Impacts of K-12 school reopening on the COVID-19 epidemic in Indiana, USA

Guido España[#], Sean Cavany^{*}, Rachel Oidtman^{*}, Carly Barbera, Alan Costello, Anita Lerch, Marya Poterek, Quan Tran, Annaliese Wieler Sean Moore, T. Alex Perkins[#] Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, United States [#]Correspondence: guido.espana@nd.edu, taperkins@nd.edu ^{*}Equal contribution

4 Abstract

3

In the United States, schools closed in March 2020 due to COVID-19 and began reopening in August 2020, despite continuing transmission of SARS-CoV-2. In states where in-person instruction resumed at that time, two major unknowns were the capacity at which schools would operate, which depended on the proportion of families opting for remote instruction, and adher-8 ence to face-mask requirements in schools, which depended on cooperation from students and enforcement by schools. To determine the impact of these conditions on the statewide burden of 10 COVID-19 in Indiana, we used an agent-based model calibrated to and validated against mul-11 tiple data types. Using this model, we quantified the burden of COVID-19 on K-12 students, 12 teachers, their families, and the general population under alternative scenarios spanning three 13 levels of school operating capacity (50%, 75%, and 100%) and three levels of face-mask adher-14 ence in schools (50%, 75%, and 100%). Under a scenario in which schools operated remotely, 15 we projected 45,579 (95% CrI: 14,109-132,546) infections and 790 (95% CrI: 176-1680) deaths 16 statewide between August 24 and December 31. Reopening at 100% capacity with 50% face-17 mask adherence in schools resulted in a proportional increase of 42.9 (95% CrI: 41.3-44.3) and 18 9.2 (95% CrI: 8.9-9.5) times that number of infections and deaths, respectively. In contrast, 19 our results showed that at 50% capacity with 100% face-mask adherence, the number of infec-20 tions and deaths were 22% (95% CrI: 16%-28%) and 11% (95% CrI: 5%-18%) higher than the 21 scenario in which schools operated remotely. Within this range of possibilities, we found that 22 high levels of school operating capacity (80-95%) and intermediate levels of face-mask adherence 23 (40-70%) resulted in model behavior most consistent with observed data. Together, these results 24 underscore the importance of precautions taken in schools for the benefit of their communities. 25

²⁶ Keywords

27 COVID-19, agent-based model, public health, school reopening, face masks

28 1 Introduction

The United States has been the country most severely impacted by the COVID-19 pandemic 29 in terms of total reported cases and deaths, with over 28 million reported cases and more than 30 500 thousand deaths by March, 2021 [1]. This severity led to social interventions on an unprece-31 dented scale, including restrictions on mass gatherings, bans on non-essential travel, and school 32 closures [2, 3, 4, 5]. While such restrictions were initially successful in reducing transmission, the 33 subsequent relaxation of restrictions on mass gatherings and movement were followed by large 34 increases in notified cases and deaths [1, 3, 6, 7]. By the time the 2020-2021 school year began 35 in August, transmission was at its highest point in the epidemic vet in some parts of the US. In 36 Indiana, for example, the maximum number of daily cases was around 1,200 by then, which was 37 higher than the previous maximum of fewer than 800 in late April [8]. 38

This context of intense community transmission raised numerous questions about how schools 39 should approach reopening for the start of the school year in August [9, 10, 11]. During influenza 40 epidemics, school closures have been estimated to reduce transmission community-wide [12, 13, 41 14]. In general, schools are seen as key drivers of the transmission of respiratory pathogens due to 42 close contact among children at school [15, 16, 17]. However, several factors complicated the effect 43 of school reopenings on SARS-CoV-2 transmission. In particular, children and adolescents appear 44 less susceptible to infection and are much less likely to experience severe outcomes following 45 infection [18, 19, 20, 21, 22, 23]. It is also still unclear what their contribution to transmission is, 46 but several studies suggest they can play an important role [18, 24, 25, 26]. There are additional 47 economic and social factors to consider, too, such as the economic costs of school closures for 48 families that must then stay home from work, and the nutritional benefits of school reopening 49 for children who rely on free and subsidized school meals [27, 28, 29]. 50

Our objective in this study was to explore how different conditions for school reopening during 51 the fall semester of 2020 could have impacted the statewide burden of COVID-19 in Indiana. 52 Specifically, we focused on the effects of school operating capacity and adherence to wearing face 53 masks in schools. This focus was motivated by the fact that Indiana and other US states reopened 54 their schools for in-person instruction in August with only minimal interventions of requiring face 55 masks and physical distancing in schools, despite uncertainty about the proportion of students 56 who would elect to attend in person and the degree to which they would adhere to face-mask 57 and physical-distancing requirements. We approached this question with an agent-based model 58 originally developed for pandemic influenza [30], which we tailored to SARS-CoV-2 [19, 31, 32, 59 33, 21, 34] and applied to a geographically and demographically realistic synthetic population 60 representing Indiana. In addition to presenting outcomes across a range of hypothetical scenarios, 61 we calibrated our model to data from the fall to assess the plausibility of K-12 school reopening 62 as a driver of the observed resurgence of SARS-CoV-2 in Indiana during fall 2020. 63

$_{64}$ 2 Methods

65 2.1 Approach

Our approach to modeling SARS-CoV-2 transmission was based on the Framework for Recon-66 structing Epidemic Dynamics (FRED) [30], an agent-based model that offers the ability to explore 67 the impacts of complex, non-pharmaceutical interventions in a natural way through modifications 68 to individual behaviors. Using this model, we simulated the spread of SARS-CoV-2 in Indiana 69 using a synthetic population with demographic and geographic characteristics of the state's real 70 population, including age, household composition, household location, and occupation [35]. We 71 analyzed the impact of school reopening from August 24 (first day of classes in Marion County, 72 the most populous) to December 31, 2020 in the population of Indiana as a whole, as well as in 73 students, teachers, and their cohabitants. We quantified impact as the difference in the num-74 ber of COVID-19 infections, symptomatic infections, and deaths between each scenario and a 75 baseline scenario, the details of which differed according to which comparisons were of interest. 76

77 2.2 Agent-based model

Our model was based on a synthetic population of the entire state of Indiana [35], which included 78 1.3 million K-12 students, 1.7 million people living with students, and 6.3 million people in total 79 (Table S10). Each of these agents visits a set of places defined by their activity space, which can 80 include houses, schools, workplaces, long-term care facilities, and various neighborhood locations. 81 Transmission can occur when an infected agent visits the same location as a susceptible agent on 82 the same day, with numbers of contacts per agent specific to each location type. For example, 83 school contacts depend not on the size of the school but on the age of the student and their 84 assigned school grade, given that students have a higher number of contacts with students in 85 their classroom than with those in other classrooms. Every day of the week, students and 86 teachers visit their school, and students are assigned to classrooms based on their age. Given 87 that schools are closed during the weekends, community contact is increased by 50% for students 88 and teachers on weekends, unless they are sheltering in place [30]. Community contacts are 89 modeled as contacts with other agents in the same neighborhood. For both schools and other 90 locations, we adopted contact rates for each location type that were previously calibrated to 91 influenza attack rates specific to each location type [30, 36]. 92

Once infected, each agent had a latent period (mean = 3.35 days, standard deviation = 1.1693 d) and an infectious period (mean = 3.7 d, s.d. = 1.2 d) drawn from distributions calibrated 94 so that the average generation interval distribution matched estimates from Singapore (mean =95 5.20 d, s.d. = 1.72 d [33]. The absolute risk of transmission depended on the number and loca-96 tion of an infected agent's contacts and a parameter that controls SARS-CoV-2 transmissibility 97 upon contact, which we calibrated. A proportion of the infections were asymptomatic, with the 98 probability of symptoms increasing with age [37, 19, 38] (Table S9, Fig. S4). We assumed that 99 these infections were as infectious as symptomatic infections and had identical incubation and 100 infectious period distributions [39, 40, 26, 41]. Furthermore, we assumed that children were less 101 susceptible to infection than adults, which we modeled with a logistic function calibrated to 102

¹⁰³ model-based estimates of this relationship by Davies et al. [19]. We assumed that severity of ¹⁰⁴ disease increased with age, consistent with statistical analyses described elsewhere [32, 31, 38, ¹⁰⁵ 21].

Agent behavior in FRED has the potential to change over the course of an epidemic. Following 106 the onset of symptoms, infected agents self-isolate at home according to a fixed daily probability, 107 whereas others continue their daily activities [42, 43]. This probability was chosen so that, on 108 average, 68% of agents will self-isolate at some point during their symptoms. This figure of 68%109 was based on the proportion of symptomatic infections among healthcare workers in the USA 110 that developed fever during the course of their infection [43], and assuming that those with fever 111 are likely to self-isolate. Agents can also engage in a variety of non-pharmaceutical interventions, 112 including school closure, sheltering in place, and a combination of mask-wearing and physical 113 School closures occur on specific dates across the state [44], resulting in students distancing. 114 limiting their activity space to household and neighborhood locations on those days. Within 115 households, agents interacted with their cohabitants on a daily basis. We assumed that agents 116 did not wear face masks inside their homes, nor did they isolate from their household members 117 if infected. To capture temporal changes in overall mobility and community contact over the 118 course of the epidemic, we modeled a time-varying probability of sheltering in place. On days 119 when an agent sheltered in place, they reduced their activity spaces to their home only, whereas 120 other agents continued with their normal routines. 121

We modeled protection from face masks by reducing the probability of transmission when 122 an infected agent wore a face mask. Our default assumption followed a median estimate of an 123 adjusted odds ratio of 0.3 against SARS-CoV in non-healthcare settings [45, 34]. A meta-analysis 124 by Chu et al. [34] included studies in healthcare settings in addition to those from non-healthcare 125 settings, but we included only estimates that referred to the latter. The proportion of agents 126 wearing face masks in workplace and community settings changed over time, and we assumed 127 that agents did not wear face masks inside their households. In addition, students and teachers 128 wore face masks according to probabilities specified as part of school reopening scenarios in our 129 analysis. Further details about the model are available in the Supplementary Text. 130

¹³¹ 2.3 Model calibration and validation

Two of the time-varying drivers of transmission in our model were informed by time-varying data 132 inputs. First, we informed the daily probability of sheltering in place with mobility reports from 133 Google [46]. In doing so, this sheltering-in-place probability in our model accounts for both the 134 effects of shelter-in-place orders and some people deciding to continue staying at home after those 135 orders are lifted [47]. Second, we informed the daily proportion of agents wearing face masks in 136 workplace and community settings with Google Trends data for Indiana using the terms "face 137 mask" and "social distancing" [48]. To inform the magnitude of the proportion of people wearing 138 face masks in workplace and community settings, we used survey data [49] on face-mask usage 139 from a single point in time. 140

The values of nine model parameters (listed in Table S9) were informed by calibrating the model to three time-varying epidemiological data streams corresponding to the state of Indiana daily deaths, daily hospitalizations, and daily test positivity at the state level through August

10—and to the age distribution of cumulative deaths through July 13. We obtained daily in-144 cidence of reported cases and deaths from the New York Times COVID-19 database [1]. Daily 145 hospitalizations and the age distribution of cumulative deaths were obtained from the Indiana 146 COVID-19 dashboard [8]. Daily numbers of tests performed in the state were available from The 147 Covid Project [50]. To calibrate the model to these data, we first used a Sobol design sampling 148 algorithm [51, 52] to draw 6,000 combinations of the nine calibrated parameters. We then calcu-149 lated the likelihood of each parameter combination given the data, and resampled the parameter 150 combinations proportional to their likelihoods to obtain an approximation of the posterior dis-151 tribution of parameter values. Additional details about the calibration procedure are described 152 in the Supplementary Material. 153

To validate the model, we compared its predictions to data withheld from model calibration. Specifically, we compared the calibrated model's predictions of the infection attack rate, both overall and by age, to results from two statewide serological surveys undertaken in late April and early June [53, 54]. We assessed the model's success in this validation exercise by visual comparison of the model's predictions with the data, with a focus on overlap between the 95% prediction intervals of the model and the 95% credible intervals associated with the empiricallyderived estimates from the serological surveys.

