Impacts of COVID-19 interventions: Health, economics, and inequality

X. Flora Meng (xflorameng@gmail.com)
Massachusetts Institute of Technology https://orcid.org/0000-0002-2998-5101
Dalton Jones
Massachusetts Institute of Technology
Roberto Rigobon
Massachusetts Institute of Technology
Munther Dahleh
Massachusetts Institute of Technology

Article

Keywords: Coronavirus disease 2019 (COVID-19), inequalities, economics

Posted Date: April 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-406480/v1

Impacts of COVID-19 interventions: Health, economics, and inequality

X. Flora Meng[a,b,1], **Dalton J. Jones**[a,b], **Roberto Rigobon**[b,c], **and Munther A. Dahleh**[a,b,c]

[a]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA; [b]Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA, USA; [c]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA

This manuscript was compiled on April 9, 2021

Coronavirus disease 2019 (COVID-19) is exacerbating inequalities in the US. We build an agent-based model to elucidate the differential causal effects of nonpharmaceutical interventions on different communities and validate the results with US data. We simulate viral transmission and the consequent deterioration of economic conditions on socioeconomically disadvantaged and privileged populations. As found in data, our model shows that the trade-off between COVID-19 deaths and deaths of despair, dependent on the lockdown level, only exists in the socioeconomically disadvantaged population. Moreover, household overcrowding is a strong predictor of the infection rate. The model also yields new insights that fill in the gaps of our data analysis. While subsidisation narrows the socioeconomic gap in deaths of despair, the combination of testing and contact tracing alone is effective at reducing disparities in both types of death. Our results contribute to policy modelling and evaluation for reducing inequality during a pandemic.

As the world continues to battle coronavirus disease 2019 (COVID-19), growing evidence indicates that the pandemic is exacerbating inequalities in the US (1–9). Many studies have focused on racial and ethnic disparities in health outcomes (1–7). For example, non-Hispanic African American patients were more than twice likely to be hospitalized than non-Hispanic white patients in a large health care system in California (1). Overrepresentation of Non-Hispanic black patients among COVID-19 hospitalizations has also been found in Louisiana (2) and Georgia (3). Not only are racial and ethnic minorities at increased risk of comorbidities that are associated with severe illness (10–13), but also they are disadvantaged by structural factors such as residential segregation and employment in essential services (4).

In addition to minority status, COVID-19 case and death rates are often higher in urban counties that rank lower in socioeconomic status, housing, and transportation (5). According to studies of New York City, the test positivity rate was high in neighbourhoods that were characterized by poverty, big households, and a large non-Hispanic black or immigrant population (6, 8). Similar conclusions were drawn about Massachusetts (7). The high positivity rate was partially explained by insufficient testing that was available to people in poverty and minority groups (14). Moreover, poorer areas across the US exhibited less physical distancing (9, 15). This is particularly disturbing because adequate testing and quarantining have been shown to effectively stop the spread of the disease (16).

The growing body of evidence that the COVID-19 pandemic is worsening inequalities in the US stresses the importance of understanding how public policy influences different communities (17). Socioeconomically disadvantaged regions not only tend to have higher COVID-19 death rates but also are less resilient to economic distress (4, 5, 18, 19). The latter, which can be manifested in unemployment, may further lead to deaths of despair from suicide, drug overdose, and alcoholism (20, 21). This link between economic hardship and deaths of despair suggests a possible trade-off between recession-related deaths and COVID-19 deaths.

Although obtaining a clear definition of vulnerability amid the pandemic remains elusive (22, 23), it is clear that certain features correlate with bad outcomes. Our analysis of census, mobility, and COVID-19 data of the US confirms disparities in both COVID-19 deaths and deaths of despair. The strongest predictors for the regional COVID-19 death rate are income, age, race, and household overcrowding. Moreover, we find that regions with worse health outcomes also tend to have higher unemployment and eviction rates. We further investigate the effects of income and household overcrowding on health and economic outcomes. Our analysis confirms the widely believed trade-off between COVID-19 deaths and economic distress-related deaths as the level of lockdown changes (22, 24–26). However, we find that this trade-off only exists among socioeconomically disadvantaged counties. Furthermore, the percent of overcrowded households and the COVID-19 death rate are positively correlated. Although our data analysis is inconclusive on whether the identified effects are causal, we answer this question affirmatively by reproducing similar results using agent-based modelling.

While it is crucial that government interventions reduce inequality during the pandemic, designing good interventions is challenging (17, 22, 24). First, multiple criteria, such as health and economic impacts, can be used for policy evaluation, which may give conflicting advice (24, 25). Second, it is often hard to estimate the causal effects of a single intervention from data (9). We can consider society fighting a pandemic as a complex system that has time-varying nonlinear interactions. Moreover, multiple interventions are usually at work simultaneously (27, 28). Third, data only exists for the policies that have been implemented (25). Finally, granular data needed for definitive conclusions are sometimes scarce and incomplete (29).

In order to overcome the limitations of observational data, we develop an agent-based model that simulates the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the consequent rise in deaths of despair. The model incorporates key elements including socioeconomic status, age-dependent risks, household transmission, asymptomatic transmission, and hospital capacity. Agent-based modelling enables analysis of causal links between various

[1]To whom correspondence should be addressed. E-mail: xmeng@mit.edu.

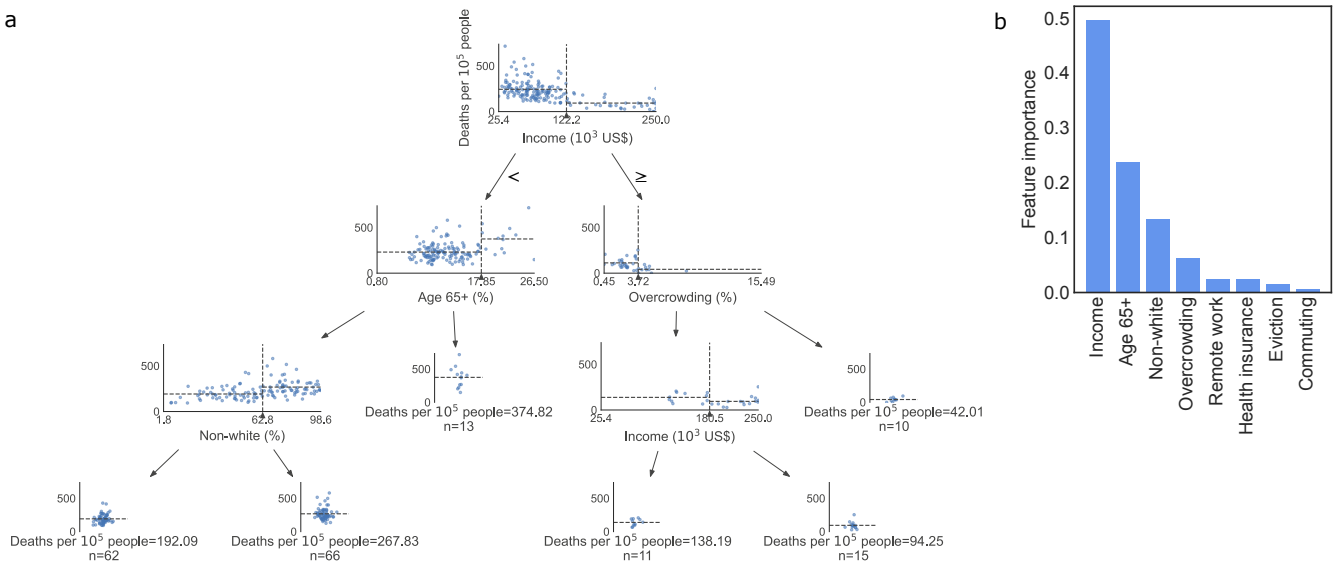


Fig. 1 | Regional features associated with local COVID-19 death rates. a, We build a decision tree that predicts the COVID-19 death rate of New York City by ZCTA. We show a pruned tree here to illustrate the method and provide the full tree in Extended Data Fig. 1. The x and y-axes of each scatterplot are the feature used for the split and the number of deaths per $100,000$ people, respectively. ZCTAs are divided into two subsets at the vertical lines so that the death rates are close to the average (marked by horizontal lines) within each group. **b**, We compute the importance of a feature in the decision tree as the normalized total reduction of the mean squared error that is attributable to the feature.

policies and metrics of interest (30–35). We investigate the effects of four nonpharmaceutical interventions (NPIs) on inequality, namely, lockdown, testing along with contact tracing, government subsidisation, and housing provision. We use the COVID-19 death number to measure health outcomes and deaths of despair as a proxy for economic consequences. Our model generates a stylized population that comprises socioeconomically disadvantaged and privileged people, referred to as poor and rich, respectively, for brevity. As shown with US data, we find that the trade-off between COVID-19 deaths and deaths of despair, hinging on the lockdown level, only exists in the poor community. While subsidisation narrows the socioeconomic gap in deaths of despair, the combination of testing, contact tracing, and home isolation alone is effective at reducing disparities in both types of death. Similar to our data findings, our model also suggests a strong link between household overcrowding and the COVID-19 infection rate, which we quantify with mathematical analysis.

Our simulation not only reflects patterns observed in US data but also yields new insights that fill in the gaps of our data analysis. Our findings demonstrate the importance of targeted intervention design to relieve both health-related and economic pressure on socioeconomically disadvantaged populations. Our model suggests a moderate lockdown, adequate testing combined with contact tracing and home isolation, sufficient targeted subsidies, and mitigation of overcrowding in housing. Our results contribute to policy modelling and evaluation for reducing inequalities during a pandemic. The paper focuses on the US, but our approach and results can be extended to other regions in the world.

Results

Data Analysis. We start by building a decision tree to identify the strongest predictors for the regional COVID-19 death rate. A decision tree is a predictive model that sequentially partitions an input dataset into subsets so that prediction accuracy improves after each split (36). Decision tree learning provides a natural method of feature selection by quantifying the contribution of each feature to the prediction task (37). We use census (38, 39) and eviction data (40) from 2019, and COVID-19 death data (41) from 2020 in New York City by ZIP Code Tabulation Area (ZCTA). The census data contains many factors including household overcrowding, the percent of 65-and-older population, the percent of home-based workers, commuting, health insurance coverage, median income, and race.

