Modeling for COVID-19 College Reopening Decisions: Cornell, A Case Study

Peter I. Frazier[a,1,3], **J. Massey Cashore**[a,2], **Ning Duan**[a,2], **Shane G. Henderson**[a,2], **Alyf Janmohamed**[a,2], **Brian Liu**[a,2], **David B. Shmoys**[a,2], **Jiayue Wan**[a,2], **and Yujia Zhang**[b,2]

[a]School of Operations Research & Information Engineering, Cornell University, Ithaca, NY 14850; [b]Center for Applied Mathematics, Cornell University, Ithaca, NY 14850

This manuscript was compiled on July 8, 2021

We consider the use of epidemiological models to support college reopening decisions during the COVID-19 pandemic: whether to reopen for in-person instruction and, if so, what interventions to implement. The central challenge in this use of mathematical modeling is sensitivity of predictions to input parameters coupled with uncertainty about these parameters. Substantial uncertainty remains today and was even larger when decisions were made for the fall 2020 semester. Moreover, universities' unique characteristics hinder translation of outcomes, models and parameters from the general population: a high fraction of young people, who have higher rates of asymptomatic disease and social contact, intermixing of these young people with more vulnerable individuals, and an enhanced ability to implement behavioral and testing interventions. We describe how epidemiological models supported Cornell University's decision to reopen for in-person instruction in fall 2020 and supported the design of an asymptomatic screening program instituted concurrently to prevent viral spread. We demonstrate how the structure of these decisions allowed risk to be minimized despite parameter uncertainty and how this generalizes to other university settings. We find, in particular, that twice-per-week asymptomatic screening of undergraduates provides robust protection against COVID-19 across a range of parameter settings, including parameters derived from a retrospective analysis based on data collected in Cornell's fall 2020 semester. In this retrospective analysis, parameter uncertainty did indeed cause substantial differences between predicted and realized outcomes but the quality of the decisions made was nonetheless high when compared to the best decisions that could have been made in hindsight.

COVID-19 | Epidemiological Modeling | Parameter Uncertainty | Asymptomatic Screening

Can universities safely reopen for in-person instruction during the COVID-19 pandemic? If so, then how? Universities across the globe faced this question in summer 2020 and face similar questions today as they contemplate in-person instruction with partially vaccinated student populations and more transmissible variants with the potential for immune escape in fall 2021. This question has significant consequences since virtual instruction degrades educational and mental health outcomes (1) but virus outbreaks in student populations threaten the health of students, more vulnerable staff and faculty that interact with them, and community members. This question was challenging to answer in summer 2020, in part because experiences at the city, state and national level do not easily generalize to university populations. Indeed, university populations are younger than the general population and thus have increased rates of contact (2) that may elevate virus transmission (3, 4). Mitigating this risk but adding to the question's difficulty, universities could choose from a broad collection of interventions, many of which would be difficult or impossible to implement for the general population: mandatory testing of students upon arrival to campus (5, 6); social distancing measures in and out of the classroom (7); behavioral contracts (8); travel restrictions; contact tracing; and asymptomatic screening (9, 10).

Universities have responded to this central question in dramatically different ways. Many went completely online in fall 2020 and spring 2021 while some opened with a few modest interventions (11). Others, like Cornell University's Ithaca campus (12), opened with an aggressive set of interventions including social distancing, asymptomatic surveillance and travel restrictions. This diversity in approaches reflects, in part, a diversity of circumstance, such as proximity and interaction with population centers, prevalence in those population centers, availability of housing to quarantine students, and the desires of the surrounding community, all of which should be considered (13). However, it also reflects a fundamental lack of knowledge about how policy translates into outcomes. Even today, with the opportunity to look back on the 2020-2021 academic year, the extent to which university outcomes are explained by interventions, circumstance, differences in under-reporting bias, or luck is not completely clear. Such uncertainty and diversity in approach among universities reflects the larger response to the pandemic, in which US states and national governments adopted dramatically different responses

Significance Statement

- The decision of whether to reopen universities directly impacts 7% of the US population (students, staff) and indirectly impacts tens of millions more (families, communities).
- After witnessing large COVID-19 outbreaks among students in the 2020-21 academic year, universities are planning for future semesters despite uncertainty about vaccine hesitancy, more transmissible variants with the potential for immune escape, and community prevalence. They must decide whether to bring students back for instruction and, if they do, what interventions to implement.
- While uncertainty in input parameters required by epidemiological models introduces substantial uncertainty in outcomes, we show that models can nonetheless provide insight into college reopening decisions that minimize risk.

Frazier led the study. Henderson and Shmoys provided additional modeling and research guidance. Cashore, Duan, Janmohamed, Liu, Wan and Zhang provided additional modeling and computation.

[1]Lead author. [2]All others contributed equally.

[3]To whom correspondence should be addressed. E-mail: pf98cornell.edu

to the pandemic despite apparently similar circumstances.

Epidemic models would seem to offer the power to resolve this uncertainty in support of high-quality decisions. They allow prediction, customized to the circumstances of a university, city, state, or nation. By varying the interventions *in silico* and observing predicted outcomes one can hope to choose the best course of action.

Ostensibly, this strategy requires predictions to be accurate. Unfortunately, epidemic models only approximate reality (14). Ever-present uncertainty in model input parameters coupled with the potential for exponential growth significantly limit accuracy. Small differences in behavioral and biological parameters can cause huge differences in predicted case counts. As a consequence, epidemic models have been maligned for producing inaccurate point estimates (14, 15).

In this article we demonstrate that, perhaps paradoxically, simulation models can support effective selection of COVID-19 interventions even when they are unable to provide accurate point estimates of epidemic outcomes. (The use of epidemic models in the presence of significant parameter uncertainty is also discussed in, e.g., (16). In such settings, clear communication of uncertainties is key; see, e.g., (17).) We demonstrate this in the specific context of deciding whether Cornell University's Ithaca campus should reopen for the fall 2020 semester. In this context, we conducted a simulation-based analysis in summer 2020 using a compartmental SEIR model with multiple subpopulations; see (7, 18, 19) for closely related models. Our work contributed heavily to the decision to reopen the campus for residential instruction (12) and was used in myriad other ways including the choice of screening frequencies within an asymptomatic screening program that was a critical part of Cornell's strategy, and sizing of quarantine capacity.

Our approach hinges on delineating those simulation model input parameters yielding epidemics that can be successfully controlled from those that cannot. If the set of plausible input parameters are contained within the set of safe parameters, then we can be highly confident, though never certain, that the epidemic can be controlled. At Cornell in summer 2020, we demonstrated this to the university administration for a suite of interventions available with in-person instruction: frequent asymptomatic screening, testing students on arrival to campus, contact tracing, social distancing on campus, limits on student and employee travel, masking requirements, and a behavioral compact curtailing student social gatherings. It was also possible that we would have found that plausible ranges of the input parameters overlapped the portion of parameter space where epidemics would grow out of control, in which case we would not have been able to recommend reopening.

We found that access to regular asymptomatic screening (9, 20), with an ability to increase testing frequency if needed, was critical. Asymptomatic surveillance was enabled at Cornell through a major effort to support large-scale sample collection, to retool a veterinary lab to enable high-volume PCR testing, and the use of pooled testing. Our analysis and interventions were customized to Cornell's environment, but the results would likely generalize to other universities and other settings. Indeed, those few universities employing a similar asymptomatic screening approach succeeded by and large in controlling campus outbreaks (21–24). See also (25–30) for explorations of the interaction of pooled testing and asymptomatic surveillance for controlling epidemics.

We also found it was critical to analyze epidemic growth if in-person instruction were not offered, to quantify the relative merits of the alternative to in-person instruction. Survey results (12, 31) suggested that a significant number of students would return to the Ithaca area even if in-person instruction were not offered. Without the benefits of the legal framework offered by in-person instruction, frequent asymptomatic screening would have been difficult to mandate for this population. Moreover, our analysis suggested that many of those parameter settings in which asymptomatic screening would not ensure safe in-person instruction would also be ones in which a significant outbreak would occur in the local student population under virtual instruction. This resulted in the decision to reopen Cornell's Ithaca campus (12).

The purpose of this paper is to explain how epidemic modeling, coupled with careful analysis of input parameters, can provide key insights to those deciding how to keep campuses and their surrounding communities safe. More broadly, these insights are of value to those using epidemic models to support decisions in other contexts during the COVID-19 pandemic and future epidemics. We view the key contributions of this paper to be as follows.