¹⁶¹ 2.4 Model outputs

The main outputs from our model were the numbers of infections, symptomatic cases, hospital-162 izations, and deaths at the state level. For different comparisons, these outputs were examined 163 either on a daily basis, cumulatively between August 24 and December 31, or stratified by place 164 of infection (school, home, other) or affiliation with schools (student, teacher, none). Another 165 output that we examined was the daily reproduction number, R(t), which was calculated in the 166 model for each infected agent and averaged across the modeled population. We defined R(t) as 167 the average number of secondary infections caused by any agent who was infected on day t-168 i.e., the case reproduction number a la Fraser [55]. Finally, to account for uncertainty in these 169 outcomes due to uncertainty in the calibrated parameters, we sampled parameter sets from our 170 approximation of the posterior distribution of parameter values. 171

172 2.5 Model scenarios

173 2.5.1 Effects of conditions in schools

To explore how alternative conditions for school reopening could have impacted the statewide 174 burden of COVID-19 in Indiana, we performed simulations that spanned a range of assumptions 175 about school operating capacity and adherence of students and teachers to face masks while in 176 school. We chose to focus on these parameters given that they were two of the major unknowns as 177 the state proceeded with its plans for in-person instruction beginning in August 2020. Given that 178 students were offered the option of either in-person or remote instruction, we evaluated scenarios 179 in which school operating capacity was either 50%, 75%, or 100%. Specifically, this parameter 180 represents the daily probability that a student would go to school, such that all students could go 181

to school at some point during the simulation but the average number of students in attendance on a given day is determined by the school operating capacity parameter. For each of these scenarios about school operating capacity, we also considered scenarios in which the adherence of students and teachers to face masks was either 50%, 75%, or 100%. In addition, we considered a scenario in which schools reopened normally (100% capacity, 0% face-mask adherence) and a scenario in which schools operated remotely (0% capacity, face-mask adherence irrelevant since no one in school).

189 2.5.2 Sensitivity analyses

To explore the sensitivity of our results to model uncertainties, we performed three sets of sensi-190 tivity analysis. First, for each of the nine scenarios comprising our primary analysis, we analyzed 191 the sensitivity of cumulative infections and the proportion of infections acquired in schools to 192 each of the nine calibrated parameters by calculating partial rank correlation coefficients [56]. 193 Second, we considered the sensitivity of our results to values of two assumed parameters not 194 included in the calibration: protection afforded by face masks and the probability of isolation 195 given symptoms. This included a total of four scenarios exploring lower and higher values of 196 each of those parameters. Third, we considered alternative scenarios about select model assump-197 tions that we regarded as potentially important unknowns about transmission of SARS-CoV-2 198 by children. These included scenarios in which asymptomatic infections (which are more likely 199 to occur among children) are half as infectious as symptomatic infections, a scenario in which 200 children aged 0-10 years have lower susceptibility (0.1), and a scenario in which individuals of all 201 ages are equally susceptible to SARS-CoV-2 infection. For each of the alternative scenarios in 202 the second and third sets of sensitivity analyses, we re-calibrated the model under that scenario 203 and simulated it forward for the fall semester under the nine primary scenarios about school 204 operating capacity and adherence to face-masks in schools. These latter two sets of sensitivity 205 analyses all focused on cumulative infections statewide between August 24 and December 31, 206 2020. 207

208 2.5.3 Retrospective analysis

To understand what conditions for school reopening resulted in model behavior consistent with 209 data from fall 2020, we calibrated parameters for school operating capacity and face-mask adher-210 ence in schools to data from this time period. To do so, we held all other calibrated parameters 211 at the values we estimated based on calibration to data from January 1 to August 24, 2020. 212 Whereas we used Google mobility reports to drive a time-varying probability of sheltering in 213 place during that initial period, we opted to assume fixed levels of mobility from August 24 214 through December 31. The main reason for this choice was that the Google data we used for 215 the initial period showed a decrease in mobility during the fall, despite an increase in incidence, 216 making it difficult to explain the epidemiological data under that assumption. To overcome that 217 problem and to make this analysis more directly comparable to our other analyses, we held the 218 probability of sheltering in place fixed at either of two levels of mobility during the fall: 1) mo-219 bility remained at summer levels, or 2) mobility increased to pre-pandemic levels. Under these 220

assumptions, we calibrated the two focal parameters about conditions in schools to data from the fall using the same calibration procedure as we used during the initial period.

223 **3** Results

224 3.1 Model calibration and validation

Our model was generally consistent with the data to which it was calibrated, capturing trends over 225 time in daily deaths, hospitalizations, and test positivity at the state level (Fig. 1A-C), as well as 226 greater proportions of deaths among older age groups (Fig. 1D). Some trade-offs in the model's 227 ability to recreate different data types were apparent, such as a recent increase in hospitalizations 228 that the model failed to capture (Fig.1C), likely due to the predominance of data on deaths in the 229 likelihood. Similarly, the model underestimated the proportion of deaths in people older than 80 230 (Fig. 1B), indicating a possible underestimation of the contacts in this age group in the overall 231 community or in long-term care facilities. Even so, the model's predictions reproduced the range 232 of variability in the data, as assessed by the coverage probabilities of its 95% posterior predictive 233 intervals (daily deaths: 0.85; daily hospitalizations: 0.93; daily test positivity: 0.95; cumulative 234 deaths by age: 1.0). The model was also consistent with data withheld from fitting. Across 235 all ages, the model's 95% posterior predictive intervals of the cumulative proportion infected 236 through late April (median: 0.017; 95% CrI: 0.0045-0.051) and early June (median: 0.022; 95% 237 CrI: 0.0058-0.069) spanned estimates from two state-wide serological surveys [53] (Fig. 2A). Our 238 model's predictions also overlapped with age-stratified estimates from those surveys (Fig. 2B), 239 although it underpredicted infections among individuals aged 40-60 years. 240

Calibration of the parameter that scaled the magnitude of SARS-CoV-2 importations [57, 58] 241 in our model resulted in a median of 1.30 (95% PPI: 0.50-1.46) imported infections per day from 242 February 1 to August 10. To ensure that the model reliably reproduced the high occurrence of 243 deaths observed in long-term care facilities, we seeded infections into those facilities at a daily 244 rate proportional to the prevalence of infection on that day; this calibrated proportion was 0.037 245 (95% PPI: 0.022-0.092). On the opposite end of the age spectrum, our calibration resulted in 246 a median estimate of susceptibility among children of 0.346 (95% CrI: 0.311-0.506), compared 247 to 0.834 (95% CrI: 0.652-0.946) in adults (Fig. S1). Our calibration resulted in an estimate of 248 transmissibility (median: 0.593; 95% CrI: 0.501-0.788) that corresponded to values of R(t) during 249 the initial phase of the epidemic in Indiana of 1.73 (95% CrI: 1.11-2.34), which represents an 250 average of daily values across the first two weeks of March (Fig. 1A). This estimate is within the 251 range of other estimates for R(t) that include the state of Indiana [59, 60]. Driven by a calibrated 252 estimate that the proportion of people sheltering in place rose in early March and peaked at a 253 median of 32.1% (95% CrI: 28.8-66.9%) on April 7 (Fig. S2A), our estimates of R(t) dropped to 254 a low of 0.57 (95% CrI: 0.42-0.71) on April 7 and remained below 1 thereafter (Fig. 1A). Also 255 impacting our estimates of R(t) was the increasing use of face masks in the community, which 256 we estimated at 53.4% (95% CrI: 46.1-54.0%) as of July 19 (Fig. S2B). Note that this estimated 257 distribution of community face-mask adherence does not differ between scenarios and is not the 258 same as the level of adherence in schools, which we imposed at different levels depending on the 259

260 scenario.



Figure 1. Model calibration to statewide data: A) daily incidence of death; B) proportion of deaths through July 13 in decadal age bins; C) daily incidence of hospitalization; and D) daily proportion of tests administered that are positive for SARS-CoV-2. In all panels, blue diamonds represent data. In A, C, and D, the gray line is the median, the dark shaded region the 50% posterior predictive interval, and the light shaded region the 95% posterior predictive interval.

261 3.2 Effects of conditions in schools

262 3.2.1 Effects on statewide burden

Under a scenario in which schools reopened at full capacity and without any use of face masks, 263 our model projected that R(t) across the state as a whole would have increased to 1.72 (95%) 264 CrI: 1.43-2.17) by mid-September (Fig. 3A). Given our assumption that levels of sheltering in 265 place and face-mask adherence in the community remained constant during the fall (Fig. S2B), 266 this increase in transmission was driven by infections arising in schools (Fig. 3B). As a result, 267 new infections statewide would have risen to levels in the fall far exceeding those from the spring 268 (Fig. 3C). Under this scenario, our model projected a total of 2.57 million (95% CrI: 2.36-2.88 269 million) infections (Fig. 4A) and 10,246 (95% CrI: 7,862-13,794) deaths (Fig. 4B) from Indiana's 270 population as a whole between August 24 and December 31. 271

Under a scenario in which schools went to remote instruction and all children remained at home, our model projected that R(t) would have remained near levels from August for the remainder of 2020 (Fig. S3A). This was a result of our assumption that sheltering in place and face-mask adherence in the community would have remained at their estimated levels as of August 13. Under this scenario, transmission would have continued through contacts at workplaces, within homes, and elsewhere in the community (Fig. S3B), resulting in a total of



Figure 2. Model comparison with data withheld from fitting. We validated the model's predictions against statewide data withheld from fitting on A) the cumulative proportion of the population of Indiana infected through late April and early June, and B) the cumulative proportion infected among individuals aged 12-40, 40-60, and 60+. Data are shown in navy and come from a random, statewide serological survey [53]. Model predictions are shown in gray. In A, the line and band indicate the median and 95% posterior predictive interval. In B, lines, boxes, and error bars indicate median, interquartile range, and 95% posterior predictive interval.

²⁷⁸ 45,579 (95% CrI: 14,651-132,546) infections (Fig. S3C) and 790 (95% CrI: 176-1,680) deaths ²⁷⁹ from Indiana's population as a whole between August 24 and December 31.

