of the study population. More formally, let $Z^{-i}$ be the $1 \times N - 1$ vector containing the instrumental variable assignments of each $j = 1...N$ such that $j \neq i$. Likewise $V^{-i}$ is the $1 \times N - 1$ vector of vaccination outcomes for each $j \neq i$. Now let $C_i(Z_i, V_i, Z^{-i}, V^{-i})$ be the potential outcome that person $i$ would experience under a specific combination of own instrument and vaccine exposures **and** peer instrument and vaccine exposures. Under SUTVA $C_i(Z_i, V_i, Z^{-i}, V^{-i}) = C_i(Z_i, V_i)$ so that each person's potential outcomes do not depend on the vaccine status or instrumental variable status of any other member of the study population.

**A2 Independence** – The instrument is statistically independent of potential vaccine take up and potential Covid-19 infection outcomes. Formally, independence implies $Pr(Z_i = 1|V_i(z), C_i(z, v)) = Pr(Z_i = 1)$ for all combinations of $z$ and $v$.

**A3 Exclusion** – The instrument has no causal effect on Covid-19 infection outcomes. This implies that $C_i(z, v) = C_i(v)$ for all $i = 1...N$.

**A4 Monotonicity** – The causal effect of the instrument on vaccine take up is non-negative for any individual in the sample. In other words $V_i(1) - V_i(0) \geq 0$ for all $i = 1...N$.

**A5 First Stage** – The instrument has a non-zero causal effect on vaccine take up for at least some members of the study population so that $E[V_i(1) - V_i(0)] \neq 0$.

## B.2 Treatment Effects

### B.2.1 Additive Effects

At the person level, the additive causal effect of the vaccine on Covid-19 infections is $\beta_i = C_i(1) - C_i(0)$. Since the infection variable is binary, the treatment effect for any single individual can only take on three different values. When $\beta_i = -1$, the person would have been infected with Covid-19 if not for the vaccine. When $\beta_i = 1$ the person is infected with Covid-19 if she is vaccinated but not infected if she is not vaccinated. Finally $\beta_i = 0$ if the person would either be infected in both vaccination states of the world or uninfected in both states of the world.

Treatment effect heterogeneity across subjects may occur for a variety of reasons, including: (i) behavioral responses to vaccination (i.e. Peltzman effects) that lead some people to engage in riskier behaviors (Peltzman effects) or safer behaviors (health complementarity); (ii) biological differences in the immune response generated by the vaccine across subjects; and (iii) differences in epidemiological conditions (exposures) experienced by subjects in different times, places, and social settings.

The average treatment effect of the vaccine is

$$ATE = E[C_i(1) - C_i(0)].$$

The ATE is the difference in Covid-19 infection rates between counterfactual states in which the population is universally vaccinated or universally unvaccinated. It's straightforward to defined conditional average treatment effects. Standard examples are the average treatment effect on the treated: $ATT = E[C_i(1) - C_i(0)|V_i = 1]$, which represents the average effect of the vaccine on Covid-19 infection among people who are actually vaccinated. If $ATT > ATE$ , vaccinated people benefit more from the vaccine than unvaccinated people. If $ATT < ATE$ then vaccination would have larger effects on the unvaccinated population.

## B.3 Vaccine Efficacy Effects

The literature on vaccine trials often focuses on measures of vaccine efficacy rather than on additive average treatment effects. Usin the notation developed so far, vaccine efficacy is

$$\delta = 1 - \frac{Pr(C_i(1))}{Pr(C_i(0))}$$

With a vaccine that is perfectly effective, vaccinating the entire population eliminate 100% of the infections that would occur in the absence of the vaccine. Note, however, that vaccine efficiency is undefined when there is infection risk in the absence of the vaccine so that $Pr(C_i(0)) = 0$. In addition, it is less sensible to define efficacy at the person level the way we do for the additive treatment effect. For instance, $\delta_i = 1 - \frac{C_i(1)}{C_i(0)}$ will equal 0 for people who get infected regardless of vaccination status, 1 for people who

avoid an infection due to vaccination, and is undefined for people who are are not infected in the absence of vaccination. That's unappealing since the vaccine could – in theory – increase infection risk among some people due to Peltzman type risk adjustment responses. The vaccine efficacy concept makes sense at a group level as long as there is a non-zero prevalence of cases of disease in the absence of vaccination.

## B.4 Treatment Effects With Non-compliance

The Covid-19 vaccine trials for the Pfizer, Moderna, and Johnson and Johnson vaccines used randomized experimental designs Polack et al. (2020); Baden et al. (2020); Sadoff et al. (2021). People were randomly assigned to a vaccine group and a placebo group. Covid-19 infections were measured at follow up and the infection rates in the two groups were used to estimate the causal effects of the vaccine. For example, Baden et al. (2020) report that at the end point of the Moderna trial, there were about $131.5$ Covid-19 cases per 10,000 people in the placebo group and about $7.8$ Covid-19 cases per 10,000 in the vaccine group. The average treatment effect implies that the vaccine reduced Covid-19 infection rates by $7.8 - 131.5 \approx 123.7$ cases per 10,000. The efficacy of the vaccine was $1 - \frac{7.8}{123.7} \times 100 \approx 94.1\%$.

### B.4.1 Complier Average Treatment Effects

The Covid-19 vaccine trials experienced a small amount of non-compliance with the study protocol. Some subjects were lost to follow up, did not receive both doses of the vaccine, or experienced other events that made them ineligible. The main analysis in the trials used some form of per-protocol analysis in which these subjects were discarded, although various types of intent-to-treat samples were also considered.

In empirical economics, non-compliance with assigned treatments is often handled using instrumental variables analysis, providing a bridge between randomized experiments and quasi-experimental designs. A pair of papers by Imbens and Angrist (1994) and Angrist et al. (1996) show that in settings with a binary treatment and a binary instrumental variable satisfying assumptions A1-A5, the Wald-IV estimator identifies a parameter called the "Complier Average Treatment Effect" (CATE). Using the notation developed above, these papers show that

$$\frac{E[C_i|Z_i = 1] - E[C_i|Z_i = 0]}{E[C_i|Z_i = 1] - E[C_i|Z_i = 0]} = E[C_i(1) - C_i(0)|V_i(1) > V_i(0)]$$

The right hand side is the CATE, which is the average treatment effect in the sub-population of people who are induced to be vaccinated because of the instrumental variable. Given a valid instrumental variable, it is straightforward to estimate the CATE parameter from observed data. We report estimates of the CATE in our study of the own effects of the vaccine in Table 1.

