Mobility network models of COVID-19 explain inequities and inform reopening

https://doi.org/10.1038/s41586-020-2923-3

Received: 15 June 2020

Accepted: 21 October 2020

Published online: 10 November 2020

Check for updates

Serina Chang[1,9], Emma Pierson[1,2,9], Pang Wei Koh[1,9], Jaline Gerardin[3], Beth Redbird[4,5], David Grusky[6,7] & Jure Leskovec[1,8]✉

The coronavirus disease 2019 (COVID-19) pandemic markedly changed human mobility patterns, necessitating epidemiological models that can capture the effects of these changes in mobility on the spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)[1]. Here we introduce a metapopulation susceptible–exposed–infectious–removed (SEIR) model that integrates fine-grained, dynamic mobility networks to simulate the spread of SARS-CoV-2 in ten of the largest US metropolitan areas. Our mobility networks are derived from mobile phone data and map the hourly movements of 98 million people from neighbourhoods (or census block groups) to points of interest such as restaurants and religious establishments, connecting 56,945 census block groups to 552,758 points of interest with 5.4 billion hourly edges. We show that by integrating these networks, a relatively simple SEIR model can accurately fit the real case trajectory, despite substantial changes in the behaviour of the population over time. Our model predicts that a small minority of 'superspreader' points of interest account for a large majority of the infections, and that restricting the maximum occupancy at each point of interest is more effective than uniformly reducing mobility. Our model also correctly predicts higher infection rates among disadvantaged racial and socioeconomic groups[2–8] solely as the result of differences in mobility: we find that disadvantaged groups have not been able to reduce their mobility as sharply, and that the points of interest that they visit are more crowded and are therefore associated with higher risk. By capturing who is infected at which locations, our model supports detailed analyses that can inform more-effective and equitable policy responses to COVID-19.

In response to the COVID-19 crisis, stay-at-home orders were enacted in many countries to reduce contact between individuals and slow the spread of the SARS-CoV-2[9]. Since then, public officials have continued to deliberate over when to reopen, which places are safe to return to and how much activity to allow[10]. Answering these questions requires epidemiological models that can capture the effects of changes in mobility on virus spread. In particular, findings of COVID-19 superspreader events[11–14] motivate models that can reflect the heterogeneous risks of visiting different locations, whereas well-reported disparities in infection rates among different racial and socioeconomic groups[2–8] require models that can explain the disproportionate effect of the virus on disadvantaged groups.

To address these needs, we construct fine-grained dynamic mobility networks from mobile-phone geolocation data, and use these networks to model the spread of SARS-CoV-2 within 10 of the largest metropolitan statistical areas (hereafter referred to as metro areas) in the USA. These networks map the hourly movements of 98 million people from census block groups (CBGs), which are geographical units that typically contain 600–3,000 people, to specific points of interest (POIs). As shown in Supplementary Table 1, POIs are non-residential locations that people visit such as restaurants, grocery stores and religious establishments. On top of each network, we overlay a metapopulation SEIR model that tracks the infection trajectories of each CBG as well as the POIs at which these infections are likely to have occurred. This builds on prior research that models disease spread using aggregate[15–19], historical[20–22] or synthetic mobility data[23–25]; separately, other studies have analysed mobility data in the context of COVID-19, but without an underlying model of disease spread[26–30].

Combining our epidemiological model with these mobility networks allows us to not only accurately fit observed case counts, but also to conduct detailed analyses that can inform more-effective and equitable policy responses to COVID-19. By capturing information about individual POIs (for example, the hourly number of visitors and median visit duration), our model can estimate the effects of specific reopening strategies, such as only reopening certain POI categories or restricting the maximum occupancy at each POI. By modelling movement from CBGs, our model can identify at-risk populations and correctly predict, solely from mobility patterns, that disadvantaged racial and

[1]Department of Computer Science, Stanford University, Stanford, CA, USA. [2]Microsoft Research, Cambridge, MA, USA. [3]Department of Preventive Medicine, Northwestern University, Chicago, IL, USA. [4]Department of Sociology, Northwestern University, Evanston, IL, USA. [5]Institute for Policy Research, Northwestern University, Evanston, IL, USA. [6]Department of Sociology, Stanford University, Stanford, CA, USA. [7]Center on Poverty and Inequality, Stanford University, Stanford, CA, USA. [8]Chan Zuckerberg Biohub, San Francisco, CA, USA. [9]These authors contributed equally: Serina Chang, Emma Pierson, Pang Wei Koh. ✉e-mail: jure@cs.stanford.edu

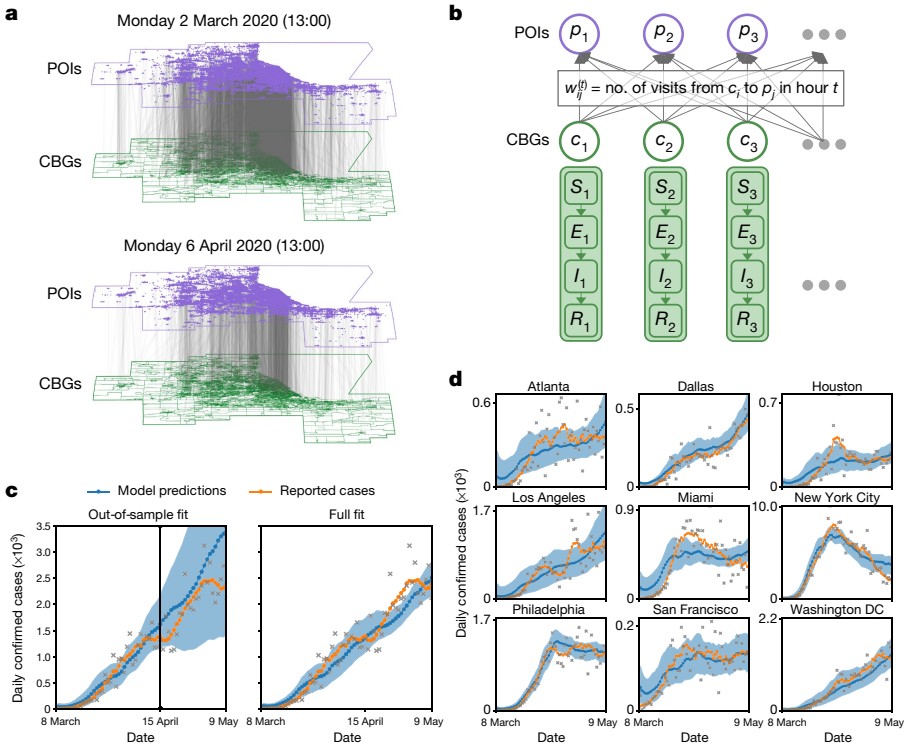


Fig. 1 | Model description and fit. a, The mobility network captures hourly visits from each CBG to each POI. The vertical lines indicate that most visits are between nearby POIs and CBGs. Visits dropped markedly from March to April, as indicated by the lower density of grey lines. Mobility networks in the Chicago metro area are shown for 13:00 on two Mondays, 2 March 2020 (top) and 6 April 2020 (bottom). **b**, We overlaid a disease-spread model on the mobility network, with each CBG having its own set of SEIR compartments. New infections occur at both POIs and CBGs, with the mobility network governing how subpopulations from different CBGs interact as they visit POIs. **c**, Left, to test the out-of-sample prediction, we calibrated the model on data before 15 April 2020 (vertical black line). Even though its parameters remain fixed over time, the model accurately predicts the case trajectory in the Chicago metro area after 15 April using the mobility data (r.m.s.e. on daily cases = 406 for dates ranging from 15 April to 9 May). Right, model fit was further improved when we calibrated the model on the full range of data (r.m.s.e. on daily cases = 387 for the dates ranging from 15 April to 9 May). **d**, We fitted separate models to 10 of the largest US metro areas, modelling a total population of 98 million people; here, we show full model fits, as in **c** (right). In **c** and **d**, the blue line represents the model predictions and the grey crosses represent the number of daily reported cases; as the numbers of reported cases tend to have great variability, we also show the smoothed weekly average (orange line). Shaded regions denote the 2.5th and 97.5th percentiles across parameter sets and stochastic realizations. Across metro areas, we sample 97 parameter sets, with 30 stochastic realizations each (n = 2,910); see Supplementary Table 6 for the number of sets per metro area.

socioeconomic groups face higher rates of infection. Our model thus enables the analysis of urgent health disparities; we use it to highlight two mobility-related mechanisms that drive these disparities and to evaluate the disparate effect of reopening on disadvantaged groups.