Less extreme scenarios about school operating capacity and face-mask adherence in schools 280 also resulted in a wide range of variation in the projected statewide burden of COVID-19 during 281 fall 2020. Under a scenario in which schools operated at 50% capacity and achieved 100% face-282 mask adherence, the cumulative numbers of infections and deaths that our model projected were 283 similar to projections under a scenario in which schools operated remotely (Fig. 4, Tables S1 284 & S5). In general, cumulative infections and deaths statewide in fall 2020 were more sensitive 285 to school operating capacity than to face-mask adherence in schools, with the worst outcomes 286 projected to occur under a scenario with 100% school operating capacity and 50% face-mask 287 adherence. Under this scenario, cumulative infections statewide were projected to have been 288 42.8 (95% CrI: 41.3-44.3) times greater than if schools had operated remotely, and cumulative 289 deaths statewide were projected to have been 9.2 (95% CrI: 8.9-9.5) times greater (Table S1). 290

²⁹¹ 3.2.2 Effects on risk for individuals affiliated with schools

Relative to a scenario with remote instruction, risk of infection and symptomatic infection was 292 greatest for students (Figs. 5 & S5, left column), with a hundred-fold or greater increase in the 293 risk of infection under a scenario with 100% school operating capacity and 50-75% face-mask 294 adherence in schools (Tables S2 & S6). The risk of symptomatic disease in school-aged children 295 was two-fold lower for children under 10 years of age (Fig. S6). Compared to students, the risk 296 of infection was slightly lower for teachers, and much lower for students' families. Due to their 297 older ages, however, teachers and families experienced a much higher risk of death than students 298 (Figs. 5 & S5, center & right columns). The highest risk of death was for teachers under scenarios 299 with 100% school operating capacity (Figs. 5 & S5, center column). Compared to a scenario 300 with remote instruction, the relative risk of death for teachers under these scenarios ranged from 301 a 41-fold increase when face-mask adherence was 100% to a 166-fold increase when face-mask 302 adherence was 50% (Table S3, S7). Under scenarios with 75% school operating capacity, those 303 same relative risks dropped to a four-fold increase when face-mask adherence was 100% and a 304 22-fold increase when face-mask adherence was 50%. This again illustrates the overall greater 305 effect of school operating capacity than face-mask adherence in schools. 306

307 3.3 Sensitivity analyses

Our sensitivity analysis of the model's nine calibrated parameters quantified the partial rank 308 correlation coefficient (PRCC) of each of two model outputs: cumulative infections statewide 309 from August 24 to December 31, and the proportion of infections acquired in schools. In general, 310 these outputs were most sensitive to parameters controlling the age-susceptibility relationship 311 (Figs. S7 & S8). Under some scenarios, the minimum susceptibility parameter, which applied to 312 young children, had a PRCC as high as 0.6. The transmissibility parameter also had a PRCC 313 that high, but only in some scenarios. For example, in scenarios with lower school operating 314 capacity, the two parameters most relevant to community transmission—transmissibility and 315 face-mask adherence in community settings—had a greater influence on cumulative infections 316

317 statewide (Fig. S7), given that the contribution of schools to transmission was diminished in 318 those scenarios.

Our sensitivity analysis of parameters for protection afforded by face masks and the proba-319 bility of isolation given symptoms focused on cumulative infections statewide in fall 2020. For 320 each alternative value of these parameters, we re-calibrated the model and generated projec-321 tions under our nine primary scenarios about school operating capacity and face-mask adherence 322 in schools. In general, the relative effects of differences in face-mask adherence in schools and 323 school operating capacity were similar under alternative values of protection afforded by face 324 masks (Table S11). However, relative to a baseline with schools operating remotely, the magni-325 tude of the proportional increase in cumulative infections was sensitive to the level of protection 326 afforded by face masks. For example, under a scenario with 75% face-mask adherence and 75%327 school operating capacity, the increase in cumulative infections ranged from 23.5-fold to 1.6-fold 328 across the range of values of protection afforded by face masks that we explored (adjusted odds 329 ratio $= 0.12 \cdot 0.73$). Proportional increases in cumulative infections were generally insensitive to 330 the probability of isolation given symptoms (Table S12). For example, under a scenario with 331 75% face-mask adherence and 75% school operating capacity, the increase in cumulative infec-332 tions ranged from 4.9-fold to 6.9-fold across the range of values of isolation probability that we 333 explored (0.5-0.9). 334

Our sensitivity analysis of assumptions related to the role of children in transmission also 335 focused on cumulative infections statewide in fall 2020. Relative to a baseline with schools 336 operating remotely, the proportional increase in cumulative infections was very similar to our 337 default assumptions under a scenario with lower susceptibility among children aged 0-10 years and 338 a scenario with lower infectiousness of asymptomatic infections (Table S13). This was the case 339 for all scenarios about face-mask adherence in schools and school operating capacity, except for 340 the most extreme case in which face-mask adherence in schools was 50% and school operating 341 capacity was 100%. In that case, our default assumptions resulted in a 42.9-fold increase in 342 cumulative infections, whereas the two alternative scenarios resulted in a 26-fold increase. Under 343 a scenario with equal susceptibility for all ages, proportional increases in cumulative infections 344 were much higher than under our default assumptions (Table S13). Because susceptibility in 345 children was higher under this scenario, transmission statewide was much higher when school 346 reopened, especially in scenarios with higher school operating capacity (Fig. S9). By the same 347 token, prevalence dropped to levels over the summer when school was not in session that were 348 much lower than the data suggest (Fig. S10), which raises doubts about the plausibility of this 349 scenario. 350

351 3.4 Retrospective analysis

The model successfully reproduced statewide data from fall 2020 under relatively high values of school operating capacity and intermediate values of face-mask adherence in schools (Fig. 6). Under an assumption that the probability of sheltering in place in the state as a whole was fixed at levels from summer 2020 (Fig. 6, gray), the model calibration resulted in a median school operating capacity of 0.88 (95% CrI: 0.8-0.95) and a median face-mask adherence in schools of 0.6 (95% CrI: 0.4-0.7). Under an assumption that the probability of sheltering in place was fixed

- at pre-pandemic levels (Fig. 6, red), the model calibration resulted in a median school operating
- $_{359}$ capacity of 0.85 (95% CrI: 0.77-0.92) and a median face-mask adherence in schools of 0.61 (95%
- ³⁶⁰ CrI: 0.41-0.75). Overall, there was a wide range of values of face-mask adherence in schools that
- ³⁶¹ were consistent with the data, and the range of values of school operating capacity consistent
- ³⁶² with the data depended somewhat on face-mask adherence in schools (Fig. 6, bottom left).



Figure 3. The impact of school reopening on August 24 under a scenario with 100% school operating capacity and 0% face-mask adherence in schools. Model outputs shown include: A) the reproduction number, R(t), over time; B) the proportion of infections acquired in different location types (colors) over time; and C) the daily incidence of infection statewide over time. In A and C, the line represents the median, and the shaded region represents the 50% posterior predictive interval.



Figure 4. The impact of different scenarios about conditions for school reopening on A) cumulative infections and B) cumulative deaths in Indiana between August 24 and December 31. Scenarios are defined by school operating capacity (x-axis) and face-mask adherence in schools (shading). Orange lines represent projections under a scenario of school reopening at full capacity without masks (solid: median; dotted: 95% posterior predictive interval). Blue lines represent a scenario where schools operate remotely. Error bars indicate inter-quartile ranges.



Figure 5. The impact of different scenarios about conditions for school reopening on infections (top row), symptomatic infections (middle row), and deaths (bottom row) per 1,000 people. These outcomes are presented separately for students (left column), teachers (middle column), and school-affiliated families (right column). Scenarios are defined by school operating capacity (x-axis) and face-mask adherence in schools (shading). Orange lines represent projections under a scenario of school reopening at full capacity without masks (solid: median; dotted: 95% posterior predictive interval). Blue lines represent a scenario where schools operate remotely. Error bars indicate inter-quartile ranges.



Figure 6. Retrospective analysis of the model calibrated to statewide data from fall 2020. Under two alternative scenarios about the daily probability of sheltering in place (summer vs. pre-pandemic mobility level in gray and red, respectively), we calibrated the parameters for school operating capacity and face-mask adherence in schools to data from August 24 through December 31, 2020 (blue diamonds). The calibrated model's correspondence to daily incidence of death statewide is shown in the top two panels, and values of the calibrated parameters are shown in the bottom panels.

363 4 Discussion

Our model provides a detailed, demographically realistic representation of SARS-CoV-2 trans-364 mission in Indiana that is consistent both with data to which it was calibrated and to data that 365 was withheld from calibration. In contrast to models that rely on assumptions about intervention 366 impacts or estimate them statistically [2, 61], our model makes predictions about intervention 367 impacts based on first-principles assumptions about individual-level behavior and contact pat-368 terns. Consistent with results from other analyses [2, 61, 20], the inputs and assumptions in 369 our model led to a prediction that schools made a considerable contribution to SARS-CoV-2 370 transmission in February and early March, prior to large-scale changes in behavior. Extending 371 that, a primary result of our analysis is that K-12 school reopening was capable of making a 372 considerable contribution to SARS-CoV-2 transmission during fall of 2020, with the degree of 373 that contribution dependent on conditions in schools. 374

The burden of COVID-19 associated with in-person school operation was predicted by our 375 model to fall unevenly across the state's population. In scenarios of school operating capacity of 376 100% with 50% face-mask adherence in schools, our model predicted that hundreds of thousands 377 of children could have been infected during the fall semester, with very few of those resulting in 378 deaths. In contrast, our results show that hundreds of deaths in teachers and school-affiliated 379 families could have occurred. Our model indicates that the burden of COVID-19 in schools, 380 teachers, and school-affiliated families across the state could have been reduced by operating 381 at reduced capacity and achieving high face-mask adherence in schools. Under the relatively 382 optimistic scenario of 50% school operating capacity and 100% face-mask adherence in schools. 383 our model predicted that infections and deaths statewide would have been only 22% greater 384 than under a scenario with fully remote instruction. In contrast, our model results suggest 385 that if schools would have operated at full capacity, infections and deaths statewide could have 386 been one to two orders of magnitude greater than the scenario with fully remote instruction, 387 especially with poor face-mask adherence in schools. When we extended our model calibration to 388 account for data from fall 2020, we found that school operating capacity of 80-95% and face-mask 389 adherence in schools of 40-70% resulted in model predictions most consistent with the observed 390 data. For reference, data from the National COVID-19 School Response Dashboard [62] indicate 391 that school operating capacity in Indiana was 77-83% during the fall. Although conditions 392 elsewhere in the community likely played a role in statewide trends during the fall and were not 393 accounted for fully by our model, these results demonstrate that transmission associated with 394 K-12 school reopening was capable of driving the statewide resurgence of COVID-19 observed in 395 Indiana in fall 2020. 396

The impacts associated with reduced school operating capacity result from reductions in both the number of contacts within the school and the probability that an infected student would be in attendance in the first place, similar to the logic behind why smaller gatherings are associated with reduced risk of transmission [3, 63, 64]. The magnitude of our results was most sensitive to the degree of protection afforded by face masks, which remains uncertain in school and other community settings for SARS-CoV-2 [34]. That uncertainty can be reduced as more studies are conducted. Recently, some studies have shown that face masks offer significant protection in

community settings similar to what we assumed. For example, Payne et al. followed 382 U.S.
Navy service members who reported wearing face masks and found a reduced risk of 70% in
those who did [65]. Similar values were found in studies of 124 households in China [66] and 839
close contacts of 211 index cases in Thailand [67].