- We provide a modeling and decision-making framework for helping universities decide whether and how to open. Central is a careful accounting for uncertainty. A recommendation to reopen relies on plausible parameter ranges falling within a portion of the input parameter space corresponding to epidemics that can be controlled with the intended interventions. The same methods apply to modifying interventions midway through a semester.
- Central to this question is: what is the alternative? At Cornell, an analysis suggesting health risks would be higher if the campus opened for online-only instruction led to the decision to reopen for fall in-person instruction.
- We highlight the value of asymptomatic surveillance, which in addition to enabling early isolation of positive cases, also provides near-real-time information on infection levels. This permits recourse decisions such as increased testing frequency, stricter social-distancing, or in an extreme case, moving to all-online instruction.
- Our modeling structure can be, and was, employed to provide insight on a host of additional decisions, e.g., what frequency should be used to test each campus subpopulation, what is the needed quarantine capacity, and after the end of the semester should asymptomatic surveillance continue for the remaining campus population.
- We evaluate this framework by providing a retrospective analysis of what happened when Cornell reopened and revisit which aspects of our model were accurate, which aspects were not, and what consequences this had for the quality of the decisions that were made.
- We highlight the impact of sensitivity analyses that showed the efficacy of contact tracing. This led to the development and use at Cornell of *adaptive testing*, i.e., testing all individuals in similar social circles as a positive case, and not just those identified as close contacts.
- We explore those qualities of Cornell and the surrounding Ithaca community that enabled a successful fall semester.

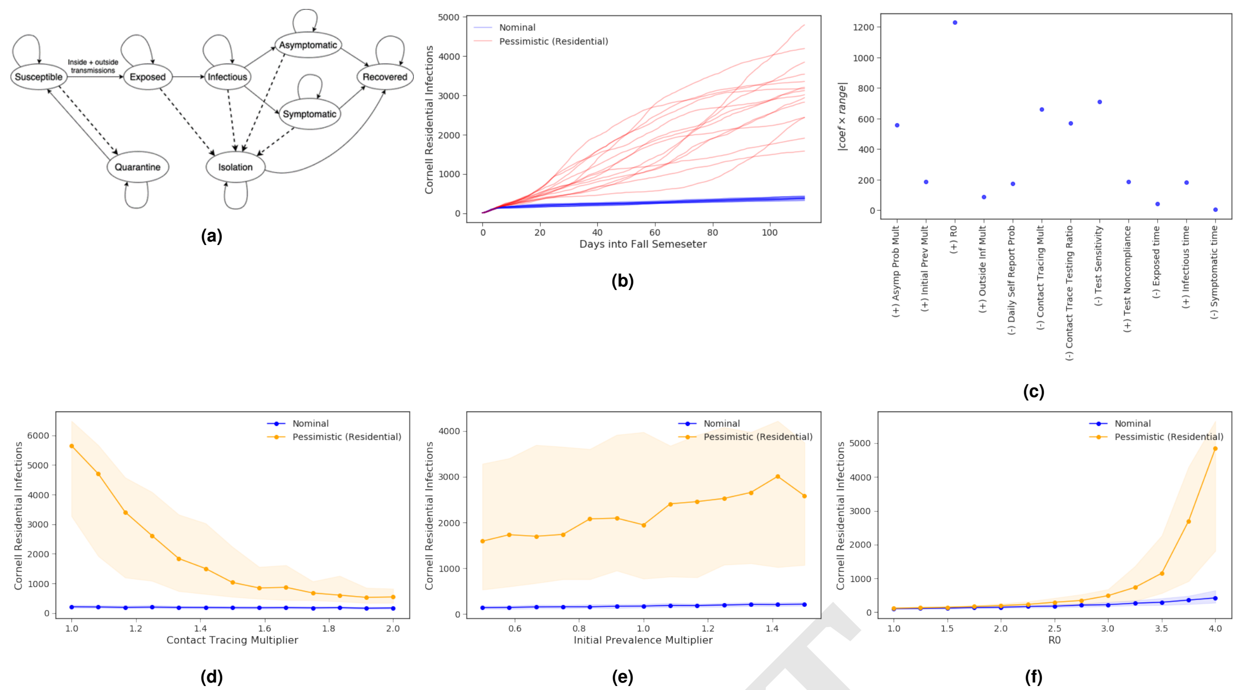


Fig. 1. (a) The dynamics of our compartmental simulation across compartment categories (ellipses). Population counts are maintained for each compartment on each day, and compartments comprise a category, e.g., "susceptible", a demographic group, e.g., undergraduates in high-density housing, and the elapsed time in that compartment category. Solid lines represent virus transmission, disease progression, and the end of quarantine/isolation. Dashed lines represent the effect of testing, self-reported symptoms and contact-tracing, which put individuals testing positive into isolation and their contacts into quarantine. (b) Sample cumulative infection counts over time, under the nominal and pessimistic (for residential infections) parameter scenarios. (c) The first-order effect of parameter uncertainty on predicted infections, using a linear model to estimate sensitivity of infections to each parameter and a range of plausible parameter values. Each dot shows the estimated effect of uncertainty on predicted infections (absolute value of the regression coefficient times the uncertainty range's width) with the sign of the regression coefficient indicated in the label. The uncertainty of this estimated effect (derived from the regression coefficient's 95% confidence interval) is approximately 60 for all parameters. (d-f) Plots depict the number of infections (lines provide the median; shading indicates the 10-90th percentile range across simulation replications) as three key parameters vary while holding the others fixed, under nominal and pessimistic scenarios.

Our work adds to the broader literature using epidemic modeling in the context of universities. See, e.g., (32) for a perspective on the challenges of reopening as informed by a variety of epidemic models, (33, 34) for the use of agent-based modeling to evaluate mitigation strategies to enable safe in-person instruction, (35) for probabilistic modeling of strategies to suppress virus spread in dorms and classrooms, and (36) for a study of interventions for generic small residential campuses.

1. Results

A. Model Predictions and Parameter Sensitivity. We focus on Cornell's main campus, located in Ithaca, New York (37). Ithaca is located in Tompkins County with a population of approximately 102,000. Approximately 12,000 undergraduates (of which 35% live on campus), 6,800 graduate students, and 10,000 employees study or work on the Ithaca campus. Ithaca is 4-5 hours by car from major city centers such as New York, Boston, Philadelphia, and Toronto.

In June 2020 we developed a compartmental simulation model to predict infections and hospitalizations for Cornell's fall 2020 semester (Fig 1a). The model contains compartments by subpopulation, stages of symptom development and infectivity, and quarantine / isolation status (Methods A-C). Fig 1b illustrates the model's output under two different sets of input parameters, simulating a surveillance testing program similar to the one Cornell used in fall 2020. In the simulation results pictured, all undergraduates and student-facing employees were tested twice a week, grad students and non-student facing employees once a week, and off-campus employees monthly. Infections in the Cornell population are roughly proportional to other outcomes of importance (SI 3.E) and therefore figures here focus on this outcome. Since our model is stochastic, there is variance in case trajectories with fixed parameters.

A core challenge was significant uncertainty about underlying input parameters, given that critical epidemiological quantities had not been measured accurately in college populations for SARS-CoV-2 in June 2020. To represent this uncertainty, we identified ranges containing plausible values for each simulation parameter (SI 3.A) based on information available in June 2020.