Mobility network model

We use data from SafeGraph, a company that aggregates anonymized location data from mobile applications, to study mobility patterns from 1 March to 2 May 2020. For each metro area, we represent the movement of individuals between CBGs and POIs as a bipartite network with time-varying edges, in which the weight of an edge between a CBG and POI represents the number of visitors from that CBG to that POI during a given hour (Fig. 1a). SafeGraph also provides the area in square feet of each POI, as well as its category in the North American industry classification system (for example, fitness centre or full-service restaurant) and median visit duration in minutes. We validated the SafeGraph mobility data by comparing the dataset to Google mobility data (Supplementary Fig. 1 and Supplementary Tables 2, 3) and used iterative proportional fitting[31] to derive POI–CBG networks from the raw SafeGraph data. Overall, these networks comprise 5.4 billion hourly edges between 56,945 CBGs and 552,758 POIs (Extended Data Table 1).

We overlay a SEIR model on each mobility network[15,20], in which each CBG maintains its own susceptible (S), exposed (E), infectious (I) and removed (R) states (Fig. 1b). New infections occur at both POIs and CBGs, with the mobility network governing how subpopulations from different CBGs interact as they visit POIs. We use the area, median visit duration and time-varying density of infectious individuals for each POI to determine the hourly infection rate of that POI. The model has only three free parameters that scale: (1) transmission rates at POIs, (2) transmission rates at CBGs and (3) the initial proportion of exposed individuals (Extended Data Table 2); all three parameters remain constant over time. We calibrate a separate model for each metro area using the confirmed case counts from *The New York Times* by minimizing the root mean square error (r.m.s.e.) to daily incident cases[32]. Our model accurately fits observed daily case counts in all 10 metro areas from 8 March to 9 May 2020 (Fig. 1c, d). In addition, when calibrated on only the case counts up to 14 April, the model predicts case counts reasonably well on the held-out time period of 15 April–9 May 2020 (Fig. 1c and Extended Data Fig. 1a). Our key technical finding is that the dynamic mobility network allows even our relatively simple SEIR model with just three static parameters to accurately fit observed cases, despite changing policies and behaviours during that period.

Mobility reduction and reopening plans

We can estimate the impact of mobility-related policies by constructing a hypothetical mobility network that reflects the expected effects of

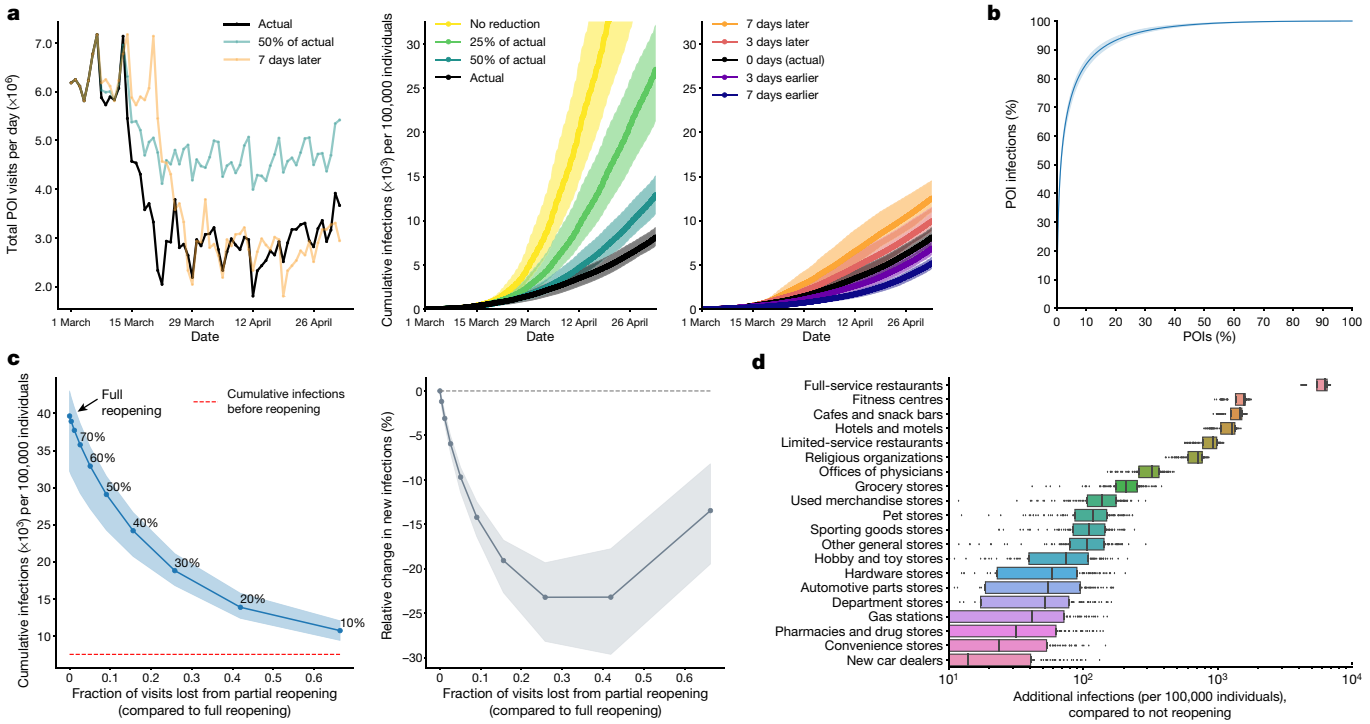


Fig. 2 | Assessing mobility reduction and reopening. The Chicago metro area is used as an example; results for all metro areas are included in Extended Data Figs. 3, 4, Supplementary Figs. 10, 15–24 and Supplementary Tables 4, 5, as indicated. **a**, Counterfactual simulations (left) of past reductions in mobility illustrate that the magnitude of the reduction (middle) was at least as important as its timing (right) (Supplementary Tables 4, 5). **b**, The model predicts that most infections at POIs occur at a small fraction of superspreader POIs (Supplementary Fig. 10). **c**, Left, the cumulative number of predicted infections after one month of reopening is plotted against the fraction of visits lost by partial instead of full reopening (Extended Data Fig. 3); the annotations within the plot show the fraction of maximum occupancy that is used as the cap and the horizontal red line indicates the cumulative number of predicted infections at the point of reopening (on 1 May 2020). Compared to full reopening, capping at 20% of the maximum occupancy in Chicago reduces the number of new infections by more than 80%, while only losing 42% of overall visits. Right, compared to uniformly reducing visits, the reduced maximum occupancy strategy always results in a smaller predicted increase in infections for the same number of visits (Extended Data Fig. 4). The horizontal grey line at 0% indicates when the two strategies result in an equal number of infections, and we observe that the curve falls well below this baseline. The *y* axis plots the relative difference between the predicted number of new infections under the reduced occupancy strategy compared to a uniform reduction. **d**, Reopening full-service restaurants has the largest predicted impact on infections, due to the large number of restaurants as well as their high visit densities and long dwell times (Supplementary Figs. 15–24). Colours are used to distinguish the different POI categories, but do not have any additional meaning. All results in this figure are aggregated across 4 parameter sets and 30 stochastic realizations ($n = 120$). Shaded regions in **a**–**c** denote the 2.5th to 97.5th percentiles; boxes in **d** denote the interquartile range and data points outside this range are shown as individual dots.

each policy, and running our SEIR model forward with this hypothetical network. Using this approach, we assess a wide range of mobility reduction and reopening strategies.