Although the scenarios we considered resulted in projected impacts spanning nearly the full 408 range between fully remote instruction and fully in-person instruction with no face masks, they 409 are a simplification of the complexities of how schools likely operated in fall 2020. Scenarios 410 that we did not explore include different groups of students attending in person or remotely [68], 411 varying degrees of modularization within schools [69], and the implementation of testing-based 412 control strategies in schools [70]. In the event that infectiousness is lower for asymptomatic 413 infections, the impact of school reopening on lower grades could have been lower than our results 414 suggest. A related simplification of our statewide analysis is that the state, in reality, consists 415 of a patchwork of policies across districts. In light of this complexity that our model does not 416 capture, our results should be interpreted with caution in relation to specific counties or school 417 districts below the state level. Across all scenarios though, our results illustrate the importance 418 of reduced school operating capacity and maximal face-mask adherence in schools, as do other 419 modeling studies [68, 69, 70, 71, 72]. 420

A critical assumption of our analysis is that children are capable of being infected with SARS-421 CoV-2 and transmitting it to others at meaningful levels. Although the burden of severe disease 422 skews strongly towards older ages [22, 73, 8], there are other lines of evidence that support 423 our assumption. These include a contact-tracing study that found no distinguishable difference 424 between infectivity of children and adults [26], several studies that found no distinguishable 425 difference in viral load between children and adults [74, 40, 39, 75], a study that observed a 426 greater secondary attack rate among children in homes [26], and a modeling study that found 427 no evidence that children were less infectious [76]. More direct evidence comes from COVID-19 428 outbreaks that have been observed in schools, such as one in a high school in Israel in which 429 13.2% of students and 16.6% of staff were infected in just 10 days [77]. Even more pertinent, since 430 schools reopened in September 2020, 31,658 COVID-19 cases have been reported in students and 431 13,240 cases have been reported in teachers and school staff across the state, as of April 2021 [8]. 432 There is now a growing body of evidence that school closures contributed to mitigating the 433 first wave of the epidemic [6, 70] and, as we have shown, may have contributed to the resurgence 434 of SARS-CoV-2 during fall 2020. Our study adds to this evidence, and suggests an even greater 435 impact of school reopening than several other studies [72, 70, 69, 78, 68]. This is due in part to our 436 assumption that asymptomatic and symptomatic infections contribute similarly to transmission 437 [26, 74, 40, 39, 75], and in part to our model's ability to capture chains of transmission within 438 schools and extending out into the community. Our study echoes several modeling studies in 439 emphasizing the importance of reducing school operating capacity to impede transmission [69, 440 70, 71, 72, 78. As schools grapple with COVID-19 going forward, results such as these provide 441 an important basis for motivating the adoption and sustainment of reduced school operating 442 capacity and adherence to face-mask requirements in schools. As we demonstrated, these actions 443 are highly consequential for those directly linked to schools and for the communities in which 444 they are embedded. 445

446 5 Acknowledgements

This work was supported by an NSF RAPID grant to TAP (DEB 2027718), an Arthur J. Schmitt
Fellowship and Eck Institute for Global Health Fellowship to RJO, and a Richard and Peggy
Notabaert Premier Fellowship to MP. We thank the University of Notre Dame Center for Research
Computing for computing resources.

451 References

- The New York Times. "Coronavirus in the U.S.: Latest Map and Case Count". en-US. In:
 The New York Times (July 2020). ISSN: 0362-4331.
- 454 [2] H Juliette T Unwin et al. "State-Level Tracking of COVID-19 in the United States". In:
 455 medRxiv (2020). DOI: 10.1101/2020.07.13.20152355. eprint: https://www.medrxiv.org/
 456 content/early/2020/07/14/2020.07.13.20152355.full.pdf.
- [3] Seth Flaxman et al. "Estimating the Effects of Non-Pharmaceutical Interventions on COVID19 in Europe". In: *Nature* 584.7820 (Aug. 2020), pp. 257–261. ISSN: 1476-4687. DOI: 10.
 1038/s41586-020-2405-7.
- [4] Coronavirus Government Response Tracker. en. https://www.bsg.ox.ac.uk/research/research/
 461 projects/coronavirus-government-response-tracker.
- [5] Hamada S Badr et al. "Association between Mobility Patterns and COVID-19 Transmission
 in the USA: A Mathematical Modelling Study". In: *The Lancet Infectious Diseases* (2020).
 ISSN: 1473-3099. DOI: 10.1016/S1473-3099(20)30553-3.
- [6] Katherine A. Auger et al. "Association between Statewide School Closure and COVID-19
 Incidence and Mortality in the US". In: JAMA (July 2020). ISSN: 0098-7484. DOI: 10.1001/
 jama.2020.14348.
- [7] Charles Courtemanche et al. "Strong Social Distancing Measures In The United States
 Reduced The COVID-19 Growth Rate: Study Evaluates the Impact of Social Distancing
 Measures on the Growth Rate of Confirmed COVID-19 Cases across the United States."
 In: Health Affairs (2020), pp. 10–1377.
- [8] ISDH Novel Coronavirus: Indiana COVID-19 Dashboard. https://www.coronavirus.in.gov/2393.htm.
- [9] Kenne A. Dibner, Heidi A. Schweingruber, and Dimitri A. Christakis. "Reopening K-12
 Schools during the COVID-19 Pandemic: A Report from the National Academies of Sciences, Engineering, and Medicine". In: JAMA (July 2020). ISSN: 0098-7484. DOI: 10.1001/
 jama.2020.14745.
- President Donald J. Trump Is Supporting America's Students and Families By Encouraging
 the Safe Reopening of America's Schools. en-US. https://www.whitehouse.gov/briefings statements/president-donald-j-trump-supporting-americas-students-families-encouraging-safe reopening-americas-schools/.

- [11] Indiana State Department of Health, Indiana Department of Education, and Indiana Fam ily and Social Services Administration. IN-CLASS: Indiana's Considerations for Learning
 and Safe Schools. COVID-19 Health and Safety Re-Entry Guidance. June 2020.
- ⁴⁸⁴ [12] Charlotte Jackson et al. "School Closures and Influenza: Systematic Review of Epidemio⁴⁸⁵ logical Studies". In: *BMJ Open* 3.2 (2013). ISSN: 2044-6055. DOI: 10.1136/bmjopen-2012⁴⁸⁶ 002149. eprint: https://bmjopen.bmj.com/content/3/2/e002149.full.pdf.
- ⁴⁸⁷ [13] Shoko Kawano and Masayuki Kakehashi. "Substantial Impact of School Closure on the
 ⁴⁸⁸ Transmission Dynamics during the Pandemic Flu H1N1-2009 in Oita, Japan". In: *PLOS*⁴⁸⁹ ONE 10.12 (Dec. 2015), pp. 1–15. DOI: 10.1371/journal.pone.0144839.
- [14] Simon Cauchemez et al. "Closure of Schools during an Influenza Pandemic". In: *The Lancet Infectious Diseases* 9.8 (2009), pp. 473–481. ISSN: 1473-3099. DOI: 10.1016/S1473-3099(09)
 70176-8.
- [15] Niel Hens et al. "Estimating the Impact of School Closure on Social Mixing Behaviour
 and the Transmission of Close Contact Infections in Eight European Countries". In: *BMC infectious diseases* 9.1 (2009), p. 187.
- [16] Simon Cauchemez et al. "Estimating the Impact of School Closure on Influenza Transmission from Sentinel Data". In: *Nature* 452.7188 (Apr. 2008), pp. 750–754. ISSN: 1476-4687.
 DOI: 10.1038/nature06732.
- ⁴⁹⁹ [17] Charlotte Jackson et al. "The Effects of School Closures on Influenza Outbreaks and Pan⁵⁰⁰ demics: Systematic Review of Simulation Studies". In: *PLOS ONE* 9.5 (May 2014), pp. 1–
 ⁵⁰¹ 10. DOI: 10.1371/journal.pone.0097297.
- [18] Sten H Vermund and Virginia E Pitzer. "Asymptomatic Transmission and the Infection Fatality Risk for COVID-19: Implications for School Reopening". In: *Clinical Infectious Diseases* (June 2020). ISSN: 1058-4838. DOI: 10.1093/cid/ciaa855.
- ⁵⁰⁵ [19] Nicholas G. Davies et al. "Age-Dependent Effects in the Transmission and Control of COVID-19 Epidemics". In: *Nature Medicine* (June 2020). ISSN: 1546-170X. DOI: 10.1038/ s41591-020-0962-9.
- Juanjuan Zhang et al. "Changes in Contact Patterns Shape the Dynamics of the COVID-19 Outbreak in China". In: Science 368.6498 (2020), pp. 1481–1486. ISSN: 0036-8075. DOI: 10.1126/science.abb8001. eprint: https://science.sciencemag.org/content/368/6498/1481.
 full.pdf.
- Joseph T Wu et al. "Estimating Clinical Severity of COVID-19 from the Transmission
 Dynamics in Wuhan, China". In: *Nature medicine* (2020), pp. 1–5.

Robert Verity et al. "Estimates of the Severity of Coronavirus Disease 2019: A ModelBased Analysis". eng. In: *The Lancet. Infectious Diseases* (Mar. 2020). ISSN: 1474-4457.
DOI: 10.1016/S1473-3099(20)30243-7.

- [23] Willem et al. "The Impact of Contact Tracing and Household Bubbles on Deconfinement Strategies for COVID-19: An Individual-Based Modelling Study". In: medRxiv (2020). DOI: 10.1101/2020.07.01.20144444. eprint: https://www.medrxiv.org/content/early/2020/07/ 06/2020.07.01.20144444.full.pdf.
- [24] Christine M Szablewski. "SARS-CoV-2 Transmission and Infection among Attendees of an
 Overnight Camp—Georgia, June 2020". In: *MMWR. Morbidity and mortality weekly report*69 (2020), pp. 1023–1025. DOI: http://dx.doi.org/10.15585/mmwr.mm6931e1.
- Young Joon Park et al. "Contact Tracing during Coronavirus Disease Outbreak, South Korea, 2020". In: *Emerging Infectious Disease journal* 26.10 (2020). ISSN: 1080-6059. DOI: 10.3201/eid2610.201315.
- Pirous Fateh-Moghadam et al. "Contact Tracing during Phase I of the COVID-19 Pandemic in the Province of Trento, Italy: Key Findings and Recommendations". In: medRxiv (2020).
 DOI: 10.1101/2020.07.16.20127357. eprint: https://www.medrxiv.org/content/early/2020/ 07/29/2020.07.16.20127357.full.pdf.
- [27] George Psacharopoulos et al. "Lost Wages: The COVID-19 Cost of School Closures". In:
 World Bank Policy Research Working Paper 9246 (2020).
- Danielle G. Dooley, Asad Bandealy, and Megan M. Tschudy. "Low-Income Children and Coronavirus Disease 2019 (COVID-19) in the US". In: JAMA Pediatrics (May 2020). ISSN: 2168-6203. DOI: 10.1001/jamapediatrics.2020.2065. eprint: https://jamanetwork.com/
 journals/jamapediatrics/articlepdf/2766115/jamapediatrics\textbackslash_dooley\textb
 ackslash_2020\textbackslash_vp\textbackslash_200021.pdf.
- ⁵³⁸ [29] Wim Van Lancker and Zachary Parolin. "COVID-19, School Closures, and Child Poverty:
 ⁵³⁹ A Social Crisis in the Making". In: *The Lancet Public Health* 5.5 (2020), e243–e244.
- John J Grefenstette et al. "FRED (A Framework for Reconstructing Epidemic Dynamics):
 An Open-Source Software System for Modeling Infectious Diseases and Control Strategies
 Using Census-Based Populations". In: *BMC public health* 13.1 (2013), p. 940.
- [31] Qifang Bi et al. "Epidemiology and Transmission of COVID-19 in 391 Cases and 1286 of
 Their Close Contacts in Shenzhen, China: A Retrospective Cohort Study". In: *The Lancet Infectious Diseases* (2020).
- [32] Stephen A. Lauer et al. "The Incubation Period of Coronavirus Disease 2019 (COVID-19)
 From Publicly Reported Confirmed Cases: Estimation and Application". en. In: Annals of Internal Medicine (Mar. 2020). ISSN: 0003-4819. DOI: 10.7326/M20-0504.
- [33] Tapiwa Ganyani et al. "Estimating the Generation Interval for COVID-19 Based on Symptom Onset Data". en. In: medRxiv (Mar. 2020), p. 2020.03.05.20031815. ISSN: 10.1101/2020.03.05.20031815.
 ⁵⁵¹ DOI: 10.1101/2020.03.05.20031815.
- ⁵⁵² [34] Derek K Chu et al. "Physical Distancing, Face Masks, and Eye Protection to Prevent
 ⁵⁵³ Person-to-Person Transmission of SARS-CoV-2 and COVID-19: A Systematic Review and
 ⁵⁵⁴ Meta-Analysis". In: *The Lancet* (2020).
- 555 [35] W D Wheaton. "2005-2009 U.S. Synthetic Population Ver.2. RTI International". In: (2012).