Fig. 1a shows a pruned tree that is fitted to the ZCTA-level data (see the complete decision tree in Extended Data Fig. 1). The top scatterplot contains all ZCTAs in the dataset. Income is identified as the feature that best splits the set with a threshold at US$122,200. The percent of 65-and-older population is the best variable to further split the lower-income group (at 17.85%), whereas the percent of household overcrowding is chosen to divide the higher-income group (at 3.72%). The decision tree is built iteratively this way. Although our goal with the dataset is to evaluate feature importance rather than predict the death rate, the decision tree sheds light on the link between regional characteristics and local health outcomes. High COVID-19 death rates are often associated with low income, a large population of seniors and racial minorities, lack of health insurance, high eviction rates, household overcrowding, commuting, and uncommonness of working from home. An exception to this pattern is the first appearance of overcrowding in the decision tree as shown in Fig. 1a. Surprisingly, the ZCTAs with more household overcrowding had lower death rates. It turns out that these ZCTAs are mostly in Lower and Midtown Manhattan where single young professionals with high salaries tend to live.

We also compare the best and worst segments in the decision tree and find economic inequality in addition to health

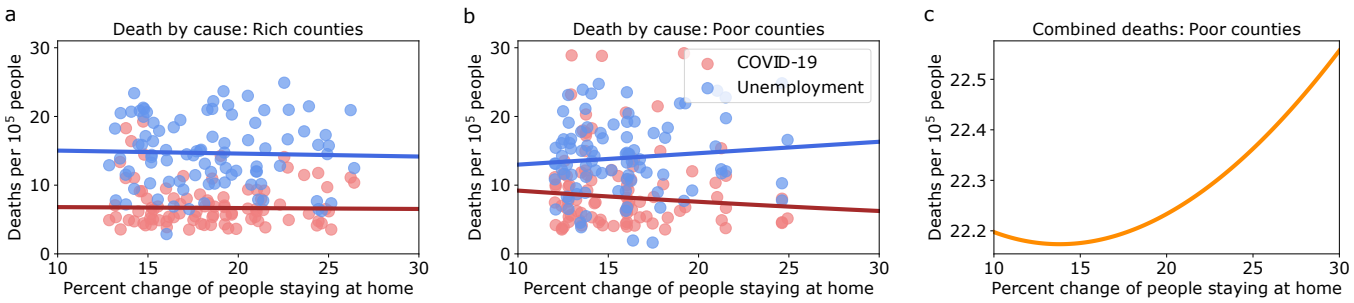


Fig. 2 | Lockdown and social distancing measures that are meant to curb the spread of COVID-19 can exacerbate inequalities. We compare the richest (**a**) and poorest (**b**) counties in the US as measured by median income. **a**, Affluent counties are resilient to the economic shock of lockdown and social distancing measures. **b**, In contrast, poor counties face the dilemma of whether to die from COVID-19 infection or economic distress. **c**, Combining estimates from both regression reveals the health and economic trade-off for poor counties.

disparities. Not only did the worst segment have a higher unemployment rate (3.03%) than the best one (2.88%) in 2019, but the former group also had a steeper increase (5.69%) in 2020 than the best segment (4.22%). The 2020 unemployment rates are projected at the ZCTA level by calculating the percent change in unemployment of the county containing the ZCTA and applying this change to the ZCTA level data from 2019. The unemployment gap coincides with the differential eviction rates, which are 0.37% and 0.30% for the worst and the best segments, respectively.

Having learned a decision tree, we then compute the importance of a feature as the normalized total reduction of the mean squared error in estimating the COVID-19 death rate of a ZCTA that is attributable to the feature. As shown in Fig. 1b, the highest-scoring features are income (0.50), the percent of 65-and-older population (0.24), the percent of non-white population (0.13), and household overcrowding (0.06).

We further investigate the effects of income on regional health and economic outcomes. We compare the poorest and richest counties in the US as measured by median income and find that the widely believed health and economic trade-offs of lockdowns only exist in poor counties (Fig. 2). The annual median personal income is less than US$70,000 for the poorest counties, in contrast to above US$80,000 for the richest counties. For this analysis, we combine datasets that measure median income, the unemployment rate, the size of the labour force, the percent change from baseline of people staying at home (as a measure of lockdown severity), and COVID-19 death counts (38, 42–44).

One puzzle presented by the data is that the level of lockdown appears to be positively correlated with the COVID-19 death rate. Our hypothesized reason is that locations with the most severe outbreaks responded with the most drastic measures. After accounting for disease progression and reporting delays, we observe that stricter policies correspond to lower death rates in poor counties whereas the correlation is weak for rich counties, with the latter possibly due to residual effects from the first wave of COVID-19 (Fig. 2a,b). Specifically, we perform linear regression of the logarithmic transformation of the COVID-19 death rate on the mobility change, delaying the death data by 62 days. Fig. 2a,b indicate that there is indeed a damping effect of lockdown and social distancing measures on COVID-19 transmission, which is consistent with conclusions in (27, 45–47).

In order to compare economic impacts with health outcomes, we project excess deaths caused by economic downturns. Prior work has shown that unemployment increases an individual's mortality hazard by at least 73% (48). Although the aggregate mortality effects of economic stagnation are open to question, the increased hazard of death associated with individual joblessness has been well established (26, 48–50). Using the individual risk inferred in (48), we project the one-year death count attributable to the pandemic-related unemployment shock (Fig. 2a,b). Specifically, we estimate the total number of newly unemployed workers in each county using the size of the labour force and the increase in the unemployment rate in 2020 compared to 2019. We then use the all-cause mortality rate from 2019 of each county to calculate the mortality rate of the newly unemployed workers. Finally, we perform linear regression of the projected death rate associated with unemployment on the mobility change. As shown in Fig. 2a,b, the unemployment shock affects poor counties more than the rich ones. One explanation is that the reduction in mobility was significantly more in wealthier areas than poorer areas during the pandemic (9), which indicates that the affluent can weather the economic repercussions of lockdowns partially because their jobs allow for flexibility in terms of working remotely. Prior work has drawn similar conclusions that excess mortality is disproportionately high in disadvantaged groups such as African Americans and people with low educational attainment (26, 51).

Fig. 2a,b suggest that the widely believed health and economic trade-offs of lockdowns only exist in poor counties. Fig. 2c illustrates this trade-off by summing regression estimates of COVID-19 deaths and projected excess deaths attributable to unemployment. Our findings confirm marked differences in the way that social distancing and lockdown measures impact different groups.

We also explore the association between household overcrowding and regional health outcomes. Household overcrowding is the condition where there is more than one person per room (52), which may accelerate the spread of respiratory diseases such as COVID-19. We use the Comprehensive Housing Affordability Strategy (CHAS) data prepared by the US Census Bureau for the 2013–2017 period (52). We focus on the largest four states for the number of urban counties, which are California, Florida, New Jersey, and New York. A county is urban if at least 95% of the population live in urban areas. The rurality data is published by the US Census Bureau for the year 2010 (53). We restrict the death data (44) to the end of July 2020 to take into account roughly the first six months since the first recorded US case. The qualitative results re-

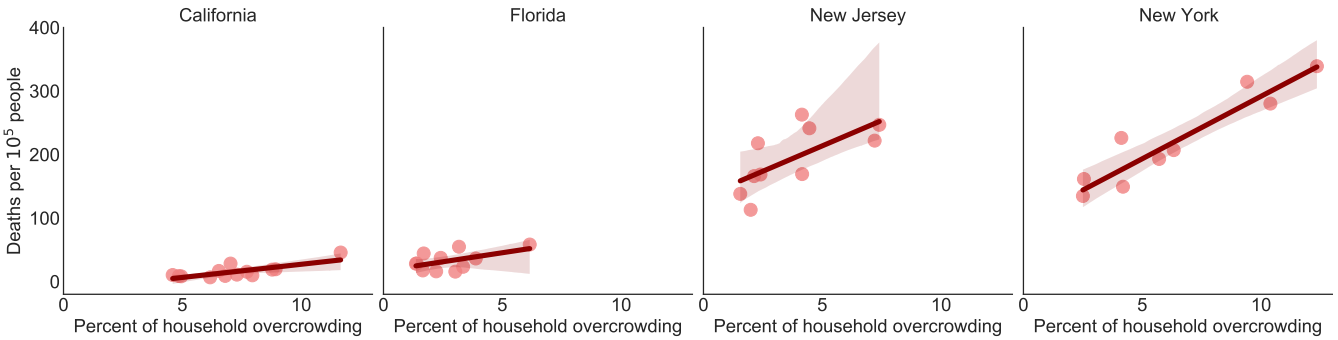


Fig. 3 | The COVID-19 death rate is positively correlated with household overcrowding in urban counties. California, Florida, New Jersey, and New York are the largest four states for the number of counties of which at least 95% of the population live in urban areas. For each state, the solid line and the shaded area represent robust linear regression that downweights outliers with a 95% confidence interval.

main the same as the time window considered changes. Fig. 3 indicates a positive correlation between the percent of household overcrowding and the COVID-19 death rate. However, data of rural areas appears particularly noisy (Extended Data Fig. 2). This may be explained by several reasons including low population density, large regional variations in infection patterns, and disease outbreaks at different times.

Our findings imply an underlying mechanism at play that causes worse health and economic outcomes for poorer communities. Although our data analysis is inconclusive on whether the identified effects are causal, we answer this question affirmatively by reproducing similar results using agent-based modelling.

Model. We develop an agent-based model that simulates the transmission of SARS-CoV-2 and the consequent rise in deaths of despair. The model takes into account key factors such as socioeconomic status, age-dependent risks, household transmission, asymptomatic transmission, and hospital capacity. We examine the effects of four NPIs on inequality, which are lockdown, testing along with contact tracing, government subsidisation, and housing provision. We overview the model in this section, providing details in Methods.

The model initializes a population where each individual has their own attributes that influence their state transitions during simulation. We sample each individual's age from the distribution as specified by the US Census Bureau's 2019 national population estimates (54). Our stylized model considers people under age 20 as students, those aged 20 to 69 years as workers, and people aged over 69 as retirees. Moreover, everyone is economically active at the start of a simulation. An active individual's output is the sum of the personal output and the connection output, the latter being a measure of the benefits of staying connected to society. Once infected, an individual progresses stochastically from asymptomatic or presymptomatic, to symptomatic, hospitalized, admitted to the ICU, and deceased, with the possibility of recovery at any stage if not deceased (Extended Data Fig. 3a). Epidemiological parameters and their sources (55–59) are in Extended Data Fig. 4. An individual is economically inactive during hospitalization and at death. Moreover, an individual loses connection output while in quarantine or staying home (Extended Data Fig. 3a). Taking into account factors that vary across communities such as the comorbidity rate and health care quality (1, 4, 5), we assume that a small fraction of the population are vulnerable to severe illness, exclusive to the poor community. Once infected, vulnerable people are more likely to experience worsening symptoms than an average person.