### B.4.2 Complier Vaccine Efficacy

In this section, we show how to identify a conditional version of the overall vaccine efficacy parameter, which we refer to as the "Complier Vaccine Efficacy" (CAE). The CAE is analogous to the CATE in the sense that it is a measure of vaccine efficacy in the sub-population of people who are induced to be vaccinated because of a binary instrumental variable. The CAE parameter that we focus on in this section is defined as:

$$\delta_{complier} = 1 - \frac{Pr(C_i(1)|V_i(1) > V_i(0))}{Pr(C_i(0)|V_i(1) > V_i(0))}.$$

$\delta_{complier}$ is a function of two counterfactual quantities. $Pr(C_i(0)|V_i(1) > V_i(0))$ is the complier base rate: it represents the Covid-19 infection rate among compliers in the absence of vaccination. $Pr(C_i(1)|V_i(1) > V_i(0))$ is the complier breakthrough rate. It represents the complier infection rate when the compliers are vaccinated.

In this section, we show that both of these quantities are identified under assumptions A1-A5. The CAE is identified under the additional restriction that $Pr(C_i(0)|V_i(1) > V_i(0)) > 0$.

**The First Stage**

Under A1-A5, the first stage comparison identifies the fraction of compliers in the population:

$$\begin{aligned} F &= E[V_i|Z_i = 1] - E[V_i|Z_i = 0] \\ &= E[V_i(1)|Z_i = 1] - E[V_i(0)|Z_i = 0] \\ &= E[V_i(1)] - E[V_i(0)] \\ &= P[V_i(1) > V_i(0)] \end{aligned}$$

The second equality follows after substitution of the potential vaccine take up expression for the observed vaccine take up outcomes. The third equality imposes the independence assumption. And the fourth equality imposes the monotonicity condition. This shows that the first stage difference in vaccine take up rates identifies the prevalence of compliers.

**The Complier Base Rate**

The logical challenge in identifying the complier base rate is that complier status is unknown at the individual level, and unvaccinated Covid-19 potential outcomes are not observed for the full population. We can apply the standard instrumental variables analysis to an adjusted/censored outcome variable to uncover complier averages of the individual outcomes.

Let $R_i^{base} = (1 - V_i)C_i$ to be an adjusted outcome that is set to 0 for people who are vaccinated and set to the value of $C_i$ for people who are unvaccinated. The reduced form difference in (adjusted) Covid-19 outcomes across levels of the instrument is:

$$\begin{aligned} ITT_{base} &= E[R_i^{base}|Z_i = 1] - E[R_i^{base}|Z_i = 0] \\ &= E[(1 - V_i)C_i|Z_i = 1] - E[(1 - V_i)C_i|Z_i = 0] \\ &= E[(1 - V_i(1))C_i(0)|Z_i = 1] - E[(1 - V_i(0))C_i(0)|Z_i = 0] \\ &= E[C_i(0)(V_i(0) - V_i(1))] \\ &= -E[C_i(0)|V_i(1) > V_i(0)]Pr(V_i(1) > V_i(0)). \end{aligned}$$

The second equality substitutes the definition of the adjusted outcome, and the third equality introduces the potential outcomes structure, invoking the exclusion restriction. The fourth quality imposes the independence assumption to drop conditioning on the instrument. The fifth line decomposes the ex-

pectation using the fact that $V_i(0) - V_i(1)$ can only take on the values 1, 0, and -1. Two of the three terms drop out: the zero term is multiplied by zero and $Pr(V_i(0) - V_i(1) = 1) = 0$ under Under A4 (monotonicity). Thus $ITT_{base}$ is equal to the negative of the complier base rate multiplied by the prevalence of compliers. Dividing by the negative of the complier share using a standard Wald Ratio gives:

$$\begin{aligned}
W_{base} &= \frac{ITT_{base}}{-F} \\
&= \frac{E[R_i^{base}|Z_i = 1] - E[R_i^{base}|Z_i = 0]]}{-(E[V_i|Z_i = 1] - E[V_i|Z_i = 0])} \\
&= \frac{-E[C_i(0)|V_i(1) > V_i(0)]Pr(V_i(1) > V_i(0))}{-Pr(V_i(1) > V_i(0))} \\
&= E[C_i(0)|V_i(1) > V_i(0)] \\
&= Pr[C_i(0) = 1|V_i(1) > V_i(0)].
\end{aligned}$$

**The Complier Breakthrough Rate**

Following a parallel approach for the complier breakthrough rate, define the adjusted outcome $R_i^{break} = V_i C_i$, which is set to 0 for people who are unvaccinated and set to $C_i$ for people who are vaccinated. The reduced form comparison in this case is:

$$\begin{aligned}
ITT_{break} &= E[R_i^{break}|Z_i = 1] - E[R_i^{break}|Z_i = 0] \\
&= E[V_i C_i|Z_i = 1] - E[V_i C_i|Z_i = 0] \\
&= E[V_i(1)C_i(1)|Z_i = 1] - E[(V_i(0)C_i(1)|Z_i = 0] \\
&= E[C_i(1)(V_i(1) - V_i(0))] \\
&= E[C_i(1)|V_i(1) > V_i(0)]Pr(V_i(1) > V_i(0)).
\end{aligned}$$

Here, the second equality uses the definition of the adjusted outcome and the third line equality introduces the potential outcomes structure, invoking the SUTVA condition and the exclusion restriction. The fourth equality imposes the independence assumption and collects terms. The fifth line decomposes the expected value of the product of $C_i(1)$ and $V_i(1) - V_i(0)$ and imposes the monotonicity assumption. The result shows that $ITT_{break}$ is the complier breakthrough infection rate multiplied by the prevalence of compliers. The Wald ratio isolates the complier breakthrough rate:

$$
\begin{aligned}
W_{break} &= \frac{ITT_{break}}{F} \\
&= \frac{E[R_i^{break}|Z_i = 1] - E[R_i^{break}|Z_i = 0]]}{E[V_i|Z_i = 1] - E[V_i|Z_i = 0]} \\
&= \frac{E[C_i(1)|V_i(1) > V_i(0)]Pr(V_i(1) > V_i(0))}{Pr(V_i(1) > V_i(0))} \\
&= E[C_i(1)|V_i(1) > V_i(0)] \\
&= Pr[C_i(1) = 1|V_i(1) > V_i(0)].
\end{aligned}
$$