- Joël Mossong et al. "Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases". In: *PLOS Medicine* 5.3 (Mar. 2008), pp. 1–1. DOI: 10.1371/journal.pmed.
 0050074.
- Kenji Mizumoto et al. "Estimating the Asymptomatic Proportion of Coronavirus Disease
 2019 (COVID-19) Cases on Board the Diamond Princess Cruise Ship, Yokohama, Japan,
 2020". en. In: *Eurosurveillance* 25.10 (Mar. 2020), p. 2000180. ISSN: 1560-7917. DOI: 10.
 2807/1560-7917.ES.2020.25.10.2000180.
- ⁵⁶³ [38] The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. "The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-⁵⁶⁵ 19) — China, 2020". en. In: China CDC Weekly 2.8 (Feb. 2020), pp. 113–122. ISSN: 2096-⁵⁶⁶ 7071.
- Lael M. Yonker et al. "Pediatric SARS-CoV-2: Clinical Presentation, Infectivity, and Immune Responses". In: *The Journal of Pediatrics* (). ISSN: 0022-3476. DOI: 10.1016/j.jpeds.
 2020.08.037.
- Taylor Heald-Sargent et al. "Age-Related Differences in Nasopharyngeal Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Levels in Patients with Mild to Moderate Coronavirus Disease 2019 (COVID-19)". In: JAMA Pediatrics (July 2020). ISSN: 2168-6203. DOI: 10.1001/jamapediatrics.2020.3651. eprint: https://jamanetwork.com/journals/jamapediatrics/articlepdf/2768952/jamapediatrics/textbackslash/_healdsargent/textbac
 kslash/_2020/textbackslash/_ld/textbackslash/_200036.pdf.
- Shixiong Hu et al. "Infectivity, Susceptibility, and Risk Factors Associated with SARS-CoVTransmission under Intensive Contact Tracing in Hunan, China". In: medRxiv (2020).
 DOI: 10.1101/2020.07.23.20160317. eprint: https://www.medrxiv.org/content/early/2020/
 08/07/2020.07.23.20160317.full.pdf.
- T Alex Perkins et al. "Calling in Sick: Impacts of Fever on Intra-Urban Human Mobility".
 In: Proceedings of the Royal Society of London B: Biological Sciences 283.1834 (2016). ISSN:
 0962-8452. DOI: 10.1098/rspb.2016.0390.
- [43] Team CDC COVID et al. "Characteristics of Health Care Personnel with COVID-19-United
 States, February 12-April 9, 2020." In: MMWR. Morbidity and Mortality Weekly Report
 (2020).
- ⁵⁸⁶ [44] "State of Indiana". Executive Order 20-16.
- Jiang Wu et al. "Risk Factors for SARS among Persons without Known Contact with SARS
 Patients, Beijing, China". In: *Emerging Infectious Disease journal* 10.2 (2004), p. 210. ISSN:
 1080-6059. DOI: 10.3201/eid1002.030730.
- ⁵⁹⁰ [46] COVID-19 Community Mobility Report. https://www.google.com/covid19/mobility?hl=en.
- [47] Christopher J Cronin and William N Evans. Private Precaution and Public Restrictions:
 What Drives Social Distancing and Industry Foot Traffic in the COVID-19 Era? Working
- ⁵⁹³ Paper 27531. National Bureau of Economic Research, July 2020. DOI: 10.3386/w27531.
- ⁵⁹⁴ [48] *Google Trends*. en-US. https://trends.google.com/trends/?geo=US.

- [49] New York Times and Dynata. Estimates from The New York Times, Based on Roughly
 250,000 Interviews Conducted by Dynata from July 2 to July 14. Tech. rep.
- ⁵⁹⁷ [50] The COVID Tracking Project. *Michigan*. en. https://covidtracking.com/data/state/michigan.
 ⁵⁹⁸ 2020.
- [51] Aaron A. King et al. pomp: Statistical Inference for Partially Observed Markov Processes.
 Manual. 2020.
- [52] Aaron A. King, Dao Nguyen, and Edward L. Ionides. "Statistical Inference for Partially
 Observed Markov Processes via the R Package pomp". In: *Journal of Statistical Software* 603 69.12 (2016), pp. 1–43. DOI: 10.18637/jss.v069.i12.
- [53] Nir Menachemi et al. "Population Point Prevalence of SARS-CoV-2 Infection Based on
 a Statewide Random Sample—Indiana, April 25–29, 2020". In: Morbidity and Mortality
 Weekly Report 69.29 (2020), p. 960.
- Andrea Zeek. COVID-19 Study's Phase 2 Results Show Spread of Virus in Indiana. en.
 https://news.iu.edu/stories/2020/06/iupui/releases/17-fairbanks-isdh-second-phase-covid 19-testing-indiana-research.html. June 2020.
- ⁶¹⁰ [55] Christophe Fraser. "Estimating Individual and Household Reproduction Numbers in an
 ⁶¹¹ Emerging Epidemic". In: *PLOS ONE* 2.8 (Aug. 2007), pp. 1–12. DOI: 10.1371/journal.
 ⁶¹² pone.0000758.
- [56] Mark Stevenson with contributions from Telmo Nunes et al. epiR: Tools for the Analysis
 of Epidemiological Data. Manual. 2020.
- [57] T. Alex Perkins et al. "Estimating Unobserved SARS-CoV-2 Infections in the United
 States". In: Proceedings of the National Academy of Sciences (2020). ISSN: 0027-8424. DOI:
 10.1073/pnas.2005476117. eprint: https://www.pnas.org/content/early/2020/08/20/
 2005476117.full.pdf.
- [58] MIDAS Network, Midas-Network/COVID-19. GitHub. https://github.com/midas-net-work/COVID 19.
- [59] Jesús Fernández-Villaverde and Charles I Jones. Estimating and Simulating a SIRD Model
 of COVID-19 for Many Countries, States, and Cities. Working Paper 27128. National
 Bureau of Economic Research, May 2020. DOI: 10.3386/w27128.
- [60] George Mohler et al. "Analyzing the Impacts of Public Policy on COVID-19 Transmission: A Case Study of the Role of Model and Dataset Selection Using Data from Indiana". In: *Statistics and Public Policy* 8.1 (2021), pp. 1–8. DOI: 10.1080/2330443X.2020.1859030.
 eprint: https://doi.org/10.1080/2330443X.2020.1859030.
- [61] Susanna Esposito and Nicola Principi. "School Closure during the Coronavirus Disease
 2019 (COVID-19) Pandemic: An Effective Intervention at the Global Level?" In: JAMA *Pediatrics* (May 2020). ISSN: 2168-6203. DOI: 10.1001/jamapediatrics.2020.1892. eprint:
 https://jamanetwork.com/journals/jamapediatrics/articlepdf/2766114/jamapediatrics/te
 xtbackslash_esposito\textbackslash_2020\textbackslash_vp\textbackslash_200019.
 pdf.

[62] COVID-19 Resource Hub: STATS Indiana. https://www.stats.indiana.edu/topic/covid19.asp.

- [63] Nor Fazila Che Mat et al. "A Single Mass Gathering Resulted in Massive Transmission
 of COVID-19 Infections in Malaysia with Further International Spread". In: *Journal of Travel Medicine* 27.3 (Apr. 2020). ISSN: 1708-8305. DOI: 10.1093/jtm/taaa059. eprint:
 https://academic.oup.com/jtm/article-pdf/27/3/taaa059/33226391/taaa059.pdf.
- ⁶³⁹ [64] Sukbin Jang, Si Hyun Han, and Ji-Young Rhee. "Cluster of Coronavirus Disease Associated
 with Fitness Dance Classes, South Korea". eng. In: *Emerging infectious diseases* 26.8 (Aug.
 ⁶⁴¹ 2020), pp. 1917–1920. ISSN: 1080-6059. DOI: 10.3201/eid2608.200633.
- [65] Daniel C Payne et al. "SARS-CoV-2 Infections and Serologic Responses from a Sample
 of US Navy Service Members—USS Theodore Roosevelt, April 2020". In: Morbidity and
 Mortality Weekly Report 69.23 (2020), p. 714.
- [66] Yu Wang et al. "Reduction of Secondary Transmission of SARS-CoV-2 in Households by
 Face Mask Use, Disinfection and Social Distancing: A Cohort Study in Beijing, China". In: *BMJ global health* 5.5 (2020), e002794.
- [67] Pawinee Doung-Ngern et al. "Case-Control Study of Use of Personal Protective Measures
 and Risk for SARS-CoV 2 Infection, Thailand". In: *Emerging Infectious Diseases* 26.11
 (2020), p. 2607.
- [68] Matt J Keeling et al. "The Impact of School Reopening on the Spread of COVID-19 in
 England". In: medRxiv (2020). DOI: 10.1101/2020.06.04.20121434. eprint: https://www.
 medrxiv.org/content/early/2020/06/05/2020.06.04.20121434.full.pdf.
- [69] Jennifer R Head et al. "The Effect of School Closures and Reopening Strategies on COVID19 Infection Dynamics in the San Francisco Bay Area: A Cross-Sectional Survey and Modeling Analysis". In: medRxiv (2020). DOI: 10.1101/2020.08.06.20169797. eprint: https:
 //www.medrxiv.org/content/early/2020/08/07/2020.08.06.20169797.full.pdf.
- [70] Jasmina Panovska-Griffiths et al. "Determining the Optimal Strategy for Reopening Schools,
 the Impact of Test and Trace Interventions, and the Risk of Occurrence of a Second COVID19 Epidemic Wave in the UK: A Modelling Study". In: *The Lancet Child & Adolescent Health* (2020). ISSN: 2352-4642. DOI: 10.1016/S2352-4642(20)30250-9.
- [71] Benjamin Lee et al. "Modeling the Impact of School Reopening on SARS-CoV-2 Transmission". In: Research Square (2020).
- 664[72]Alfonso Landeros et al. "An Examination of School Reopening Strategies during the SARS-665CoV-2 Pandemic". In: medRxiv (2020). DOI: 10.1101/2020.08.05.20169086. eprint: https:666//www.medrxiv.org/content/early/2020/08/06/2020.08.05.20169086.full.pdf.
- 667 [73] COVID-19 Hospitalizations. https://gis.cdc.gov/grasp/COVIDNet/COVID19_3.html.
- [74] Enrico Lavezzo et al. "Suppression of a SARS-CoV-2 Outbreak in the Italian Municipality of
 Vo". In: Nature 584.7821 (Aug. 2020), pp. 425–429. ISSN: 1476-4687. DOI: 10.1038/s41586 020-2488-1.