We incorporate in our model random graphs to simulate virus transmission and economic activities. In consideration of the high transmission rate in households (55, 60), we construct a collection of complete graphs to represent households where any pair of members in the same household are connected. To capture socioeconomic disparities, we assume that 90% of the population are poor and the rest are rich in expectation. A rich person is characterized by a high output and a small household size. In addition, we overlay the household network with an economic network that represent economic activities which rely on in-person contact (Extended Data Fig. 3b). We generate economic networks using the Watts–Strogatz random graph (61), a classic model that produces the small-world phenomenon as observed in many real-world networks.

Our model considers dynamics at both household and aggregate levels, which include deaths of despair, recession, and undertreatment. We take into consideration deaths of despair that are linked to financial stressors. Specifically, the probability that an individual dies from despair is a function that decreases with per capita output in the household. At the aggregate level, with government subsidies taken into account, a drop in the total output leads to more workers becoming economically inactive. In addition, our model incorporates the scenario in which hospitals are overwhelmed and poor patients are undertreated. Undertreatment increases the chance of deterioration in patients.

Impacts of NPIs on Inequality. It has been widely accepted by now that there is a trade-off between saving lives from the pandemic and saving lives from recession. What has been less scrutinized, however, is how this trade-off varies in different communities and under various policies (24–26). As we have observed in US data, poorer counties not only have had more COVID-19 deaths but also will see more recession-induced deaths. We investigate the effects of four NPIs on inequality, which are lockdown, testing along with contact tracing, government subsidisation, and housing provision. Our model suggests that, for most NPIs considered, the poor community suffers significantly more than the rich counterpart in terms of both types of death.

Unless stated otherwise, we simulate the dynamics within

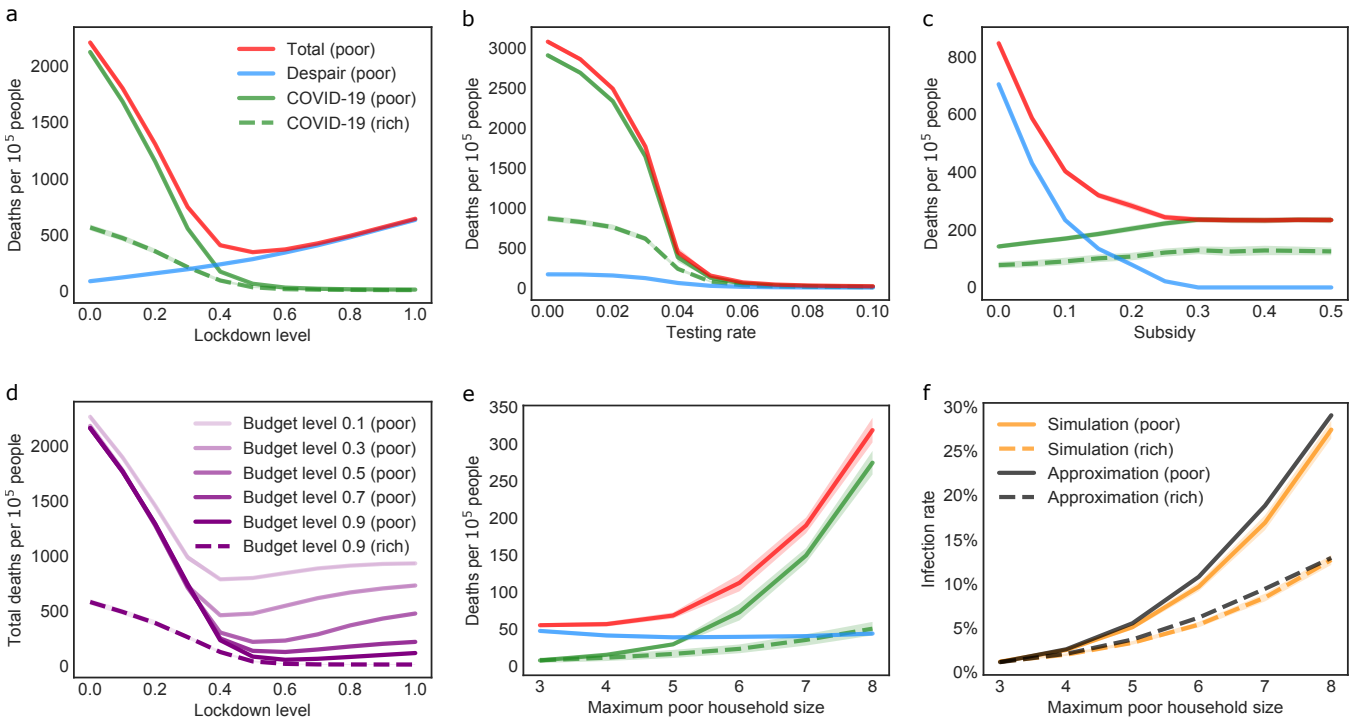


Fig. 4 | Impacts of COVID-19 NPIs on socioeconomic inequality. The fatality rate is calculated within each socioeconomic group. Since the rate of death of despair is close to zero for the rich community, we only show COVID-19 deaths for this group. **a**, The trade-off between COVID-19 deaths and deaths of despair only exists in the poor community. **b**, The combination of testing and contact tracing alone is sufficient for eliminating socioeconomic disparities in both types of death. **c**, Increasing subsidies effectively reduces the gap in deaths of despair. **d**, For the strategy of prioritizing the neediest people for subsidies, a larger budget narrows disparities in the total death rate and enables stricter lockdown before economic consequences exceed marginal health benefits. Since the rate of death of despair is almost the same for the rich community at all budget levels, we only show this group's results at a budget level of 0.9. **e**, Household overcrowding exacerbates COVID-19 in the poor community. **f**, The effect of household overcrowding can be explained by mean-field approximation. Curves and shades are the averages and the standard deviations of 100 trials, respectively.

the population for 180 days, initializing the percentage of infections to 0.1%. We assume that retirees stay at home in all simulations, as this policy has been commonly recommended for reducing COVID-19 hospitalizations and deaths (62). In order to unravel the causal effects of NPIs on inequality, we design experiments so that only one NPI is altered at a time. The baseline setting comprises a lockdown starting on the sixth day at the 0.4 level, a daily testing rate of 0.00145 (0.145% of the population), contact tracing with a success rate of 0.7, need-based subsidies of 0.1, and maximum sizes of rich and poor households at 3 and 5, respectively. We are interested in the potentials of testing, contact tracing, and home isolation, so we set aside lockdown and subsidies while varying the testing rate.

For a lockdown level of $0 \leq \psi \leq 1$, each worker stays at home with probability ψ, independently of the others. Fig. 4a shows that the trade-off between COVID-19 deaths and deaths of despair, dependent on the lockdown level, is specific to the poor community, which is consistent with our conclusion from US data (Fig. 2). Tightening lockdown from mild ($\psi = 0$, only retirees staying at home) to moderate ($\psi = 0.4$) significantly reduces COVID-19 deaths for both groups. With further lockdown restrictions, marginal health benefits decline, while more poor people die from despair. By contrast, the rich community has almost no deaths of despair and only benefits from a strict lockdown.

Our model uses reverse transcription polymerase chain reaction (RT-PCR) tests with 90% sensitivity and 100% specificity (63). Given a testing rate, we conduct random testing among susceptible, asymptomatic, and presymptomatic individuals. Once someone tests positive, the person will self-isolate at home until recovery. The person's household members and other contacts will subsequently be prioritized in testing, with the latter being found by contact tracing with a probability of 0.7. Fig. 4b suggests that, even without any other NPI, the combination of testing, contact tracing, and home isolation alone is effective at reducing disparities in both types of death. Our findings corroborate the conclusion in (64) that increased testing and contact tracing capacity enables reopening at a larger scale.

We consider government subsidies that are given to anyone in need regardless of socioeconomic status. On each day of simulation, the model looks for and gives money to low-output people who may die from despair. The subsidy is measured as a fraction of an economically active poor individual's personal output. Fig. 4c indicates that need-based subsidies no less than 0.3 effectively eliminate the gap in deaths of despair. We also explore the efficacy of greedy subsidisation subject to budget constraints. Specifically, given a budget, individuals with the lowest outputs are the ones that are most likely to be impacted by economic volatility and hence prioritized for payment. The budget level is measured as the fraction of the population that can be supported if each subsidy is 0.3. Fig. 4d suggests that increasing the budget level reduces disparities in the total death rate and enables stricter lockdown before economic consequences exceed marginal health benefits.

We investigate the effects of household overcrowding by varying the maximum size of poor households. The configura-

tions of rich households are kept at a maximum size of three and 10% of the population. For ease of mathematical analysis, lockdown starts at initialization, and simulation runs for 60 days. As shown in Fig. 4e, a larger difference in household size leads to higher inequality in COVID-19 deaths. This result confirms the causal link between household overcrowding and the COVID-19 death rate suggested by US data (Fig. 3). Inspired by (16), we quantify the dependence of the infection rate on household size using mean-field approximation. We denote the average size of poor and rich households by n_p and n_r, respectively. Let I_0 be the number of infections at initialization. Let I_t^p be the number of newly infected poor individuals at time step t. We define I_t^r similarly. Let η_t be the estimated susceptible fraction of the population at time step t. Let ϵ be the secondary attack rate. We use Φ to represent the power of secondary infections that originate from economic connections. We can derive mean-field approximation by

$$\mathbb{E}\left[I_t^p\right] \approx n_p I_0 \left[\eta_t \Phi \left(1+\epsilon \eta_t \Phi\right)\right]^t, \quad \mathbb{E}\left[I_t^r\right] \approx \frac{n_r}{n_p}\mathbb{E}\left[I_t^p\right]. \qquad [1]$$

We provide detailed derivation in Methods. It is noteworthy that the ratio between poor and rich communities' infection rates is almost equal to the ratio of average household size. Fig. 4f shows that Eq. 1 approximates simulation results well.

Since no definitive conclusions have been drawn about the possible link between intergenerational coresidence and the fatality rate (29, 33), we test the robustness of our results against transmission within multigenerational households by letting household members be in the same age group. All qualitative observations remain the same (Extended Data Fig. 5).