The magnitude of mobility reduction is at least as important as its timing

Mobility in the USA dropped sharply in March 2020: for example, overall POI visits in the Chicago metro area fell by 54.7% between the first week of March and the first week of April 2020. We constructed counterfactual mobility networks by scaling the magnitude of mobility reduction down and by shifting the timeline earlier and later, and applied our model to the counterfactual networks to simulate the resulting infection trajectories. Across metro areas, we found that the magnitude of mobility reduction was at least as important as its timing (Fig. 2a and Supplementary Tables 4, 5): for example, if the mobility reduction in the Chicago metro area had been only a quarter of the size, the predicted number of infections would have increased by 3.3× (95% confidence interval, 2.8–3.8×), compared with a 1.5× (95% confidence interval, 1.4–1.6×) increase had people begun reducing their mobility one full week later. Furthermore, if no mobility reduction had occurred at all, the predicted number of infections in the Chicago metro area would have increased by 6.2× (95% confidence interval, 5.2–7.1×). Our results are in accordance with previous findings that mobility reductions can markedly reduce infections[18,19,33,34].

A minority of POIs account for the majority of the predicted infections

We next investigated whether it matters how we reduce mobility—that is, to which POIs. We computed the number of infections that occurred at each POI in our simulations from 1 March to 2 May 2020, and found that the majority of the predicted infections occurred at a small fraction of superspreader POIs; for example, in the Chicago metro area, 10% of POIs accounted for 85% (95% confidence interval, 83–87%) of the predicted infections at the POIs (Fig. 2b and Supplementary Fig. 10). Certain categories of POIs also contributed far more to infections (for example, full-service restaurants and hotels), although our model predicted time-dependent variation in how much each category contributed (Extended Data Fig. 2). For example, restaurants and fitness centres contributed less to the predicted number of infections over time, probably because of lockdown orders to close these POIs, whereas grocery stores remained steady or even grew in their contribution, which is in agreement with their status as essential businesses.

Reopening with a reduced maximum occupancy

If a minority of POIs produce the majority of infections, then reopening strategies that specifically target high-risk POIs should be especially effective. To test one such strategy, we simulated reopening on 1 May, and modelled the effects of reducing the maximum occupancy in which

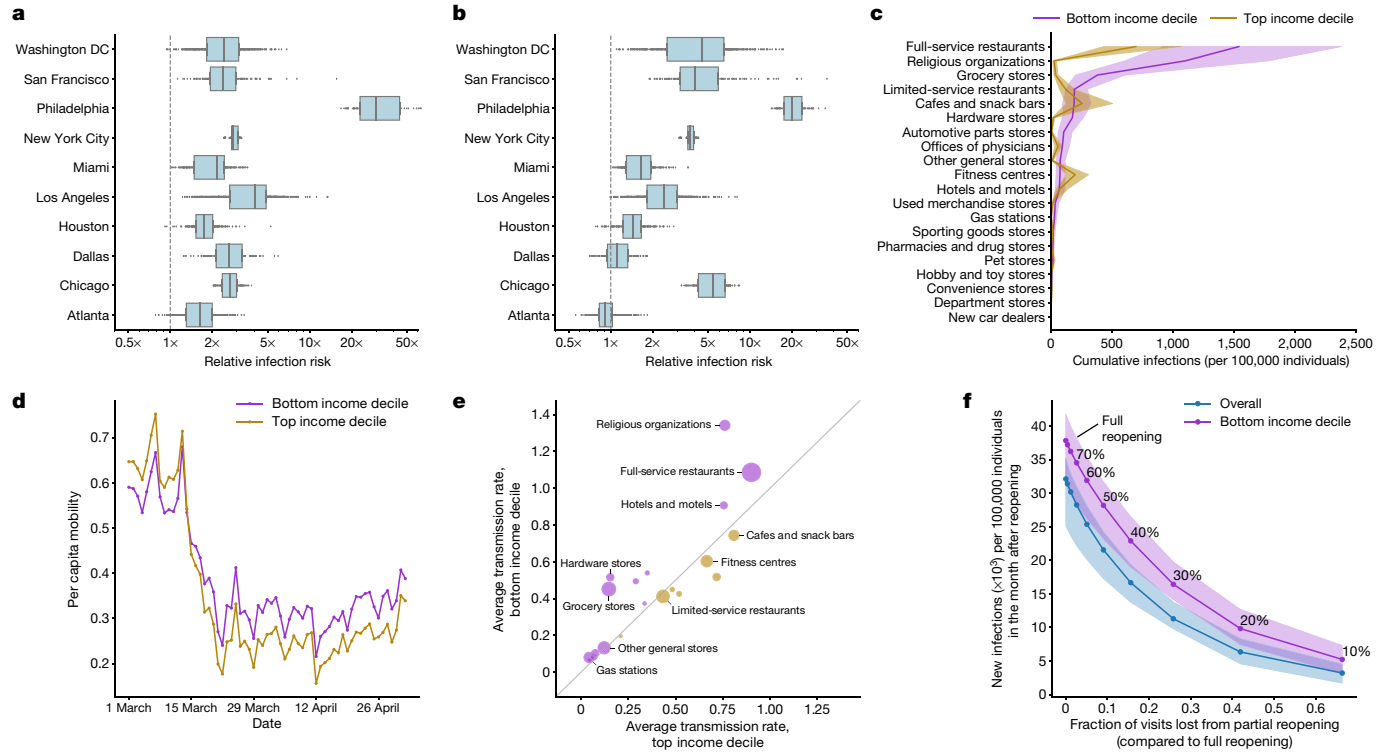


Fig. 3 | Mobility patterns give rise to infection disparities. a, In every metro area, our model predicts that people in lower-income CBGs are likelier to be infected. **b**, People in non-white CBGs area are also likelier to be infected, although results are more variable across metro areas. For **c**–**f**, the Chicago metro area is used as an example, but references to results for all metro areas are provided for each panel. **c**, The overall predicted disparity is driven by a few POI categories such as full-service restaurants (Supplementary Fig. 2). **d**, One reason for the predicted disparities is that higher-income CBGs were able to reduce their mobility levels below those of lower-income CBGs (Extended Data Fig. 6). **e**, Within each POI category, people from lower-income CBGs tend to visit POIs that have higher predicted transmission rates (Extended Data Table 3). The size of each dot represents the average number of visits per capita made to the category. The top 10 out of 20 categories with the most visits are labelled, covering 0.48–2.88 visits per capita (hardware stores–full-service restaurants). **f**, Reopening (at different levels of reduced maximum occupancy) leads to more predicted infections in lower-income CBGs than in the overall population (Extended Data Fig. 3). In **c**–**f**, purple denotes lower-income CBGs, yellow denotes higher-income CBGs and blue represents the overall population. Aside from **d** and **e**, which were directly extracted from mobility data, all results in this figure represent predictions aggregated over model realizations. Across metro areas, we sample 97 parameter sets, with 30 stochastic realizations each (n = 2,910); see Supplementary Table 6 for the number of sets per metro area. Shaded regions in **c** and **f** denote the 2.5th–97.5th percentiles; boxes in (**a**, **b**) denote the interquartile range; data points outside the range are shown as individual dots.