- [75] Terry C Jones et al. "An Analysis of SARS-CoV-2 Viral Load by Patient Age". In: medRxiv
 (2020). DOI: 10.1101/2020.06.08.20125484. eprint: https://www.medrxiv.org/content/
 early/2020/06/09/2020.06.08.20125484.full.pdf.
- [76] Itai Dattner et al. "The Role of Children in the Spread of COVID-19: Using Household
 ⁶⁷⁵ Data from Bnei Brak, Israel, to Estimate the Relative Susceptibility and Infectivity of
 ⁶⁷⁶ Children". In: medRxiv (2020). DOI: 10.1101/2020.06.03.20121145. eprint: https://www.
 ⁶⁷⁷ medrxiv.org/content/early/2020/06/05/2020.06.03.20121145.full.pdf.
- 678 [77] Chen Stein-Zamir et al. "A Large COVID-19 Outbreak in a High School 10 Days after
 679 Schools' Reopening, Israel, May 2020". In: *Eurosurveillance* 25.29 (2020), p. 2001352. DOI:
 680 10.2807/1560-7917.ES.2020.25.29.2001352.
- [78] Elaheh Abdollahi et al. "Simulating the Effect of School Closure during COVID-19 Outbreaks in Ontario, Canada". In: *BMC Medicine* 18.1 (July 2020), p. 230. ISSN: 1741-7015.
 DOI: 10.1186/s12916-020-01705-8.
- 684 [79] World Population Prospects Population Division United Nations. https://population.un.org/wpp/DataQu
- [80] Jon C Emery et al. "The Contribution of Asymptomatic SARS-CoV-2 Infections to Trans-
- mission on the Diamond Princess Cruise Ship". In: *eLife* 9 (Aug. 2020). Ed. by Marc Lipsitch, Eduardo Franco, and Marc Lipsitch, e58699. ISSN: 2050-084X. DOI: 10.7554/ eLife.58699.
- [81] IHME COVID-19 health service utilization forecasting Team and Christopher JL Murray.
 "Forecasting COVID-19 Impact on Hospital Bed-Days, ICU-Days, Ventilator-Days and Deaths by US State in the next 4 Months". en. In: medRxiv (Mar. 2020), p. 2020.03.27.20043752.
 DOI: 10.1101/2020.03.27.20043752.

¹ Supplementary material



Figure S1. Susceptibility to SARS-CoV-2 infection by age. The black line shows the median of the estimated curve of susceptibility by age, which took the form of a modified logistic function defined by four parameters: minimum, maximum, inflection point, and slope. The gray band represents the 95% credible interval. Navy lines show estimates from Davies et al. [19] that were used to inform our estimates. The dashed green line shows an alternative scenario with lower susceptibility for children under 10 years of age.



Figure S2. Changes in A) mobility patterns over time and B) face-mask adherence in the community as a whole. In A, the gray line shows the fitted pattern for the Google mobility index related to residential locations (navy line). Adherence to shelter-in-place in the model follows the same trend, but with its magnitude estimated through the model calibration process. In B, the gray line is informed by fitting to Google search data on "face mask" and assuming that values plateau from July 19 onward at a value informed by survey data [48, 49].



Figure S3. The impact of remote instruction (i.e., 0% school operating capacity). Model outputs shown include: A) the reproduction number, R(t), over time; B) the proportion of infections acquired in different location types (colors) over time; and C) the daily incidence of infection statewide over time. In A and C, the line represents the median, and the shaded region represents the 50% posterior predictive interval.



Figure S4. Distributional representations of select model parameters: A) probability of symptoms by age group; B) cumulative probability of the duration of the incubation period; C) cumulative probability of the duration of the symptomatic period; D) relative probability of death by age; E) cumulative probability of death by day after symptom onset; and F) density of duration of the pre-infectious and infectious periods.



Figure S5. The impact of different scenarios about conditions for school reopening on infections (top row), symptomatic infections (middle row), and deaths (bottom row) cumulatively between August 24 and December 31 across the state of Indiana. These outcomes are presented separately for students (left column), teachers (middle column), and school-affiliated families (right column). Scenarios are defined by school operating capacity (x-axis) and face-mask adherence in schools (shading). Orange lines represent projections under a scenario of school reopening at full capacity without masks (solid: median; dotted: 95% posterior predictive interval). Blue lines represent a scenario where schools operate remotely. Error bars indicate inter-quartile ranges.



Figure S6. The impact of different scenarios about conditions for school reopening on cumulative symptomatic infections per 1,000 people for ages 0-10 (left column) and 10-20 (right column). Results are shown under: (top row) the baseline assumption for susceptibility and infectiousness of asymptomatic infections; (middle row) lower susceptibility for children under 10 years of age; and (bottom row) the baseline assumption for susceptibility but with lower infectiousness of asymptomatic infections. Scenarios are defined by school operating capacity (x-axis) and face-mask adherence in schools (shading). Orange lines represent projections under a scenario of school reopening at full capacity without masks (solid: median; dotted: 95% posterior predictive interval). Blue lines represent a scenario where schools operate remotely. Error bars indicate inter-quartile ranges.



Figure S7. Sensitivity analysis of cumulative infections statewide between August 24 and December 31 to variation in the model's nine calibrated parameters. Bars indicate values of the partial rank correlation coefficient. Results are presented separately by school operating capacity (panels) and face-mask adherence in schools (gray shading).



Figure S8. Sensitivity analysis of the proportion of infections acquired in schools between August 24 and December 31 to variation in the model's nine calibrated parameters. Bars indicate values of the partial rank correlation coefficient. Results are presented separately by school operating capacity (panels) and face-mask adherence in schools (gray shading).



Figure S9. Comparison of median daily incidence of infection statewide in Indiana under alternative model assumptions and conditions in schools. Default model assumptions (gray) are contrasted with an alternative assumption of equal susceptibility for all ages (blue). A baseline scenario of schools operating remotely (light) is contrasted with two scenarios with schools reopened (dark): school operating capacity at 50% (left) or 75% (right). Face-mask adherence in schools was assumed to be 100% in both scenarios. The date on which schools reopened in the fall (August 24) is indicated by the vertical line.



Figure S10. Model calibration to statewide data under an alternative assumption of equal susceptibility for all ages: A) daily incidence of death; B) proportion of deaths through July 13 in decadal age bins; C) daily incidence of hospitalization; and D) daily proportion of tests administered that are positive for SARS-CoV-2. In all panels, blue diamonds represent data. In A, C, and D, the gray line is the median, the dark shaded region the 50% posterior predictive interval, and the light shaded region the 95% posterior predictive interval.

Capacity	Facemask	Cumulative infections	Cumulative deaths
0.50	0.50	1.56(1.48-1.64)	1.19(1.14-1.24)
0.50	0.75	1.34(1.28-1.41)	1.13(1.09-1.18)
0.50	1.00	1.22(1.16-1.28)	1.09(1.04-1.13)
0.75	0.50	10.66 (10.06-11.28)	2.64(2.53-2.75)
0.75	0.75	$4.94 \ (4.63-5.28)$	$1.81 \ (1.74 - 1.89)$
0.75	1.00	2.39(2.26-2.52)	1.39(1.33-1.45)
1.00	0.50	42.81 (41.44-44.44)	$9.22 \ (8.91 - 9.55)$
1.00	0.75	30.92 (29.82-32.10)	6.18(5.96-6.43)
1.00	1.00	14.31 (13.58-15.10)	3.30(3.18-3.44)

Table S1. Proportional increase in risk of infection and death for the overall population of Indiana under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the proportional increase in risk as the ratio of the cumulative number of events (infections or deaths) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate remotely.

Capacity	Facemask	Student infections	Student symptoms
0.50	0.50	3.3(3.1-3.5)	3.1(2.9-3.3)
0.50	0.75	2.4(2.3-2.5)	2.3(2.2-2.4)
0.50	1.00	1.8(1.7-1.9)	1.8(1.7-1.9)
0.75	0.50	56.5(53.1-60.1)	49.0 (45.6-52.3)
0.75	0.75	21.8(20.3-23.6)	18.9(17.6-20.3)
0.75	1.00	7.4(6.9-7.9)	6.6(6.2-7.1)
1.00	0.50	251.0(241.2-260.9)	236.4(227.8-246.7)
1.00	0.75	178.2 (170.1-186.1)	162.8 (155.5-169.7)
1.00	1.00	76.6 (72.1-81.3)	67.7 (63.6-71.7)

Table S2. Proportional increase in risk of infection and symptomatic infection for students under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the proportional increase in risk as the ratio of the cumulative number of events (infections or symptomatic infections) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate remotely.

Capacity	Facemask	Teacher infections	Teacher symptoms	Teacher deaths
0.50	0.50	2.1 (2.0-2.2)	2.0 (1.9-2.2)	2.0 (1.9-2.2)
0.50	0.75	1.7(1.6-1.8)	1.7(1.6-1.8)	1.8(1.7-2.0)
0.50	1.00	1.5(1.4-1.6)	1.5(1.4-1.6)	1.6(1.5-1.7)
0.75	0.50	24.8 (23.2-26.5)	23.6(22.1-25.1)	22.0(20.2-23.7)
0.75	0.75	10.7 (10.0-11.5)	10.1 (9.4-10.9)	9.8 (9.1-10.6)
0.75	1.00	4.5(4.2-4.8)	4.4(4.1-4.7)	4.4(4.1-4.7)
1.00	0.50	119.5 (114.6-125.0)	126.7 (120.8-132.9)	$166.1 \ (157.4-175.6)$
1.00	0.75	88.8 (84.9-93.1)	90.2 (86.0-94.8)	102.7 (96.6-109.6)
1.00	1.00	41.8(39.5-44.3)	41.0(38.5-43.7)	41.1 (38.5-44.1)

Table S3. Proportional increase in risk of infection, symptomatic infection, and death for teachers under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the proportional increase in risk as the ratio of the cumulative number of events (infections, symptomatic infections, or deaths) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate remotely.

Capacity	Facemask	Family infections	Family symptoms	Family deaths
0.50	0.50	3.5 (3.3-3.7)	3.3 (3.1-3.5)	2.8 (2.6-3.0)
0.50	0.75	2.5(2.3-2.6)	2.4(2.2-2.5)	2.3(2.2-2.5)
0.50	1.00	1.9(1.8-2.0)	1.9(1.8-2.0)	1.9(1.7-2.0)
0.75	0.50	55.8(52.2-59.5)	49.3(46.0-52.5)	32.5(30.0-35.5)
0.75	0.75	$22.1 \ (20.5-23.6)$	19.4(18.0-20.8)	13.3(12.3-14.4)
0.75	1.00	7.7 (7.2-8.3)	7.0 (6.6-7.6)	5.7(5.3-6.1)
1.00	0.50	253.1 (242.4-263.4)	249.1 (239.9-259.3)	222.7 (211.2-236.5)
1.00	0.75	177.7 (169.9-186.8)	168.6 (161.4-176.2)	$132.0\ (124.4-140.0)$
1.00	1.00	77.0 (72.6-81.7)	69.5(65.4-73.8)	48.4 (44.8-52.2)

Table S4. Proportional increase in risk of infection, symptomatic infection, and death for family members of students and teachers under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the proportional increase in risk as the ratio of the cumulative number of events (infections, symptomatic infections, or deaths) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate remotely.