Discussion

Although medical science has advanced by leaps and bounds since a century ago when the 1918 influenza pandemic claimed tens of millions of lives worldwide, many challenges remain in the face of a pandemic respiratory illness (65, 66). It is crucial that we learn from the past and the present in order to prepare for future pandemics. In this paper, we have focused on modelling and evaluating NPIs during the initial stage of a pandemic, taking into account the specifics of COVID-19. We have investigated the differential causal effects of NPIs on different communities using both US data and agent-based modelling. We have identified a socioeconomic gap in both health and economic measures in most situations. Both our data analysis and our simulations have demonstrated that the widely believed health and economic trade-offs of lockdowns only exist in the socioeconomically disadvantaged population. Moreover, household overcrowding leads to increased rates of infection. We have further shown using mean-field approximation that the ratio between two communities' infection rates is almost equal to the ratio of average household size. Our model has suggested that, even without any other NPI, the combination of testing, contact tracing, and home isolation alone is effective at reducing disparities in COVID-19 and recession-related deaths. Our simulations have also shown the efficacy of targeted subsidies in mitigating the negative economic effects of strict lockdowns, which disproportionately impact disadvantaged groups.

There are several important implications from this work. Our results underline the importance of intervention design in a pandemic as socioeconomically disadvantaged populations bear the brunt of suboptimal policies, which will worsen existing inequalities. Our findings suggest that an effective methodology for confronting COVID-19 is a combination of a moderate lockdown with targeted and sufficient subsidies to mitigate the economic consequences, adequate testing along with contact tracing and home isolation, and easing overcrowding in housing. These measures should be coordinated in order to reduce inequalities under fiscal and logistical constraints. Although we have focused on the US in this paper, our approach and results can be extended to other regions in the world. For example, the deaths of despair phenomenon in the US can be instead considered as mortality associated with food deprivation in low-income countries. Based on the estimates of (67), 22% of the adult population in Ethiopia, Malawi, Nigeria, and Uganda face severe food insecurity during the pandemic, with higher prevalence in poorer households. Understanding the differential impacts of NPIs on various demographic groups continues to be a pressing issue for low-income countries as COVID-19 vaccine shortages are expected to persist in these regions. Another contribution of this study is to identify factors that make a community more vulnerable to COVID-19 and elucidate their effects under various NPIs, which is closely related to the work on defining vulnerability indices. There is a growing number of vulnerability indices that help guide resource allocation during a pandemic, which, however, may give divergent recommendations (23). In order to determine an appropriate index, it is essential to understand how policies impact communities differently.

Our study has several limitations. First, the conclusions drawn from our analysis rely on aggregate data at the ZCTA and county levels. Ideally, comprehensive data at the individual or household level which encompass many aspects such as socioeconomic status, medical conditions, and behaviour in response to COVID-19 are used to infer the differential causal effects of NPIs on different demographic groups. In practice, such granular data rarely exist due to challenges in collection and privacy. The limitations of the data are partially addressed by our work on agent-based modelling. Second, our simulations are based on a stylized model that captures key elements to the topic studied, including socioeconomic status, age-dependent risks, and household transmission, but leaves out other details. We have chosen to build a medium-sized model in order to obtain qualitative insights. Detailed agent-based models that typically require high-performance computing are needed for drawing quantitative conclusions. Finally, we have only considered lockdowns and testing that are conducted uniformly across the population. In reality, low-income areas across the US have faced obstacles to testing and physical distancing (9, 14, 15). For this reason, the socioeconomic gap in COVID-19 deaths identified by our model is a conservative estimate.

There are several interesting directions for future research. One extension is to investigate interventions that are adjusted over time according to feedback and how such adaptive measures affect inequalities. Another interesting avenue of research is exploring how to incentivise safe behaviour that can lessen the need for drastic lockdowns. Given the national variations in the vaccine rollout strategy, it is also urgent to understand how to design vaccine programmes that reduce inequalities. Additionally, it is important to take into consideration fiscal

and logistical constraints for the task of policy evaluation. These questions are not only of much practical relevance to COVID-19 but also fascinating research problems that call for multidisciplinary efforts. Progress towards these goals will have a lasting impact on policy responses to future pandemics.

Methods

Vulnerable Group. We assume that, on average, 1% of the population are at increased risk for severe illness from SARS-CoV-2, all of whom are poor. We define a vulnerability factor $v > 0$ as the extent to which a vulnerable person is more likely to experience worsening symptoms than the average rate. For example, let μ be the hospitalization rate for people in their 50s who are infected and symptomatic. The probability that someone symptomatic in this age group needs to be hospitalized is $(1 + v)\mu$ if the person is vulnerable. For non-vulnerable individuals, the probability is $(1 - v/99)\mu$. In general, a vulnerable person, once infected, is more likely to move through the disease stages of symptoms, hospital admission, ICU admission, and death by a factor of v than the age-specific average rate.

Networks. Let m_p be the maximum number of people living in a poor household. Similarly, we define m_r as the maximum size of a rich household. Unless stated otherwise, we use $m_p = 5$ and $m_r = 3$ in simulations. Let h_p and h_r be the number of poor and rich households, respectively. To construct a household network, we generate h_p complete graphs where the number of nodes in each complete graph is sampled uniformly at random between 1 and m_p. Similarly, we create h_r rich households. We set $h_r = 34000$ and calculate h_p such that poor people constitute 90% of the population on average. We subsequently use the Watts–Strogatz random graph (61) to generate an economic network on the nodes of the household network. Intuitively, the nodes are first arranged into a ring, and then each node is connected with its k nearest neighbours. Finally, each edge in the economic network is rewired with probability p, independently of other edges. Networks constructed as such are known to exhibit the small-world phenomenon (61). Unless stated otherwise, we use $k = 20$ and $p = 0.5$ in simulations.

Individual Output. For simplicity, we assume that all the income inequality in society is explained by differences in individual productivity. Other sources of inequality are not addressed. For a poor individual who is economically active, let x_p be the output per economic connection and y_p be the personal output. Thus, the average output of an active poor individual is $O_p = y_p + kx_p$ at initialization. Similarly, we define x_r, y_r, and O_r for rich individuals. Let λ be the rich-to-poor output ratio where $x_r = \lambda x_p$ and $y_r = \lambda y_p$. To capture the wealth inequality in the US (68), we suppose that rich people account for only $0 < \theta \ll 1$ of the population but 45% of the total output. In other words, $\theta O_r = 0.45[\theta O_r + (1 - \theta)O_p]$. For $\theta = 0.1$, solving the equation gives $\lambda = 81/11$. For an economic connection to be counted in an individual's output, we require both persons to be (i) economically active, (ii) not staying at home because of lockdown, and (iii) not in isolation due to COVID-19 symptoms. Assuming that half of the workers staying at home leads to a drop in the total output by 15%, we can get $y_p = 4kx_p$. Without loss of generality, we let $y_p = 1$ and calculate other variables as discussed.

Deaths of Despair. Taking into consideration deaths caused by financial stress, we suppose that a household with a low per capita output is at increased risk for death of despair. Let O be the per capita output in a household, including subsidies received and excluding members who are hospitalized or deceased. Let $z = O_p - O$ be the difference between the household's per capita output and the average value for poor individuals at initialization. For a despair coefficient $0 < \delta \ll 1$, we define the probability of death of despair by a generalized logistic function $q(z) = \delta[1 + \nu^{z/\omega}]^{-1/\nu}$ where $\omega = kx_p/2$ and $\nu = 0.001$ set the inflection point at $z = kx_p/2$, which equals a poor individual's output loss if economic connections are halved. We use $\delta = 5.5 \times 10^{-5}$ in simulations. Extended Data Fig. 6 plots the probability of death of despair with respect to output loss. On each day, we calculate q for every household. Each member of the household then dies from despair on the day with probability q, independently of each other.

Recession. Let O_0 be the total output at initialization. Let O_t be the total output on day t, taking into account subsidies distributed on the day. We define $0 < \beta \ll 1$ as an inactive coefficient. If $O_t < O_0$, then we assume that a worker becomes economically inactive on day t with probability $\beta(1 - O_t/O_0)$, independently of each other. We use $\beta = 0.01$ in simulations.

Undertreatment. If the number of people hospitalized with COVID-19 exceeds the hospital capacity, then hospitalized patients will be at increased risk for severe illness. Let $0 < \gamma \leq 1$ and $\lambda \geq 0$ be the coefficients of hospital capacity and undertreatment effects, respectively. We denote the population size by N and the current number of COVID-19-associated hospitalizations by H. If $H > \gamma N$, then hospitalized poor patients will be more likely to be admitted to the ICU and possibly die later by a factor of $\lambda[H/(\gamma N) - 1]$ than their age-specific risks. By contrast, we assume that rich patients are not affected by overwhelmed hospitals. We use $\gamma = 0.0025$ and $\lambda = 0.5$ in simulations.

Mean-Field Approximation. We denote the average size of poor and rich households by $n_p = (1 + m_p)/2$ and $n_r = (1 + m_r)/2$, respectively. Since the majority of the population are poor, our approximation first assumes that all households are poor and then considers rich households at the end.

Let d be the average number of days that an infected person is asymptomatic or pre-symptomatic. Since only a small fraction of infections lead to hospitalization and more severe outcomes, we ignore these cases in our approximation. The same as in simulation, we assume that anyone symptomatic quarantines at home. In other words, $d \approx ad_a + (1 - a)d_s$, where a is the asymptomatic infection rate, d_a is the average number of days of illness until recovery for asymptomatic patients, and d_s is the average number of days of illness until symptom onset for symptomatic patients. We define one time step as d days.

Let I_0 be the number of infections at initialization. Let I_t^p be the number of newly infected poor individuals at time step t. We suppose that, each day, an infected person who has no symptoms spreads the virus to any of her connections from a different household with probability $\rho > 0$. Given the high risk of household transmission, we suppose that, once someone is infected, everyone else in the same household is immediately infected. Thus, the effective number of initial infections is $n_p I_0$. We consider that everyone stays at home with probability $1 - \alpha$, independently of any other event. In order for an infected person to infect someone from a different household, both persons need to leave home, which occurs with probability α^2. Under the assumption that each person's economic connections are from different households, every infected individual spreads the disease to $\alpha^2 k\rho d$ economic connections in one time step on average. Infected connections then immediately infect their household members. Moreover, we suppose that a fraction $0 < \epsilon < 1$ of these new infections further spread the disease to their economic connections and hence their household. Thus, I_1^p is roughly equal to $n_p I_0 \Phi(1 + \epsilon\Phi)$ on average where $\Phi = n_p \alpha^2 k\rho d$. Several assumptions such as immediate household transmission and uniqueness of economic connections' households make our estimated number of infections an overestimation, which becomes more marked as time goes on. We adjust for the error by taking into account the susceptible population that shrinks over time. Let $\eta_t = 1 - \sum_{\tau=0}^{t-1} \mathbb{E}\left[I_\tau^p\right]/N$ be the estimated susceptible fraction of the population at time step t. We then have $\mathbb{E}\left[I_1^p\right] \approx n_p I_0 \eta_1 \Phi(1 + \epsilon\eta_1\Phi)$. By induction on time, we have

$$\mathbb{E}\left[I_t^p\right] \approx n_p I_0 \left[\eta_t \Phi \left(1 + \epsilon\eta_t \Phi\right)\right]^t .$$

Let I_t^r be the number of newly infected rich individuals at time step t. The probability that someone from a rich household gets infected can be approximated by n_r/n_p times the infection rate of poor people. Therefore,

$$\mathbb{E}\left[I_t^r\right] \approx \frac{n_r}{n_p} \mathbb{E}\left[I_t^p\right] .$$

Data availability

All data used in this paper are publicly available as cited.