**Estimation**

The complier average vaccine efficiency can be estimated using the ratio of the two Wald ratios:

$$
\begin{aligned}
\delta_{complier} &= 1 - \frac{W_{break}}{W_{base}} \\
&= 1 - \frac{ITT_{break} \times F^{-1}}{-ITT_{base} \times F^{-1}} \\
&= 1 + \frac{ITT_{break}}{ITT_{base}} \\
&= 1 - \frac{Pr(C_i(1)|V_i(1) > V_i(0))}{Pr(C_i(0)|V_i(1) > V_i(0))}.
\end{aligned}
$$

Interestingly, the first stages cancel and so the efficiency is equal to 1 plus the ratio of the reduced forms. In practice, you could estimate the complier efficiency by computing the two IV estimates (complier base rate and complier breakthrough rate) directly and then computing the ratio of the two. Or your could compute the two ITT effects and compute their ratio. In both cases, it would be sensible to do things in a stacked framework so that you could produce a joint covariance matrix. This is pretty straightforward though.

# C Details on assigning students to schools

This appendix provides a detailed explanation of our approach to imputing school assignment, and validates the approach, in the sense of showing that measurement error from imperfect address information largely does not propagate to measurement error in school assignment.

## C.1 Detailed assignment process

We assign patients in the Regenstrief data to schools based on the geography reported in the vaccine and test registry. We obtain data on grades and catchment area of schools from National Center for Education Statistics (2022). Our goal is to assign each Regenstrief geography—census tract by zip code—to a sixth-grade serving school. We focus on schools serving sixth graders only, because our identification strategy compares sixth graders with vaccine-eligible peers to sixth graders without such peers.

Using geography to assign patients to schools faces three challenges. First, some school districts have school choice program which means that middle school assignment is not determined by geography alone. For example Indianapolis Public Schools has both "neighborhood schools" and "choice schools" (Indianapolis Public Schools, 2022), and the NCES shapefiles report that the catchment area for each school is the entire school district. Second, school catchment areas do not completely cover the state, either because of incomplete data or because no one lives in some census tracts of the state (e.g. state forests). Third, the geography reported in the test and vaccine registry data, zip codes and census tract, does not necessarily align with school catchment areas or even school districts. As a result, a given zip code and census tract can contain multiple catchment areas, or even multiple school districts (which happens for example when a census tract contains multiple towns).

Roughly speaking, to overcome these challenges, we assign patients to the school whose catchment area covers the largest share of their geography, and we limit the sample to patients whose assigned school is reasonably likely to be their school. To explain the procedure in detail, denote a given zip code-census tract combination (a "geography") as $g$, and a school catchment area $s$. Using GIS, we overlay $\{g\}$ and $\{s\}$. For each $g$, $s$ we calculate the share of the area of $g$ that $s$ covers, $p_{gs}$. If each school catchment areas covered the entire state and never overlapped, the covered share of each geography, $covered_g \equiv \sum_s p_{gs}$, would always equal 1. However, geographies are multiply covered because students can choose among multiple schools (as in the Indianapolis Public School district), which implies that $covered_g > 1$. Other geographies are undercovered because of holes in the school catchment area coverage map. Appendix Figure C.1 shows the distribution of $covered_g$ across geographies in the registry data. While the modal area is 100 percent covered, the spikes at 200 percent reflects areas served by multiple schools, and the mass at non-integer amounts reflects undercoverage.

Accounting for both over- and under-coverage, we say that the adjusted coverage share of school catchment area $s$ for geography $g$ is

$$\tilde{p}_{gs} = \frac{p_{gs}}{\max\{1, covered_g\}}.$$

$\tilde{p}_{gs}$ is equal to $p_{gs}$ as long as $g$ does not contain overlapping catchment areas. If $g$ contains overlapping catchment areas, $\tilde{p}_{gs}$ is scaled down by the sum of covered areas. If students were uniformly distributed within a geography, and chose schools at random from among the schools they were eligible to attend,

then $\tilde{p}_{gs}$ would be the probability that a student in $g$ attends $s$.

We set the assigned school for $g$ to be the school with the largest $\tilde{p}_{gs}$.[10] Appendix Figure C.2 shows the distribution of $\tilde{p}_{gs}$ for the modal school. A majority of geographies have modal school coverage above 99.9 percent, but there is a long left tail. Our analysis sample restricts attention to students in " high coverage" geographies, defined as geographies where $\tilde{p}_{gs} \geq 0.7$, meaning at least 70 percent of the area of $g$ is assigned to $s$, accounting for overlapping catchment areas.

## C.2 Illustrative cases

We illustrate the process in Appendix Figure C.3. Each panel zooms in on a different geography. Panel (a) shows the simplest case. The geography defined by zipcode 46057 and census tract 18023950200 is contained entirely within the Clinton Central Elementary School catchment area, so we assign that as the sixth grade school to all students in this geography. Panel (b) illustrates the case that multiple non-overlapping catchment areas overlay a single geography; 60 is covered by Clinton Prairie's catchment area, and 40 percent by Rossville Elementary. We assign students in this geography to Clinton Prairie, the modal school, but because its covered share is less than 70 percent, we do not include this geography in our main spillover analysis sample. Panel (c) shows a case of undercoverage; part of the geography is not covered by any school catchment area. Covington Middle School covers 72 percent of the area, so we assign Covington Middle school as the school, and include this geography in our main spillover analysis sample. Panel (d) shows the final problematic case: the catchment areas of Cleveland Elementary School and West Side Middle School both overlay the entire geography. In this case the area is 200 percent covered, and the modal school–West Side Middle School—covers 100 percent. The normalized coverage share is 50 percent (i.e. 100/200), below 70 percent, so we exclude this geography from our main spillover analysis sample.

## C.3 Measurement error likely plays a limited role in our estimates

Our approach to imputing school assignment suffers from three sources of measurement error: first, we impute public school assignment, but students may attend private school; second, a given address may be eligible to attend multiple public schools; and third, we do not observe exact address, only census tract and zip code. We argue that each of these sources is unlikely to be important in our application.

