Modeling the impact of racial and ethnic disparities on COVID-19 epidemic dynamics

Kevin C. Ma[1], Tigist F. Menkir[2], Stephen Kissler[1], Yonatan H. Grad[1,3], Marc Lipsitch[1,2*]

[1]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, USA

[2]Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, USA

[3]Division of Infectious Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, USA

*Corresponding author. Email: mlipsitc@hsph.harvard.edu

Abstract

The impact of variable infection risk by race and ethnicity on the dynamics of SARS-CoV-2 spread is largely unknown. Here, we fit structured compartmental models to seroprevalence data from New York State and analyze how herd immunity thresholds (HITs), final sizes, and epidemic risk changes across racial and ethnic groups. A proportionate mixing model reduced the overall HIT, but more realistic levels of assortative mixing increased the threshold. Across all models, the burden of infection fell disproportionately on minority populations: in an assortative mixing model fit to Long Island census data, 80% of Hispanics or Latinos were infected when the HIT is reached compared to 33% of non-Hispanic whites. Our findings, which are meant to be illustrative and not best estimates, demonstrate how racial and ethnic disparities can impact epidemic trajectories and result in a disproportionate distribution of the burden of SARS-CoV-2 infection.

Introduction

The dynamics of SARS-CoV-2 spread are influenced by population heterogeneity. This is especially true for herd immunity, which occurs when susceptible individuals in a population are indirectly protected from infection due to immunity in others. The herd immunity threshold (HIT) is the fraction of the population that is non-susceptible when the epidemic reaches its peak, and estimating the HIT for SARS-CoV-2 is important for forecasting the cost associated with letting the epidemic spread in the absence of interventions [1]. In a population with homogeneous mixing, the HIT is 1-1/R_0, where R_0 is the basic reproduction number; this translates to an HIT of 66.7% using an R_0 of 3.

However, population homogeneity is an unrealistic assumption, and models incorporating heterogeneity in social exposure and infection susceptibility (defined as the probability of infection given exposure) generally result in lowered HITs [2–6]. The key idea behind these models is that sub-populations important for epidemic spread (i.e., those with substantially increased susceptibility or exposure) become infected – and thus develop immunity – early on in an epidemic's course. Herd immunity for the population overall is then achieved earlier because once these individuals are no longer susceptible to infection, further epidemic spread is slowed.

Importantly, these models also imply that in locations where SARS-CoV-2 has spread, there may be demographic sub-populations with particularly high cumulative incidences of infection due to increased exposure, susceptibility, or both. Seroprevalence studies can identify these sub-populations and are more reliable and unbiased than reported case data [7]. Identifying and building structured models with these groups in mind is important for understanding how variation in exposure or susceptibility and social disparities are interconnected. These models are also useful for designing interventions that can both reduce disparities and disrupt overall transmission by focusing efforts on groups affected by high transmission rates [8, 9]. Many models in this space have incorporated subpopulation structure by accounting for transmission variation by age [3, 10, 11]. Supporting this approach, susceptibility to infection, contact rates, and cumulative incidence in some locations all appear to vary by age [10, 12]. Nonetheless, serosurveys in Belgium, Spain, Iran, New York City, Brazil, and other places exhibit relatively low variation in seropositivity by age [13–17], indicating additional factors that govern transmission spread.

Substantial racial and ethnic disparities in infection rates, hospitalizations, and deaths have been characterized across the US [18–24], but it is unclear how these heterogeneities in risk are expected to change over time, and what implications – if any – they have on overall epidemic dynamics. Here,

we aim to address these questions by fitting compartmental SEIR transmission models structured by race and ethnicity to seroprevalence data from New York City and Long Island [16]. We focus primarily on building and analyzing variable exposure models because observed disparities in infection rates in US cities have been strongly attributed to differences in mobility and exposure [25–27]. Because of the challenges in acquiring racial and ethnic COVID-19 data [28], including social contact data that can be used in transmission models, we analyzed a range of model structures that are compatible with the seroprevalence data and analyze how those assumptions affect estimates of the HIT and final epidemic size. We also characterize which groups have the highest risk for infection and how that changes over time. These results highlight the importance of developing COVID-19 transmission models that incorporate patterns of epidemic spread across racial and ethnic groups.

Methods

SEIR model

We initially modeled transmission dynamics in a homogeneous population using an SEIR compartmental SARS-CoV-2 infection model:

$$\frac{dS}{dt} = -\beta IS \tag{1}$$

$$\frac{dE}{dt} = \beta IS - rE \tag{2}$$

$$\frac{dI}{dt} = rE - \gamma I \tag{3}$$

$$\frac{dR}{dt} = \gamma I \tag{4}$$

where S, E, I, R refer to the number of people in susceptible, latently infected, infectious, and recovered compartments respectively. Given a mean incubation period and mean serial interval of 5 days as suggested by empirical studies [29, 30], we set the mean latent period $1/r$ to be 3 days to allow for pre-symptomatic transmission and the mean infectious period $1/\gamma$ to be 4 days to coincide with the observed serial interval. The per capita transmission rate is given by $\beta = R_0\gamma/N$, where N is the total number of people in the population. The herd immunity threshold (HIT) was defined as the fraction of non-susceptible people when for some time t, the instantaneous reproductive number

$R_t = S(t)\beta/\gamma$ equals 1, which occurs when the fraction of non-susceptible individuals equals 1-1/R_0 for a homogeneous model.

We extended this model to incorporate multiple racial and ethnic groups by including SEIR compartmental variables for each group, which interact through a social contact matrix that governs the interactions between and within groups. In matrix form, the structured SEIR model is given by:

$$\frac{d\mathbf{S}}{dt} = -(\mathbf{BI}) \circ \mathbf{S} \tag{5}$$

$$\frac{d\mathbf{E}}{dt} = (\mathbf{BI}) \circ \mathbf{S} - r\mathbf{E} \tag{6}$$

$$\frac{d\mathbf{I}}{dt} = r\mathbf{E} - \gamma\mathbf{I} \tag{7}$$

$$\frac{d\mathbf{R}}{dt} = \gamma\mathbf{I} \tag{8}$$

where $\circ$ denotes element-wise multiplication and $\mathbf{S}$, $\mathbf{E}$, $\mathbf{I}$, $\mathbf{R}$ are column vectors comprising the compartmental variables for each group (e.g., $\mathbf{S} = [S_0, ..., S_p]^T$ for p demographic groups). We let S_0 denote non-Hispanic whites, S_1 denote Hispanics or Latinos, S_2 denote non-Hispanic African-Americans, S_3 denote non-Hispanic Asians, and S_4 denote multiracial or other demographic groups, with similar ordering for elements in vectors $\mathbf{E}$ through $\mathbf{R}$.

Following the convention for age-structured transmission models [31], we defined the $p \times p$ per capita social contact matrix $\mathbf{C}$ to consist of elements $c_{i\leftarrow j}$ at row i and column j, representing the per capita rate that individuals from group i are contacted by individuals of group j. Letting N_i be the total number of individuals in group i, the social contact matrix $\mathbf{M}$ consists of elements $m_{i\leftarrow j} = c_{i\leftarrow j} * N_i$, which represents the average number of individuals in group i encountered by an individual in group j. The susceptibility to infection can vary between groups, which we modeled by allowing the probability of infection given contact with an infected individual to vary: $\mathbf{q} = [q_0, ..., q_p]^T$. The transmission matrix $\mathbf{B}$ is then given by $(\mathbf{q}1^T) \circ \mathbf{C}$, where 1^T is a 1 by p vector of 1s:

$$\mathbf{B} = \begin{bmatrix} q_0 & \cdots & q_0 & \cdots & q_0 \\ \vdots & & \vdots & & \vdots \\ q_4 & \cdots & q_4 & \cdots & q_4 \end{bmatrix} \circ \begin{bmatrix} c_{0\leftarrow 0} & \cdots & c_{0\leftarrow 2} & \cdots & c_{0\leftarrow 4} \\ \vdots & & \vdots & & \vdots \\ c_{4\leftarrow 0} & \cdots & c_{4\leftarrow 2} & \cdots & c_{4\leftarrow 4} \end{bmatrix} \tag{9}$$

Given mean duration of infectiousness $1/\gamma$, the next-generation matrix $\mathbf{G}$, representing the average number of infections in group i caused by an infected individual in group j, is given by $(q1^T) \circ \mathbf{M}/\gamma$.

R_0 for the overall population under this structured model was calculated by computing the dominant eigenvalue of matrix $\mathbf{G}$, and R_t at time t was calculated by computing the dominant eigenvalue of $(\mathbf{q}1^T) \circ \mathbf{M_t}/\gamma$, where the elements in $\mathbf{M_t}$ are given by $c_{i\leftarrow j} * S_i(t)$. To hold R_0 values across model types constant, we re-scaled transmission matrices to have the same dominant eigenvalue. We also calculated the instantaneous incidence rate of infection at some time t for all groups by calculating the force of infection $\lambda(t) = (\mathbf{BI}(t)) \circ \mathbf{S}(t)$.

Structured model variants

Simplifying assumptions are needed to constrain the number of variables to estimate in $\mathbf{B}$ given limited data. Under the *variable susceptibility* model, we set the contact rates $c_{i\leftarrow j}$ to all be 1, indicating no heterogeneity in exposure, but allowed the q_j in the susceptibility vector to vary (i.e., $\mathbf{B} = \mathbf{q}1^T$).

Under each of two variable exposure models, in contrast, we set the susceptibility factors q_j to be equal. The simplest variable exposure model we analyzed was the *proportionate mixing* model, which assumes that the contact rate for each pair of groups is proportional to the size and activity level (i.e., total number of contacts per unit time) of the two groups [32]. Denoting a_i as the activity level for a member of group i and $\mathbf{a}$ as the $1 \times p$ vector of a_is, the ijth entry in the transmission matrix is given by:

$$\beta_{i\leftarrow j} = q\frac{a_i a_j}{\sum_k a_k N_k} \tag{10}$$

and the overall transmission matrix $\mathbf{B}$ can be written as:

$$\mathbf{B} = \frac{q}{\sum_k a_k N_k}\mathbf{a}\mathbf{a}^T \tag{11}$$

Finally, under the *assortative mixing* assumption, we extended this model by partitioning a fraction ϵ of contacts to be exclusively within-group and distributed the rest of the contacts according to proportionate mixing (with $\delta_{i,j}$ being an indicator variable that is 1 when $i = j$ and 0 otherwise):

$$\beta_{i\leftarrow j} = (1-\epsilon)q\frac{a_i a_j}{\sum_k a_k N_k} + \epsilon\delta_{i,j}q\frac{a_i}{N_i} \tag{12}$$

$$\mathbf{B} = \frac{(1-\epsilon)q}{\sum_k a_k N_k}\mathbf{a}\mathbf{a}^T + \epsilon\, q\, diag(\mathbf{a} \circ 1/\mathbf{N}) \tag{13}$$

Model fitting and data sources

SEIR differential equations were solved using the lsoda function in R (version 3.6.1). We estimated the a_i in the variable exposure models and the q_i in the variable susceptibility models using maximum likelihood fits to seroprevalence data from New York, which was collected from over 15,000 adults in grocery stores from April 19-28th [16]. We assumed that the seroprevalence data were collected via a binomial sampling process: at a given time point t_s representing the time of the serosurvey, the number of seropositive cases $Y_i(t_s)$ in group i is distributed $Bin(m_i, R_i(t_s)/N_i(t_s))$, where m_i is the number of people tested from group i in the serosurvey and R_i/N_i is the fraction of recovered people from the SEIR model. The likelihood was calculated jointly for all demographic groups, with t_s set to 100 days and the initial number of infected individuals set to 1 in each demographic group; see Supplementary Information for sensitivity analyses on these assumptions.