Code availability

Code has been deposited in GitHub, available at https://github.com/xflorameng/covid-19_interventions.

1. KMJ Azar, et al., Disparities in outcomes among COVID-19 patients in a large health care system in California. *Heal. Aff.* **39**, 1–10 (2020).
2. EG Price-Haywood, J Burton, D Fort, L Seoane, Hospitalization and mortality among black patients and white patients with Covid-19. *New Engl. J. Medicine* **382**, 2534–2543 (2020).
3. JAW Gold, et al., Characteristics and clinical outcomes of adult patients hospitalized with COVID-19 — Georgia, March 2020. *Morb. Mortal. Wkly. Rep.* **69**, 545–550 (2020).
4. A van Dorn, RE Cooney, ML Sabin, COVID-19 exacerbating inequalities in the US. *Lancet* **395**, 1243–1244 (2020).
5. R Khazanchi, et al., County-level association of social vulnerability with COVID-19 cases and deaths in the USA. *J. Gen. Intern. Medicine* **35**, 2784–2787 (2020).
6. GJ Borjas, Demographic determinants of testing incidence and Covid-19 infections in New York City neighborhoods. *Covid Econ. Vetted Real-Time Pap.* **3**, 12–39 (2020).
7. N Krieger, PD Waterman, JT Chen, COVID-19 and overall mortality inequities in the surge in death rates by zip code characteristics: Massachusetts, January 1 to May 19, 2020. *Am. J. Public Heal.* **110**, 1850–1852 (2020).
8. S Schmitt-Grohé, K Teoh, M Uribe, Covid-19: Testing inequality in New York City. *Covid Econ. Vetted Real-Time Pap.* **8**, 27–43 (2020).
9. JA Weill, M Stigler, O Deschenes, MR Springborn, Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proc. Natl. Acad. Sci.* **117**, 19658–19660 (2020).
10. R Kabarriti, et al., Association of race and ethnicity with comorbidities and survival among patients with COVID-19 at an urban medical center in New York. *JAMA Netw. Open* **3**, e2019795 (2020).
11. MJ Townsend, TK Kyle, FC Stanford, Outcomes of COVID-19: Disparities in obesity and by ethnicity/race. *Int. J. Obes.* **44**, 1807–1809 (2020).
12. KM Flegal, D Kruszon-Moran, MD Carroll, CD Fryar, CL Ogden, Trends in obesity among adults in the United States, 2005 to 2014. *JAMA* **315**, 2284–2291 (2016).
13. K Arasteh, Prevalence of comorbidities and risks associated with COVID-19 among Black and Hispanic populations in New York City: An examination of the 2018 New York City community health survey. *J. Racial Ethn. Heal. Disparities*, 1–7 (2020).
14. W Lieberman-Cribbin, S Tuminello, RM Flores, E Taioli, Disparities in COVID-19 testing and positivity in New York City. *Am. J. Prev. Medicine* **59**, 326–332 (2020).
15. J Jay, et al., Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. *Nat. Hum. Behav.* **4**, 1294–1302 (2020).
16. SC Fay, DJ Jones, MA Dahleh, AE Hosoi, Simple control for complex pandemics (2020).
17. PJ Arena, M Malta, AW Rimoin, SA Strathdee, Race, COVID-19 and deaths of despair. *EClinicalMedicine* **25**, 100485 (2020).
18. DA Glei, N Goldman, M Weinstein, A growing socioeconomic divide: Effects of the Great Recession on perceived economic distress in the United States. *PLoS One* **14**, e0214947 (2019).
19. A Macintyre, D Ferris, B Gonçalves, N Quinn, What has economics got to do with it? The impact of socioeconomic factors on mental health and the case for collective action. *Palgrave Commun.* **4**, 1–5 (2018).
20. A Case, A Deaton, *Deaths of Despair and the Future of Capitalism*. (Princeton University Press), (2020).
21. T Tanaka, S Okamoto, Increase in suicide following an initial decline during the COVID-19 pandemic in Japan. *Nat. Hum. Behav.* **5**, 229–238 (2021).
22. The Lancet, Redefining vulnerability in the era of COVID-19. *The Lancet* **395**, 1089 (2020).
23. Columbia Center for Spatial Research and Yale Global Health Partnership, Mapping the new politics of care (https://newpoliticsofcare.net/) (2020).
24. I Bavli, B Sutton, S Galea, Harms of public health interventions against covid-19 must not be ignored. *BMJ* **371**, m4074 (2020).
25. TJ VanderWeele, Challenges estimating total lives lost in COVID-19 decisions: Consideration of mortality related to unemployment, social isolation, and depression. *JAMA* **324**, 445–446 (2020).
26. F Bianchi, G Bianchi, D Song, The long-term impact of the COVID-19 unemployment shock on life expectancy and mortality rates, (National Bureau of Economic Research), Working Paper 28304 (2020).
27. N Haug, et al., Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat. Hum. Behav.* **4**, 1303–1312 (2020).
28. S Flaxman, et al., Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 251–261 (2020).
29. B Arpino, V Bordone, M Pasqualini, No clear association emerges between intergenerational relationships and COVID-19 fatality rates from macro-level analyses. *Proc. Natl. Acad. Sci.* **117**, 19116–19121 (2020).
30. A Aleta, et al., Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat. Hum. Behav.* **4**, 964–971 (2020).
31. B Wilder, et al., Modeling between-population variation in COVID-19 dynamics in Hubei, Lombardy, and New York City. *Proc. Natl. Acad. Sci.* **117**, 25904–25910 (2020).
32. SM Moghadas, et al., The implications of silent transmission for the control of COVID-19 outbreaks. *Proc. Natl. Acad. Sci.* **117**, 17513–17515 (2020).
33. A Esteve, In Permanyer, D Boertien, JW Vaupel, National age and coresidence patterns shape COVID-19 vulnerability. *Proc. Natl. Acad. Sci.* **117**, 16118–16120 (2020).
34. JA Firth, et al., Using a real-world network to model localized COVID-19 control strategies. *Nat. Medicine* **26**, 1616–1622 (2020).
35. PCL Silva, et al., COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. *Chaos, Solitons Fractals* **139**, 110088 (2020).
36. C Bishop, *Pattern Recognition and Machine Learning*. (Springer), (2006).
37. SJ Kazemitabar, AA Amini, A Bloniarz, AS Talwalkar, Variable importance using decision trees in *Advances in Neural Information Processing Systems*, eds. I Guyon, et al. (Curran Associates, Inc.), Vol. 30, pp. 425–434 (2017).
38. United States Census Bureau, Selected economic characteristics, 2015–2019 American Community Survey 5-year estimates, Data file (Available at: https://data.census.gov/cedsci/table?q=economiccharacteristics&g=0400000US36.860000&tid=ACSDP5Y2019.DP03&hidePreview=false [Accessed 15 March 2021]) (2020).
39. City University of New York, New York City census data: Neighborhood profiles, Data file (Available at: https://guides.newman.baruch.cuny.edu/nyc_data/nbhoods) (2020).
40. New York University, School of Law, Furman Center for Real Estate and Urban Policy, Eviction filings by ZIP Code, Data file (Available at: https://furmancenter.org/stateofthecity/view/eviction-filings) (2020).
41. New York City Department of Health and Mental Hygiene, COVID-19: Data by ZIP, Data file (Available at: https://www1.nyc.gov/site/doh/covid/covid-19-data.page) (2020).
42. United States Census Bureau, Comparative economic characteristics, 2019 American Community Survey 1-year estimates, Data file (Available at: https://data.census.gov/cedsci/table?q=economic%20characteristics&g=0100000US.050000&tid=ACSCP5Y2019.CP03&hidePreview=false [Accessed 15 March 2021]) (2020).
43. Google LLC, Google COVID-19 community mobility reports, Data file (Available at: https://www.google.com/covid19/mobility/) (2020).
44. Hopkins Population Center, COVID-19 SES data hub, Data file (Available at: https://popcenter.jhu.edu/data-hub/ [Accessed 25 November 2020]) (2020).
45. HS Badr, et al., Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet Infect. Dis.* **20**, 1247–1254 (2020).
46. P Nouvellet, et al., Reduction in mobility and COVID-19 transmission. *Nat. Commun.* **12**, 1090 (2021).
47. EL Glaeser, C Gorback, SJ Redding, How much does COVID-19 increase with mobility? Evidence from New York and four other U.S. cities, (National Bureau of Economic Research), Working Paper 27519 (2020).
48. JA Tapia Granados, JS House, EL Ionides, S Burgard, RS Schoeni, Individual joblessness, contextual unemployment, and mortality risk. *Am. J. Epidemiol.* **180**, 280–287 (2014).
49. DJ Roelfs, E Shor, A Blank, JE Schwartz, Misery loves company? A meta-regression examining aggregate unemployment rates and the unemployment-mortality association. *Annals Epidemiol.* **25**, 312–322 (2015).
50. CJ Ruhm, Recessions, healthy no more? *J. Heal. Econ.* **42**, 17–28 (2015).
51. EC Matthay, KA Duchowny, AR Riley, S Galea, Projected all-cause deaths attributable to COVID-19-related unemployment in the United States. *Am. J. Public Heal.* **111**, 696–699 (2021).
52. United States Census Bureau, 2013–2017 CHAS data, Data file (Available at: https://www.huduser.gov/portal/datasets/cp.html#2006-2017_data [Accessed 7 December 2020]) (2020).
53. United States Census Bureau, Percent urban and rural in 2010 by state and county, Data file (Available at: https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2010-urban-rural.html [Accessed 8 December 2020]) (2012).
54. United States Census Bureau, Annual estimates of the resident population by single year of age and sex for the United States: April 1, 2010 to July 1, 2019 (NC-EST2019-AGESEX-RES), Data file (Available at: https://www2.census.gov/programs-surveys/popest/technical-documentation/file-layouts/2010-2019/nc-est2019-agesex-res.csv [Accessed 4 June 2020]) (2020).
55. HY Cheng, et al., Contact tracing assessment of COVID-19 transmission dynamics in Taiwan and risk at different exposure periods before and after symptom onset. *JAMA Intern. Medicine* **180**, 1156–1163 (2020).
56. The United States Centers for Disease Control and Prevention, the Office of the Assistant Secretary for Preparedness and Response, COVID-19 pandemic planning scenarios, Technical Report May 20 (2020).
57. NM Ferguson, et al., Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, (WHO Collaborating Centre for Infectious Disease Modelling, MRC Centre for Global Infectious Disease Analysis, Abdul Latif Jameel Institute for Disease and Emergency Analytics, Imperial College London), Report 9 (2020).
58. JA Lewnard, et al., Incidence, clinical outcomes, and transmission dynamics of severe coronavirus disease 2019 in California and Washington: prospective cohort study. *BMJ* **369**, m1923 (2020).
59. R Verity, et al., Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infect. Dis.* **20**, 669–677 (2020).
60. ZJ Madewell, Y Yang, IM Longini, ME Halloran, NE Dean, Household transmission of SARS-CoV-2: A systematic review and meta-analysis. *JAMA Netw. Open* **3**, e2031756 (2020).
61. DJ Watts, SH Strogatz, Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
62. D Acemoglu, V Chernozhukov, I Werning, MD Whinston, Optimal targeted lockdowns in a multi-group SIR model, (National Bureau of Economic Research), Working Paper 27102 (2020).
63. PB van Kasteren, et al., Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. *J. Clin. Virol.* **128**, 104412 (2020).
64. WA Chiu, R Fischer, ML Ndeffo-Mbah, State-level needs for social distancing and contact tracing to contain COVID-19 in the United States. *Nat. Hum. Behav.* **4**, 1080–1090 (2020).
65. GA Soper, The lessons of the pandemic. *Science* **49**, 501–506 (1919).
66. A Bradshaw, Lessons from the past. *Nat. Hum. Behav.* **4**, 448 (2020).
67. A Josephson, T Kilic, JD Michler, Socioeconomic impacts of COVID-19 in low-income countries. *Nat. Hum. Behav.* (2021).