First, while we cannot observe private school attendance, private school attendance represents about 8 percent of K-12 enrollment in Indiana.[11] Thus roughly 8 percent of our treatment and control group is potentially miss-assigned. Even if all treatment and control students were miss-assigned, the attenuation bias would be fairly small (16 percent). It is unlikely that all private school students are miss-assigned, of course. Second, our approach excludes students eligible to attend multiple schools, since their modal school will cover 50 percent of the normalized area.

Third, it turns out that our geography information is sufficiently rich to capture most of the variation in school assignment, after imposing our sample restrictions. To show this, we first assign *census blocks*

[10] In our high-coverage sample, ties are impossible. In broader samples ties are possible because, for example, two schools both cover 100 percent of an area. In those cases we break ties first by assigning students to the school with more sixth grade enrollment, and then by assigning students to the school with the first NcES id.

[11] See `https://www.in.gov/doe/about/news/indiana-k-12-school-enrollment-grows-for-2021-2022-school-year/`

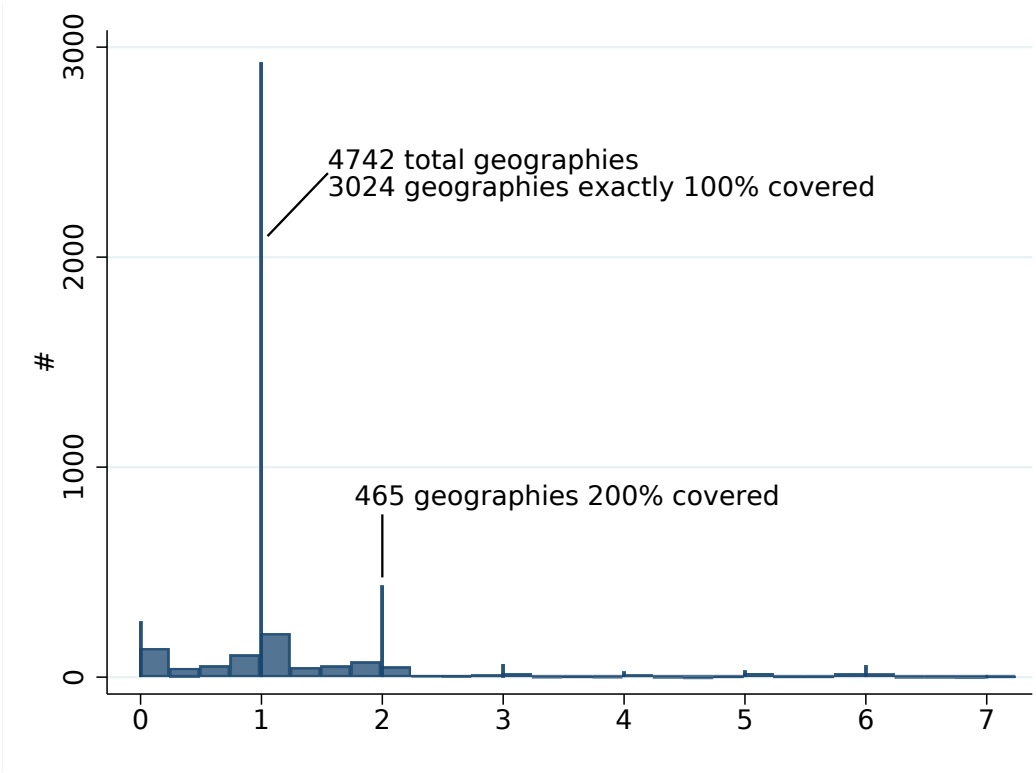
to both schools and census tracts and zip codes. Census blocks are sufficiently small that there is essentially no ambiguity about school assignment outside of "school choice" districts where each location has multiple options. We then compare the census block's true school to the one our algorithm assigned.

Appendix Table C.1 shows the concordance between true school assignment and imputed assignment, weighting by census-block population. Looking at all census blocks, we see that the true assigned and imputed assigned school agree in 83 percent of cases, and the type agrees in 96 percent. The mean absolute error in school-wide vaccination rate (August-September 2021) is 0.7 percentage points. Thus without any sample restrictions the agreement rate is very high. In the remaining columns we impose restrictions to improve the agreement rate. Limiting to high coverage schools brings the agreement rate to 95 percent, and the type agreement rate to 98.5 percent (in column 5 we report the statistics for our analysis sample). These restrictions limit the analysis sample to 62 percent of the population. We can further improve the coverage by restricting the sample to geographies where there is essentially no ambiguity in treatment/control status among possible schools. Doing so reduces the coverage by a further 8 percent, so we opted not to.[12] Thus more detailed address data would not improve our school assignment substantially.

report: - school match rate - type match rate - mean absolute error - share of poplation

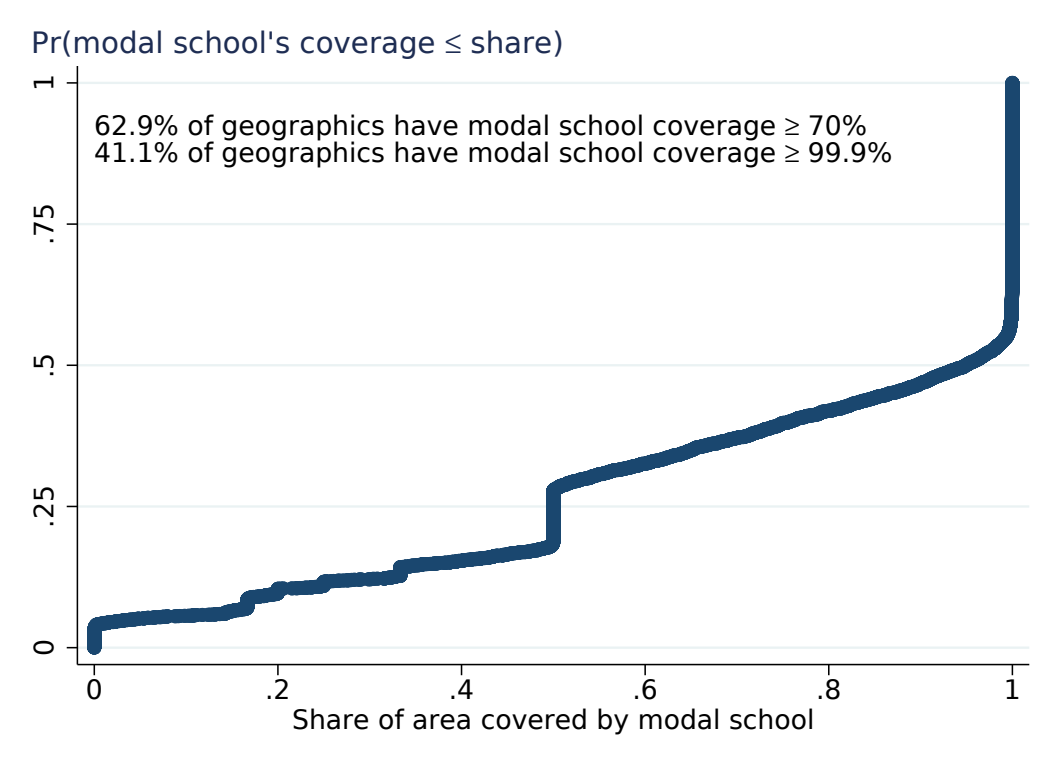
[12] The treatment status agreement rate is less than 100 percent because we allow for a slight amount of ambiguity; the most common type must cover at least 99 percent of the geography's area, not 100 percent.