Capacity	Facemask	Cumulative infections	Cumulative deaths
0.50	0.50	$0.03 \ (0.03 - 0.03)$	$0.09 \ (0.09-0.10)$
0.50	0.75	$0.03 \ (0.03 - 0.03)$	$0.09 \ (0.09 \text{-} 0.09)$
0.50	1.00	$0.02 \ (0.02 - 0.02)$	$0.08 \ (0.08-0.09)$
0.75	0.50	$0.21 \ (0.20 - 0.22)$	$0.21 \ (0.20 - 0.21)$
0.75	0.75	$0.10 \ (0.09-0.10)$	$0.14 \ (0.14 - 0.15)$
0.75	1.00	$0.05\ (0.04-0.05)$	0.11 (0.11 - 0.11)
1.00	0.50	0.84(0.83-0.85)	0.72(0.71-0.73)
1.00	0.75	$0.61 \ (0.60-0.62)$	$0.48 \ (0.47 - 0.49)$
1.00	1.00	$0.28 \ (0.27 - 0.29)$	$0.26 \ (0.25 - 0.26)$

Table S5. Risk ratio for the overall population of Indiana under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the risk ratio as the ratio of the cumulative number of events (infections or deaths) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate at full capacity and without face masks.

Capacity	Facemask	Student infections	Student symptoms
0.50	0.50	$0.011 \ (0.011 - 0.012)$	0.011 (0.010-0.011)
0.50	0.75	$0.008 \ (0.008 - 0.008)$	0.008 (0.008-0.008)
0.50	1.00	$0.006 \ (0.006 - 0.006)$	0.006 (0.006-0.006)
0.75	0.50	$0.186\ (0.176-0.196)$	0.169(0.161-0.178)
0.75	0.75	$0.072 \ (0.068 - 0.077)$	0.065 (0.061 - 0.070)
0.75	1.00	$0.024 \ (0.023 - 0.026)$	0.023 (0.022-0.024)
1.00	0.50	0.828(0.823 - 0.834)	0.817 (0.811-0.822)
1.00	0.75	$0.587 \ (0.577 - 0.599)$	0.562(0.550-0.573)
1.00	1.00	$0.253 \ (0.241 - 0.265)$	0.234 (0.223-0.245)

Table S6. Risk ratio for students under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the risk ratio as the ratio of the cumulative number of events (infections or symptomatic infections) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate at full capacity and without face masks.

Capacity	Facemask	Teacher infections	Teacher symptoms	Teacher deaths
0.50	0.50	$0.015 \ (0.014 - 0.016)$	0.014 (0.013-0.014)	0.009 (0.008-0.009)
0.50	0.75	$0.012 \ (0.012 - 0.013)$	$0.011 \ (0.011 - 0.012)$	$0.008 \ (0.008 - 0.008)$
0.50	1.00	0.011 (0.010-0.011)	$0.010 \ (0.010 - 0.010)$	0.007 (0.007 - 0.007)
0.75	0.50	0.178(0.170-0.187)	0.157 (0.149 - 0.164)	0.095 (0.090 - 0.100)
0.75	0.75	0.077 (0.072 - 0.082)	$0.067 \ (0.063 - 0.071)$	$0.043 \ (0.041 - 0.045)$
0.75	1.00	0.032(0.031-0.034)	0.029(0.027 - 0.031)	0.019 (0.018-0.020)
1.00	0.50	$0.858 \ (0.852 - 0.863)$	0.839(0.832 - 0.844)	0.723(0.711-0.733)
1.00	0.75	0.637 (0.626 - 0.648)	$0.598 \ (0.585 - 0.610)$	0.447 (0.433 - 0.461)
1.00	1.00	0.300(0.288-0.313)	$0.271 \ (0.259 - 0.283)$	0.179(0.170-0.188)

Table S7. Risk ratio for teachers under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the risk ratio as the ratio of the cumulative number of events (infections, symptomatic infections, or deaths) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate at full capacity and without face masks.

Capacity	Facemask	Family infections	Family symptoms	Family deaths
0.50	0.50	$0.005 \ (0.005 - 0.005)$	$0.005 \ (0.004 - 0.005)$	$0.004 \ (0.004 - 0.005)$
0.50	0.75	$0.003 \ (0.003 - 0.003)$	$0.003 \ (0.003 - 0.003)$	$0.003 \ (0.003 - 0.003)$
0.50	1.00	$0.002 \ (0.002 - 0.002)$	$0.002 \ (0.002 - 0.002)$	$0.003 \ (0.003 - 0.003)$
0.75	0.50	$0.081 \ (0.077 - 0.086)$	$0.069 \ (0.066 - 0.072)$	$0.042 \ (0.039 - 0.044)$
0.75	0.75	$0.015 \ (0.014 - 0.015)$	0.013 (0.013 - 0.014)	$0.011 \ (0.011 - 0.012)$
0.75	1.00	$0.004 \ (0.004 - 0.004)$	$0.004 \ (0.004 - 0.004)$	$0.004 \ (0.004 - 0.005)$
1.00	0.50	0.705 (0.694 - 0.716)	0.673 (0.661 - 0.685)	0.533 (0.519 - 0.546)
1.00	0.75	$0.241 \ (0.232 - 0.251)$	0.210 (0.201-0.218)	0.128(0.122 - 0.133)
1.00	1.00	$0.015 \ (0.015 - 0.016)$	$0.014 \ (0.013 - 0.015)$	$0.012 \ (0.012 - 0.013)$

Table S8. Risk ratio for family members of students and teachers under alternative scenarios about school operating capacity and face-mask adherence in schools. We define the risk ratio as the ratio of the cumulative number of events (infections or deaths) between August 24 and December 31 under the scenario indicated by the left two columns as compared to a scenario in which schools operate at full capacity and without face masks.

Parameter	Value
Importations scaling factor	1.299 (95% CrI: 0.502-1.461)
Susceptibility slope	0.757 (95% CrI: 0.578-3.344)
Susceptibility minimum	0.346 (95% CrI: 0.311-0.506)
Susceptibility inflection point	18.640 (95% CrI: 18.521-19.804)
Susceptibility maximum	$0.834 \ (95\% \ CrI: \ 0.652-0.946)$
Shelter in place adherence	0.321 (95% CrI: 0.288-0.669)
Face-mask adherence	$0.534 \ (95\% \ CrI: \ 0.461-0.540)$
LTC importations scaling factor	0.037 (95% CrI: 0.022-0.092)
Transmissibility	0.593 (95% CrI: 0.501-0.788)
Probability of symptoms (by age)	Fig. S4A
Incubation period	Fig. S4B
Symptomatic period	Fig. S4C
Probability of death (by age)	Fig. S4D
Probability of death after symptom onset	Fig. S4E
Pre-infectious period	meanlog = 23.7, sdlog = 50 (Fig. S4F)
Infectious period	meanlog = 33.8, sdlog = 43.1 (Fig. S4F)
Adjusted odds ratio for face-mask efficacy	0.30 [34]
School contacts	0.6 [30, 36]
Classroom contacts	1.2 [30, 36]
Workplace contacts	0.06 [30, 36]
Office contacts	0.13 [30, 36]
Household contacts	0.14 [30, 36]
Neighborhhood contacts	0.78 [30, 36]

Table S9. List of model parameters. The first nine were estimated through the model calibration process, and the subsequent fourteen were assumed based on the literature, as described in the Methods. Values of the first nine parameters reflect marginal posterior estimates.

Group	Size
Schools of size 0-100	945
Schools of size 100-500	703
Schools of size $1000+$	291
Schools of size 500-1000	830
State population	$6,\!286,\!795$
Students in grades 1 to 12	1,094,916
Students in kinder	72,210
Students in prek	$73,\!695$
Students in ungraded	57,222
Total 0 - 10	$903,\!156$
Total 10 - 20	$907,\!120$
Total 20+	$4,\!476,\!519$
Total living with students	$1,\!682,\!904$
Total students	$1,\!298,\!043$

Table S10. Summary of key features of the synthetic population of Indiana.

		Face-mask adjusted odds ratio			
Capacity	Facemask	0.12	0.30	0.73	
0.50	0.50	1.3(1.2-1.3)	1.6 (1.5-1.6)	3.7(3.6-3.9)	
0.50	0.75	1.1 (1.1-1.2)	1.3(1.3-1.4)	3.4(3.2-3.5)	
0.50	1.00	1.1 (1.0-1.1)	1.2(1.2-1.3)	3.1 (2.9-3.2)	
0.75	0.50	3.3(3.1-3.4)	10.7 (10.1-11.3)	24.9(24.0-25.7)	
0.75	0.75	1.6 (1.5-1.6)	4.9(4.6-5.3)	23.5(22.7-24.2)	
0.75	1.00	1.2 (1.1-1.2)	2.4(2.3-2.5)	21.9(21.1-22.7)	
1.00	0.50	29.5(28.7-30.4)	42.8(41.5-44.3)	33.5(32.3-34.7)	
1.00	0.75	5.6(5.3-5.9)	31.0(29.8-32.1)	33.1 (32.0-34.2)	
1.00	1.00	1.5(1.4-1.5)	14.3 (13.6-15.2)	32.7 (31.7-33.9)	

Table S11. Proportional increases in cumulative infections statewide during fall 2020 under alternative values of face-mask protection, defined as the odds ratio (columns) and alternative scenarios about school operating capacity and face-mask adherence in schools (rows). These increases are relative to a scenario in which schools operated remotely in fall 2020.

		Settings related to probability of isolation			
Capacity	Facemask	0.5	0.68	0.9	
0.50	0.50	1.8 (1.8-1.9)	1.6 (1.5-1.6)	1.8 (1.6-2.0)	
0.50	0.75	1.5(1.4-1.6)	1.3(1.3-1.4)	1.5(1.4-1.7)	
0.50	1.00	1.3(1.2-1.3)	1.2(1.2-1.3)	1.3(1.2-1.5)	
0.75	0.50	$13.1\ (12.6-13.7)$	$10.6\ (10.1-11.3)$	10.6 (9.6-11.6)	
0.75	0.75	6.9(6.5-7.3)	4.9(4.7-5.2)	6.2(5.6-6.8)	
0.75	1.00	3.3(3.1-3.5)	2.4(2.3-2.5)	3.3(3.0-3.6)	
1.00	0.50	34.1 (33.1-35.2)	42.9(41.2-44.4)	31.3(29.1-33.7)	
1.00	0.75	26.7 (25.8-27.6)	30.9 (29.8-32.2)	23.3(21.5-25.0)	
1.00	1.00	14.4 (13.8-15.1)	14.3 (13.6-15.2)	12.7 (11.6-13.9)	

Table S12. Proportional increases in cumulative infections statewide during fall 2020 under alternative values of the probability of isolation given symptoms (columns) and alternative scenarios about school operating capacity and face-mask adherence in schools (rows). These increases are relative to a scenario in which schools operated remotely in fall 2020.