We acquired total population numbers (i.e. N_i for $i \in \{0, ..., 4\}$) from the 2018 ACS census 1-year estimates Table B03002 (Hispanic or Latino origin by race) subsetted to the following counties: Bronx, Kings, New York, Queens, and Richmond Counties for New York City and Nassau and Suffolk Counties for Long Island. For exposure index calculations to inform model assortativity levels, we used the approach of Richardson et al. [33] using census data from the 2018 ACS 5-year estimates Table B03002 (Hispanic or Latino origin by race) at the level of "all block groups" within the above counties. Copies of the census data we used are available at `https://github.com/kevincma/covid19-race-ethnicity-model/tree/main/data`. The exposure index $P_{a,b}$ between two demographic groups a and b was defined as in [34]:

$$P_{a,b} = \sum_{i}^{L} \left(\frac{N_{a,i}}{N_a}\right)\left(\frac{N_{b,i}}{T_i}\right) \tag{14}$$

where L is the number of census block groups, $N_{j,i}$ is the number of people from demographic group j in census block i, N_j is the total number of people from demographic group j across the city, and T_i is the total number of people in census block group i. The interpretation is that the average a individual lives in a neighborhood with $P_{a,b} * 100\%$ of people from demographic group b. To fit ϵ, we reasoned that if all activity levels a_i were equal to 1, the social contact matrix $\mathbf{M_{a=1}}$ (comprising elements $m_{i\leftarrow j,a=1} = c_{i\leftarrow j,a=1} * N_i = (1-\epsilon)\frac{N_i}{\sum_k N_k} + \epsilon\delta_{i,j}$) should also reflect the composition of the average neighborhood. Using this equivalency, we identified the value of ϵ that minimized the difference between the social contact and exposure index matrices, weighted by the population fraction of each demographic group:

$$arg\min_{\epsilon} \sum_{i}^{p} \frac{N_i}{N} \sum_{j}^{p} |(1-\epsilon)\frac{N_i}{\sum_k N_k} + \epsilon\delta_{i,j} - P_{i,j}| \tag{15}$$

Using the best-fit ϵ values, we then conducted maximum-likelihood to fit the activity levels as described above.

Code availability

Code and data to reproduce all analyses and figures is available at `https://github.com/kevincma/covid19-race-ethnicity-model`. An executable version of the Jupyter notebook is available at `https://mybinder.org/v2/gh/kevincma/covid19-race-ethnicity-model/HEAD`.

Results

We model the dynamics of COVID-19 infection allowing for social exposure to infection to vary across racial and ethnic groups. Models incorporating variable susceptibility to COVID-19 are commonly used when stratifying by age because children are thought to have decreased susceptibility to infection [10]. Variable susceptibility to infection across racial and ethnic groups has been less well characterized, and observed disparities in infection rates can already be largely explained by differences in mobility and exposure [25–27], likely attributable to social factors such as structural racism that have put racial and ethnic minorities in disadvantaged positions [35–38]. In line with the notion that variation in exposure could instead be the main driver of observed seroprevalence differences, our primary focus is on analyzing variable exposure models; we have also analyzed variable susceptibility models for comparison (see Supplementary Information).

The simplest variable exposure models assume proportionate mixing, where the contact rate between groups is set to be proportional to the size and activity level (i.e., mean number of contacts per time period) of the two groups [32]. We fit proportionate mixing models allowing for variable activity levels across racial and ethnic demographic groups to serosurvey data collected in late-April from New York City (NYC) and Long Island, comprising 5946 and 2074 adults, respectively [16]. The serosurvey data were compatible with proportionate mixing models in which Hispanics or Latinos, non-Hispanic Black people, non-Hispanic Asians, and multiracial or other people had 2.25, 1.62, 0.86, and 1.28 times the activity level relative to non-Hispanic whites in NYC. Model fits to Long Island resulted in even more pronounced exposure differences because of greater between-group differences in seropositivity (e.g., the seropositivity in Hispanics or Latinos relative to non-

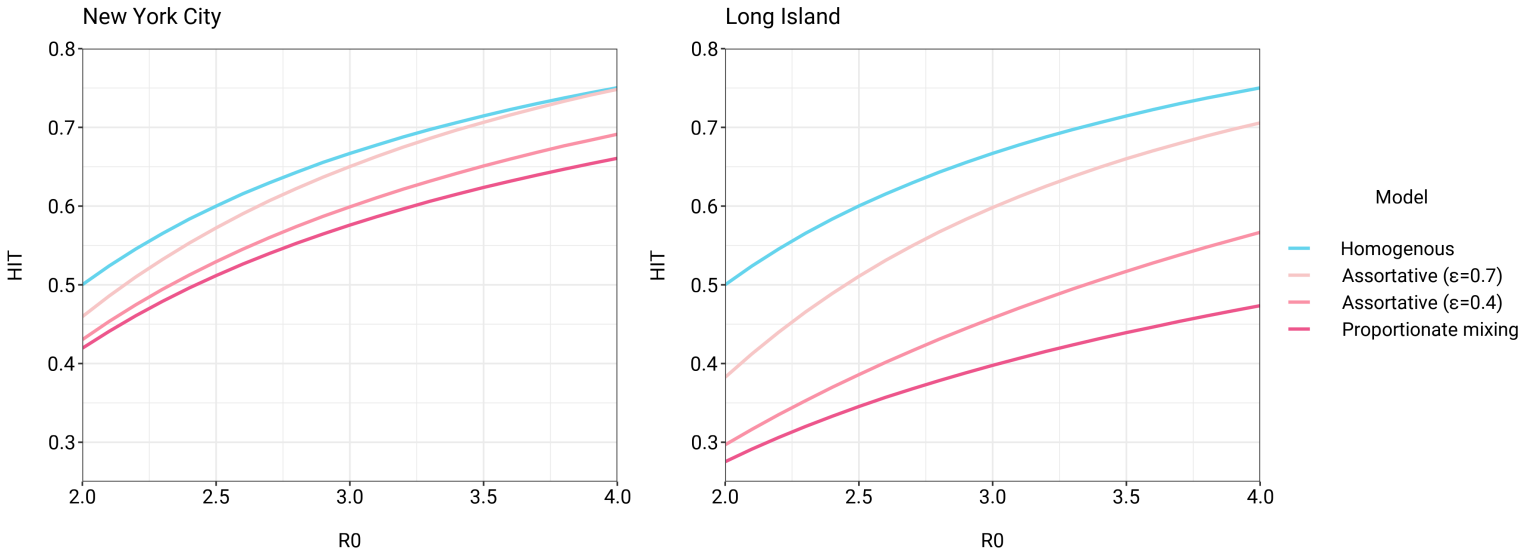


Figure 1: Incorporating assortativity in variable exposure models results in increased HITs across a range of R_0 values. Variable exposure models were fitted to NYC and Long Island serosurvey data.

Hispanic whites was 1.85 times higher in Long Island than in NYC); Hispanics or Latinos, non-Hispanic Black people, non-Hispanic Asians, and multiracial or other people had 4.31, 1.96, 0.92, and 2.48 times the activity level relative to non-Hispanic whites in Long Island, respectively. These differences in exposure impacted herd immunity levels and final epidemic sizes relative to the homogeneous model across a range of R_0 values (Figure 1 and Supplementary Figure 5), and are in line with theoretical derivations (Supplementary Figure 4). For example, for an R_0 of 3, the HIT decreases to 58% in NYC and 40% in Long Island, compared to 67% under the homogeneous model. The HIT overall is reached after cumulative incidence has disproportionately increased in certain minority groups: at the HIT, 75% of Hispanics or Latinos and 63% of non-Hispanic Black people were infected compared to 46% of non-Hispanic whites in NYC, and 77% of Hispanics or Latinos and 48% of non-Hispanic Black people were infected compared to 29% of non-Hispanic whites in Long Island (Figure 2).

The estimated activity ratios indicate higher activity levels for minority groups such as Hispanics or Latinos and non-Hispanic Black people, which is in line with studies using cell phone mobility data [25], but the magnitudes of the activity level ratios are substantially higher than expected. This may reflect some of the limitations of the proportionate mixing assumption, which does not allow for preferential within-group contacts and hence must fit observed seropositivity differences solely by scaling activity levels. To address this, we augment our model by partitioning a specified fraction ϵ of

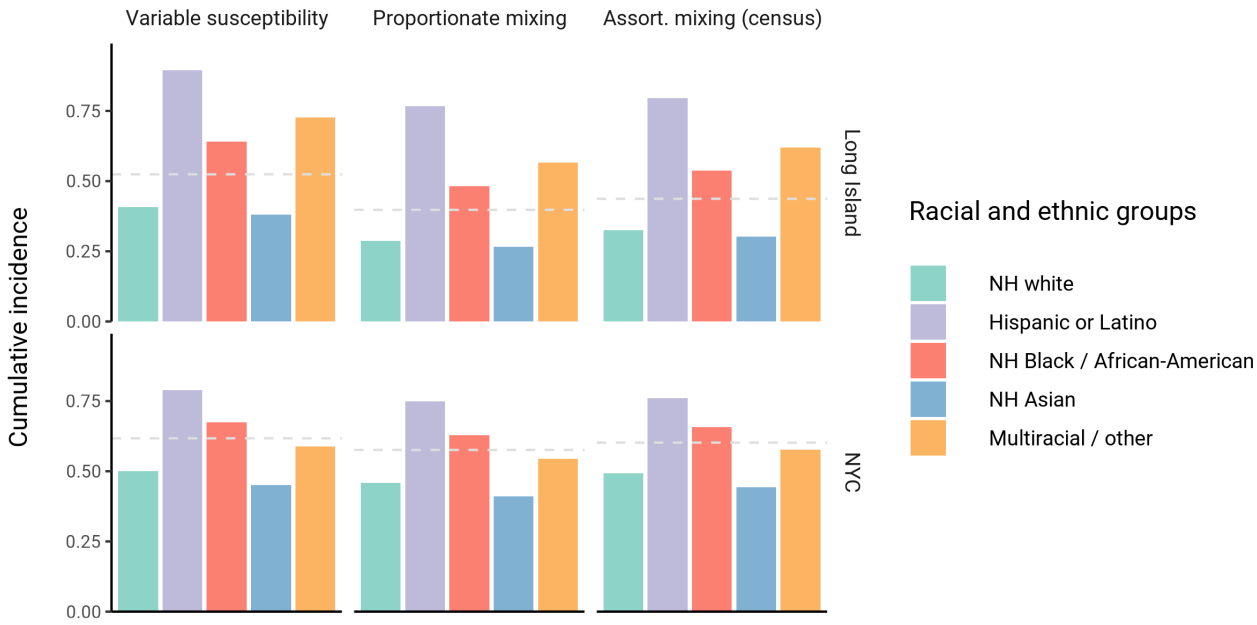


Figure 2: Cumulative incidence is disproportionately higher in some racial and ethnic minorities when the overall HIT is reached across model types and locations. Results are shown for an epidemic with R_0 = 3. The HIT for the population is indicated with a gray dashed line.

contacts to be exclusively within-group, with the remaining contacts distributed proportionately. This assortative mixing model captures more realistic patterns of interactions due to neighborhood structure. After fitting the models across a range of ϵ values, we observed that as ϵ increases, HITs and epidemic final sizes shifted higher back towards the homogeneous case (Figure 1, Supplementary Figure 5). This observation can be understood by comparing the epidemic cumulative incidence trajectories (Supplementary Figure 6) and next-generation matrices (Supplementary Figures 2 and 3): under proportionate mixing ($\epsilon = 0$), lower-risk demographic groups are protected from further infection due to built-up immunity in higher-risk demographic groups, but the magnitude of this indirect protection decreases as the proportion of exclusively within-group contacts increases and groups become more isolated.