Figure C.1: Distribution of land mass covered by school catchment areas, across geographies




Notes: Figure shows the frequency distribution, across areas, of the share of the area covered by a school catchment area. Some areas can be doubly covered, leading to shares above 1. For readability, figure uses two bin widths: 0.001 in the vicinity of integers, and 0.25 otherwise.

Figure C.2: Cumulative distribution of modal school's normalized coverage, across geographies



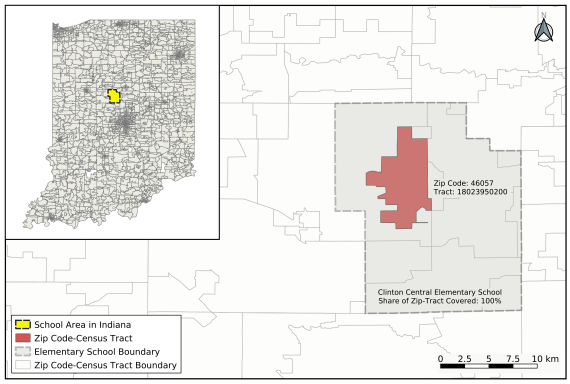

Notes: Figure shows the cumulative distribution, across areas, of the share of the area covered by the modal school's catchment area (i.e. the school with the greatest coverage), normalizing by the total covered share if that share exceeds one.

Table C.1: Coarse geography does not lead to substantial measurement error

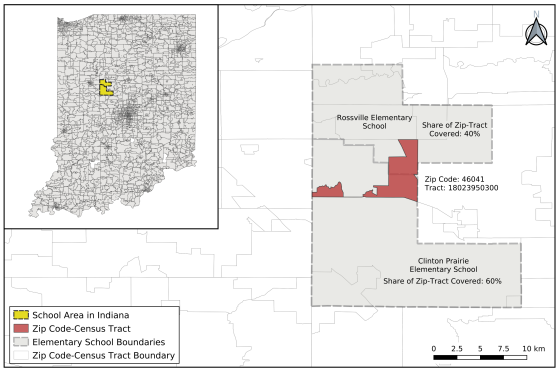
| Sample | Full | No ambiguous type | $\geq$ 60% coverage | No ambiguity, $\geq$ 60% coverage | $\geq$ 70% coverage | no ambiguity, $\geq$ 70% coverage |
|---|---|---|---|---|---|---|
| School agreement rate | .8383 | .8588 | .9267 | .9404 | .9477 | .9586 |
| Type agreement rate | .9591 | .9988 | .9801 | .9999 | .9851 | .9999 |
| Mean Absolute Error | .0065 | .0034 | .004 | .0023 | .0031 | .0018 |
| Share of Population | 1 | .7455 | .6792 | .5781 | .6189 | .5389 |

Notes: Table shows concordance between census-block level school assignment and the assignment imputed from zip code-census tract alone, for indicated samples. The first row reports the population-weighted fraction of census blocks where the actual and imputed public school assignment match, and the second row reports the agreement rate between actual and imputed school treatment status (treatment, control, neither) match. Rows 3 and 4 report the constant and slope from a regression of vaccination rate in the imputed school on vaccination rate in the true school. The final row shows the share of population included in the sample. The samples are restricted to zip-code census tract within which there is no substantial ambiguity about treatment status, and/or where the modal school's normalized coverage share exceeds the indicated threshold.

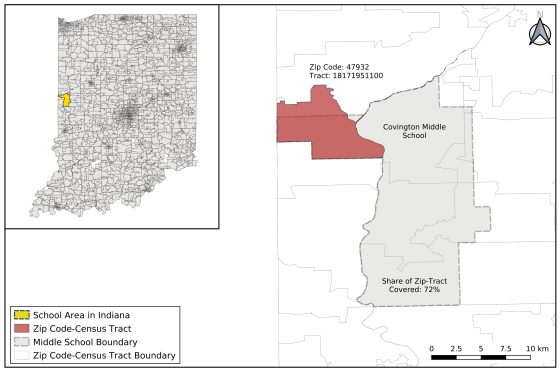
Figure C.3: Examples of geography and school assignment



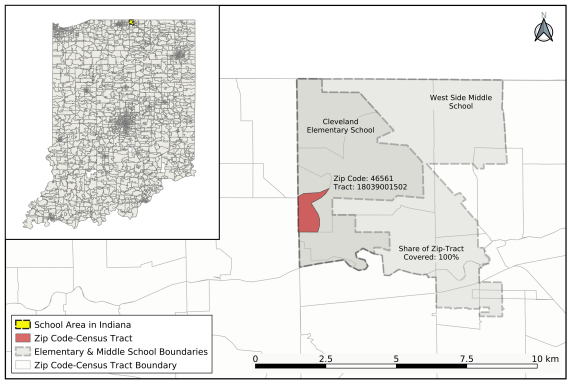

(a) Geography 100% covered, single catchment areas




(b) Geography 100% covered, non-overlapping catchment areas




(c) Geography not fully covered by catchment areas




(d) Geography covered by overlapping catchment areas

Notes: Figure illustrates the process of assigning census tract-zip code geographies to school districts. We consider four cases illustrating the ideal case (panel a) and the various ways things can go wrong. The top left of each panel shows the geography in the state. In the rest of the panel, the lightly shaded area is school catchment area (with dashed boundary) and the heavily shaded area is the intersection of the indicated zip code and census tract.