		Settings related to juvenile transmission			
Capacity	Facemask	Default	Equal Susc.	Low 0-10 Susc.	Low Asymp.
					Inf.
0.50	Low	1.6(1.5-1.6)	6.1 (5.8-6.4)	1.5 (1.4-1.6)	1.7 (1.7-1.8)
0.50	Med	1.3(1.3-1.4)	4.2(4.0-4.4)	1.3(1.3-1.4)	1.5(1.4-1.5)
0.50	High	1.2(1.2-1.3)	3.0 (2.9-3.1)	1.2(1.1-1.3)	1.3(1.3-1.4)
0.75	Low	10.7 (10.0-11.3)	348.0 (335.2-	9.3 (8.8-9.8)	10.0 (9.5-10.4)
			361.7)		
0.75	Med	4.9(4.6-5.2)	265.3 (254.7-	4.6 (4.4-4.9)	5.9(5.7-6.2)
			276.8)		
0.75	High	2.4(2.3-2.5)	181.6 (174.0-	2.3(2.2-2.4)	3.2(3.0-3.3)
			189.9)		
1.00	Low	42.9(41.4-44.5)	687.5 (670.0-	26.5(25.4-27.7)	26.2 (25.5-26.9)
			706.1)		
1.00	Med	31.0 (29.8-32.2)	669.3 (652.3-	20.8 (19.9-21.7)	21.0 (20.3-21.6)
			686.8)		
1.00	High	14.3 (13.6-15.1)	641.9 (624.5-	11.3 (10.8-11.9)	13.2(12.7-13.8)
			659.7)		

Table S13. Proportional increases in cumulative infections statewide during fall 2020 under alternative assumptions about juvenile transmission (columns) and alternative scenarios about school operating capacity and face-mask adherence in schools (rows). These increases are relative to a scenario in which schools operated remotely in fall 2020.

² Supplementary methods

³ Model description

4 People and their contacts

FRED simulates pathogen spread in a population by recreating interactions among people on a daily basis. To realistically represent the population of Indiana, we drew on a synthetic 6 population of the US that represents demographic and geographic characteristics from 2010 [35]. Each human is modeled as an agent that visits a set of places defined by their activity space. This activity space contains places such as houses, schools, workplaces, and neighborhood locations. Transmission can occur when an infected person visits the same location as a susceptible person 10 on the same day, with numbers of contacts per person specific to each location type. For instance, 11 school contacts depend not on the size of the school but on the age of the student. We adopted 12 contact rates specific to each location type that were previously calibrated to attack rates for 13 influenza in each location type [30, 36]. 14

15 Importation to seed local transmission

To initialize the model, we simulated international and domestic importations similar to Perkins 16 et al. [57]. First, we obtained data on internationally and domestically imported deaths in 17 Indiana up to March 18 [58], which we used to extrapolate total international and domestic im-18 portations based on the case fatality risk [38], the proportion of infections that are asymptomatic 19 [37], and the probability of detecting local and international symptomatic infections [57]. Second, 20 we assigned times to internationally imported infections proportional to international incidence 21 patterns, adjusted to account for the timing of a ban on travel from China. We assigned times to 22 domestically imported infections proportional to total US incidence. Drawing from uncertainty 23 distributions for each of the three aforementioned parameters, we repeated this process 1,000 24 times and averaged across replicates. We used that average curve to seed our model, scaling 25 its magnitude with a parameter that we calibrated. Although importations from outside Indi-26 ana likely continued beyond those that we were able to account for explicitly, we assumed that 27 transmission within Indiana was sufficient at that point to be the primary driver of incidence. 28 In addition to importations in the overall population, we simulated importation into long-term 29 care facilities, given the large number of deaths that took place there and the limited realism of 30 our model in simulating visitors to those facilities. We introduced infections into these facilities 31 at a constant rate that we calibrated. 32

33 Transmission and disease progression

Once infected, each individual had latent and infectious periods drawn from distributions calibrated so that the average generation interval distribution matched estimates from Singapore $(\mu = 5.20, \sigma = 1.72)$ [33]. The absolute risk of transmission depended on the number and location of an infected individual's contacts and a parameter that controls SARS-CoV-2 transmissibility upon contact, which we calibrated. We assumed asymptomatic infections were as infectious as

symptomatic infections and had identical timing of infectiousness [39, 40, 26, 41]. Following 39 exposure, we assumed that children were less susceptible to infection than adults, which we 40 modeled with a modified logistic function calibrated to results of Davies et al. [19]. We defined 41 four parameters of this function as the *minimum* susceptibility, the *maximum* susceptibility, the 42 inflection point of susceptibility with respect to age, and the slope of the age-susceptibility re-43 lationship around the inflection point. For agents that developed symptoms, we took random 44 draws from lognormal distributions for the incubation period [32] and duration of symptoms 45 [31]. Both the probabilities of developing symptoms and dving [21] were assumed to increase 46 with age. The probability of developing symptoms by age was estimated using the age distribu-47 tion of cases in the China CDC report from early in the pandemic[38] and the age distribution 48 of the population in China [79]. By assuming equal exposure by age and assuming that 50% of 49 infections proceed asymptomatically (an intermediate value of two estimates from the Diamond 50 Princess [37, 80]), we can calculate the proportion of symptomatic infections we would expect 51 in each age group as under our assumptions, the total number of infections in an age-group will 52 be proportional to the population size in that age-group. For infections that resulted in death, 53 we modeled the time to death with a gamma distribution [22] truncated at the 99th percentile. 54 These and other parameters are summarized in Table S9 and Fig. S4. 55

Changes in agent behavior during the epidemic 56

Agent behavior in FRED has the potential to change over the course of an epidemic. Following 57 the onset of symptoms, infected agents self-isolate at home according to a fixed daily rate, 58 whereas others continue their daily activities [42]. This rate is chosen so that on average 68% of 59 agents will self-isolate at some point during their symptoms, assuming that all individuals who 60 develop a fever will isolate at some point during their symptoms, based on published studies in 61 the U.S. [43]. Agents can also respond to public health interventions, including school closure, 62 shelter in place, and a combination of mask-wearing and social distancing. School closures 63 occur on specific dates [81], resulting in students limiting their activity space to household and 64 neighborhood locations. Shelter-in-place interventions reduce some agents' activity spaces to 65 their households only, whereas others continue with their daily routines. We used mobility 66 reports from Google [46] to drive daily compliance with shelter-in-place, such that shelter-in-67 place compliance in our model accounts for both the effects of shelter-in-place orders and some 68 people deciding to continue staying at home after those orders are lifted [47]. To account for 69 voluntary mask-wearing and social distancing, we used Google Trends data for Indiana using the 70 terms "face mask" and "social distancing" [48] and used estimates on face-mask adherence from 71 a New York Times analysis of a survey from Dynata [49]. 72

Model calibration 73

We selected nine parameters to estimate based on calibration of the model to four data types 74

on COVID-19 in Indiana: daily incidence of death, age distribution of deaths, daily incidence of 75 hospitalization, and daily test positivity. The initial ranges for the statewide and long-term care 76

facility importations were adjusted to cover a wide range of values. Compliance with shelter-

77

in-place was informed with changes in mobility patterns in the Google community reports [46]. 78 We fitted a GAM to the trends from the percentage change on mobility trends for places of 79 residence, and projected the compliance of shelter-in-place orders after the period for which we 80 had data by assuming a linear trend thereafter. We normalized these mobility trends from 0%81 (baseline) to 100% (everyone at home) and adjusted its magnitude with a parameter representing 82 the maximum compliance in the historical trends. The minimum, maximum, inflection point, 83 and slope of the logistic function with which we model the age-susceptibility relationship were 84 calibrated to estimates by Davies et al. [19]. 85

⁸⁶ We simulated 6,000 combinations of these nine parameters, $\vec{\theta}$, using a sobol design sampling ⁸⁷ algorithm with the sobolDesign function in R [51, 52]. For each parameter set, we calculated the ⁸⁸ likelihood of the model given the observed data on daily incidence of death, cumulative deaths ⁸⁹ in long-term care facilities through July 13, the decadal age distribution of cumulative deaths ⁹⁰ through July 13, daily incidence of hospitalization, and test positivity.

We calculated the contribution to the likelihood for daily incidence of death and cumulative deaths in long-term care facilities using a negative binomial distribution as

$$\mathcal{L}(\vec{\theta}|D_{t,k}) = \text{Negative Binomial}(r, p),$$

where $D_{t,k}$ is the daily incidence of death on day t and location k (long-term care facilities or all other locations), and r and p are size and probability parameters, respectively. We informed r and p using the conjugate prior relationship between a beta prior and negative binomial likelihood, such that $r = r_{prior} + d_{t,m}$ and $p = 1/(1 + \frac{p_{prior}}{p_{prior}+1})$, where $d_{t,m}$ is the daily incidence of death predicted by the model on day t. For the decadal age distribution of cumulative deaths through July 13, we used a multinomial distribution, such that

$$\mathcal{L}(\vec{\theta}|D_a) = \text{Multinomial}(D_a, d_{a,m} / \sum_a d_a),$$

where D_a is the observed number of deaths in the age group a, and $d_{a,m}$ are the deaths by age group obtained by the model. To fit to data on testing, we first observe, using Bayes' rule, that

$$P(C|T) = \frac{P(T|C)P(C)}{P(T|C)P(C) + P(T|\neg C)P(\neg C)} = \frac{P(C)}{P(C) + r(1 - P(C))},$$

where C refers to a symptomatic case, T refers to an administered PCR test for current infection, and $r = P(T|\neg C)/P(T|C)$. Next, we assume that non-symptomatic infections (either presymptomatic or asymptomatic) exhibit treatment-seeking behavior similar to uninfected individuals, or $P(T|I) = P(T|U) = P(T|\neg C)$, where I refers to a non-symptomatic infection and U to uninfected. We then observe, again using Bayes' rule, that

$$P(I|T) = \frac{rP(I)}{P(C) + r(1 - P(C))}$$

106 and

$$P(U|T) = \frac{rP(U)}{P(C) + r(1 - P(C))}$$

Next, we incorporate PCR sensitivity and specificity by assuming that sensitivity = (P|C) = P(P|I), where P refers to a positive test (i.e., we assume that PCR sensitivity is similar for non-symptomatic and symptomatic infections). This allows us to write

$$P(P|T) = \text{sensitivity}(P(C|T) + P(I|T)) + (1 - \text{specificity})P(U|T)$$

Then, we are in a position to write the contribution to the likelihood from the testing data, assuming that the number of positive tests in the data, T_+ , follows a binomial distribution

$$\mathcal{L}(\theta|T_+, T_-) = \text{Binomial}(T_+ + T_-, P(P|T)),$$

¹¹² where T_{-} represents the number of negative tests in the data.

¹¹³ Finally, the combined log-likelihood was obtained as

$$\log(\mathcal{L}(\vec{\theta})) = \sum_{t} \left(\log(\mathcal{L}(\vec{\theta}|D_{t,overall})) \right) + \log(\mathcal{L}(\vec{\theta}|D_{longtermcare}) + \log(\mathcal{L}(\vec{\theta}|T_{+},T_{-})) + \sum_{a} \left(\log(\mathcal{L}(\vec{\theta}|D_{a})) \right) + \log(\mathcal{L}(\vec{\theta}|D_{a})) +$$

¹¹⁴ We sampled the parameters proportional to their likelihood to obtain a set of parameter combi-

¹¹⁵ nations that constitute our approximation of the posterior distribution of parameter values.