We assessed a range of values for ϵ because the serosurvey data cannot be used to also fit the optimal ϵ value; given limited numbers of data points, all of the models can fit exactly to the single seroprevalence time point we consider. To inform plausible assortativity levels, we instead used additional data on demographic distributions from the American Community Survey US census [33]. We calculated the exposure index, which represents the average neighborhood's demographic com-

position from the perspective of an individual from a given racial or ethnic group (Supplementary Tables 1 and 2), and used these results to fit ϵ. The exposure index describes contacts based on proximity but may not capture contacts in other settings, such as work, beyond one's immediate neighborhood of residence. The census data were compatible with assortative mixing matrices in which 43% and 32% of contacts were exclusively within-group in NYC and Long Island, respectively (Supplementary Figure 7). After fitting to seroprevalence data using these optimal ϵ levels, we observed that the groups with higher activity levels remained the same as under proportionate mixing, but the magnitudes of differences were now lower and more concordant with reported mobility differences [25]: model estimates indicated Hispanics or Latinos, non-Hispanic Black people, non-Hispanic Asians, and multiracial or other people had 1.65, 1.37, 0.90, and 1.18 times the activity level relative to non-Hispanic whites in NYC, respectively, and 2.85, 1.70, 0.93, and 2 times the activity level relative to non-Hispanic whites in Long Island, respectively. For an epidemic with $R_0 = 3$, the HITs for NYC and Long Island using these census-informed assortativity values were 60% and 44%, respectively. Similar to the other models, at the HIT, 76% of Hispanics or Latinos and 66% of non-Hispanic Black people were infected compared to 49% of non-Hispanic whites in NYC, and 80% of Hispanics or Latinos and 54% of non-Hispanic Black people were infected compared to 33% of non-Hispanic whites in Long Island (Figure 2).

Using these census-informed assortative mixing models, we then considered how the relative incidence rates of infection in demographic groups could change over the course of the epidemic. Early comparisons of infection and mortality rates have helped to identify racial and ethnic groups at high risk and the risk factors for infection [18–24], but these studies often rely on cross-sectional snapshots of epidemiological patterns. The challenge is that these metrics can change over time: for instance, county level correlations of monthly incidence and mortality rates with percent people of color rose and fell in multiple regions as the epidemic progressed [39]. The reasons for these changes are multifaceted, but even independent of the effect of interventions or behavioral changes, models of epidemic spread in structured populations imply that incidence rate ratios for high-risk groups can decrease substantially as the epidemic progresses because of depletion of susceptible individuals from these groups [40, 41]. In line with this, we observe that instantaneous incidence rate ratios are elevated initially in high-activity groups relative to non-Hispanic whites, but this trend reverses after the epidemic has peaked – a consequence of the fact that a majority of individuals have already become infected (Figure 3). Similarly, cumulative incidence ratios remain elevated in high-activity racial and ethnic groups throughout the epidemic, but the magnitude decreases as the

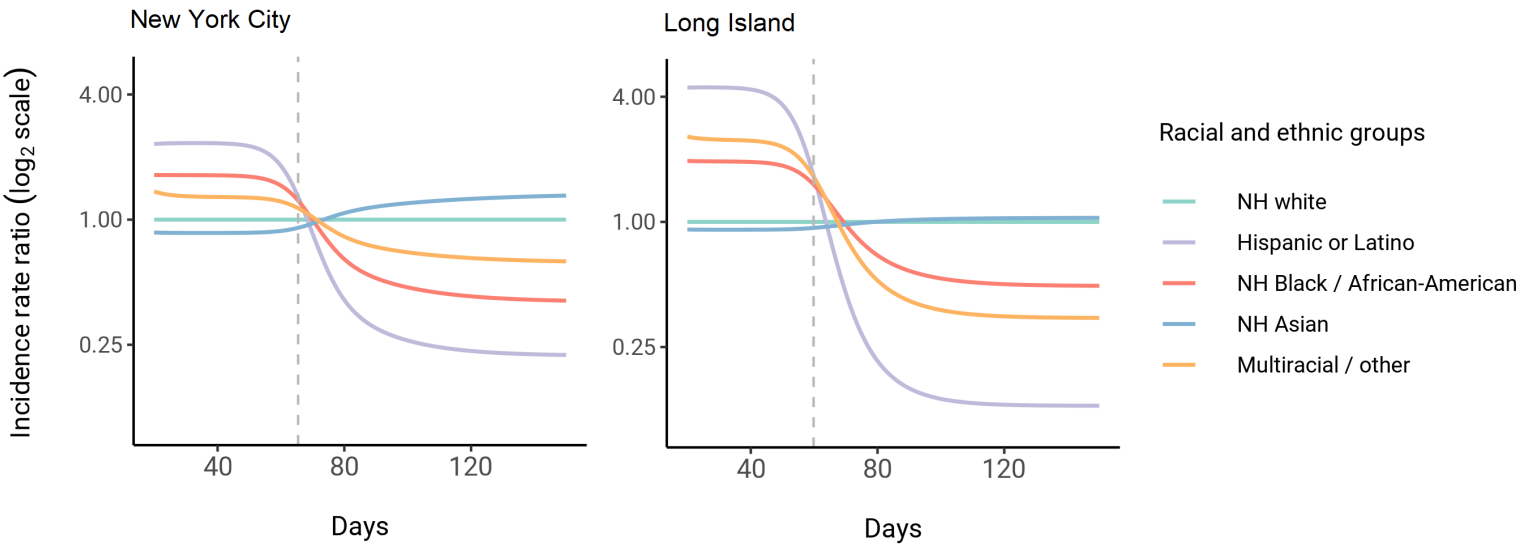


Figure 3: Dynamics of incidence rate ratios relative to non-Hispanic whites in assortative mixing models fitted to census and serosurvey data. Dashed line represents the peak overall incidence for the epidemic.

epidemic progresses (Supplementary Figure 8). These results highlight the importance of controlling for dynamics-induced changes in epidemiological measures of disease burden when evaluating the impact of interventions for reducing inequities in SARS-CoV-2 infections [42]; otherwise, ineffective interventions – depending on their timing – might still be associated with declines in relative measures of disease burden.

Discussion

Here, we explored how incorporating heterogeneity in SARS-CoV-2 spread across racial and ethnic groups could affect epidemic dynamics using deterministic transmission models. Models incorporating variable exposure generally decreased the HIT and final epidemic size, but incorporating preferential within-group contacts shifted HITs and final epidemic sizes higher, approaching the homogeneous case. Epidemiological measures of disease burden such as incidence rate ratios and cumulative incidence ratios also changed substantially over the course of the epidemic, highlighting the need to account for these trends when evaluating interventions [42]. These results illustrate the varied effects of different structured heterogeneity models, but are not meant to be best estimates given the limited seroprevalence data. Longitudinal serosurveys in this same population or incorporating additional data such as cell-phone mobility and social contact surveys [25, 43] could help to

refine these estimates.

Across all model variants, the observed higher cumulative incidence among Hispanics or Latinos and non-Hispanic Black people compared to non-Hispanic whites led to estimates of higher estimated activity levels relative to non-Hispanic whites, mirroring existing inequities in housing, education, healthcare, and beyond [20, 38, 44, 45]. The estimated activity level differences concord with reports that frontline workers, who are unable to engage in physical or social distancing to the same degree as other types of workers, are disproportionately from minority backgrounds [25, 27, 46]. The assortativity we observed in the census data has root causes in many areas, including residential segregation arising from a long history of discriminatory practices [44, 47]. All of these factors highlight the fact that incorporating heterogeneity in models in a mechanism-free manner can conceal the disparities that underlie changes in epidemic final sizes and HITs. In particular, overall lower HIT and final sizes occur because certain groups suffer not only more infection than average, but more infection than under a homogeneous mixing model; incorporating heterogeneity lowers the HIT but increases it for the highest-risk groups (Figure 2).

These results also suggest that public health interventions for reducing COVID-19 inequities can have synergistic effects for controlling the overall epidemic [33]. For instance, from a transmission-control perspective, age-structured models indicate that vaccination of high-activity age groups – such as young adults – is optimal for controlling the spread of SARS-CoV-2 [9]. Similar interventions could be explored for disadvantaged populations because of how increased exposure underlies much of the higher infection risk that racial and ethnic minorities experience. Policy proposals in this space should, of course, carefully consider the ethical dimensions of vaccinations or other interventions targeted by race or ethnicity [48].

We note several limitations with this study. First, biases in the serosurvey sampling process can substantially affect downstream results; any conclusions drawn depend heavily on the degree to which serosurvey design and post-survey adjustments yield representative samples [49]. Other sources of uncertainty, such as antibody test sensitivity and specificity, could also be incorporated into epidemic model predictions in future work [50]. Second, we have assumed that seropositivity implies complete immunity and that immunity does not wane. These are strong assumptions that can be revisited as empirical studies on the length of natural immunity are conducted. Third, we have for simplicity modeled an unmitigated epidemic, where overall transmission rates (per infected and susceptible individual) remain constant, and individual group-specific rates do as well.

That is, we did not model the impact of non-pharmaceutical interventions such as stay-at-home policies, closures, or the like, either in reducing the overall transmission rate or in the relative changes in activity levels for different groups. Empirical evidence suggests that during periods of lockdown, certain neighborhoods that are disproportionately wealthy and white tend to show greater declines in mobility than others [27, 51]. These simplifying assumptions were made to aid in illustrating the key findings of this model, but for more detailed predictive models, the extent to which activity level differences change can be evaluated using mobility data and longitudinal survey data [25, 43].

In summary, we have explored how deterministic transmission models can be extended to study the dynamics of infection in racial and ethnic groups, and how the impact of heterogeneity on the HIT and final epidemic size can depend strongly on the details of how heterogeneity is modeled. We have shown that measures of impact may decline and even reverse over the course of the epidemic due to early increased burden of infection in individuals from the most at-risk groups, even in the absence of interventions to reduce inequities. These results describe a framework that can be extended to other cities and countries in which racial and ethnic disparities in seropositivity have been observed [17, 52] and are a step towards using transmission models to design policy interventions for reducing disparities in COVID-19 and other diseases.

Acknowledgements

We thank Dr. Eli Rosenberg at the University at Albany School of Public Health and members of the Center for Communicable Disease Dynamics and Dr. Mary Bassett at the Harvard T.H. Chan School of Public Health for helpful comments. K.C.M. was supported by National Science Foundation GRFP grant DGE1745303. Y.H.G. and M.L. were funded by the Morris-Singer Foundation.