# D A quantitative model of vaccine externalities

This section presents and numerically analyzes an SIR-model with vaccination to understand whether incomplete vaccine take-up could explain the near-zero spillovers we estimate. We have three results. First, except for very high levels of infectiousness, prevalence among the unvaccinated is nearly linear in the vaccination rate, up to the herd immunity threshold. The effect of additional vaccinations on the unvaccinated therefore is not very sensitive to baseline baccination rates except at high levels of infectiousness. Our second result is that at high levels of infectiousness, a marginal vaccination provide little protection to the unvaccinated, because an unvaccinated person is likely to become infected from another source.[13]

Taken together these results imply that when spillovers exist, they are likely large enough for us to detect from a 20 percentage point increase in schoolmate vaccination rates. Our third result shows this directly: at all levels of vaccination below herd immunity, the simulated effect of a 20 percentage point increase in vaccination is much larger than what our confidence intervals rule out, except in the case of high infectiousness, when spillovers are small.

We caution that this model is particularly simple and may not capture the dynamics of COVID-19. The results here do not necessarily generalize to other disease models.

## D.1 Model set-up

We consider the simple SIR model with uniform mixing and a share $v$ of the population of size $N$ is vaccinated. The vaccine is assumed to be 100 percent effective, and we model vaccinated people as removed from the susceptible pool. Our set up is the discrete time analog of the model in Goodkin-Gold et al. (2020), except we assume perfect effectiveness and abstract from the vaccine demand phase. Thus

$$N = S + I + R + vN.$$

The equations describing infection dynamics are

$$\begin{aligned} \Delta S &= -\beta S \cdot I/N \\ \Delta I &= \beta S \cdot I/N - \gamma I \\ \Delta R &= \gamma I. \end{aligned}$$

Here $\beta$ is the transmission rate and $\gamma$ is the recovery rate.

A key parameter is the reproductive number $\mathcal{R}$, the number of new infections spawned by a single infection. The basic reproductive number $\mathcal{R}_0$ is the value of $\mathcal{R}$ in a completely susceptible population, so $\mathcal{R}_0 = \beta/\gamma$. When $\mathcal{R} < 1$, infections do not replace themselves and so disease outbreaks die out.

Because the vaccinated and unvaccinated populations mix uniformly, a vaccination rate of $v$ scales down the susceptible population by $(1 - v)$, and so reduce the reproductive number $(1 - v)$. A large enough vaccinated population ensures that $R < 1$; the so-called herd-immunity vaccination rate guaranteeing this condition is

$$v^* = 1 - 1/R_0.$$

As we will see, this threshold plays an important role in the results.

[13] This intuition is from Goodkin-Gold et al. (2020), who develop it in a series of related results.

**Simulation details** Closed-form solutions for final infection rates and infection dynamics do not exist, so we solve the model with forward simulation to obtain the final-period count of ever infected individuals, $R(T)$. We simulate for $T = 20000$ time periods, starting with $I(1) = 1$ and $R(1) = 0$. In each run we verify that the simulation converges in the sense that the number of infected people change by less than 1/1000 over the last periods.

**Parameterization** The model parameters are $N, \beta, \gamma$ and $v$. We fix $N = 100,00$ and $\gamma = 1/10$. We choose $\beta$ so that $\mathcal{R}_0 \in 1.1, 1.5, 2, 3, 5$; note that $\gamma = .1$, meaning a 10 day expected infection length. For each $\beta$ I vary the vaccination rate from 0 to 1 in increments of 0.01.

These parameters trace out a range of reasonable values for $\mathcal{R}_0$ in the context of COVID-19; 1.1 is lower than estimates; 1.5 is the estimated $\mathcal{R}_0$ for the ancestral strain (also used in Goodkin-Gold et al. (2020)), and 5 represents a very high estimate, possibly occurring with the latest strains, although it is unclear if high transmission of the latest waves reflects immune escape or high $\mathcal{R}_0$. The scale of $\gamma$ is not relevant for $\mathcal{R}_0$, but $\gamma = 1$ has multiple advantages. First, high values of $\gamma$ ensure that the epidemic concludes in relatively few iterations. However, $\beta$ must be less than 1 since it is a transmission probability. Choosing $\gamma = .1$ means that $\beta = .5$ when $\mathcal{R}_0 = 5$.

**Simulation output:** For each value of $v$ and $\mathcal{R}_0$ we calculate the fraction of the unvaccinated population that ever becomes. Since the vaccinated cannot be infected, this fraction is

$$pr(infected|unvaccinated; v, R_0) = R(T; v, R_0)/(N \cdot (1 - v)).$$

Our empirical analysis of vaccine spillovers considers a shock that increases peer vaccination rate by roughly 20 percentage points. We therefore also calculate the implied impact of such a shock on the unvaccinated infection rate:

$$\Delta pr(infected|unvacc; v, R_0) = pr(infected|unvacc; v + .2, R_0) - pr(infected|unvacc; v, R_0).$$