To understand sensitivity of model predictions to uncertain parameters, we then evaluated residential infections at 2000 parameter configurations chosen using a Latin Hypercube design (38) over the collective set of parameters defined by these plausible ranges (while also supporting a sensitivity analysis over four additional parameters used for predicting outcomes under virtual instruction, SI 3.D). We fit a linear model to quantify the first-order impact of each parameter on infections. We then multiply the estimated rate (from this linear model) at which residential infections change as we vary a parameter by the width of the range quantifying uncertainty for this parameter (Fig 1c and SI Table S17). To first order, this value is the change in predicted infections resulting from moving the parameter from the lower to the upper bound of its corresponding range. Parameters for which this value is largest are those that lead to the greatest uncertainty about infections,

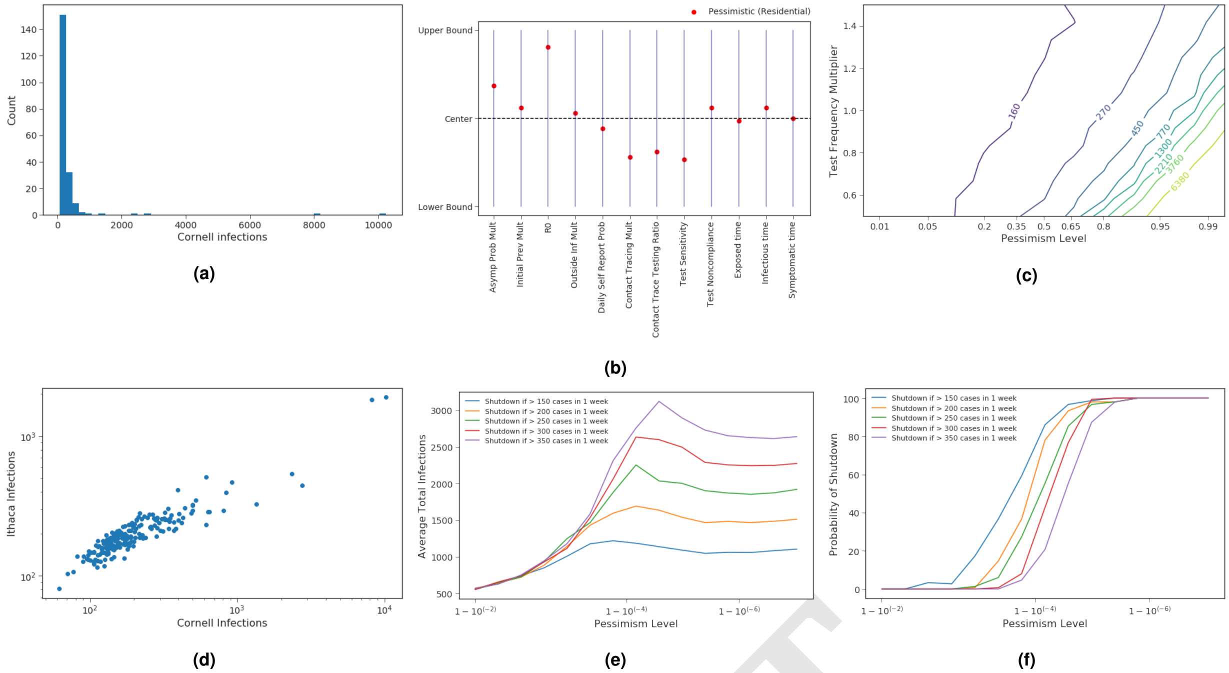


Fig. 2. (a) Histogram of median Cornell infections when parameters are sampled from the prior. (b) Parameter values of the pessimistic (for residential infections) scenario when the range of each parameter has been normalized. (c) Contour plot showing the number of Cornell infections as test frequency and the level of pessimism changes. (d) Scatter plot showing the median number of Cornell and Ithaca infections for each of the 200 points sampled from the prior. (e) The average number of infections and (f) probability of shutdown under each of a set of shutdown policies that model the more nuanced decision-making process available to leaders in practice.

whether because our uncertainty about the parameter is large or because outcomes are sensitive to it.

The effect of uncertainty is substantial, with uncertainty about several individual parameters creating uncertainties of more than 500 infections relative to a baseline of approximately 250 infections.The parameters that most drive uncertainty about infections are those that influence 1) transmission of the virus, especially R_0 as well as two contact tracing parameters and a parameter governing the likelihood an infected individual develops symptoms, and 2) our ability to control it through testing (test sensitivity). These parameters are described in detail in the SI's Section 1.

To understand nonlinear dependence on parameters and interactions across parameters, we additionally plotted outcomes as we vary each parameter individually while holding the other parameters fixed. Fig 1d-f shows three of the parameters for which uncertainty has the largest effect on outcomes, while others are shown in the SI. In each plot, the two lines correspond to two different parameter scenarios described in detail in the next section.

B. Coping with Parameter Uncertainty. Uncertainty about parameters presented the central challenge when deciding whether it would be safe to bring students back to campus. As demonstrated through sensitivity analysis (Results A), simulation outcomes are sensitive to parameters that were unknown in June 2020. This prevented accurate point estimates for the number of infections that would result from a particular set of interventions. A central theme of this article, however, is that accurate point estimates are not a prerequisite for supporting good decisions through modeling.

We show that asymptomatic screening is a powerful measure that can be brought to bear against this challenge of parameter uncertainty. Any fixed test frequency prevents widespread epidemic growth over a set of parameter settings. A larger test frequency results in a larger set of safe parameter settings.

To understand whether a candidate value of 2x / week asymptomatic screening would be enough to make residential instruction safe under plausible parameter values, we formed a Bayesian prior over parameters consisting of independent normal distributions for each parameter. The (marginal) mean and variance of the prior over each parameter was chosen so that the resulting symmetric 95% Bayesian credible interval corresponded to the previously-selected plausible ranges for each parameter (Results A, SI 3.A). The nominal scenario consists of setting each parameter to its mean.

We then drew random sets of parameters from this Bayesian prior and ran our simulation for each, to form a prior distribution over infections accounting for parameter uncertainty (Fig 2a). In most parameter settings, 2x / week testing is sufficient to achieve substantial infection control but there are some parameter settings where large outbreaks occur.

To better understand the impact of interventions like testing on robustness to parameter uncertainty, we developed a one-dimensional family of parameter configurations with varying levels of pessimism about the number of infections (Methods D, SI 3.A). This family of parameter configurations is indexed by a pessimism level between 0 and 1, with larger levels corresponding to parameter configurations with more infections. The parameter configuration at pessimism level q is the most likely configuration under the prior for which median infections is equal to the q-quantile of infections under the prior, assuming that infections for a given parameter configuration are given by the previously fitted linear model. The nominal scenario corresponds to pessimism level 0.5. We additionally define the "pessimistic scenario" as the parameter configuration corresponding to pessimism level 0.99, indicating it as "pessimistic (residential)" in plots to distinguish it from a scenario that is pessimistic about a different outcome defined in Results C. Fig 2b summarizes how the pessimistic and nominal scenarios differ. Relative to the nominal scenario, the pessimistic scenario significantly increases the asymptomatic ratio and transmission rate while decreasing test sensitivity and contact tracing effectiveness. These parameters also have the largest absolute normalized effect on infections (Fig 1c).

We then plotted infections using 2x / week testing at parameter configurations across a range of pessimism levels (Fig 2c at test frequency multiplier = 1.0; Fig 3c). This level of testing is sufficient to keep the number of infections below 1000 in all but the most pessimistic parameter configurations. Still, there are some parameter configurations where more than 1000 infections arise, and infections grow rapidly as the parameter configuration grows more pessimistic.

We hypothesized that additional testing could help mitigate this risk. We performed simulation experiments varying the testing frequency and the parameter configuration's pessimism level (Fig 2c). When the test frequency is low, a pessimistic configuration results in many infections (e.g., more than 6K infections at 1x / week testing at a pessimism level of .95). When the test frequency is high enough, however, predicted infections remain low even under such pessimistic scenarios.

Plots of infections as a function of test frequency and simulation parameters were distributed in public reports (39), at several Faculty Senate meetings (40), and in other open fora. These reports also included nominal and pessimistic scenarios (SI 3.B) similar to the ones detailed here, though highlighting more concerning outcomes. They were an important component of deliberation at Cornell on whether testing could allow a safe residential reopening. They resulted in the decision, at Cornell, that if the campus were to open for residential instruction, we would use a test frequency that was as large as could be provided reliably, to maximize the range of parameter settings under which Cornell's strategy would provide effective infection control.

While we focus here on a single outcome, infections in the campus population, other outcomes are important: infections in the surrounding community created by clusters in the campus population; and hospitalization and deaths. As shown in Fig 2d and SI 3.E these outcomes tend to move together when varying parameters and the overall frequency of testing. Thus, for the purposes of understanding the overall level of risk and deciding whether to reopen, considering only on-campus infections tends to produce the same decisions as would more holistic consideration of on-campus and community outcomes examining infections, hospitalizations and deaths.