Supplementary Information

Variable susceptibility versus variable exposure models

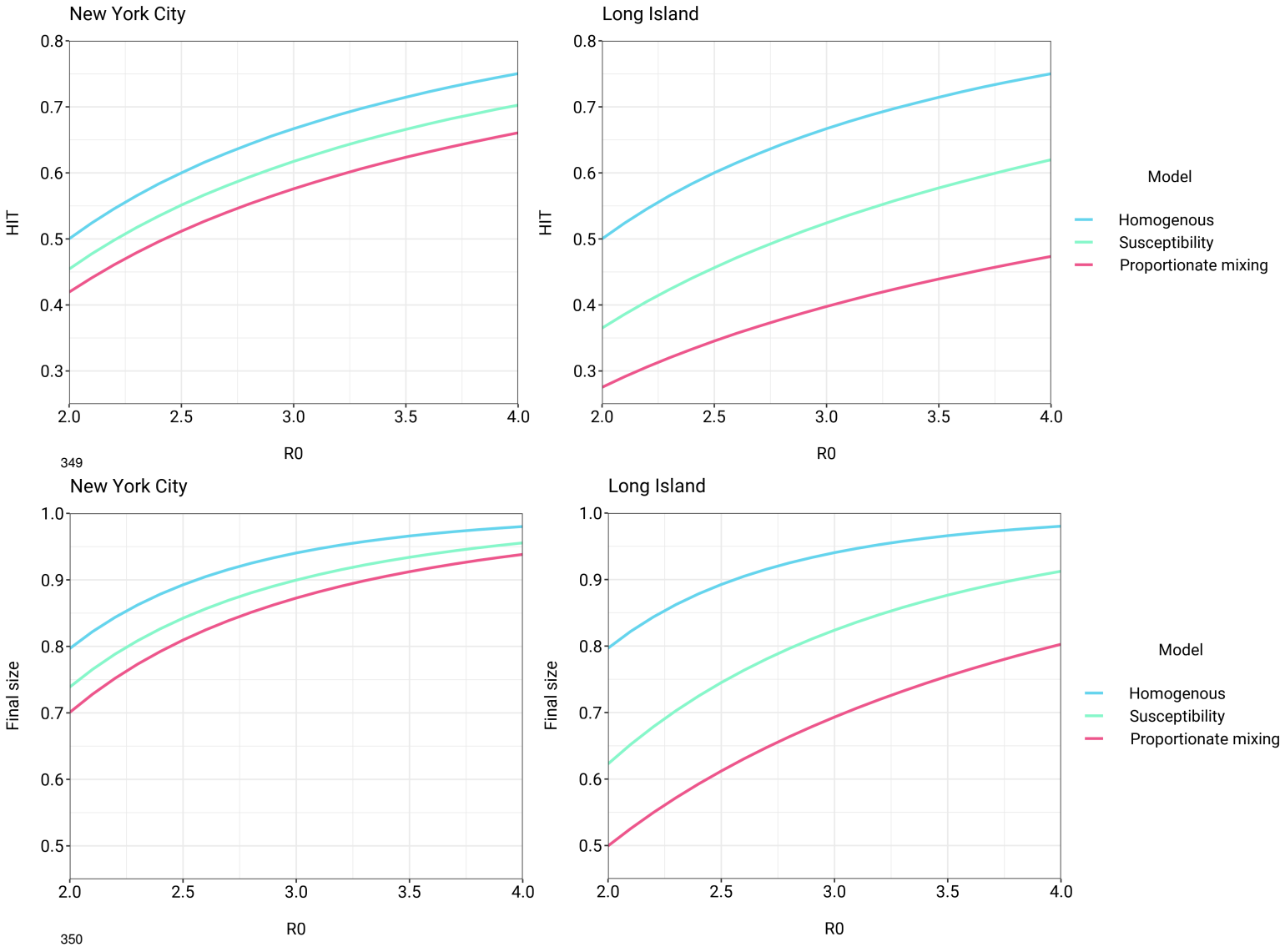
The serosurvey data were compatible with variable susceptibility models in which Hispanics or Latinos, non-Hispanic Black people, non-Hispanic Asians, and multiracial or other people had 2.25, 1.62, 0.86, and 1.28 times the susceptibility to infection relative to non-Hispanic whites in NYC, respectively, and 4.32, 1.96, 0.92, and 2.48 times the susceptibility to infection relative to non-Hispanic whites in Long Island, respectively. As with variable exposure models, these differences in susceptibility lowered herd immunity levels and final epidemic sizes relative to the homogeneous model (Supplementary Figure 1), but to a lesser extent; for instance, variable susceptibility models resulted in HITs $\sim$10% greater than HITs under proportionate mixing for Long Island.

The difference between these models is that incorporating heterogeneity in susceptibility only affects susceptible individuals, but heterogeneity in activity levels impacts both susceptible and infectious individuals: individuals from racial and ethnic groups with high activity levels are both more likely to be infected, and when infected, to infect a greater number of secondary cases. This contrast is clear when comparing the next-generation matrices for each model, which lists the average number of secondary infections caused by an infected individual from a given demographic group (Supplementary Figures 2 and 3). The epidemic resolves at an earlier stage in variable exposure models once these key transmission groups become immune because of this additional compound effect on transmission.

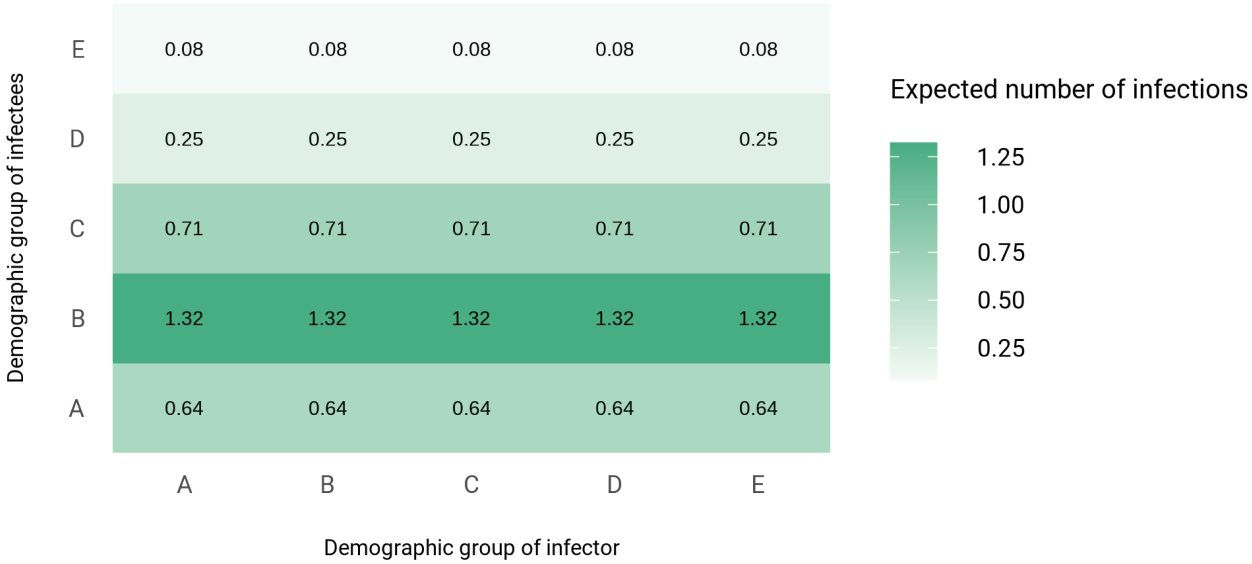
Our results contrasting mechanistic variable exposure and susceptibility models are in line with theoretical studies, which also indicate that models incorporating heterogeneity in exposure have more pronounced effects on HITs than models incorporating heterogeneity in susceptibility, assuming comparable continuous distributions of exposure and susceptibility [4, 6]. Tkachenko et al. showed that the $HIT = 1 - (1/R_0)^{(1/\lambda)}$, where λ is either $1 + CV^2$ for variable susceptibility models or $1 + CV^2(2 + \gamma_s CV)/(1 + CV^2)$ for variable exposure models, and CV is the coefficient of variation and γ_s is the skewness for the exposure distribution [6]. We calculated CV and skewness using the susceptibility and activity ratios and substituted those values into the HIT formula, which is an approximation because our exposure and susceptibility distributions are discrete; nonetheless, the approximations result in similar HIT curves to the simulation results (Supplementary Figure 4).

Sensitivity analyses

We conducted sensitivity analyses to assess whether assumptions on epidemic timing, and number and distribution of initial infected individuals affected parameter and HIT estimates. Varying the timing of epidemic start did not substantially affect HIT estimates, as long as the time between epidemic start and serosurvey t_s was reasonable large (e.g., $>$ 20 days) and assortativity was low ($\epsilon < 0.8$) (Supplementary Figure 9). The distribution and number of initial infected individuals also did not substantially affect HIT estimates for low levels of assortativity ($\epsilon < 0.8$) (Supplementary Figures 10 and 11). We limited our analyses to models with ϵ less than 0.8.

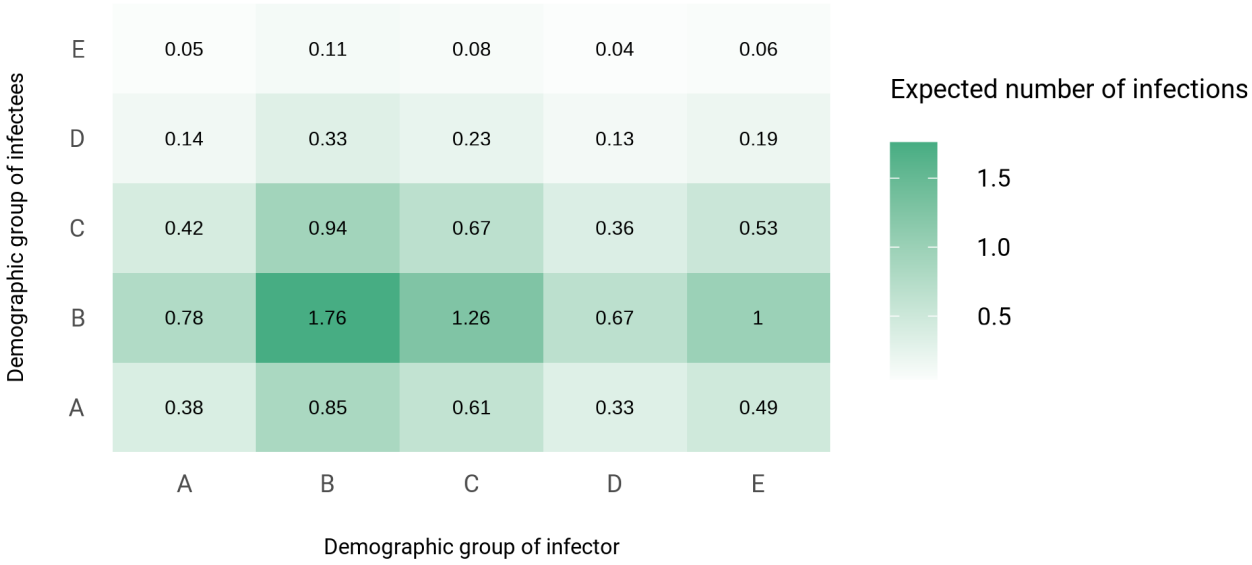


Supplementary Figure 1: Models incorporating variable susceptibility to COVID-19 fitted to NYC and Long Island serosurvey data result in reduced HITs (top) and final epidemic sizes (bottom across a range of R_0 values.