## D.2 Model results

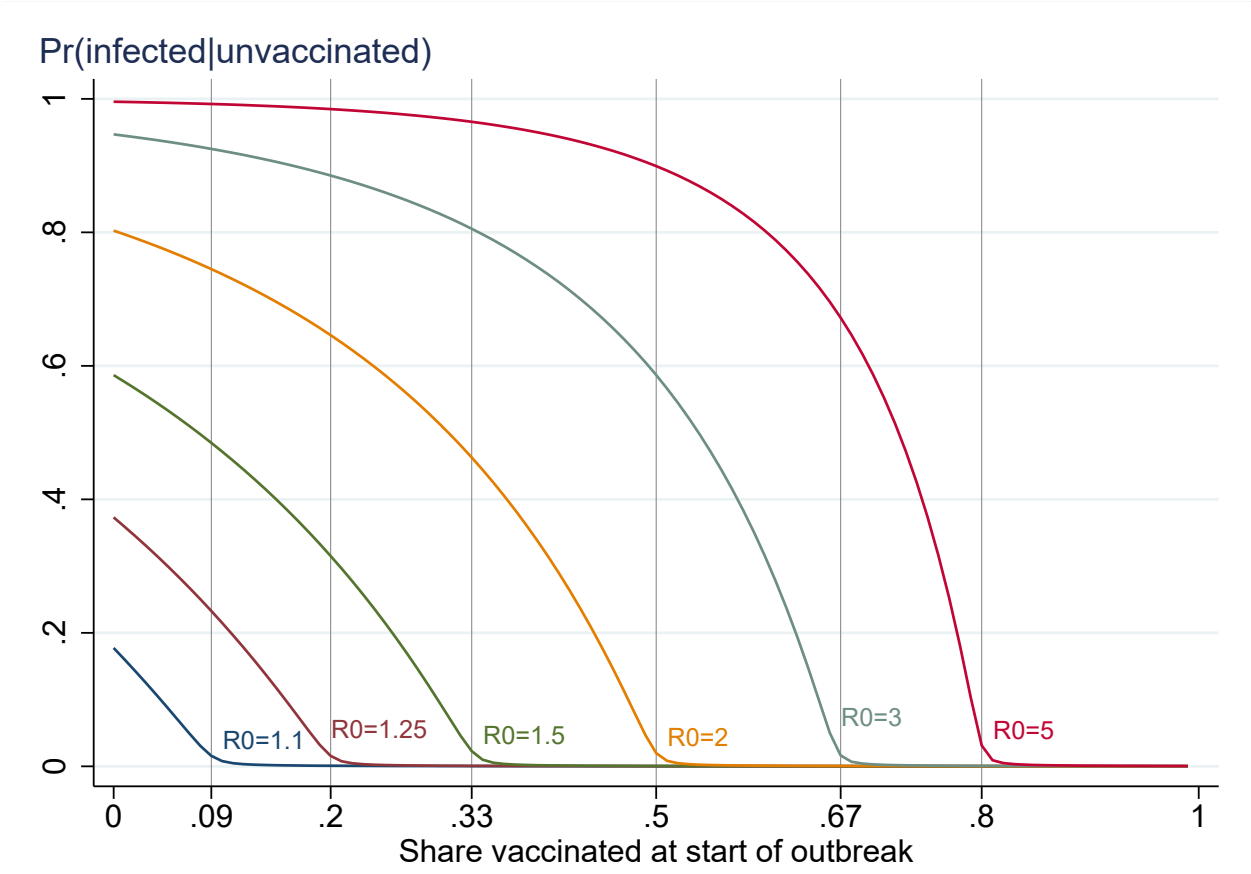
We begin by showing the fraction of the unvaccinated population that ever becomes infected, as a function of the vaccination rate, for various $\mathcal{R}_0$, in Appendix Figure D.1. Several patterns are clear in the figure. Most obviously, marginal vaccinations beyond the herd immunity threshold (the vertical lines) have only very small impacts on the unvaccinated, because at the the herd immunity threshold and beyond, infections die out and nearly all unvaccinated would not become infected even absent greater vaccination.

More importantly, for low levels of infectiousness—$\mathcal{R}_0 < 3$—the relationship between $Pr(infected|unvacc)$ and $v$ is approximately linear, up to the herd immunity threshold. Thus the marginal benefit of vaccinations is roughly constant in $v$; it does not depend on the starting level of vaccination. Estimates of the impact of greater peer vaccination on own infections, if this model were true, would not be too sensitive to baseline vaccination rate. For high levels of infection, the nonlinearity is stronger. However it is also true at these high levels of infection, especially when $\mathcal{R}_0 = 5$, there is very little external benefit of vaccines; even large increases in the vaccination rate do not produce large reductions in unvaccinated incidence, except at very high levels of vaccination.

We show this more specifically in Appendix Figure D.2. The figure shows the simulated impact of a 20 percentage point increase in the vaccination rate on incidence among the unvaccinated, as a function of the initial vaccination rate, for different levels of $\mathcal{R}_0$. While the effect size does vary with $v$, it is always

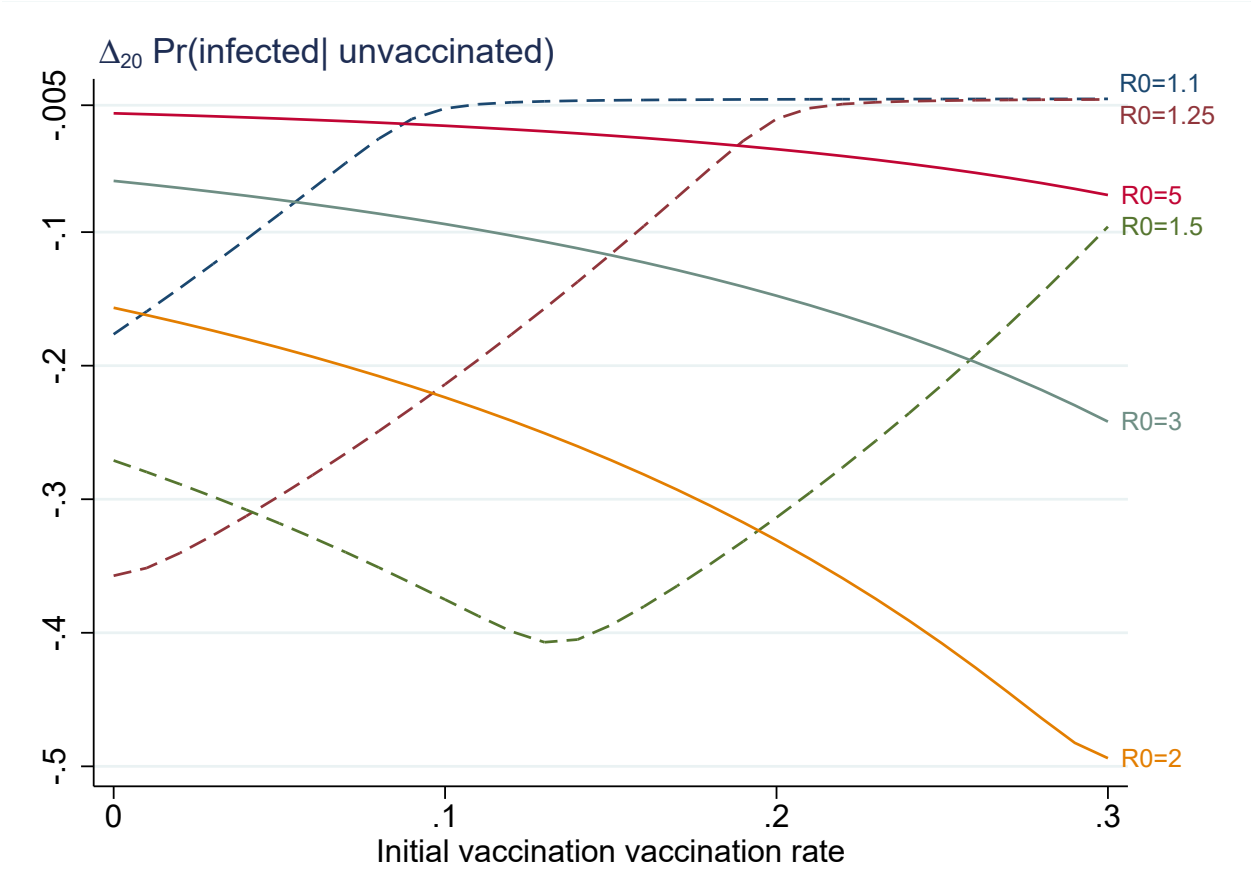
large when $\mathcal{R}_0 < 5$. Indeed the lower bound of the confidence interval from our main estimates—about -0.005 — easily lets us rule out any effect size in the figure, except when either (a) herd immunity is reached, or (b) infectiousness is so high that the spillover is small for a wide range of initial vaccination levels. Even in the case, however, the implied effect is on the order of a few percentage points, an order of magnitude larger (in absolute value) than the lower bound of our confidence interval.