the numbers of hourly visits to each POI returned to their 'normal' levels from the first week of March but were capped if they exceeded a fraction of the maximum occupancy of that POI[35]. Full reopening without reducing the maximum occupancy produced a spike in the predicted number of infections: in the Chicago metro area, our models projected that an additional 32% (95% confidence interval, 25–35%) of the population would be infected by the end of May (Fig. 2c). However, reducing the maximum occupancy substantially reduced the risk without sharply reducing overall mobility: capping at 20% of the maximum occupancy in the Chicago metro area reduced the predicted number of new infections by more than 80% but only lost 42% of overall visits, and we observed similar trends across other metro areas (Extended Data Fig. 3). This result highlights the nonlinearity of the predicted number of infections as a function of the number of visits: one can achieve a disproportionately large reduction in infections with a small reduction in visits. Furthermore, in comparison to a different reopening strategy, in which the number of visits to each POI was uniformly reduced from their levels in early March, reducing the maximum occupancy always resulted in fewer predicted infections for the same number of total visits (Fig. 2c and Extended Data Fig. 4). This is because reducing the maximum occupancies takes advantage of the time-varying visit density within each POI, disproportionately reducing visits to the POI during the high-density periods with the highest risk, but leaving visit counts unchanged during periods with lower risks. These results support previous findings that precise interventions, such as reducing the maximum occupancy, may be more effective than less targeted measures, while incurring substantially lower economic costs[36].

Relative risk of reopening different categories of POIs

Because we found that certain POI categories contributed far more to predicted infections in March (Extended Data Fig. 2), we also expected that reopening some POI categories would be riskier than reopening others. To assess this, we simulated reopening each category in turn on 1 May 2020 (by returning its mobility patterns to early March levels, as above), while keeping all other POIs at their reduced mobility levels from the end of April. We found large variation in predicted reopening risks: on average across metro areas, full-service restaurants, gyms, hotels, cafes, religious organizations and limited-service restaurants produced the largest predicted increases in infections when reopened (Extended Data Fig. 5d). Reopening full-service restaurants was associated with a particularly high risk: in the Chicago metro area, we predicted an additional 595,805 (95% confidence interval, 433,735–685,959) infections by the end of May, more than triple that of the POI category with the next highest risk (Fig. 2d). These risks are summed over all POIs in the category, but the relative risks after normalizing by the number of POIs were broadly similar (Extended Data Fig. 5c). These categories were predicted to be have a higher risk because, in the mobility data, their POIs tended to have higher visit densities and/or visitors stayed there longer (Supplementary Figs. 15–24).

Demographic disparities in infections

We characterize the differential spread of SARS-CoV-2 along demographic lines by using US census data to annotate each CBG with its racial composition and median income, then tracking predicted infection rates in CBGs with different demographic compositions: for example, within each metro area, comparing CBGs in the top and bottom deciles for income. We use this approach to study the mobility mechanisms behind disparities and to quantify how different reopening strategies affect disadvantaged groups.

Predicting disparities from mobility data

Despite having access to only mobility data and no demographic information, our models correctly predicted higher risks of infection among disadvantaged racial and socioeconomic groups[2–8]. Across all metro areas, individuals from CBGs in the bottom decile for income had a substantially higher likelihood of being infected by the end of the simulation, even though all individuals began with equal likelihoods of infection (Fig. 3a). This predicted disparity was driven primarily by a few POI categories (for example, full-service restaurants); far greater proportions of individuals from lower-income CBGs than higher-income CBGs became infected in these POIs (Fig. 3c and Supplementary Fig. 2). We similarly found that CBGs with fewer white residents had higher predicted risks of infection, although results were more variable across metro areas (Fig. 3b). In the Supplementary Discussion, we confirm that the magnitude of the disparities that our model predicts is generally consistent with real-world disparities and further explore the large predicted disparities in Philadelphia, that stem from substantial differences in the POIs that are frequented by higher- versus lower-income CBGs. In the analysis below, we discuss two mechanisms that lead higher predicted infection rates among lower-income CBGs, and we show in Extended Data Fig. 6 and Extended Data Table 4 that similar results hold for racial disparities as well.

Lower-income CBGs saw smaller reductions in mobility

A first mechanism producing disparities was that, across all metro areas, lower-income CBGs did not reduce their mobility as sharply in the first few weeks of March 2020, and these groups showed higher mobility than higher-income CBGs for most of March–May (Fig. 3d and Extended Data Fig. 6). For example, in April, individuals from lower-income CBGs in the Chicago metro area had 27% more POI visits per capita than those from higher-income CBGs. Category-level differences in visit patterns partially explained the infection disparities within each category: for example, individuals from lower-income CBGs made substantially more visits per capita to grocery stores than did those from higher-income CBGs (Supplementary Fig. 3) and consequently experienced more predicted infections for that category (Supplementary Fig. 2).

POIs visited by lower-income CBGs have higher transmission rates

Differences in visits per capita do not fully explain the infection disparities: for example, cafes and snack bars were visited more frequently by higher-income CBGs in every metro area (Supplementary Fig. 3), but our model predicted that a larger proportion of individuals from lower-income CBGs were infected at cafes and snack bars in the majority of metro areas (Supplementary Fig. 2). We found that even within a POI category, the predicted transmission rates at POIs frequented by individuals fom lower-income CBGs tended to be higher than the corresponding rates for those from higher-income CBGs (Fig. 3e and Extended Data Table 3), because POIs frequented by individuals from lower-income CBGs tended to be smaller and more crowded in the mobility data. As a case study, we examined grocery stores in further detail. In eight of the ten metro areas, visitors from lower-income CBGs encountered higher predicted transmission rates at grocery stores than visitors from higher-income CBGs (median transmission rate ratio of 2.19) (Extended Data Table 3). We investigated why one visit to the grocery store was predicted to be twice as dangerous for an individual from a lower-income CBG: the mobility data showed that the average grocery store visited by individuals from lower-income CBGs had 59% more hourly visitors per square foot, and their visitors stayed 17% longer on average (medians across metro areas). These findings highlight how fine-grained differences in mobility patterns—how often people go out and which POIs that they go to—can ultimately contribute to marked disparities in predicted infection outcomes.

Reopening plans must account for disparate effects

Because disadvantaged groups suffer a larger burden of infection, it is critical to not only consider the overall impact of reopening plans but also their disparate effects on disadvantaged groups specifically. For example, our model predicted that full reopening in the Chicago metro area would result in an additional 39% (95% confidence interval, 31–42%) of the population of CBGs in the bottom income decile being infected within a month, compared to 32% (95% confidence interval, 25–35%) of the overall population (Fig. 3f; results for all metro areas are shown in Extended Data Fig. 3). Similarly, Supplementary Fig. 4 illustrates that reopening individual POI categories tends to have a larger predicted effect on lower-income CBGs. More stringent reopening plans produce smaller absolute disparities in predicted infections—for example, we predict that reopening at 20% of the maximum occupancy in Chicago would result in additional infections for 6% (95% confidence interval, 4–8%) of the overall population and 10% (95% confidence interval, 7–13%) of the population in CBGs in the bottom income decile (Fig. 3f)—although the relative disparity remains.