In addition to testing, which was instituted at the start of the semester, a second measure combating parameter uncertainty is *recourse*: the opportunity to change decisions based on data as it is observed. Among the decisions that can be made in this way, smaller interventions include changes to behavioral interventions and modifications to test frequency. A few large interventions are also available: shutting the campus down, taking strong temporary measures to reduce social contact, and even sending students home. The effectiveness of the above-mentioned temporary measures was of primary interest in supporting the decision of whether to reopen Cornell's campus. Indeed, a reopening decision that was *likely* to be safe but not guaranteed to be so under a fixed set of interventions (testing frequency) can be made more safe by developing a plan up front. Whether such safety options exist in the event that estimates are wrong is an important component of deciding whether to bring students back at all.

Modeling decision-makers' ability to halt in-person instruction and shut down the campus upon observing substantial transmission, Fig 2e-f shows outcomes under plans of the form "initiate campus shutdown if confirmed cases exceed more than Y cases in X weeks" as a function of the location on the nominal-pessimistic line. While simpler than the more nuanced decision-making process available to leaders in practice, these plans take a simple interpretable form to support broad understanding, and are reminiscent of CUSUM charts used for monitoring industrial processes for defects (41). They essentially "learn" whether reality is such that infection control is not being provided, and then respond when a threshold is reached. By varying X and Y, one achieves different tradeoffs between acting quickly and "false alarms" in which the campus would be shut down when infection control would have been achieved if it had stayed open. Fig 2e-f demonstrates that adding recourse through shutting down can reduce the risk associated with reopening: in the unlikely event that transmission is high enough that the chosen test frequency does not prevent spread, then a prompt shutdown can mitigate the negative health consequences.

C. Virtual vs. Residential Instruction. While substantial focus was given to what would happen if universities reopened for residential instruction in the late summer of 2020 (32–36), an equally important consideration is what would happen if they did *not*. If a university chooses to offer only virtual instruction in place of in-person classes, many students may still elect to return to the university's local area. This was the case at Cornell, where a significant number of students had signed leases with local landlords for the fall before the severity of the pandemic became clear, and where a survey of students revealed a significant fraction were planning to return to Ithaca even if residential instruction and on-campus housing were not offered (12, 31).

For universities like Cornell located in college towns, where the student population represents a significant fraction of the overall population during normal semesters, this influx of students may represent a significant increase in the number of young people in the area. This could be dangerous because these young people may be socially cohesive and, during normal times, young people have elevated rates of social contact (42, 43). Moreover, a university may have reduced ability to mandate and enforce behavioral restrictions and asymptomatic screening for students taking classes virtually even if they are currently in the university's local area.

Thus, when deciding whether to reopen, in addition to whether the number of infections can be kept reliably low during a residential semester, an additional key consideration is the risk of an outbreak among virtual instruction students and from there to the community. In other words, a key tradeoff is whether to invite back all students and have stronger behavioral and screening interventions, or to have a smaller number of students return but have weaker interventions.

To study this tradeoff, we extended our model to capture virtual instruction at Cornell's Ithaca campus. Under virtual

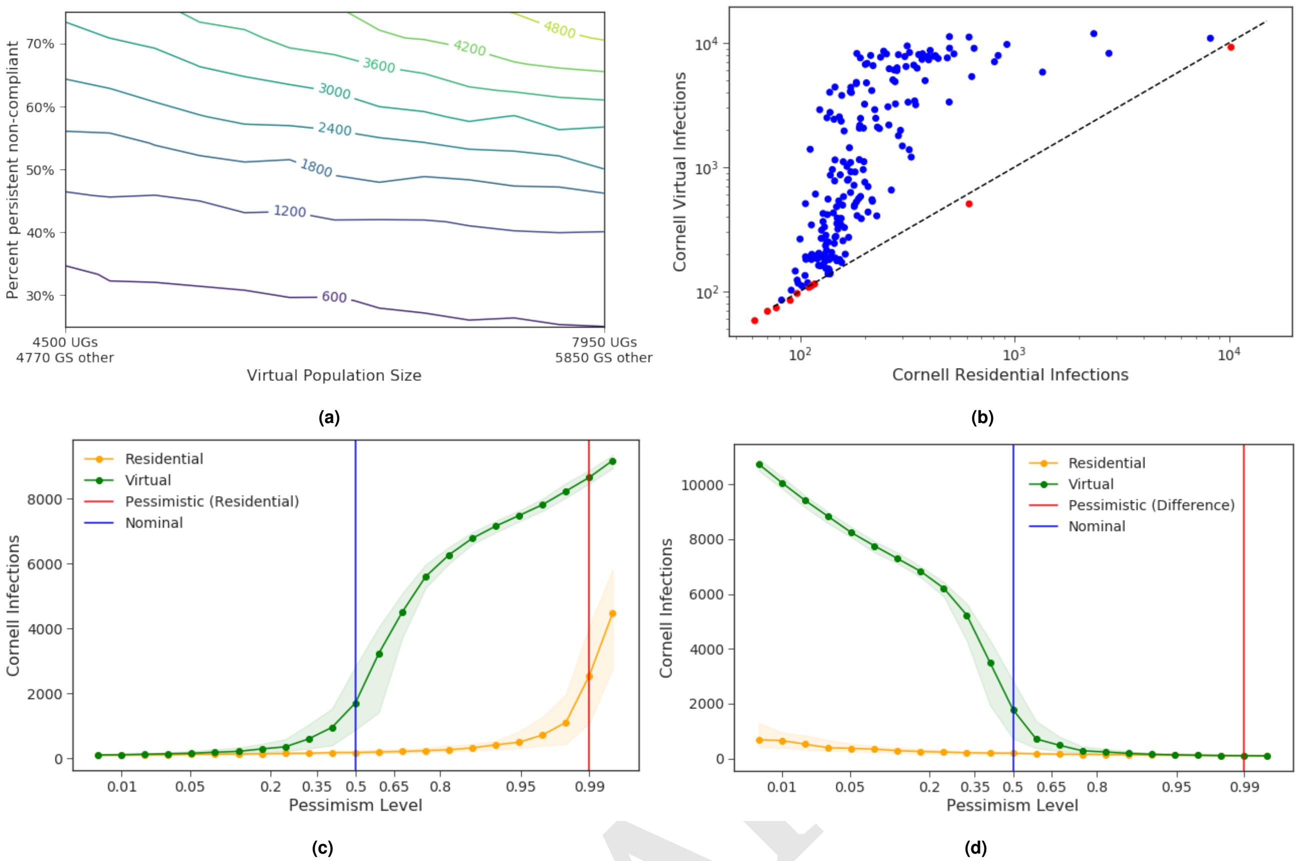


Fig. 3. (a) The number of infections in the Cornell community under virtual instruction under the nominal scenario, varying the fraction of off-campus students engaged in virtual instruction that do not use the offered testing ("persistent non-compliance") and the total number of off-campus students. If off-campus students' willingness to comply with 2x / week optional testing is not sufficiently high, a large number of infections result. (b) Under 200 parameter configurations drawn from the prior, the number of infections under virtual and residential instruction. Infections are smaller under residential instruction than under virtual instruction in most parameter configurations (blue dots), and when they are not (red dots) they are not substantially larger (c-d) Number of infections under virtual and residential instruction under a range of parameter configurations with different levels of pessimism. (c) uses a pessimistic configuration that maximizes residential infections while (d) maximizes residential - virtual infections.

instruction, staff and faculty along with some research-focused graduate students stay on campus. They are tested 2x / week and are subject to the same behavioral compact governing student behavior under residential instruction. Based on survey results and discussions with local landlords, we also model some other students returning to Ithaca to live while taking classes virtually, outside of the control of the university. We assume that Cornell offers twice-weekly testing to these students, but noncompliance is higher than in a residential semester.

This extended model included four additional parameters, about which we also had uncertainty (SI Table S16). We thus extended our parameter uncertainty framework to understand the range of outcomes possible under virtual instruction. We first generated ranges for these additional parameters based on information available in June 2020 (SI 3.A), extending our prior probability distribution to include independent normal priors for these new parameters. We also extend our nominal scenario for residential instruction to set each of these additional parameters to the center of its range.

In this extended model, reduced population density lowers transmission for unmonitored off-campus students relative to off-campus students during residential instruction, but lower compliance with social distancing and masks (which Cornell cannot mandate to these students) raises transmission. These effects can offset each other, or one can be more dominant, depending on parameters. Reduced use of now-optional testing reduces the benefits of this intervention. Fig 3a shows that reduced use of testing by the off-campus population can lead to a substantial number of infections, while outcomes are less sensitive to the number of off-campus students.