Figure D.1: Infections among the unvaccinated fall with the vaccination rate, up to the herd immunity threshold




Notes: Figure shows the simulated infections per unvaccinated capita, over the course of a pandemic, as a function of the vaccination rate, for various levels of infectiousness given by $R_0$.

Figure D.2: A 20 percentage point increase in the vaccination rate causes a large reduction in infections among the unvaccinated, regardless of starting level




Notes: Figure shows the change in the share of the unvaccinated that are ever infected, when the vaccination rate increases by 20 percentage points, from a given initial vaccination rate for various levels of infectiousness given by $R_0$.

# E No confounding from school modality or mask mandates

Here we show that two key school mitigation policies—mask mandates and in-person instruction—do not confound our estimates. We also show that our estimated effect sizes do not differ by the presence or absence of mask mandates. Mask mandates were in effect for the entire state in the 2020/21 school year, but they were at school districts' discretion during the 2021-22 year. Our mask mandate data come from Waldron (2022b), who collected data on school mask policies for most Indiana school districts through September, 2021, by visiting district websites and monitoring Google alerts (Waldron, 2022a). We focus on mask mandates for students (as opposed to recommendations, or staff mandates). Our school modality data come from COVID-19 School Data Hub (2022), who report instruction modality by school and week for the 2020-21 school year. We focus on in-person instruction as a measure of modality, and we assume all schools are in person for the 2021-22 year.

To show that these mitigation policies do not confound our treatment effect estimates, we re-estimate our main DID spillover models, but (a) putting the policy as the dependent variable, and (b) including the policy as a covariate. We limit the sample to cases with non-missing policy information. The results for mask mandates are in Appendix Table E.1, and for in-person instruction they are in Appendix Table E.2. Neither policy changes differently for the treatment group, and adjusting for the policies makes no difference to our main estimate. We further show in column (4) of Appendix Table E.1 that mask mandates do not mediate our effect; we include an interaction between mask mandate, post ,and treatment, and find a statistically insignificant coefficient. (We cannot study such interactions for instruction modality because they do not vary across schools in the post period, so the coefficient would not be identified.)

Table E.1: Mask mandates neither confound not interact with the spillover effect

| Outcome | Mandate (1) | COVID-19 (2) | COVID-19 (3) | COVID-19 (4) |
|---|---|---|---|---|
| PostXtreat | -0.0426 | 0.0047 | 0.0044 | 0.0033 |
| | (0.0749) | (0.0022) | (0.0021) | (0.0030) |
| Mask mandate in effect | | | -0.0068 | -0.0079 |
| | | | (0.0020) | (0.0027) |
| mandate X post X treat | | | | 0.0018 |
| | | | | (0.0038) |
| # Observations | 93,860 | 93,860 | 93,860 | 93,860 |
| # Students | 46,930 | 46,930 | 46,930 | 46,930 |
| # Schools | 319 | 319 | 319 | 319 |

Notes: Table reports difference-in-differences estimates of the effect of vaccine eligible schoolmates on the indicated outcome. Columns (3) and (4) regression adjust for differences in mask mandates, and column (4) allows the effect of vaccine-eligible schoolmates to vary with mask mandates. The sample consists of monthly observations of sixth grade students with reliable school assignment, in August-September (the dates for which mask mandate policy information exists). The pre-period is 2020 and the post-period 2021. The treatment group is students who go to school with older grades (who are vaccine eligible in May 2021 and later the post-period) and the control group students who go to school with younger students (who are not vaccine-eligible until November 2021). Robust standard errors, clustered on school, in parentheses.

Table E.2: In-person schooling neither confounds nor interacts with spillover effect

| Outcome | In-person (1) | COVID-19 (2) | COVID-19 (3) |
|---|---|---|---|
| PostXtreat | 0.0412 | 0.0026 | 0.0027 |
| | (0.0807) | (0.0012) | (0.0011) |
| In person instruction | | | -0.0026 |
| | | | (0.0006) |
| # Observations | 230,615 | 230,615 | 230,615 |
| # Students | 46,123 | 46,123 | 46,123 |
| # Schools | 325 | 325 | 325 |

Notes: Table reports difference-in-differences estimates of the effect of vaccine eligible schoolmates on the indicated outcome. Columns (3) regression adjusts for differences in in-person schooling. The sample consists of monthly observations of sixth grade students with reliable school assignment, in August-December. The pre-period is 2020 and the post-period 2021. The treatment group is students who go to school with older grades (who are vaccine eligible in May 2021 and later the post-period) and the control group students who go to school with younger students (who are not vaccine-eligible until November 2021). Robust standard errors, clustered on school, in parentheses.