Discussion

The mobility dataset that we use has limitations: it does not cover all populations, does not contain all POIs and cannot capture sub-CBG heterogeneity. Our model itself is also parsimonious, and does not include all real-world features that are relevant to disease transmission. We discuss these limitations in more detail in the Supplementary Discussion. However, the predictive accuracy of our model suggests that it broadly captures the relationship between mobility and transmission, and we thus expect our broad conclusions—for example, that people from lower-income CBGs have higher infection rates in part because they tend to visit denser POIs and because they have not reduced mobility by as much (probably because they cannot work from home as easily[4])—to hold robustly. Our fine-grained network modelling approach naturally extends to other mobility datasets and models that capture more aspects of real-world transmission, and these represent interesting directions for future work.

Our results can guide policy-makers that seek to assess competing approaches to reopening. Despite growing concern about racial and socioeconomic disparities in infections and deaths, it has been difficult for policy-makers to act on those concerns; they are currently operating without much evidence on the disparate effects of reopening policies, prompting calls for research that both identifies the causes of observed disparities and suggests policy approaches to mitigate them[5,8,37,38]. Our fine-grained mobility modelling addresses both these needs. Our results suggest that infection disparities are not the unavoidable consequence of factors that are difficult to address in the short term, such as differences in preexisting conditions; on the contrary, short-term policy decisions can substantially affect infection outcomes by altering the overall amount of mobility allowed and the types of POIs reopened. Considering the disparate effects of reopening plans may lead policy-makers to adopt policies that can drive down infection densities in disadvantaged neighbourhoods by supporting, for example, more stringent caps on POI occupancies, emergency food distribution centres to reduce densities in high-risk stores, free and widely available testing in neighbourhoods predicted to be high risk

(especially given known disparities in access to tests[2]), improved paid leave policy or income support that enables essential workers to curtail mobility when sick, and improved workplace infection prevention for essential workers, such as high-quality personal protective equipment, good ventilation and physical distancing when possible. As reopening policies continue to be debated, it is critical to build tools that can assess the effectiveness and equity of different approaches. We hope that our model, by capturing heterogeneity across POIs, demographic groups and cities, helps to address this need.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2923-3.

1. Buckee, C. O. et al. Aggregated mobility data could help fight COVID-19. *Science* **368**, 145–146 (2020).
2. Wilson, C. These graphs show how COVID-19 is ravaging New York City's low-income neighborhoods. *Time* (15 April 2020).
3. Garg, S. et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019 — COVID-NET, 14 states, March 1—30, 2020. *MMWR Morb. Mortal. Wkly Rep.* **69**, 458–464 (2020).
4. Reeves, R. V. & Rothwell, J. *Class and COVID: How the Less Affluent face Double Risks*. https://www.brookings.edu/blog/up-front/2020/03/27/class-and-covid-how-the-less-affluent-face-double-risks/ (The Brookings Institution, 2020).
5. Pareek, M. et al. Ethnicity and COVID-19: an urgent public health research priority. *Lancet* **395**, 1421–1422 (2020).
6. Dorn, A. V., Cooney, R. E. & Sabin, M. L. COVID-19 exacerbating inequalities in the US. *Lancet* **395**, 1243–1244 (2020).
7. Yancy, C. W. COVID-19 and African Americans. *J. Am. Med. Assoc.* **323**, 1891–1892 (2020).
8. Chowkwanyun, M. & Reed, A. L. Jr. Racial Health Disparities and Covid-19 — caution and context. *N. Engl. J. Med.* **383**, 201–203 (2020).
9. Flaxman, S. et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
10. Rojas, R. & Delkic, M. As states reopen, governors balance existing risks with new ones. *The New York Times* (17 May 2020).
11. Endo, A., Abbott, S., Kucharski, A. J. & Funk, S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res.* **5**, 67 (2020).
12. Adam, D. C. et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat. Med.* https://doi.org/10.1038/s41591-020-1092-0 (2020).
13. Park, S. Y. et al. Coronavirus disease outbreak in call center, South Korea. *Emerg. Infect. Dis.* **26**, 1666–1670 (2020).
14. Bi, Q. et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919 (2020).
15. Chinazzi, M. et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
16. Jia, J. S. et al. Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* **582**, 389–394 (2020).
17. Pei, S., Kandula, S. & Shaman, J. Differential effects of intervention timing on COVID-19 spread in the United States. Preprint at https://doi.org/10.1101/2020.05.15.20103655 (2020).
18. Lai, S. et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).
19. Badr, H. S. et al. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect. Dis.* **20**, 1247–1254 (2020).
20. Li, R. et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
21. Pei, S. & Shaman, J. Initial simulation of SARS-CoV2 spread and intervention effects in the continental US. Preprint at https://doi.org/10.1101/2020.03.21.20040303 (2020).
22. Aleta, A. 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).
23. Duque, D. et al. Timing social distancing to avert unmanageable COVID-19 hospital surges. *Proc. Natl Acad. Sci. USA* **177**, 19873–19878 (2020).
24. Block, P. et al. Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nat. Hum. Behav.* **4**, 588–596 (2020).
25. Karin, O. et al. Adaptive cyclic exit strategies from lockdown to suppress COVID-19 and allow economic activity. Preprint at https://doi.org/10.1101/2020.04.04.20053579 (2020).
26. Gao, S. et al. Mapping county-level mobility pattern changes in the United States in response to COVID-19. *SIGSPATIAL Special* **12**, 16–26 (2020).
27. Klein, B. et al. *Assessing Changes in Commuting and Individual Mobility in Major Metropolitan Areas in the United States during the COVID-19 Outbreak*. https://www.networkscienceinstitute.org/publications/assessing-changes-in-commuting-and-individual-mobility-in-major-metropolitan-areas-in-the-united-states-during-the-covid-19-outbreak (2020).
28. Benzell, S. G., Collis, A. & Nicolaides, C. Rationing social contact during the COVID-19 pandemic: transmission risk and social benefits of US locations. *Proc. Natl Acad. Sci. USA* **117**, 14642–14644 (2020).
29. Baicker, K. et al. Is it safer to visit a coffee shop or a gym? *The New York Times* (6 May 2020).
30. Hsiang, S. et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267 (2020).
31. Deming, W. E. & Stephan, F. F. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.* **11**, 427–444 (1940).
32. The New York Times. *Coronavirus (COVID-19) Data in the United States*. https://github.com/nytimes/covid-19-data (2020).
33. Tian, H. et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).
34. Watts, D. J., Muhamad, R., Medina, D. C. & Dodds, P. S. Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proc. Natl Acad. Sci. USA* **102**, 11157–11162 (2005).
35. California Department of Public Health. *COVID-19 Industry Guidance: Retail*. https://covid19.ca.gov/pdf/guidance-retail.pdf (2020).
36. Birge, J., Candogan, O. & Feng, Y. Controlling epidemic spread: reducing economic losses with targeted closures. *BFI Working Paper No. 2020-57* (8 May 2020).
37. Webb Hooper, M., Nápoles, A. M. & Pérez-Stable, E. J. COVID-19 and racial/ethnic disparities. *J. Am. Med. Assoc.* **323**, 2466–2467 (2020).
38. Laurencin, C. T. & McClinton, A. The COVID-19 pandemic: a call to action to identify and address racial and ethnic disparities. *J. Racial Ethn. Health Disparities* **7**, 398–402 (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

The Methods is structured as follows. We describe the datasets that we used in the 'Datasets' section and the mobility network that we derived from these datasets in the 'Mobility network' section. In the 'Model dynamics' section, we discuss the SEIR model that we overlaid on the mobility network; in the 'Model calibration' section, we describe how we calibrated this model and quantified uncertainty in its predictions. Finally, in the 'Analysis details' section, we provide details on the experimental procedures used for our analyses of mobility reduction, reopening plans and demographic disparities.