68. M Batty, et al., *The Distributional Financial Accounts of the United States*. (University of Chicago Press), (2020).

Acknowledgements

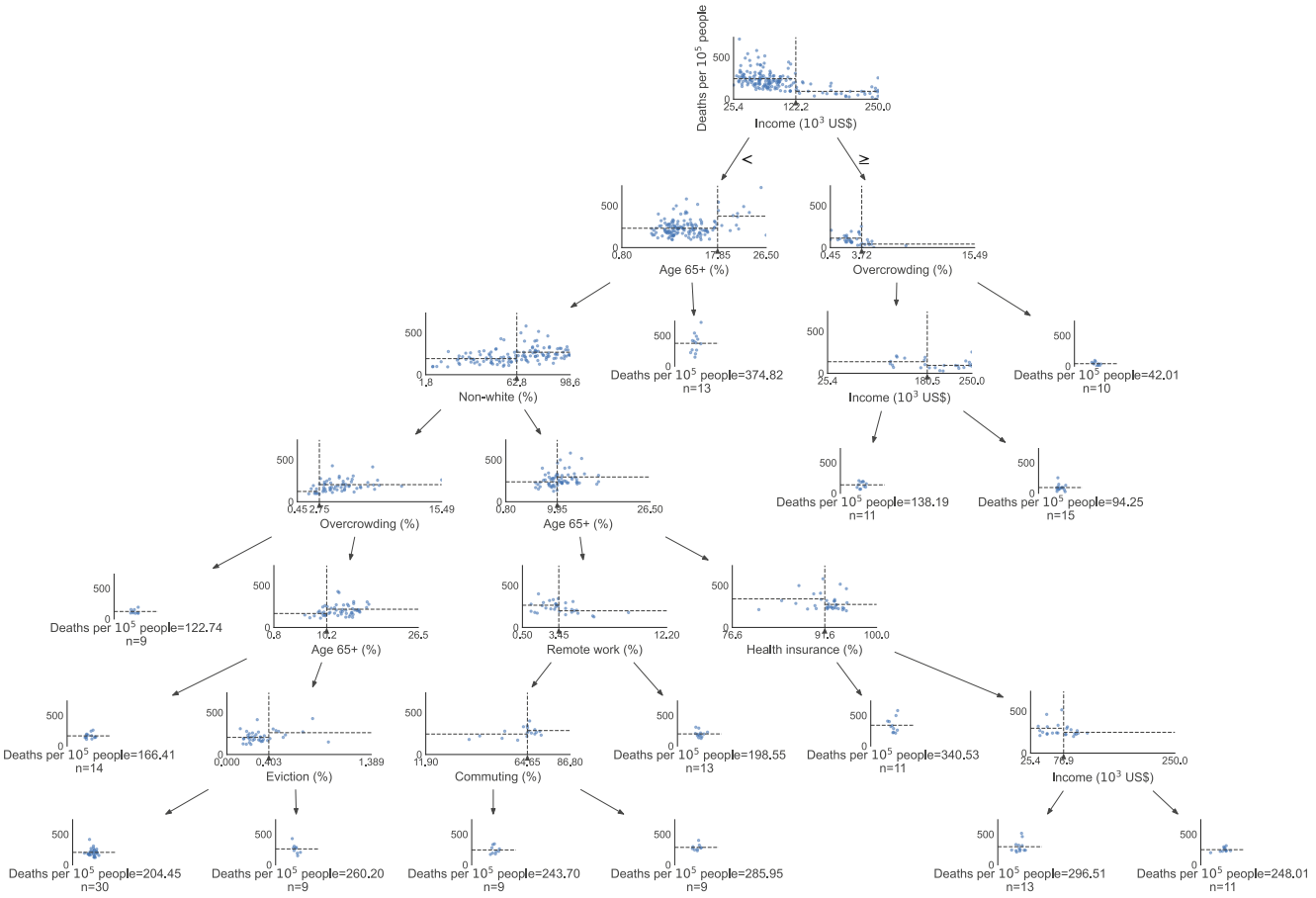
This research was supported by the C3.ai Digital Transformation Institute. The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank Emma Tegling for her help with collecting epidemiological parameters of COVID-19. We also thank Devavrat Shah for granting us access to computing for running extensive simulation. We acknowledge Anette E. (Peko) Hosoi, Devavrat Shah, Anish Agarwal, Yash Deshpande, and other members of the IDSS COVID-19 Collaboration (Isolat) at MIT for helpful comments.

Author contributions

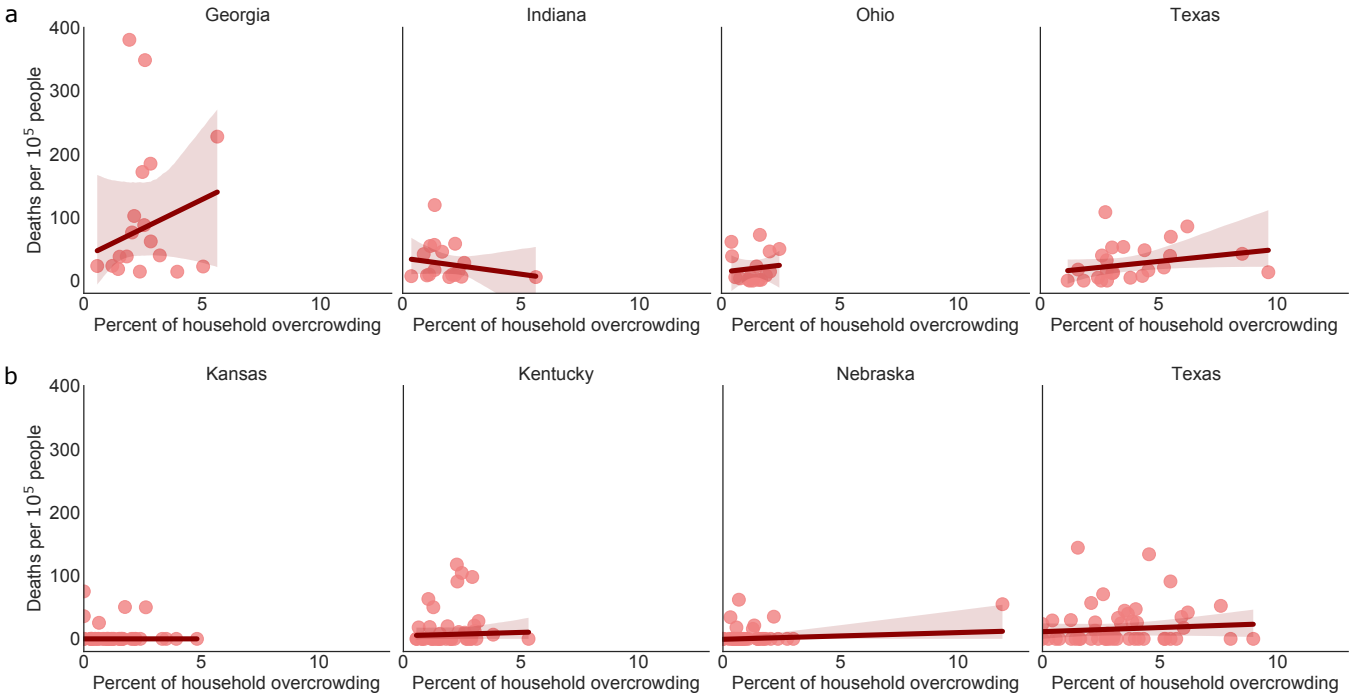
M.A.D., R.R., D.J.J., and X.F.M. designed research; D.J.J. collected and analysed the data for decision tree learning and impacts of lockdown; X.F.M. collected and analysed the data for household overcrowding; D.J.J. and X.F.M. built the agent-based model and wrote the manuscript. All authors revised the manuscript.