To understand the relative safety of residential and virtual instruction under plausible parameter configurations, we sampled 200 parameter configurations from the prior. Fig 3b then plots the number of infections under residential and virtual instruction for each of these configurations. We see that infections are fewer under residential instruction in almost all parameter configurations. In those parameter configurations where there are more infections under residential instruction, the number of additional infections is small. This suggests that residential instruction is a safer strategy than virtual instruction, given the information available in June 2020.

Exploring further, we extended each of the 2000 parameter configurations used for our residential infection sensitivity analysis (Results B) to include the four additional virtual-instruction parameters (SI 3.D) and used simulation to predict virtual instruction infections (in the on-campus students and employees and the off-campus virtual instruction students in Ithaca) under each parameter configuration, enabling a comparison with predicted residential instruction infections for each configuration.

We then set out to identify a new collection of parameter configurations of varying pessimism about the *relative* safety of residential instruction compared to virtual instruction. To do so, we adopt the same approach used to identify configurations of varying pessimism about residential infections, but taking our primary outcome as the difference in infections between residential instruction and virtual instruction (a positive value indicates residential has more infections) (SI 3.A). The nominal scenario remains the same and corresponds to pessimism level 0.5. We obtain a new pessimistic scenario, corresponding to pessimism level 0.99. Unlike the pessimistic scenario for residential infections, this new pessimistic scenario (SI Table S17, Fig S7) decreases R_0 relative to nominal. This is because the most likely parameter configurations with small residential - virtual infections (according to the fitted linear model) are those in which transmission is small regardless of instruction method.

Fig 3c-d plots residential and virtual instruction infections for the two families of parameter configurations, one varying our pessimism about residential infections (Fig 3c), and the other varying our pessimism about residential - virtual infections (Fig 3d). In almost all scenarios, virtual infections are larger than residential infections, and are sometimes much larger. Those few scenarios where residential infections are larger have few infections in both modes of instruction.

Thus, modeling suggests that virtual instruction presented a substantial risk of high infection counts in the student population and the surrounding community, while residential instruction would result in lower infection counts under a broad range of the most reasonable parameter settings. This was a primary basis for Cornell's decision to reopen for residential instruction (12).

D. Design of Asymptomatic Testing Protocol. Our modeling approach was also an important tool for supporting detailed design of Cornell's surveillance testing strategy.

A first key question was the testing frequency for students and employees. Operational constraints limited the total number of tests that could be done per day. We hypothesized that targeting more frequent testing to those groups likely to have higher rates of transmission would provide more robust infection control within this constraint.

We enumerated testing policies consisting of a testing frequency (1x or 2x per week) for the 6 groups spending significant time on campus (e.g., undergraduates living on campus) under both the pessimistic (for the residential infections outcome) and nominal scenarios, producing 64 testing policies. We then discarded those policies not on the Pareto frontier under the pessimistic scenario. Fig 4a shows the resulting Pareto frontier and highlights the policy that was selected.

A second key question was the sampling methodology for surveillance testing. We considered anterior nares (AN) and nasopharyngeal (NP) sampling. While NP is more sensitive (44), it is also less comfortable and we hypothesized that this discomfort might lower test compliance. Fig 4b, generated using our simulation under the pessimistic (residential) scenario, evaluates this trade-off between test sensitivity and test compliance. AN was chosen for Cornell's surveillance sampling methodology in part based on this analysis, because the risk of a substantial loss in test compliance caused by NP sampling would not be known until after the launch of the program, while the test sensitivity of AN was measured before launch and was known to be sufficient for robust infection control.

E. A Retrospective View. Students returned to Cornell's Ithaca campus for residential instruction in the fall of 2020 based in large part on the results of the analysis above (12). An asymptomatic screening program using AN samples and pooled PCR testing was implemented as recommended with undergraduates being tested twice per week, graduate students once per week, and staff and faculty at a frequency between twice per week and once every two weeks depending on how often they were on campus. In addition, students were tested on arrival using NP swabs. Of the student population, 75% returned to the Ithaca campus (45).

This provides an opportunity to evaluate with hindsight the modeling described above. Using de-identified aggregated data obtained from fall surveillance, we build and calibrate a model with a revised set of groups based on risk separation observed during the semester.

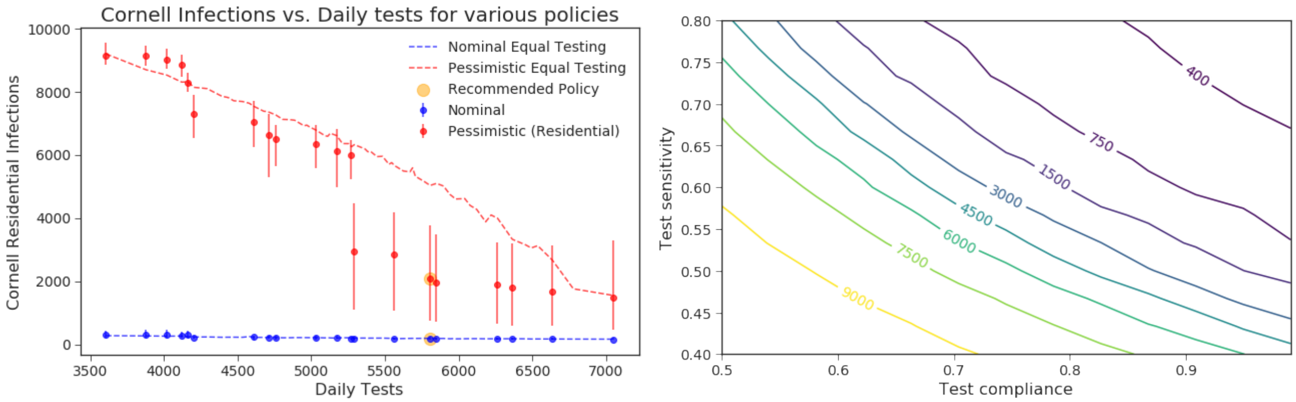
In the revised model we consider employees and students to be two completely separated groups that do not infect each other, because transmission between students and non-student employees was observed through contact tracing to be extremely rare. We separate students into three groups: undergraduates who are in social Greek-life organizations or on varsity athletic teams; other undergraduates; and graduate / professional students. During the semester, observed cases and contact tracing suggested that the greatest transmission was associated with off-campus social activity and co-habitation, especially in the Greek-life and athletics community. Our modeling in the summer of 2020 separated students by the density of their housing, but we did not see substantial transmission associated with residence halls, suggesting that this distinction could be removed.

We then used observed data to estimate key model parameters: one governing the effectiveness of contact tracing for students and employees; the rate of transmission between the 3 student groups up to a proportionality constant; and the rate of infection from outside sources for students and employees.

After, one free parameter remained for employees (transmission rate) and one for students (a proportionality constant giving the transmission rate between each student group). We estimated these parameters, separately for employees and for students, by calibrating simulation results to observed data. More precisely, for each population we varied the free parameter and calculated the sum of squared differences between the mean of the model's predictions and the observed cumulative infection count (Fig 5a-b). Fig 5c compares observed total (student + employee) cases to simulation trajectories from the calibrated model. For details see SI Section 2.

Fig 5d-f shows results from our retrospective evaluation using this calibrated model, focusing on 2 measures of quality: consistency with the range of plausible scenarios identified earlier; and quality in hindsight of the decision made.

First, we evaluate whether the calibrated model is consistent with the range of likely scenarios identified in Results B. A comparison of calibrated parameters (SI Table S20) shows that the calibrated parameters are consistent with our prior. Fig 5d compares observed infections to a range of scenarios. Observed infections were close to the nominal scenario and well



(a) **(b)**

Fig. 4. (a) Select Pareto-efficient testing policies (frequencies for each group) according to median simulated infections. Vertical bars depict the distribution of infections, ranging from the 10% to the 90% quantile over simulation outcomes. The point highlighted in yellow corresponds to the testing frequencies Cornell selected for the fall semester. Policies shown are Pareto-efficient policies from the set of policies where each on-campus group is tested either once or twice a week. The dotted lines are simulation estimates of the expected number of infections if tests are split homogeneously among the on-campus population. (b) Contour plot showing the number of infections in the Cornell community as test compliance and test sensitivity vary under the pessimistic (residential) scenario.

under the pessimistic scenario for almost the entire semester, consistent with their design.