Datasets

SafeGraph. We use data provided by SafeGraph, a company that aggregates anonymized location data from numerous mobile applications. SafeGraph data captures the movement of people between CBGs, which are geographical units that typically contain a population of between 600 and 3,000 people, and POIs such as restaurants, grocery stores or religious establishments. Specifically, we use the following SafeGraph datasets.

First, we used the Places Patterns[39] and Weekly Patterns (v1)[40] datasets. These datasets contain, for each POI, hourly counts of the number of visitors, estimates of median visit duration in minutes (the 'dwell time') and aggregated weekly and monthly estimates of the home CBGs of visitors. We use visitor home CBG data from the Places Patterns dataset: for privacy reasons, SafeGraph excludes a home CBG from this dataset if fewer than five devices were recorded at the POI from that CBG over the course of the month. For each POI, SafeGraph also provides their North American industry classification system category, as well as estimates of its physical area in square feet. The area is computed using the footprint polygon SafeGraph that assigns to the POI[41,42]. We analyse Places Patterns data from 1 January 2019 to 29 February 2020 and Weekly Patterns data from 1 March 2020 to 2 May 2020.

Second, we used the Social Distancing Metrics dataset[43], which contains daily estimates of the proportion of people staying home in each CBG. We analyse Social Distancing Metrics data from 1 March 2020 to 2 May 2020.

We focus on 10 of the largest metro areas in the United States (Extended Data Table 1). We chose these metro areas by taking a random subset of the SafeGraph Patterns data and selecting the 10 metro areas with the most POIs in the data. The application of the methods described in this paper to the other metro areas in the original SafeGraph data should be straightforward. For each metro area, we include all POIs that meet all of the following requirements: (1) the POI is located in the metro area ; (2) SafeGraph has visit data for this POI for every hour that we model, from 00:00 on 1 March 2020 to 23:00 on 2 May 2020; (3) SafeGraph has recorded the home CBGs of visitors to this POI for at least one month from January 2019 to February 2020; (4) the POI is not a 'parent' POI. Parent POIs comprise a small fraction of POIs in the dataset that overlap and include the visits from their 'child' POIs: for example, many malls in the dataset are parent POIs, which include the visits from stores that are their child POIs. To avoid double-counting visits, we remove all parent POIs from the dataset. After applying these POI filters, we include all CBGs that have at least one recorded visit to at least ten of the remaining POIs; this means that CBGs from outside the metro area may be included if they visit this metro area frequently enough. Summary statistics of the post-processed data are shown in Extended Data Table 1. Overall, we analyse 56,945 CBGs from the 10 metro areas, and more than 310 million visits from these CBGs to 552,758 POIs.

SafeGraph data have been used to study consumer preferences[44] and political polarization[45]. More recently, it has been used as one of the primary sources of mobility data in the USA for tracking the effects of the COVID-19 pandemic[26,28,46–48]. In Supplementary Methods section 1, we show that aggregate trends in SafeGraph mobility data match the aggregate trends in Google mobility data in the USA[49], before and after the imposition of stay-at-home measures. Previous analyses of SafeGraph data have shown that it is geographically representative—for example, it does not systematically overrepresent individuals from CBGs in different counties or with different racial compositions, income levels or educational levels[50,51].

US census. Our data on the demographics of the CBGs comes from the American Community Survey (ACS) of the US Census Bureau[52]. We use the 5-year ACS data (2013–2017) to extract the median household income, the proportion of white residents and the proportion of Black residents of each CBG. For the total population of each CBG, we use the most-recent one-year estimates (2018); one-year estimates are noisier but we wanted to minimize systematic downward bias in our total population counts (due to population growth) by making them as recent as possible.

***The New York Times* dataset.** We calibrated our models using the COVID-19 dataset published by the *The New York Times*[32]. Their dataset consists of cumulative counts of cases and deaths in the USA over time, at the state and county level. For each metro area that we modelled, we sum over the county-level counts to produce overall counts for the entire metro area. We convert the cumulative case and death counts to daily counts for the purposes of model calibration, as described in the 'Model calibration' section.

Data ethics. The dataset from *The New York Times* consists of aggregated COVID-19-confirmed case and death counts collected by journalists from public news conferences and public data releases. For the mobility data, consent was obtained by the third-party sources that collected the data. SafeGraph aggregates data from mobile applications that obtain opt-in consent from their users to collect anonymous location data. Google's mobility data consists of aggregated, anonymized sets of data from users who have chosen to turn on the location history setting. Additionally, we obtained IRB exemption for SafeGraph data from the Northwestern University IRB office.

Mobility network

Definition. We consider a complete undirected bipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with time-varying edges. The vertices $\mathcal{V}$ are partitioned into two disjoint sets $\mathcal{C} = \{c_1, \ldots, c_m\}$, representing m CBGs, and $\mathcal{P} = \{p_1, \ldots, p_n\}$, representing n POIs. From US census data, each CBG c_i is labelled with its population N_{c_i}, income distribution, and racial and age demographics. From SafeGraph data, each POI p_j is similarly labelled with its category (for example, restaurant, grocery store or religious organization), its physical size in square feet a_{p_j}, and the median dwell time d_{p_j} of visitors to p_j. The weight $w_{ij}^{(t)}$ on an edge (c_i, p_j) at time t represents our estimate of the number of individuals from CBG c_i visiting POI p_j at the tth hour of simulation. We record the number of edges (with non-zero weights) in each metro area and for all hours from 1 March 2020 to 2 May 2020 in Extended Data Table 1. Across all 10 metro areas, we study 5.4 billion edges between 56,945 CBGs and 552,758 POIs.

Overview of the network estimation. The central technical challenge in constructing this network is estimating the network weights $W^{(t)} = \{w_{ij}^{(t)}\}$ from SafeGraph data, as this visit matrix is not directly available from the data. Our general methodology for network estimation is as follows.

First, from SafeGraph data, we derived a time-independent estimate $\overline{W}$ of the visit matrix that captures the aggregate distribution of visits from CBGs to POIs from January 2019 to February 2020.

Second, because visit patterns differ substantially from hour to hour (for example, day versus night) and day to day (for example, before versus after lockdown), we used current SafeGraph data to capture these hourly variations and to estimate the CBG marginals $U^{(t)}$, that

is, the number of people in each CBG who are out visiting POIs at hour t, as well as the POI marginals $V^{(t)}$, that is, the total number of visitors present at each POI p_j at hour t.

Finally, we applied the iterative proportional fitting procedure (IPFP) to estimate an hourly visit matrix $W^{(t)}$ that is consistent with the hourly marginals $U^{(t)}$ and $V^{(t)}$ but otherwise 'as similar as possible' to the distribution of visits in the aggregate visit matrix $\overline{W}$, in terms of Kullback–Leibler divergence.