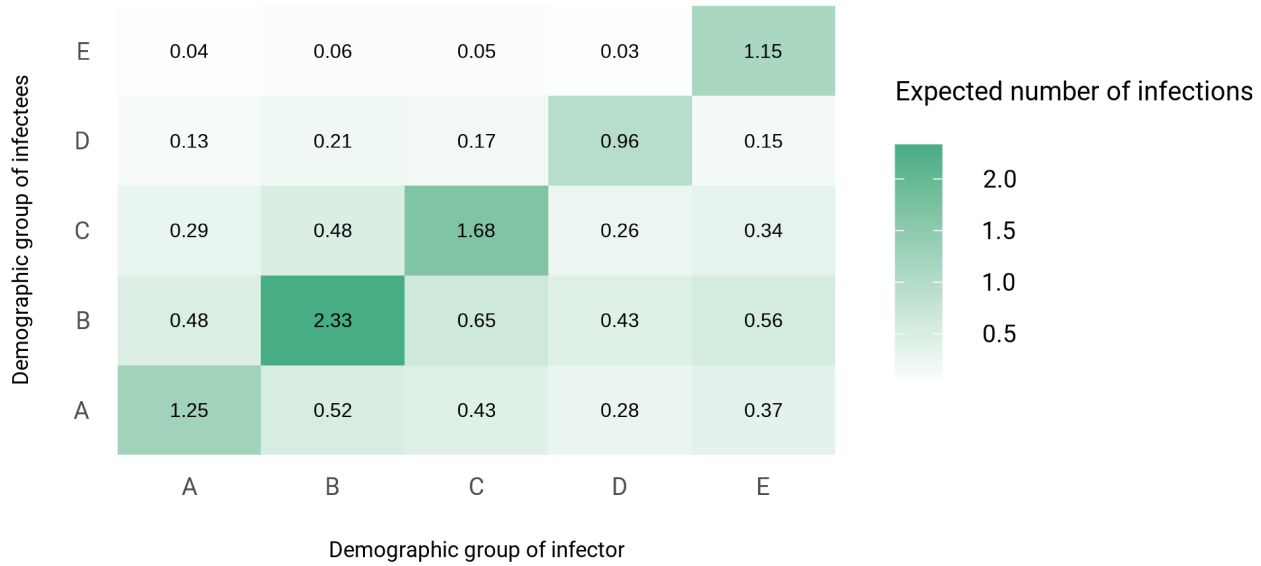


354

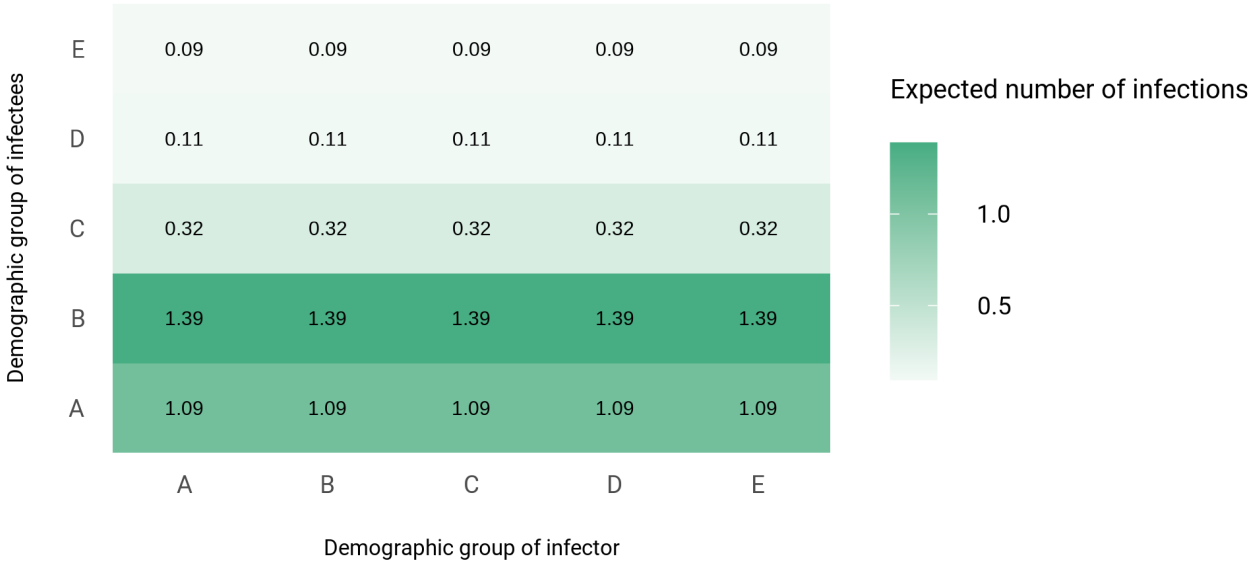
355



356

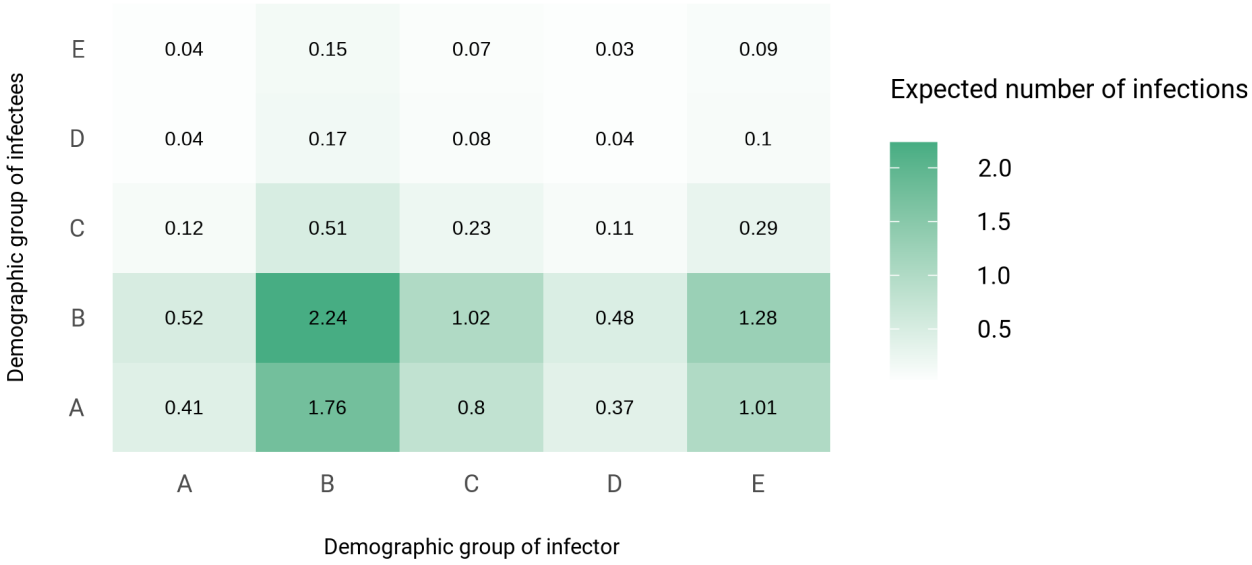


Supplementary Figure 2: Next-generation matrices for variable susceptibility (top), proportionate mixing (middle), and census-informed assortativity (bottom) models fitted to NYC seroprevalence data. Each column corresponds to an infected individual from a given demographic group and lists the expected number of secondary infections by group across the rows. Group A denotes non-Hispanic whites, B denotes Hispanics or Latinos, C denotes non-Hispanic African-Americans, D denotes non-Hispanic Asians, and E denotes multiracial or other demographic groups.