Competing interests

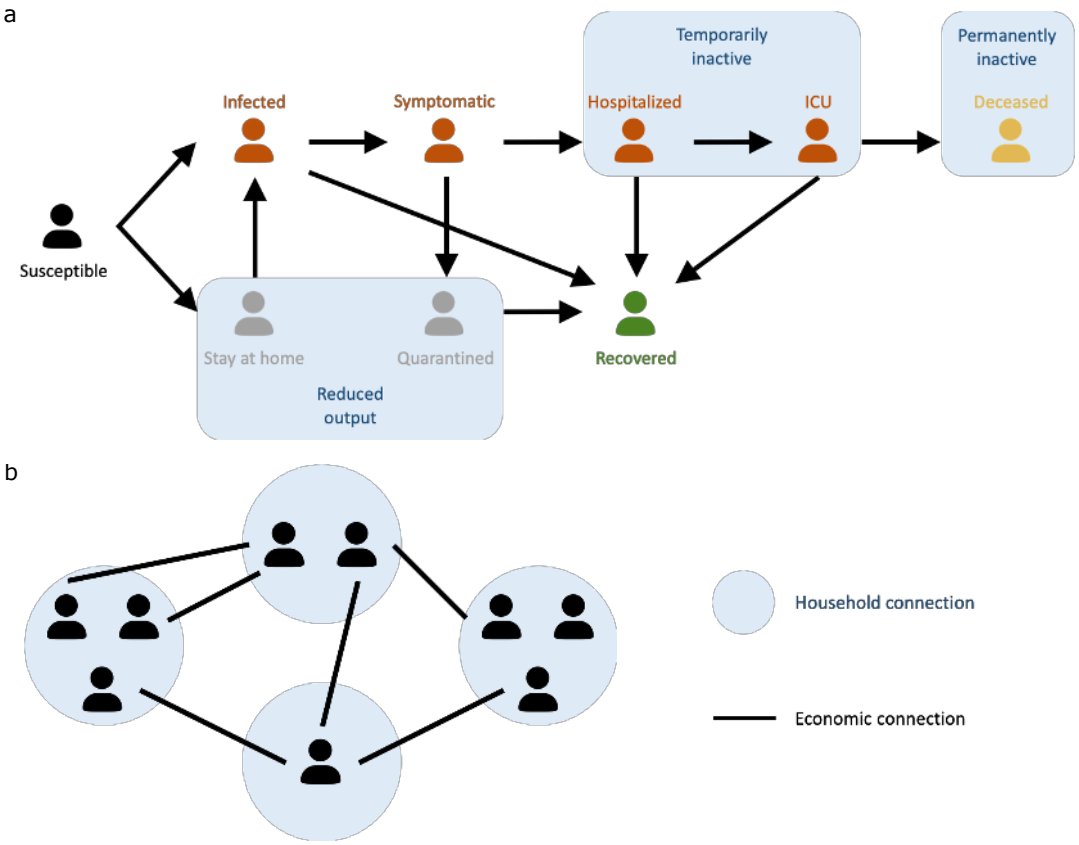
The authors declare no competing interests.



Extended Data Fig. 1 | A decision tree that predicts the COVID-19 death rate of New York City by ZCTA. The x and y-axes of each scatterplot are the feature used for the split and the number of deaths per $100,000$ people, respectively. ZCTAs are divided into two subsets at the vertical lines so that the death rates are close to the average (marked by horizontal lines) within each group.



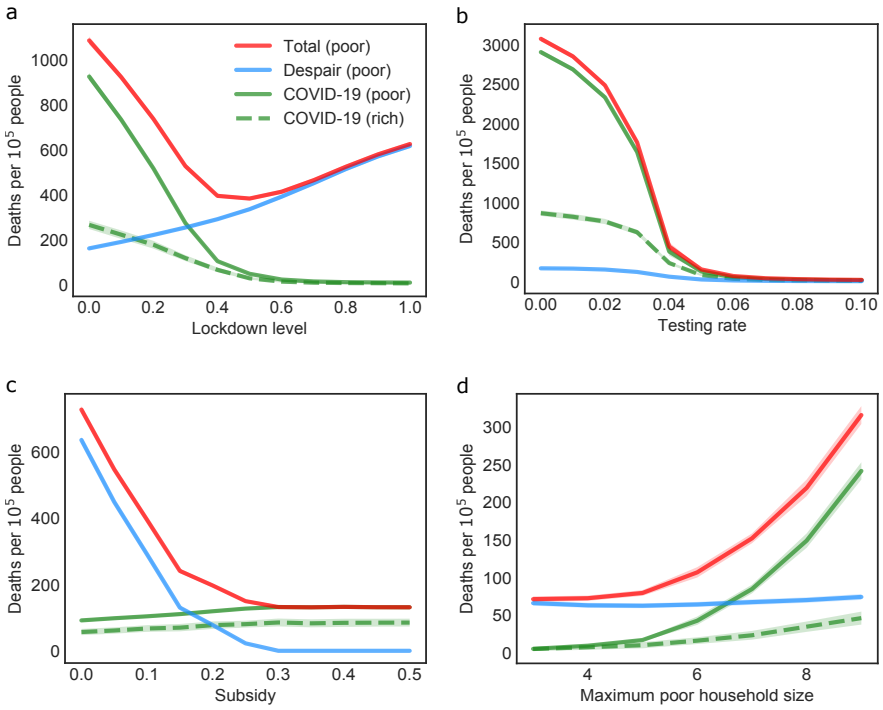
Extended Data Fig. 2 | The relationship between household overcrowding and the COVID-19 death rate are unclear in rural counties. Potential reasons include low population density, large regional variations in infection patterns, and disease outbreaks at different times. **a**, The largest four states for the number of counties of which the percent of the population living in rural areas is between 45% and 55%. **b**, The largest four states for the number of completely rural counties where the whole population live in rural areas. For each state, the solid line and the shaded area represent robust linear regression that downweights outliers with a 95% confidence interval.



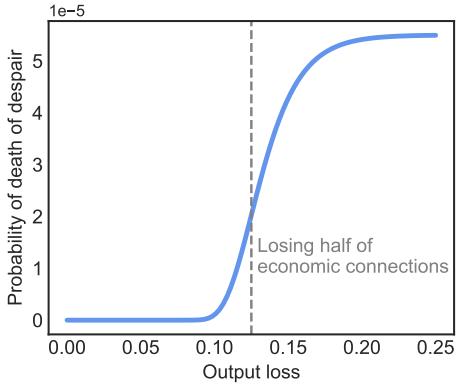
Extended Data Fig. 3 | Schematic diagrams of the agent-based model. a, Once infected, an individual progresses stochastically from asymptomatic or presymptomatic, to symptomatic, hospitalized, admitted to the ICU, and deceased, with the possibility of recovery at any stage if not deceased. While staying at home, a susceptible individual may still be infected by people in the same household. Once symptomatic, the infected individual quarantines at home until recovery unless hospitalization becomes necessary. An individual is economically inactive during hospitalization and at death. Moreover, an individual loses connection output while in quarantine or staying home, **b**, Each blue circle corresponds to a complete graph that represents a household. The economic network is generated using the Watts–Strogatz random graph.

Extended Data Fig. 4 | Epidemiological parameter definitions, baseline values, and sources. Time between different stages of infection is sampled uniformly at random from the corresponding intervals listed.

Definition	Baseline value	Source
Asymptomatic rate	35%	(56)
Probability of hospitalization conditional on symptomatic infection	$\leq$ 9 years: 0.001 10–19 years: 0.003 20–29 years: 0.012 30–39 years: 0.032 40–49 years: 0.049 50–59 years: 0.102 60–69 years: 0.166 70–79 years: 0.243 $\geq$ 80 years: 0.273	(57)
Probability of ICU admission conditional on hospitalization	$\leq$ 39 years: 0.05 40–49 years: 0.063 50–59 years: 0.122 60–69 years: 0.274 70–79 years: 0.432 $\geq$ 80 years: 0.709	(57)
Probability of mortality conditional on ICU admission	$\leq$ 19 years: 0.615 20–39 years: 0.769 40–49 years: 0.748 50–59 years: 0.742 60–69 years: 0.744 70–79 years: 0.747 $\geq$ 80 years: 0.739	(57)
Pre-symptomatic period	2–10 days	(56)
Time from symptom onset to hospitalization	1–12 days	(56)
Time from hospitalization to ICU admission	$\leq$ 14 days	(58)
Time from ICU admission to mortality	$\leq$ 14 days	(58)
Time from symptom onset to recovery	7–28 days	(59)
Probability of infection transmission per contact per day	Household: 0.25 Others: 0.005	(55)



Extended Data Fig. 5 | Robustness tests for impacts of COVID-19 NPIs on socioeconomic inequality. Each household comprises members from the same age group. All qualitative observations remain the same as those with multigenerational households (Fig. 4). The fatality rate is calculated within each socioeconomic group. Since the rate of death of despair is close to zero for the rich community, we only show COVID-19 deaths for this group. **a**, The trade-off between COVID-19 deaths and deaths of despair only exists in the poor community. **b**, The combination of testing and contact tracing alone is sufficient for eliminating socioeconomic disparities in both types of death. **c**, Increasing subsidies effectively reduces the gap in deaths of despair. **d**, Household overcrowding exacerbates COVID-19 in the poor community. Curves and shades are the averages and the standard deviations of 100 trials, respectively.



Extended Data Fig. 6 | Probability of death of despair. The probability that an individual dies from despair increases with per capita output loss in the household.