While the observed values are quite close to the nominal scenario's predictions, this was accidental. Indeed, it was clear *a priori* that model outputs are sensitive to inputs and these inputs were unknown, so the predictive accuracy of the range across scenarios, in the sense of whether it contained reality or not, is more important than whether a point prediction was close to the realized trajectory.

Fig 5d also includes predictions from the nominal scenario in our June 2020 report (46). This scenario was nearly identical to the nominal scenario reported here, except that three key parameters were set to conservative values given the urgent need to generate recommendations without enough time to identify a plausible range for these parameters (SI 3.B). This reduced the accuracy of this particular point prediction by making it much more conservative, emphasizing the importance of basing decisions based on a range of plausible scenarios rather than a single point prediction.

Second, we study the quality of the decisions made, relative to potentially better decisions that could have been made with the benefit of hindsight. We focus on two key decisions: the design of the testing policy and whether to reopen campus.

Fig 5e shows the expected number and range of Cornell infections under Pareto-optimal testing policies, testing each of the 3 student groups from the retrospective analysis at either 1x or 2x per week, and testing employees at a frequency averaged across those used in practice. Retrospectively, we selected one of many testing policies resulting in few infections, but we did not select the most efficient. The most efficient testing policies require segmenting the riskiest undergraduates (who are in Greek-life organizations or on varsity athletic teams), which we did not recognize at the beginning of the semester. The figure also shows that there would have been limited benefit in increasing the total testing capacity.

Turning to the question of whether to reopen campus, Fig 5f shows the expected Cornell infections for a *virtual* fall semester under our calibrated model as the number of returning undergraduate students and their test compliance varies. More than 80% of the returning students would need to remain test-compliant throughout the entire semester to achieve a number of infections comparable to reopening campus (where fewer than 250 occurred). As discussed previously, enforcement of test compliance would have been significantly more challenging in a virtual scenario and therefore, the decision to reopen campus was robust.

We conclude that while the benefit of hindsight would have allowed us a modest gain in the efficiency of testing, the decision we selected gave similar health outcomes to the best policies with the benefit of hindsight.

2. Discussion

A. Use at Cornell. The primary purpose of the epidemiological modeling effort at Cornell was to answer the question: is there a way to safely reopen the campus for in-person instruction? Paralleling this question was the accompanying concern: what are the comparative impacts of the "best" implementation of in-person instruction to the "best" implementation of virtual instruction? The previous section outlined the analysis derived from our models in answering these questions. Specifically, the model provided the basis for determining the overall testing frequency required to limit spread, as well as the refined policy of varying test frequency across university sub-populations.

Beyond this primary goal, the modeling framework has many secondary applications. Foremost among these is the ability to do recourse planning: having a baseline model and real-time information about the prevalence of the epidemic on campus enabled measures responding to current conditions. The most visible of these was the color-coded state of the campus, which ranged from the green ("new normal") to the red alert which called for careful closing of campus, i.e., ending in-person instruction and the emptying of the on-campus dormitories. One of the advantages of the model in its use at Cornell was that it enhanced communication with the community — as a tool for explaining the considerations in making decisions, and for explaining what might happen and why. For example, the limitations on campus behavior imposed by moving the campus alert code from green to yellow could

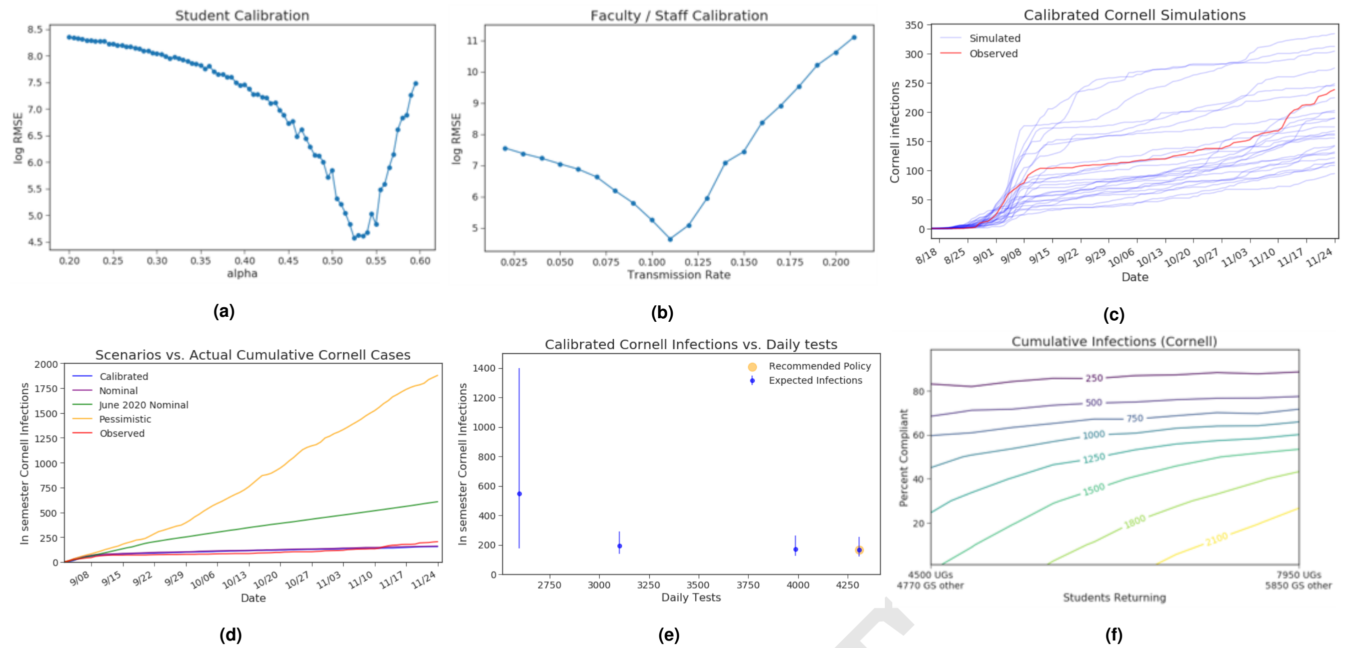


Fig. 5. (a) The log of the root mean-squared error (RMSE) of projected student infections versus α, the proportionality constant that multiplies the contact matrix to obtain the daily transmission rate. The fitted value of α is where RMSE is lowest, at $\alpha = 0.525$. (b) The log of the root mean-squared error (RMSE) of projected employee infections versus employee transmission rate. The fitted value of transmission rate that minimizes RMSE is 0.11. (c) Simulated cumulative case trajectories for all Cornell cases under calibrated parameters and observed fall 2020 cases, including the pre-semester period prior to the start of asymptomatic surveillance. (d) Cornell fall 2020 cases relative to the calibrated trajectory and to the nominal, pessimistic, and June 2020 nominal scenarios. (e) Re-creating the Pareto efficient testing frontier based on calibrated parameter values. (f) Using calibrated parameters to re-create the contour plot showing expected Cornell infections (including students) under a virtual instruction scenario as the number of returning students and their test compliance varies.

be modeled by changes in the parameters such as the number of daily contacts, and so the impact of this recourse approach could be demonstrated in a concrete way.

In addition to these strategic questions, the epidemiological modeling provided a framework for considering a broad range of tactical and operational policy questions. First, it was essential to estimate, and then plan for, quarantine and isolation capacity. In doing so, the model was used to understand both the requirements for the initial arrival period, and for the stochastic evolution of demand for these resources as the semester unfolded. Second, the model quantified the impact of shortening the time between collecting samples and isolating positive individuals, which supported the design of the testing infrastructure. Third, the Cornell academic calendar was restructured to limit impact of a potential "second wave": in-person education ended prior to Thanksgiving and the final three weeks of instruction were given virtually. Similar to the primary question of virtual vs. in-person education, the epidemiological modeling demonstrated the imperative of continued surveillance testing for the Cornell community residing in Ithaca beyond Thanksgiving.

B. Use at Other Universities. The *conclusions* of our Cornell study do not directly translate to other universities, but the key decisions informed by the model are broadly generalizable. Accordingly, our modeling framework can be readily adapted to support decision-making at other universities.