IPFP is a classic statistical method[31] for adjusting joint distributions to match prespecified marginal distributions, and it is also known in the literature as biproportional fitting, the RAS algorithm or raking[53]. In the social sciences, it has been widely used to infer the characteristics of local subpopulations (for example, within each CBG) from aggregate data[54–56]. IPFP estimates the joint distribution of visits from CBGs to POIs by alternating between scaling each row to match the hourly row (CBG) marginals $U^{(t)}$ and scaling each column to match the hourly column (POI) marginals $V^{(t)}$. Further details about the estimation procedure are provided in Supplementary Methods section 3.

Model dynamics

To model the spread of SARS-CoV-2, we overlay a metapopulation disease transmission model on the mobility network defined in the 'Mobility Network' section. The transmission model structure follows previous work[15,20] on epidemiological models of SARS-CoV-2 but incorporates a fine-grained mobility network into the calculations of the transmission rate. We construct separate mobility networks and models for each metropolitan statistical area.

We use a SEIR model with susceptible (S), exposed (E), infectious (I) and removed (R) compartments. Susceptible individuals have never been infected, but can acquire the virus through contact with infectious individuals, which may happen at POIs or in their home CBG. They then enter the exposed state, during which they have been infected but are not infectious yet. Individuals transition from exposed to infectious at a rate inversely proportional to the mean latency period. Finally, they transition into the removed state at a rate inversely proportional to the mean infectious period. The removed state represents individuals who can no longer be infected or infect others, for example, because they have recovered, self-isolated or died.

Each CBG c_i maintains its own SEIR instantiation, with $S_{c_i}^{(t)}$, $E_{c_i}^{(t)}$, $I_{c_i}^{(t)}$ and $R_{c_i}^{(t)}$ representing how many individuals in CBG c_i are in each disease state at hour t, and $N_{c_i} = S_{c_i}^{(t)} + E_{c_i}^{(t)} + I_{c_i}^{(t)} + R_{c_i}^{(t)}$. At each hour t, we sample the transitions between states as follows:

$$N_{S_{c_i}\to E_{c_i}}^{(t)} \sim \mathrm{Pois}\left(\frac{S_{c_i}^{(t)}}{N_{c_i}}\sum_{j=1}^{n}\lambda_{p_j}^{(t)}w_{ij}^{(t)}\right) + \mathrm{Binom}\left(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)}\right) \quad (1)$$

$$N_{E_{c_i}\to I_{c_i}}^{(t)} \sim \mathrm{Binom}\left(E_{c_i}^{(t)}, 1/\delta_E\right) \quad (2)$$

$$N_{I_{c_i}\to R_{c_i}}^{(t)} \sim \mathrm{Binom}\left(I_{c_i}^{(t)}, 1/\delta_I\right), \quad (3)$$

where $\lambda_{p_j}^{(t)}$ is the rate of infection at POI p_j at time t; $w_{ij}^{(t)}$, the ijth entry of the visit matrix from the mobility network (see 'Mobility network'), is the number of visitors from CBG c_i to POI p_j at time t; $\lambda_{c_i}^{(t)}$ is the base rate of infection that is independent of visiting POIs; δ_E is the mean latency period; and δ_I is the mean infectious period.

We then update each state to reflect these transitions. Let $\Delta S_{c_i}^{(t)} = S_{c_i}^{(t+1)} - S_{c_i}^{(t)}$ and likewise for $\Delta E_{c_i}^{(t)}$, $\Delta I_{c_i}^{(t)}$ and $\Delta R_{c_i}^{(t)}$. Then, we make the following updates:

$$\Delta S_{c_i}^{(t)} = -N_{S_{c_i}\to E_{c_i}}^{(t)} \quad (4)$$

$$\Delta E_{c_i}^{(t)} = N_{S_{c_i}\to E_{c_i}}^{(t)} - N_{E_{c_i}\to I_{c_i}}^{(t)} \quad (5)$$

$$\Delta I_{c_i}^{(t)} = N_{E_{c_i}\to I_{c_i}}^{(t)} - N_{I_{c_i}\to R_{c_i}}^{(t)} \quad (6)$$

$$\Delta R_{c_i}^{(t)} = N_{I_{c_i}\to R_{c_i}}^{(t)}. \quad (7)$$

The number of new exposures. We separate the number of new exposures $N_{S_{c_i}\to E_{c_i}}^{(t)}$ in CBG c_i at time t into two parts: cases from visiting POIs, which are sampled from $\mathrm{Pois}((S_{c_i}^{(t)}/N_{c_i})\sum_{j=1}^{n}\lambda_{p_j}^{(t)}w_{ij}^{(t)})$, and other cases not captured by visiting POIs, which are sampled from $\mathrm{Binom}(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)})$.

First, we calculate the number of new exposures from visiting POIs. We assume that any susceptible visitor to POI p_j at time t has the same independent probability $\lambda_{p_j}^{(t)}$ of being infected and transitioning from the susceptible (S) to the exposed (E) state. As there are $w_{ij}^{(t)}$ visitors from CBG c_i to POI p_j at time t, and we assume that a $S_{c_i}^{(t)}/N_{c_i}$ fraction of them are susceptible, the number of new exposures among these visitors is distributed as $\mathrm{binom}(w_{ij}^{(t)}S_{c_i}^{(t)}/N_{c_i}, \lambda_{p_j}^{(t)}) \approx \mathrm{Pois}(\lambda_{p_j}^{(t)}w_{ij}^{(t)}S_{c_i}^{(t)}/N_{c_i})$. The number of new exposures among all outgoing visitors from CBG c_i is therefore distributed as the sum of the above expression over all POIs, $\mathrm{Pois}((S_{c_i}^{(t)}/N_{c_i})\sum_{j=1}^{n}\lambda_{p_j}^{(t)}w_{ij}^{(t)})$.

We model the infection rate at POI p_j at time t, $\lambda_{p_j}^{(t)} = \beta_{p_j}^{(t)}(I_{p_j}^{(t)}/V_{p_j}^{(t)})$, as the product of its transmission rate $\beta_{p_j}^{(t)}$ and proportion of infectious individuals $I_{p_j}^{(t)}/V_{p_j}^{(t)}$, where $V_{p_j}^{(t)} = \sum_{i=1}^{m} w_{ij}^{(t)}$ is the total number of visitors to p_j at time t. We model the transmission rate at POI p_j at time t as

$$\beta_{p_j}^{(t)} = \psi d_{p_j}^2 \frac{V_{p_j}^{(t)}}{a_{p_j}}, \quad (8)$$

where a_{p_j} is the physical area of p_j, and ψ is a transmission constant (shared across all POIs) that we fit to data. The inverse scaling of transmission rate with area a_{p_j} is a standard simplifying assumption[57]. The dwell time fraction $d_{p_j} \in [0, 1]$ is what fraction of an hour an average visitor to p_j at any hour will spend there (Supplementary Methods section 3); it has a quadratic effect on the POI transmission rate $\beta_{p_j}^{(t)}$ because it reduces both the time that a susceptible visitor spends at p_j and the density of visitors at p_j. With this expression for the transmission rate $\beta_{p_j}^{(t)}$, we can calculate the infection rate at POI p_j at time t as

$$\lambda_{p_j}^{(t)} = \beta_{p_j}^{(t)}\frac{I_{p_j}^{(t)}}{V_{p_j}^{(t)}} = \psi d_{p_j}^2 \frac{I_{p_j}^{(t)}}{a_{p_j}}. \quad (9)$$

For sufficiently large values of ψ and a sufficiently large proportion of infected individuals, the expression above can sometimes exceed 1. To address this, we simply clip the infection rate to 1. However, this occurs very rarely for the parameter settings and simulation duration that we use.