Figures

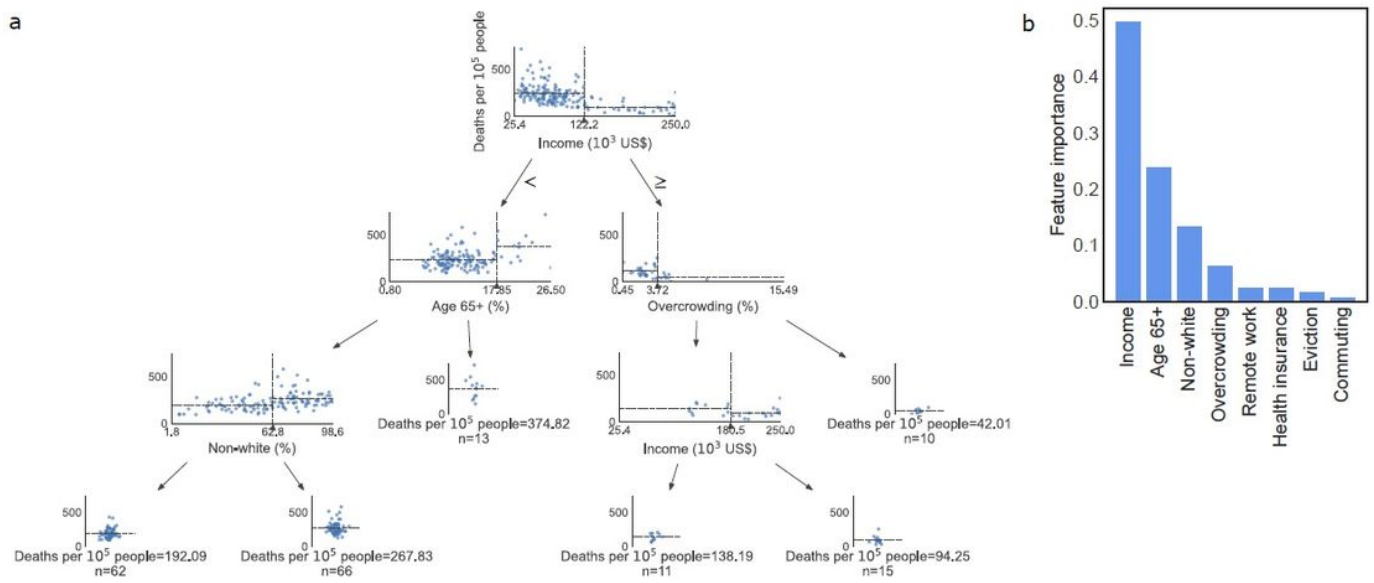


Figure 1

Regional features associated with local COVID-19 death rates. a, We build a decision tree that predicts the COVID-19 death rate of New York City by ZCTA. We show a pruned tree here to illustrate the method and provide the full tree in Extended Data Fig. 1. The x and y-axes of each scatterplot are the feature used for the split and the number of deaths per 100, 000 people, respectively. ZCTAs are divided into two subsets at the vertical lines so that the death rates are close to the average (marked by horizontal lines) within each group. b, We compute the importance of a feature in the decision tree as the normalized total reduction of the mean squared error that is attributable to the feature.

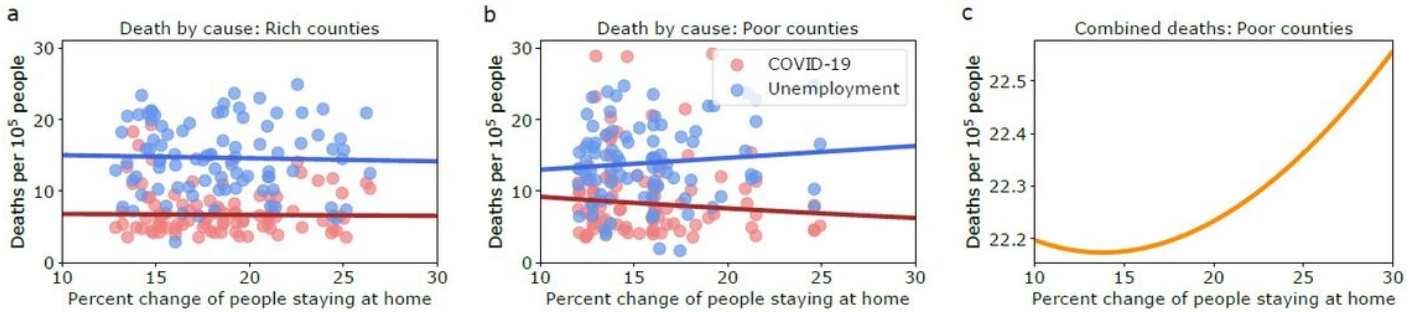


Figure 2

Lockdown and social distancing measures that are meant to curb the spread of COVID-19 can exacerbate inequalities. We compare the richest (a) and poorest (b) counties in the US as measured by median income. a, Affluent counties are resilient to the economic shock of lockdown and social distancing measures. b, In contrast, poor counties face the dilemma of whether to die from COVID-19 infection or

economic distress. c, Combining estimates from both regression reveals the health and economic trade-off for poor counties.

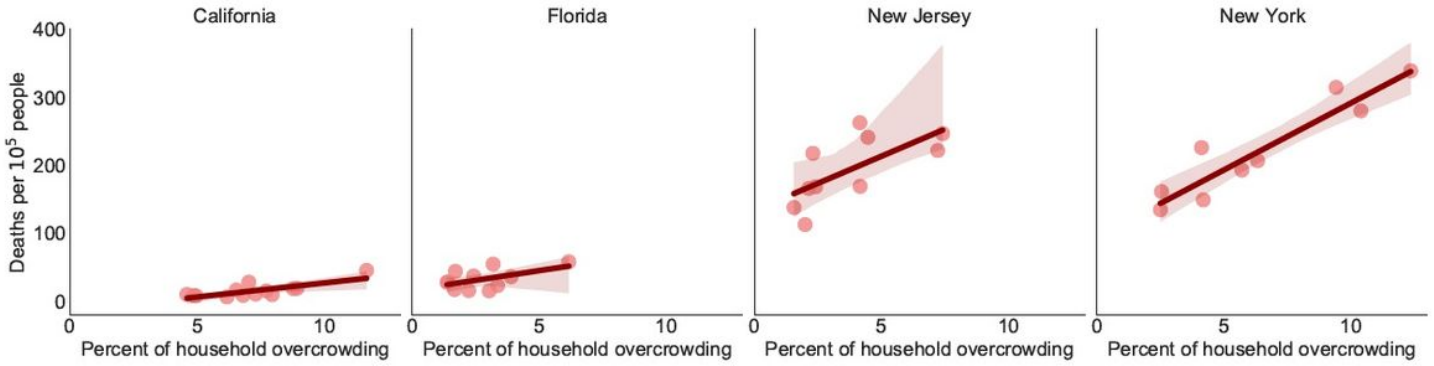


Figure 3

The COVID-19 death rate is positively correlated with household overcrowding in urban counties. California, Florida, New Jersey, and New York are the largest four states for the number of counties of which at least 95% of the population live in urban areas. For each state, the solid line and the shaded area represent robust linear regression that downweights outliers with a 95% confidence interval.

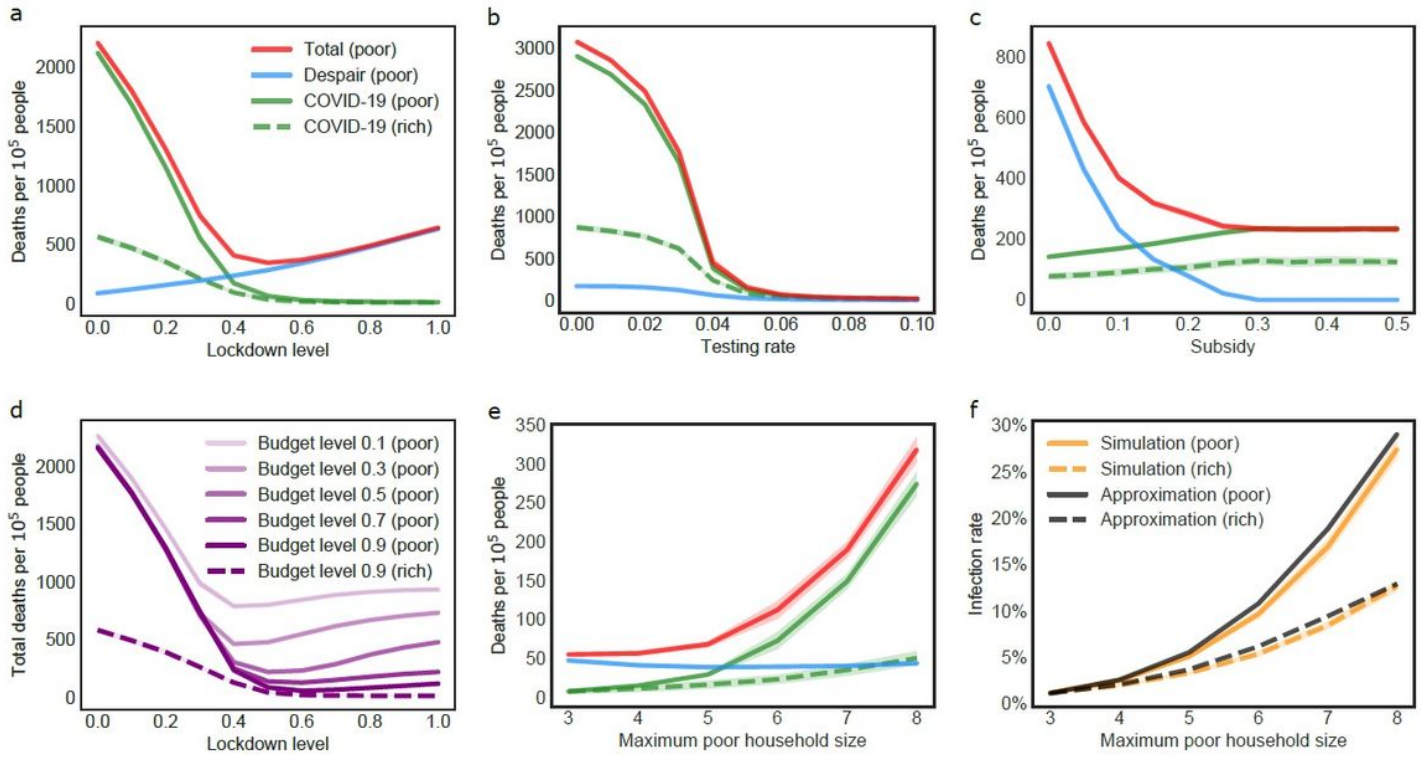


Figure 4

Impacts of COVID-19 NPIs on socioeconomic inequality. The fatality rate is calculated within each socioeconomic group. Since the rate of death of despair is close to zero for the rich community, we only show COVID-19 deaths for this group. a, The trade-off between COVID-19 deaths and deaths of despair

only exists in the poor community. b, The combination of testing and contact tracing alone is sufficient for eliminating socioeconomic disparities in both types of death. c, Increasing subsidies effectively reduces the gap in deaths of despair. d, For the strategy of prioritizing the neediest people for subsidies, a larger budget narrows disparities in the total death rate and enables stricter lockdown before economic consequences exceed marginal health benefits. Since the rate of death of despair is almost the same for the rich community at all budget levels, we only show this group's results at a budget level of 0.9. e, Household overcrowding exacerbates COVID-19 in the poor community. f, The effect of household overcrowding can be explained by mean-field approximation. Curves and shades are the averages and the standard deviations of 100 trials, respectively.