In translating this approach to other university campuses, a number of characteristics should be explicitly considered. For many of these, Cornell was particularly well-suited for in-person education. We highlight the following characteristics:

- the amount of interaction between the campus and surrounding communities, as well as the initial COVID-19 prevalence in the surrounding community; Ithaca had low prevalence and is several hours' drive from major cities;

- the extent to which campus culture provides "buy-in" for restrictions on student behavior; Cornell's administration worked closely with student leadership to create a student compact establishing behavioral norms;

- the extent to which restricting travel beyond the immediate environs is feasible; again, geography meant that travel beyond campus was generally not common, and restrictions were relatively easy to implement;

- the fraction of the campus community that will reside locally for both in-person and virtual modes; as discussed above, many students live in off-campus housing and survey data indicated a significant number would return even for virtual instruction;

- the ability to identify student cohorts with elevated social contact, and the social connectivity of those involved; case clusters at Cornell were often connected to off-campus social activity in a small segment of the student population and an understanding of this mechanism for viral spread supported contact tracing;

- the capacity of the institution to provide surveillance testing; the College of Veterinary Medicine at Cornell had experience using PCR to track viral infections in animal populations, which could be converted and rescaled to

- provide the required capacity, especially since there was the expertise to further augment capacity through pooling;
- the ability to cope with the demands of arrival testing, both in terms of contact tracing and quarantine capacity (which depends on initial prevalence); in this aspect, Cornell students' willingness to self-quarantine in advance of their return appears to have limited initial incidence to levels below anticipated values.

In addition to these considerations, there are other local conditions relevant to the adaptation of our approach that are not specifically dependent on our modeling effort. For example, we believe that it is important to have appropriate enforceable quarantine and isolation space on campus.

Beyond universities, our modeling framework can be applied to other closed (or nearly closed) communities such as cruise ships, prisons, retirement homes, homeless shelters (47), military bases, or professional sports bubbles. While differences in the behavioral and health characteristics of the populations involved, as well as the interventions available, would require adjustments to model parameters, any of these relatively closed communities could be modeled and analyzed in a similar way, coping with parameter uncertainty as discussed here.

Like most decision-making based on epidemiological models, ours has a variety of limitations. Most importantly, our results are highly sensitive to input parameters for which there was significant uncertainty. Although our results showed that our conclusions were remarkably robust, our June 2020 nominal scenario was cautious in selecting conservative values for parameters based on literature and other relevant data. Cumulatively, this caused the predictions we released in the summer of 2020 to be conservative.

Beyond parameter uncertainty, our model itself has limitations due to its simple structure. For example, we assume that all members of a subgroup are homogeneous and therefore we do not model contact network structure. As a result, while it is similar to that used in other analysis (19), our model of contact tracing has limitations, e.g., we assume that all quarantined contacts are in the pre-infectious exposed state. Further, within a single compartment, we assume that infectiousness does not vary across individuals or over time, in contrast to, e.g., (9). Moreover, we model the impact of age and its effect on the distribution of infection severity in a discretized way. We do not model overdispersion in contacts (for example, one individual having many more contacts than another) because our prior over transmission parameters is based on empirically derived R_0 estimates. When infections are driven by spread among subpopulations with the highest transmission, empirical estimates of R_0 naturally correspond to transmission rates in these subpopulations.

3. Methods

Here we provide an overview of modeling methods. Full details appear in the SI.

A. Simulation Model: Overview. We model COVID-19 transmission using a compartmental model with Susceptible, Exposed, Infectious and Recovered compartments along with additional compartments to reflect specific characteristics of COVID-19 and the interventions applied (Fig 1a). To account for asymptomatic and pre-symptomatic transmission, the infectious phase is split into 3 compartments: Infectious, Asymptomatic and Symptomatic. Individuals formerly in the Exposed compartment enter the Infectious compartment before randomly being assigned to either the Asymptomatic or Symptomatic compartments. We also add compartments for quarantined non-infected individuals (Quarantine) and isolated infected individuals (Isolation). Due to significant age and social heterogeneity in a university community, we replicate the compartments described above for each of several university groups. The groups interact via cross-group contacts. In addition, since the severity of COVID-19 varies significantly with age, the probabilities of symptom severities are determined by a group's age distribution.

B. Simulation Model: Transmission. Modeled COVID-19 transmission is governed by two variables: contact rates between groups, represented by a contact matrix, and the probability of transmission during any interaction. The contact matrix is estimated using pre-pandemic measurements of age-based socialization patterns. The probability of transmission is calibrated to match external estimates of R_0 (2.5 for the nominal scenario). Transmission is modeled as stochastic, and the number of new infections in each group has a Poisson distribution with a mean determined by the product of the contact matrix and the transmission probability and the number of non-isolated infectious individuals in each group.

C. Simulation Model: Quarantine and Isolation. We model three mechanisms for identifying and quarantining / isolating infected individuals.

1. Testing: every day, a fraction of each group is selected uniformly at random to be tested and test results are available the same day. The number of individuals selected to be tested per group per day is determined by the group's testing frequency.
2. Symptomatic self-reporting: Every day, each symptomatic individual not in Isolation has a constant probability of self-reporting their symptoms, upon which they enter Isolation the same day.
3. Contact tracing: Each case found through testing or self-reporting is contact traced. (Cases found through contact tracing are not, themselves, modeled as contact traced. This is an approximation of reality that is necessary because our simulation tracks counts within groups, not individuals.) Each contact trace moves a Poisson-distributed number of people to Quarantine or Isolation after a 1-day delay. We assume that contact tracing only finds contacts within the same group as the source case, reflecting the social dynamics of college campuses.

D. Parameter Configurations of Varying Pessimism. To summarize the effect of parameter uncertainty on an outcome (infections in a residential semester, or the difference in infections between residential and virtual instruction), we developed a one-dimensional family of parameter configurations with varying levels of risk for each outcome. For each real number y we consider the set of parameter configurations $A(y)$ whose median outcome is equal to y according to the fitted linear model. (All configurations are in exactly one $A(y)$, so this partitions the parameter configuration space.) For each y, we select a representative configuration from $A(y)$: the one that is most likely under the prior. As we increase y, outcomes under the representative configuration tend to degrade. We then graph the outcome under this configuration versus $P(\cup_{y' \leq y} A(y'))$, i.e., the probability under the prior of seeing a parameter configuration whose outcomes (under the linear model) are no worse than those in $A(y)$. We refer to this probability as the "pessimism level". It ranges from 0 to 1, with larger values corresponding to representative parameter configurations that are more pessimistic. For details see SI Section 3.A.

E. Retrospective Parameter Estimation and Model Calibration. We estimate most model parameters for the fall semester (initial prevalence, outside infection rate, a contact matrix with entries that are proportional to the inter-group transmission rates, and contact-tracing effectiveness) directly from fall semester data available in operationally-focused non-research public communication (48). These public health communications were based on de-identified positive-case, testing, and contact-tracing stored along with student life, housing, and employee data in a HIPAA-compliant database. This data was collected and analyzed pursuant to the urgent public health need presented by the pandemic. Additional aggregations of de-identified data for research purposes were provided by Cornell University and deemed by Cornell's IRB to not meet the regulatory definition of human subjects research.

Remaining parameters, i.e., the proportionality constant scaling the contact matrix for the student group and the rate of transmission for the employee group were estimated by calibrating the model output to the infection trajectories observed during the Fall 2020 semester. As already discussed, we consider employees and students to be two completely separate groups that do not infect each other. The calibration results are robust to the addition of a small but positive employee-student transmission term, although such a term can change outcomes more substantially in extreme regimes where student or employee prevalence is much higher than observed in the fall.

Acknowledgments

We gratefully acknowledge support from Cornell University. Additional support was provided by National Science Foundation grants TRIPODS+X DMS-1839346 and CMMI-2035086, and Air Force Office of Scientific Research FA9550-19-1-0283. We are also grateful to many individuals from Cornell University, the Tompkins County Health Department, and the Cayuga Health System who together kept our community safe.