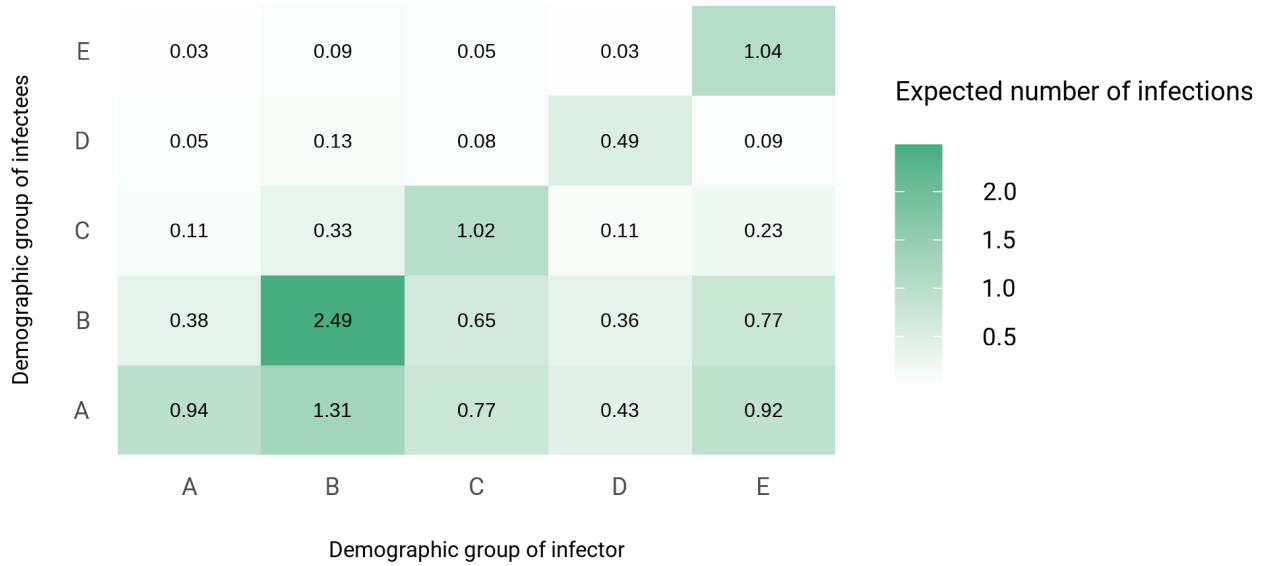


364

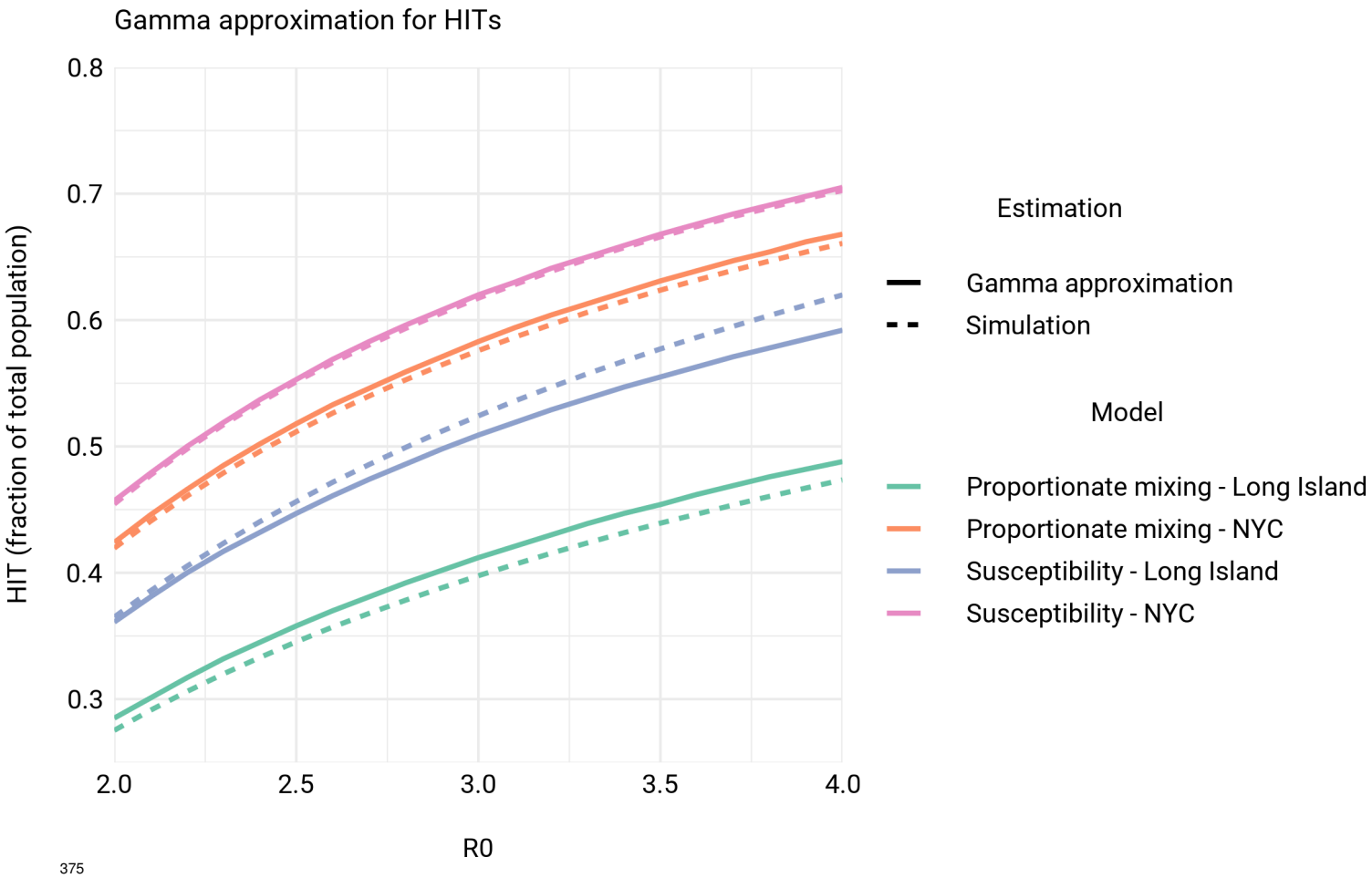
365



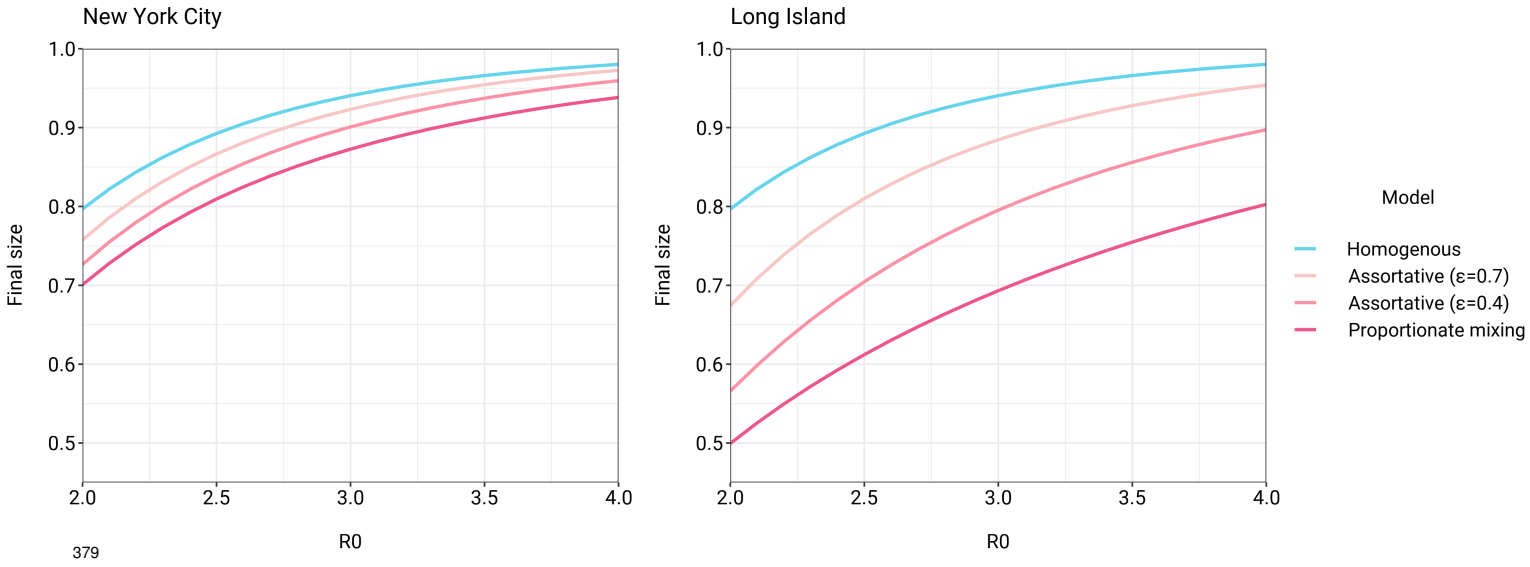
366



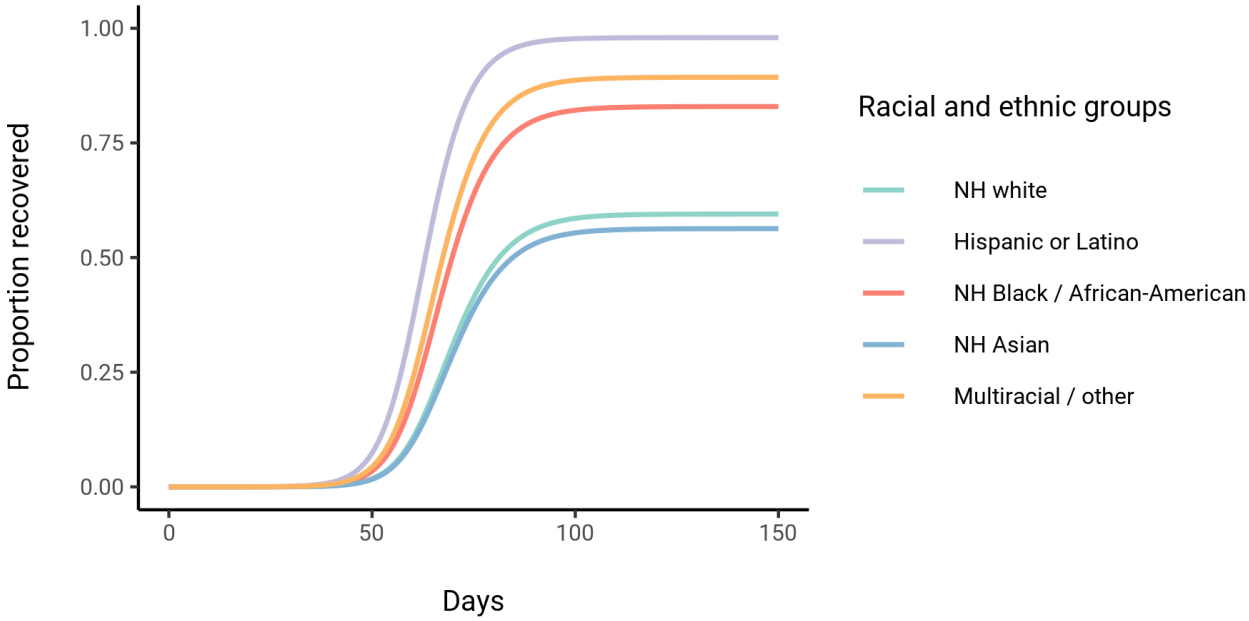
Supplementary Figure 3: Next-generation matrices for variable susceptibility (top), proportionate mixing (middle), and census-informed assortativity (bottom) models fitted to Long Island seroprevalence data. Each column corresponds to an infected individual from a given demographic group and lists the expected number of secondary infections by group across the rows. Group A denotes non-Hispanic whites, B denotes Hispanics or Latinos, C denotes non-Hispanic African-Americans, D denotes non-Hispanic Asians, and E denotes multiracial or other demographic groups.

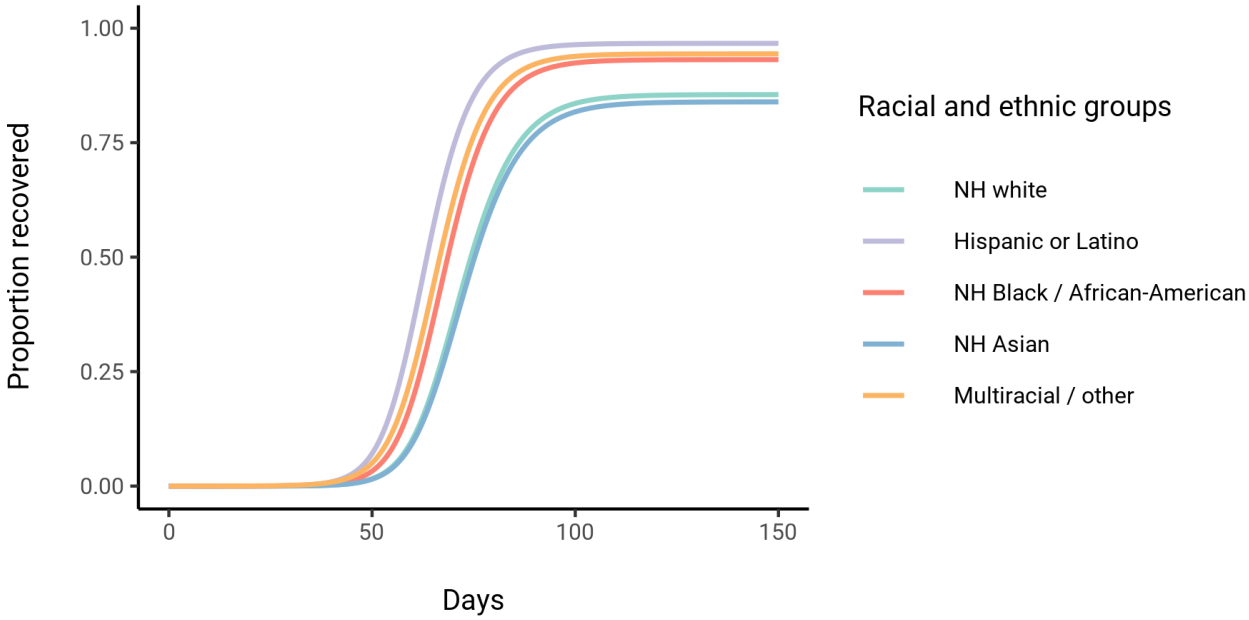


Supplementary Figure 4: Comparison of HITs from simulations to theoretical HIT curves for models with gamma distributed exposure and susceptibility [6]. See Supplementary Information for additional details on comparisons.

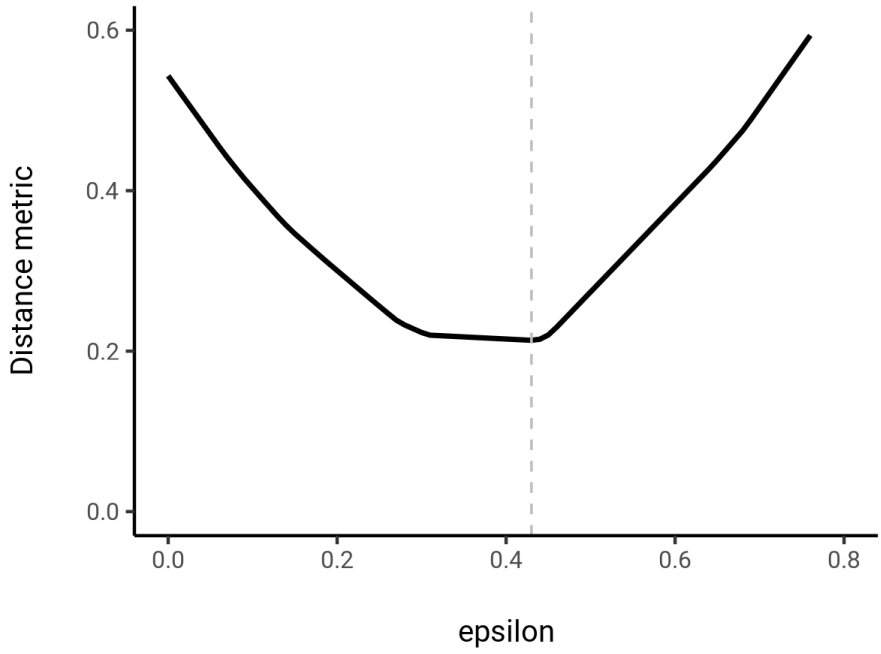


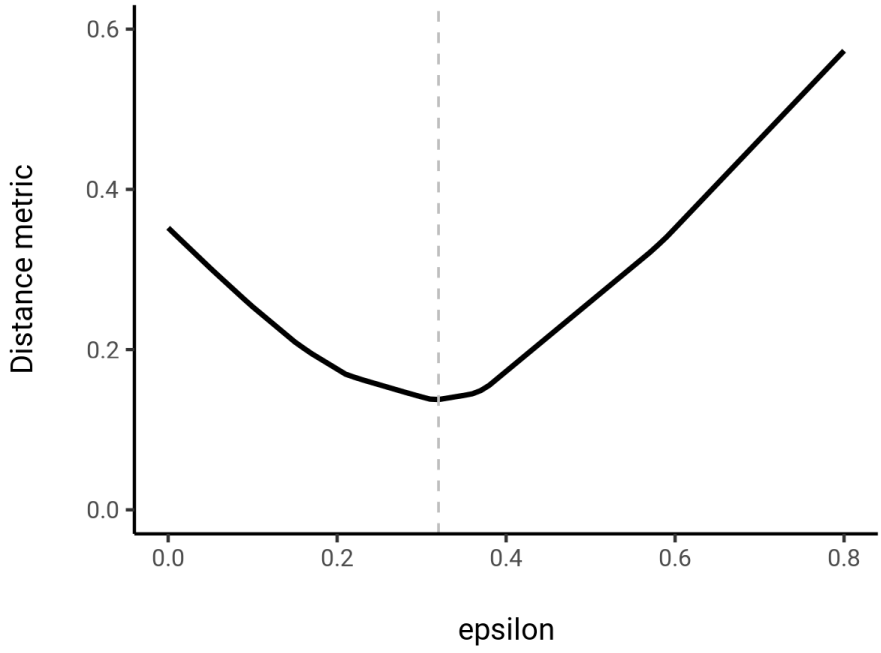
Supplementary Figure 5: Incorporating assortativity in variable exposure models results in increased final epidemic sizes across a range of R_0 values. Variable exposure models were fitted to NYC and Long Island serosurvey data.



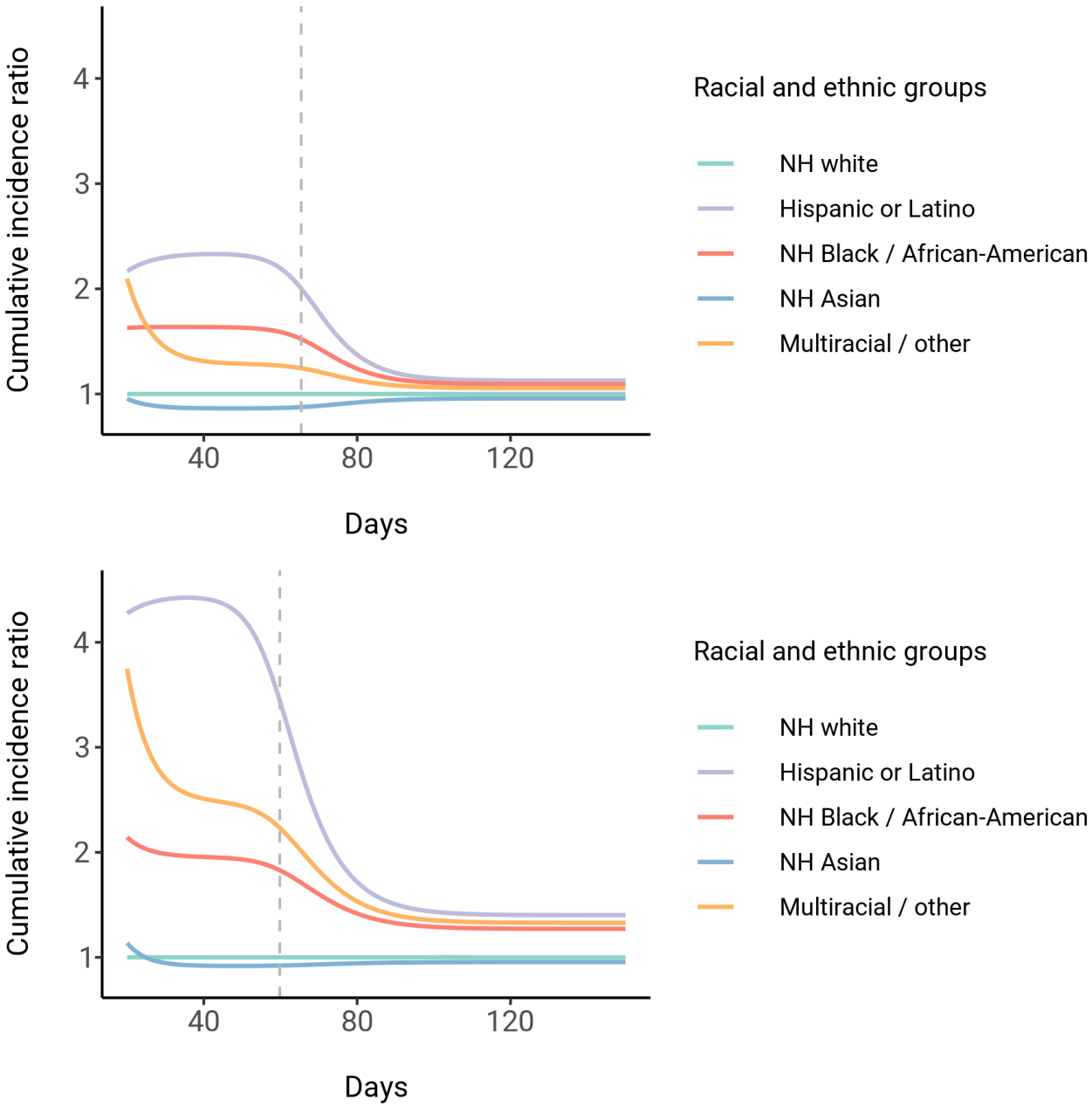


Supplementary Figure 6: Comparison of cumulative incidence trajectories for proportionate mixing (top) and assortative mixing ($\epsilon = 0.7$; bottom) models fitted to Long Island seroprevalence data.

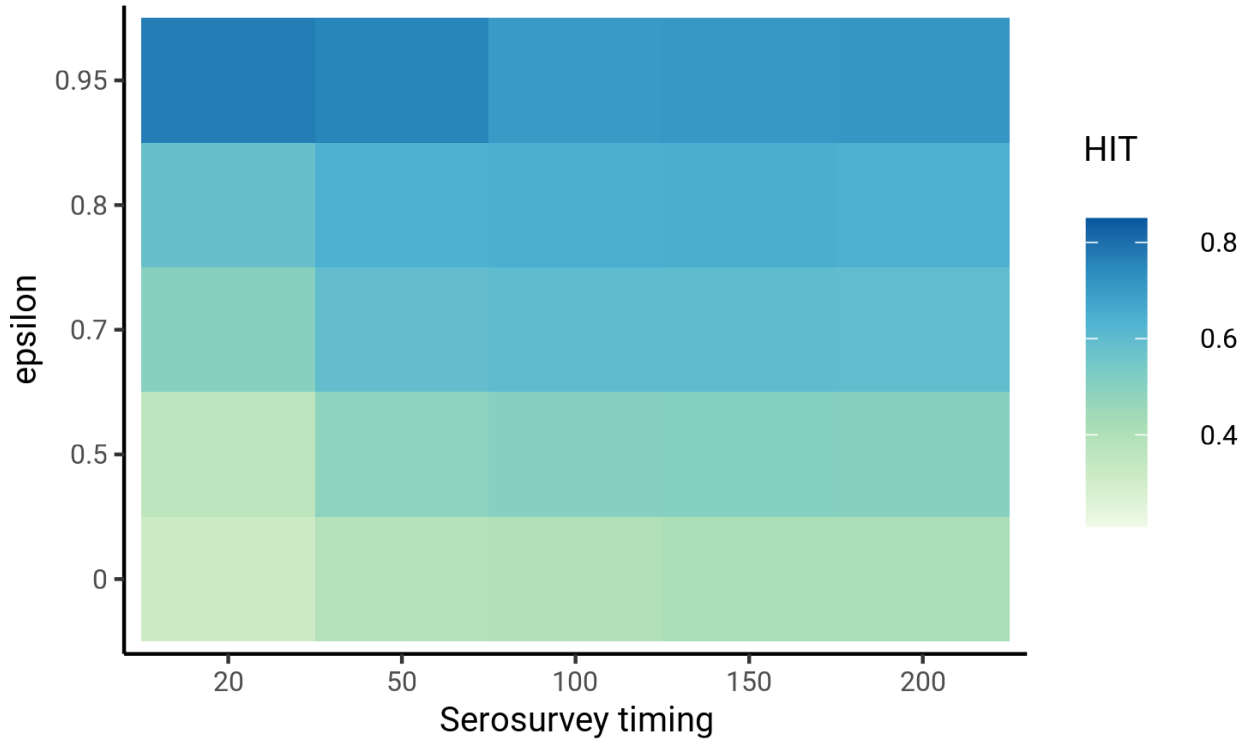




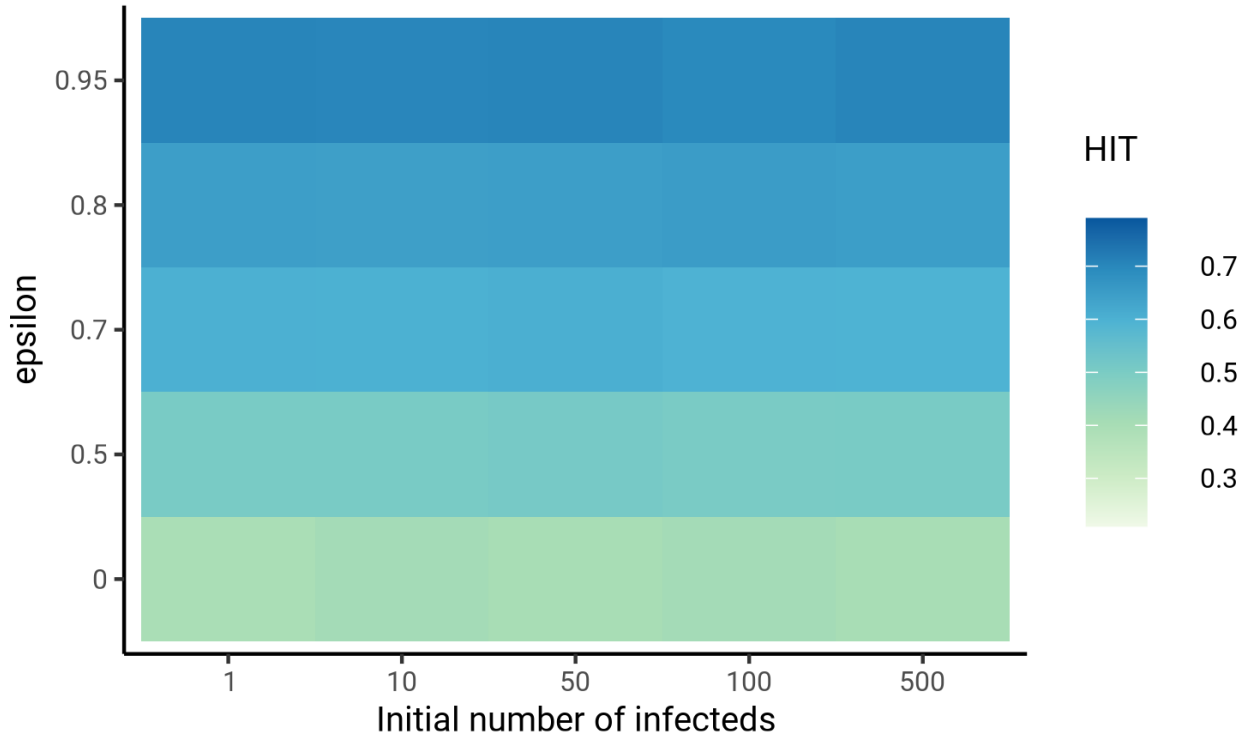
Supplementary Figure 7: Results of fitting ϵ in social contact matrices to census data for NYC (top) and Long Island (bottom).



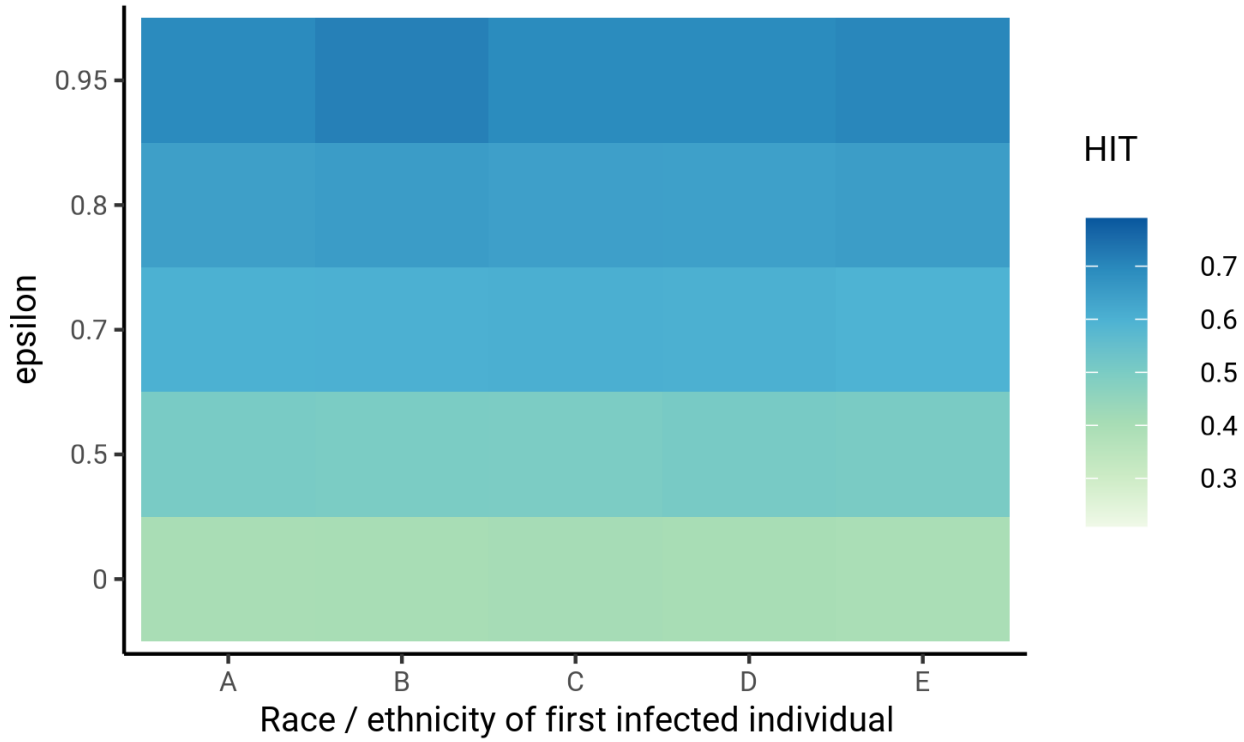
Supplementary Figure 8: Dynamics of cumulative incidence rate ratios relative to non-Hispanic whites in census-informed assortative mixing models, fitted to NYC (top) and Long Island (bottom) seroprevalence data. Dashed line represents the peak overall incidence for the epidemic.