1. E Black, R Ferdig, LA Thompson, K-12 virtual schooling, COVID-19, and student success. *JAMA Pediatr.* **175**, 119–120 (2021).
2. J Wallinga, P Teunis, M Kretzschmar, Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* **164**, 936–944 (2006).
3. G Yamey, RP Walensky, Covid-19: re-opening universities is high risk. *BMJ* **370** (2020).
4. SM Danielle Ivory, Robert Gebeloff, Young people have less COVID-19 risk, but in college towns, deaths rose fast. *The New York Times* (2020) Accessed: 2021-03-12.
5. L Rennert, C Kalbaugh, C McMahan, L Shi, CC Colenda, The urgent need for phased university reopenings to mitigate the spread of COVID-19 and conserve institutional resources: A modeling study. *medRxiv* (2020).
6. L Rennert, CA Kalbaugh, L Shi, C McMahan, Modelling the impact of presemester testing on COVID-19 outbreaks in university campuses. *BMJ open* **10**, e042578 (2020).
7. E Brooks-Pollock, et al., High COVID-19 transmission potential associated with re-opening universities can be mitigated with layered interventions. *medRxiv* (2020).
8. A Hale, The behavioral compact, explained. *Cornell Dly. Sun* (2020) Accessed: 2021-04-11.
9. JT Chang, FW Crawford, EH Kaplan, Repeat SARS-CoV-2 testing models for residential college populations. *Heal. Care Manag. Sci.* **24**, 305–318 (2021).
10. CW Stubbs, M Springer, TS Thomas, The impacts of testing cadence, mode of instruction, and student density on fall 2020 COVID-19 rates on campus. *medRxiv* (2020).
11. College Crisis Initiative, The college crisis initiative. https://collegecrisis.shinyapps.io/dashboard/ (2020) Accessed: 2021-03-12.
12. MI Kotlikoff, ME Pollack, Why Cornell will reopen in the fall, https://www.wsj.com/articles/why-cornell-will-reopen-in-the-fall-11593535516. *Wall Str. J.* (2020) Accessed: 2021-04-04.
13. EH Bradley, MW An, E Fox, Reopening colleges during the coronavirus disease 2019 (covid-19) pandemic—one size does not fit all. *JAMA Netw. Open* **3**, e2017838–e2017838 (2020).
14. WC Roda, MB Varughese, D Han, MY Li, Why is it difficult to accurately predict the COVID-19 epidemic? *Infect. Dis. Model.* **5**, 271–281 (2020).
15. T Blackwell, COVID-19 fight has relied too much on uncertain math modelling, some scientists say, https://ottawacitizen.com/news/canada/covid-19-fight-has-relied-too-much-on-uncertain-math-modelling-some-scientists-say. *Ott. Citiz.* (2021) Accessed:2021-03-12.
16. I Holmdahl, C Buckee, Wrong but useful — what COVID-19 epidemiologic models can and cannot tell us. *New Engl. J. Medicine* **383**, 303–305 (2020).
17. S Eker, Validity and usefulness of COVID-19 models. *Humanit. Soc Sci Commun* **7** (2020).
18. EM Hill, BD Atkins, MJ Keeling, M Tildesley, L Dyson, Modelling SARS-CoV-2 transmission in a UK university setting. *medRxiv* (2020).
19. FC Motta, et al., Benefits of surveillance testing and quarantine in a sars-cov-2 vaccinated population of students on a university campus. *medRxiv* (2021).
20. AD Paltiel, A Zheng, RP Walensky, Assessment of SARS-CoV-2 screening strategies to permit the safe reopening of college campuses in the United States. *JAMA network open* **3**, e2016818–e2016818 (2020).
21. M Peplow, How one university built a COVID-19 screening system, https://cen.acs.org/biological-chemistry/infectious-disease/one-university-built-COVID-19/98/i42. *Chem. Eng. News* (2020) Accessed: 2021-04-04.
22. News @ Northeastern, Here's why Northeastern is testing everyone on the Boston campus for the coronavirus, https://news.northeastern.edu/coronavirus/university-messages/heres-why-northeastern-is-testing-everyone-on-the-boston-campus-for-the-coronavirus/ (2020) Accessed: 2021-04-04.
23. TN Denny, et al., Implementation of a pooled surveillance testing program for asymptomatic SARS-CoV-2 infections on a college campus—Duke University, Durham, North Carolina, August 2–October 11, 2020. *Morb. Mortal. Wkly. Rep.* **69**, 1743 (2020).
24. MD Fox, DC Bailey, MD Seamon, ML Miranda, Response to a COVID-19 outbreak on a university campus—Indiana, August 2020. *Morb. Mortal. Wkly. Rep.* **70**, 118 (2021).
25. C Gollier, O Gossner, Group testing against COVID-19. *Covid Econ.* **2** (2020).
26. B Cleary, et al., Using viral load and epidemic dynamics to optimize pooled testing in resource constrained settings. *medRxiv* (2020).
27. Y Lin, et al., Group testing enables asymptomatic screening for COVID-19 mitigation: Feasibility and optimal pool size selection with dilution effects. *arXiv:2008.06642* (2020).
28. N Chen, M Hu, C Zhang, Capacitated SIR model with an application to COVID-19. *Available at SSRN 3692751* (2020).
29. M Müller, PM Derlet, C Mudry, G Aeppli, Testing of asymptomatic individuals for fast feedback-control of COVID-19 pandemic. *Phys. biology* **17**, 065007 (2020).
30. GD Lyng, NE Sheils, CJ Kennedy, D Griffin, EM Berke, Identifying optimal COVID-19 testing strategies for schools and businesses: Balancing testing frequency, individual test technology, and cost. *medRxiv* (2020).
31. Cornell University Committee on Teaching Reactivation Options, Committee on teaching reactivation options report, https://covid.cornell.edu/_assets/files/ctro-final-report.pdf (2020) Accessed: 2021-04-04.
32. N Ghaffarzadegan, LM Childs, UC TÄuber, Diverse computer simulation models provide unified lessons on university operation during a pandemic. *BioScience* (2020).
33. UK Mukherjee, et al., An evaluation of educational institutions reopening strategies for in-person classes safely amid the COVID-19 pandemic. *medRxiv* (2020).
34. PT Gressman, JR Peck, Simulating COVID-19 in a university environment. *Math. biosciences* **328**, 108436 (2020).
35. M Borowiak, et al., Controlling the spread of COVID-19 on college campuses. *arXiv:2008.07293* (2020).
36. R Bahl, et al., Modeling COVID-19 spread in small colleges. *arXiv:2008.09597* (2020).
37. Cornell University Real Estate, FAQ, https://realestate.fs.cornell.edu/about/faq.cfm#CP_JUMP_5682 (2021) Accessed: 2021-03-12.
38. MD McKay, RJ Beckman, WJ Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **42**, 55–61 (2000).
39. Cornell COVID-19 Modeling Team, Covid-19 Mathematical Modeling for Cornell's Fall Semester, https://covid.cornell.edu/_assets/files/covid_19_modeling_main_report.pdf (2020).
40. Cornell University Faculty Senate, June 24 Online Faculty Senate Meeting, https://theuniversityfaculty.cornell.edu/faculty-senate/archives-and-actions/archived-agenda-and-minutes/online-faculty-senate-june-24/ (2020).
41. SV Crowder, An introduction to the Bernoulli CUSUM in *2017 Annual Reliability and Maintainability Symposium (RAMS).* (IEEE), pp. 1–6 (2017).
42. J Wallinga, P Teunis, M Kretzschmar, Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. journal epidemiology* **164**, 936–944 (2006).
43. J Mossong, et al., Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* **5**, e74 (2008).
44. RA Lee, et al., Performance of saliva, oropharyngeal swabs, and nasal swabs for sars-cov-2 molecular detection: a systematic review and meta-analysis. *J. Clin. Microbiol.* **59**, e02881–20 (2021).
45. M Rosenberg, Three-quarters of enrolled students are zooming from Ithaca this fall, https://cornellsun.com/2020/11/08/three-quarters-of-enrolled-students-are-zooming-from-ithaca-this-fall/. *The Cornell*

Dly. Sun (2020) Accessed: 2021-04-04.

46. JM Cashore, et al., Covid-19 Mathematical Modeling for Cornell's Fall Semester (https://covid.cornell.edu/_assets/files/covid_19_modeling_main_report.pdf) (2020).
47. LA Chapman, et al., Comparison of infection control strategies to reduce COVID-19 outbreaks in homeless shelters in the United States: a simulation study. *medRxiv* (2020).
48. Cornell COVID-19 Modeling Team, Mathematical Modeling for Cornell's Spring Semester, https://covid.cornell.edu/_assets/files/general-audience-spring-modeling-20210216.pdf (2021).