Supplementary Figure 9: Sensitivity analysis on timing of the serosurvey relative to the start of the epidemic. Models were fit to Long Island seroprevalence data. Varying the timing of epidemic start did not substantially affect HIT estimates, as long as the time between epidemic start and serosurvey was reasonable (e.g., $>$ 20 days) and assortativity was low ($\epsilon < 0.8$).



Supplementary Figure 10: Sensitivity analysis on initial number of infected individuals in each group. Models were fit to Long Island seroprevalence data. The number of initial infected individuals did not substantially affect HIT estimates for low levels of assortativity ($\epsilon < 0.8$).



Supplementary Figure 11: Sensitivity analysis on race or ethnicity of first infected individual. Models were fit to Long Island seroprevalence data. The race or ethnicity of the first infected individual did not substantially affect HIT estimates for low levels of assortativity ($\epsilon < 0.8$). Group A denotes non-Hispanic whites, B denotes Hispanics or Latinos, C denotes non-Hispanic African-Americans, D denotes non-Hispanic Asians, and E denotes multiracial or other demographic groups.

	A	B	C	D	E
A	0.616	0.161	0.069	0.128	0.026
B	0.178	0.509	0.188	0.101	0.024
C	0.101	0.249	0.572	0.049	0.029
D	0.298	0.213	0.078	0.375	0.035
E	0.284	0.236	0.215	0.164	0.102

Supplementary Table 1: Exposure index matrix for NYC. Each row lists the proportions of people from each demographic group in an average neighborhood inhabited by an individual from the group indicated by the row index. Group A denotes non-Hispanic whites, B denotes Hispanics or Latinos, C denotes non-Hispanic African-Americans, D denotes non-Hispanic Asians, and E denotes multiracial or other demographic groups.

	A	B	C	D	E
A	0.770	0.119	0.038	0.056	0.017
B	0.432	0.358	0.142	0.047	0.022
C	0.269	0.278	0.378	0.047	0.028
D	0.57	0.130	0.067	0.208	0.026
E	0.532	0.189	0.126	0.080	0.073

Supplementary Table 2: Exposure index matrix for Long Island. Each row lists the proportions of people from each demographic group in an average neighborhood inhabited by an individual from the group indicated by the row index. Group A denotes non-Hispanic whites, B denotes Hispanics or Latinos, C denotes non-Hispanic African-Americans, D denotes non-Hispanic Asians, and E denotes multiracial or other demographic groups.

References

[1] Randolph, H. E. & Barreiro, L. B. Herd immunity: Understanding COVID-19. *Immunity* **52**, 737–741 (2020).

[2] Hill, A. N. & Longini, I. M., Jr. The critical vaccination fraction for heterogeneous epidemic models. *Math. Biosci.* **181**, 85–106 (2003).

[3] Britton, T., Ball, F. & Trapman, P. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science* **369**, 846–849 (2020).

[4] Gomes, M. G. M. *et al.* Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *medRxiv* (2020).

[5] Aguas, R. *et al.* Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. *medRxiv* 2020.07.23.20160762 (2020).

[6] Tkachenko, A. V. *et al.* Persistent heterogeneity not short-term overdispersion determines herd immunity to COVID-19 (2020). 2008.08142.

[7] Metcalf, C. J. E., Viboud, C., Spiro, D. J. & Grenfell, B. T. Using serology with models to clarify the trajectory of the SARS-CoV-2 emerging outbreak. *Trends Immunol.* **41**, 849–851 (2020).

[8] Wallinga, J., van Boven, M. & Lipsitch, M. Optimizing infectious disease interventions during an emerging epidemic. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 923–928 (2010).

[9] Bubar, K. M. *et al.* Model-informed COVID-19 vaccine prioritization strategies by age and serostatus (2020).

[10] Davies, N. G. *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211 (2020).

[11] Miller, I. F., Becker, A. D., Grenfell, B. T. & Metcalf, C. J. E. Disease and healthcare burden of COVID-19 in the united states. *Nat. Med.* **26**, 1212–1217 (2020).

[12] Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).

[13] Herzog, S. *et al.* Seroprevalence of IgG antibodies against SARS coronavirus 2 in belgium: a serial prospective cross-sectional nationwide study of residual samples. *medRxiv* 2020.06.08.20125179 (2020).

[14] Pollán, M. *et al.* Prevalence of SARS-CoV-2 in spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* **396**, 535–544 (2020).

[15] Shakiba, M. *et al.* Seroprevalence of COVID-19 virus infection in guilan province, iran. *medRxiv* 2020.04.26.20079244 (2020).

[16] Rosenberg, E. S. *et al.* Cumulative incidence and diagnosis of SARS-CoV-2 infection in new york. *Ann. Epidemiol.* **48**, 23–29.e4 (2020).

[17] Hallal, P. *et al.* Remarkable variability in SARS-CoV-2 antibodies across brazilian regions: nationwide serological household survey in 27 states. *medRxiv* 2020.05.30.20117531 (2020).

[18] Chamie, G. *et al.* SARS-CoV-2 community transmission disproportionately affects latinx population during Shelter-in-Place in san francisco. *Clin. Infect. Dis.* (2020).

[19] Moore, J. T. *et al.* Disparities in incidence of COVID-19 among underrepresented Racial/Ethnic groups in counties identified as hotspots during june 5-18, 2020 - 22 states, February-June 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1122–1126 (2020).

[20] Millett, G. A. *et al.* Assessing differential impacts of COVID-19 on black communities. *Ann. Epidemiol.* **47**, 37–44 (2020).

[21] Pan, D. *et al.* The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. *EClinicalMedicine* **23**, 100404 (2020).

[22] Chen, J. T. & Krieger, N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county vs ZIP code analyses. *Harvard Center for Population and Development Studies Working Paper Series* **19** (2020).

[23] Bassett, M. T., Chen, J. T. & Krieger, N. Variation in racial/ethnic disparities in COVID-19 mortality by age in the united states: A cross-sectional study. *PLoS Med.* **17**, e1003402 (2020).

[24] Hanage, W. P. *et al.* COVID-19: US federal accountability for entry, spread, and inequities-lessons for the future. *Eur. J. Epidemiol.* (2020).

[25] Chang, S. *et al.* Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* (2020).

[26] Zelner, J. *et al.* Racial disparities in COVID-19 mortality are driven by unequal infection risks (2020).

[27] Kissler, S. M. *et al.* Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in new york city. *Nat. Commun.* **11**, 4674 (2020).

[28] Krieger, N., Testa, C., Hanage, W. P. & Chen, J. T. US racial and ethnic data for COVID-19 cases: still missing in action. *Lancet* **396**, e81 (2020).

[29] Nishiura, H., Linton, N. M. & Akhmetzhanov, A. R. Serial interval of novel coronavirus (COVID-19) infections. *Int. J. Infect. Dis.* **93**, 284–286 (2020).

[30] Lauer, S. A. *et al.* The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).

[31] Wallinga, J., Kassteele, J. v. d. & Hens, N. Contact patterns for contagious diseases. In Held, L., Hens, N., O'Neill, P. & Wallinga, J. (eds.) *Handbook of Infectious Disease Data Analysis*, 93–110 (Chapman and Hall/CRC, 2019), 1 edn.

[32] Hethcote, H. W. Modeling heterogeneous mixing in infectious disease dynamics. In *Models for Infectious Human Diseases: Their Structure and Relation to Data*, 215–238 (Cambridge University Press, 1996).

[33] Richardson, E. T. *et al.* Reparations for black american descendants of persons enslaved in the U.S. and their estimated impact on SARS-CoV-2 transmission. *medRxiv* (2020).

[34] Bell, W. A probability model for the measurement of ecological segregation. *Soc. Forces* **32**, 357–364 (1954).

[35] Henry Akintobi, T. *et al.* Community engagement of african americans in the era of COVID-19: Considerations, challenges, implications, and recommendations for public health. *Prev. Chronic Dis.* **17**, E83 (2020).

[36] Thakur, N., Lovinsky-Desir, S., Bime, C., Wisnivesky, J. P. & Celedón, J. C. The structural and social determinants of the Racial/Ethnic disparities in the U.S. COVID-19 pandemic. what's our role? *Am. J. Respir. Crit. Care Med.* **202**, 943–949 (2020).

[37] Tai, D. B. G., Shah, A., Doubeni, C. A., Sia, I. G. & Wieland, M. L. The disproportionate impact of COVID-19 on racial and ethnic minorities in the united states. *Clin. Infect. Dis.* (2020).

[38] Khazanchi, R., Evans, C. T. & Marcelin, J. R. Racism, not race, drives inequity across the COVID-19 continuum. *JAMA Netw Open* **3**, e2019933 (2020).

[39] Krieger N, Testa C, Chen JT, Waterman PD, Hanage WP. A warning against using static US county-level community data to guide equity in COVID-19 vaccine distribution: Temporal and spatial correlations of community characteristics with COVID-19 cases and deaths vary enormously and are increasingly uninformative. *HCPDS Working Paper* **20** (2020).

[40] Goldstein, E., Pitzer, V. E., O'Hagan, J. J. & Lipsitch, M. Temporally varying relative risks for infectious diseases: Implications for infectious disease control. *Epidemiology* **28**, 136–144 (2017).

[41] Koopman, J. S. *et al.* Assessing risk factors for transmission of infection. *Am. J. Epidemiol.* **133**, 1199–1209 (1991).

[42] Kahn, R., Kennedy-Shaffer, L., Grad, Y. H., Robins, J. M. & Lipsitch, M. Potential biases arising from epidemic dynamics in observational seroprotection studies. *Am. J. Epidemiol.* (2020).

[43] Feehan, D. & Mahmud, A. Quantifying population contact patterns in the united states during the COVID-19 pandemic (2020).

[44] Millett, G. A. *et al.* White counties stand apart: The primacy of residential segregation in COVID-19 and HIV diagnoses. *AIDS Patient Care STDS* **34**, 417–424 (2020).

[45] Boyd, R., Lindo, E., Weeks, L. & McLemore, M. On racism: A new standard for publishing on racial health inequities. `https://www.healthaffairs.org/do/10.1377/hblog20200630.939347/full/`. Accessed: 2020-11-27.

[46] Blau, F. D., Koebe, J. & Meyerhofer, P. A. Who are the essential and frontline workers? Tech. Rep. w27791, National Bureau of Economic Research (2020).

[47] Yang, T.-C., Emily Choi, S.-W. & Sun, F. COVID-19 cases in US counties: roles of racial/ethnic density and residential segregation. *Ethn. Health* 1–11 (2020).

[48] Schmidt, H., Gostin, L. O. & Williams, M. A. Is it lawful and ethical to prioritize racial minorities for COVID-19 vaccines? *JAMA* **324**, 2023–2024 (2020).

[49] Clapham, H. *et al.* Seroepidemiologic study designs for determining SARS-COV-2 transmission and immunity. *Emerg. Infect. Dis.* **26**, 1978–1986 (2020).

[50] Larremore, D. B. *et al.* Estimating SARS-CoV-2 seroprevalence and epidemiological parameters with uncertainty from serological surveys (2020).

[51] Kishore, N. *et al.* Lockdown related travel behavior undermines the containment of SARS-CoV-2 (2020).

[52] Flannery, D. D. *et al.* SARS-CoV-2 seroprevalence among parturient women in philadelphia. *Sci Immunol* **5** (2020).