Finally, to compute the number of infectious individuals at p_j at time t, $I_{p_j}^{(t)}$, we assume that the proportion of infectious individuals among the $w_{kj}^{(t)}$ visitors to p_j from a CBG c_k mirrors the overall density of infections $I_{c_k}^{(t)}/N_{c_k}$ in that CBG, although we note that the scaling factor ψ can account for differences in the ratio of infectious individuals who visit POIs. This gives

$$I_{p_j}^{(t)} = \sum_{k=1}^{m}\frac{I_{c_k}^{(t)}}{N_{c_k}}w_{kj}^{(t)}. \quad (10)$$

In addition to the new exposures from infections at POIs, we model a CBG-specific base rate of new exposures that is independent of POI visit activity. This captures other sources of infections, for example, household infections or infections at POIs that are absent from the SafeGraph data. We assume that at each hour, every susceptible

individual in CBG c_i has a base probability $\lambda_{c_i}^{(t)}$ of becoming infected and transitioning to the exposed state, where

$$\lambda_{c_i}^{(t)} = \beta_{\text{base}} \frac{I_{c_i}^{(t)}}{N_{c_i}} \tag{11}$$

is the product of the base transmission rate β_{base} and the proportion of infectious individuals in CBG c_i. β_{base} is a constant (shared across all CBGs) that we fit to data.

By adding all of the above together, the expression for the distribution of the overall number of new exposures in CBG c_i at time t becomes

$$\begin{aligned} N_{S_{c_i}\to E_{c_i}}^{(t)} &\sim \text{Pois}\left(\frac{S_{c_i}^{(t)}}{N_{c_i}} \sum_{j=1}^{n} \lambda_{p_j}^{(t)} w_{ij}^{(t)}\right) + \text{Binom}\left(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)}\right) \\ &= \underbrace{\text{Pois}\left(\psi \frac{S_{c_i}^{(t)}}{N_{c_i}} \sum_{j=1}^{n} \frac{d_{p_j}^2}{a_{p_j}} \left(\sum_{k=1}^{m} \frac{I_{c_k}^{(t)}}{N_{c_k}} w_{kj}^{(t)}\right) w_{ij}^{(t)}\right)}_{\text{New infections from visiting POIs}} \\ &\quad + \underbrace{\text{Binom}\left(S_{c_i}^{(t)}, \beta_{\text{base}} \frac{I_{c_i}^{(t)}}{N_{c_i}}\right)}_{\text{Base rate of new CBG infections}}. \end{aligned} \tag{12}$$

The number of new infectious and removed cases. We model exposed individuals as becoming infectious at a rate that is inversely proportional to the mean latency period δ_E. At each time step t, we assume that each exposed individual has a constant, time-independent probability of becoming infectious, with

$$N_{E_{c_i}\to I_{c_i}}^{(t)} \sim \text{Binom}\left(E_{c_i}^{(t)}, 1/\delta_E\right). \tag{13}$$

Similarly, we model infectious individuals as transitioning to the removed state at a rate that is inversely proportional to the mean infectious period δ_I, with

$$N_{I_{c_i}\to R_{c_i}}^{(t)} \sim \text{Binom}\left(I_{c_i}^{(t)}, 1/\delta_I\right), \tag{14}$$

We estimate $\delta_E = 96$ h (refs. [20,58]) and $\delta_I = 84$ h (ref. [20]) based on previous studies.

Model initialization. In our experiments, $t = 0$ is the first hour of 1 March 2020. We approximate the infectious I and removed R compartments at $t = 0$ as initially empty, with all infected individuals in the exposed E compartment. We further assume that the same expected initial prevalence p_0 occurs in every CBG c_i. At $t = 0$, every individual in the metro area has the same independent probability p_0 of being exposed E instead of susceptible S. We thus initialize the model state by setting

$$S_{c_i}^{(0)} = N_{c_i} - E_{c_i}^{(0)} \tag{15}$$

$$E_{c_i}^{(0)} \sim \text{Binom}\left(N_{c_i}, p_0\right) \tag{16}$$

$$I_{c_i}^{(0)} = 0 \tag{17}$$

$$R_{c_i}^{(0)} = 0. \tag{18}$$

Aggregate mobility and no-mobility baseline models. To test whether the detailed mobility network is necessary, or whether our model is simply making use of aggregate mobility patterns, we tested an alternative SEIR model that uses the aggregate number of visits made to any POI in the metro area in each hour, but not the breakdown of visits between specific CBGs to specific POIs. Like our model, the aggregate mobility model also captures new cases from visiting POIs and a base rate of infection that is independent of POI visit activity; thus, the two models have the same three free parameters (ψ, scaling transmission rates at POIs; β_{base}, the base transmission rate ; and p_0, the initial fraction of infected individuals). However, instead of having POI-specific rates of infection, the aggregate mobility model captures only a single probability that a susceptible person from any CBG will become infected due to a visit to any POI at time t; we make this simplification because the aggregate mobility model no longer has access to the breakdown of visits between CBGs and POIs. This probability $\lambda_{\text{POI}}^{(t)}$ is defined as

$$\lambda_{\text{POI}}^{(t)} = \psi \underbrace{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij}^{(t)}}{nm}}_{\text{Average mobility at time } t} \frac{I^{(t)}}{N}, \tag{19}$$

where m is the number of CBGs, n is the number of POIs, $I^{(t)}$ is the total number of infectious individuals at time t, and N is the total population size of the metro area. For the base rate of infections in CBGs, we assume the same process as in our network model: the probability $\lambda_{c_i}^{(t)}$ that a susceptible person in CBG c_i will become infected in their CBG at time t is equal to β_{base} times the current infectious fraction of c_i (equation (11)). Putting it together, the aggregate mobility model defines the number of new exposures in CBG c_i at time t as

$$N_{S_{c_i}\to E_{c_i}}^{(t)} \sim \underbrace{\text{Binom}\left(S_{c_i}^{(t)}, \lambda_{\text{POI}}^{(t)}\right)}_{\text{New infections from visiting POIs}} + \underbrace{\text{Binom}\left(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)}\right)}_{\text{Base rateof new CBG infections}}. \tag{20}$$

All other dynamics remain the same between the aggregate mobility model and our network model, and we calibrated the models in the same way, which we describe in the 'Model calibration' section. We found that our network model substantially outperformed the aggregate mobility model in predictions of out-of-sample cases: on average across metro areas, the out-of-sample r.m.s.e. of our best-fit network model was only 58% that of the best-fit aggregate mobility model (Extended Data Fig. 1). This demonstrates that it is not only general mobility patterns, but specifically the mobility network that allows our model to accurately fit observed cases.

Next, to determine the extent to which mobility data could aid in modelling the case trajectory, we compared our model to a baseline SEIR model that does not use mobility data and simply assumes that all individuals within an metro area mix uniformly. In this no-mobility baseline, an individual's risk of being infected and transitioning to the exposed state at time t is

$$\lambda^{(t)} = \beta_{\text{base}} \frac{I^{(t)}}{N}, \tag{21}$$

where $I^{(t)}$ is the total number of infectious individuals at time t, and N is the total population size of the metro area. As above, the other model dynamics are identical, and for model calibration we performed a similar grid search over β_{base} and p_0. As expected, we found both the network and aggregate mobility models outperformed the no-mobility model on out-of-sample case predictions (Extended Data Fig. 1).

Model calibration and validation

Most of our model parameters can either be estimated from SafeGraph and US census data, or taken from previous studies (see Extended Data Table 2 for a summary). This leaves three model parameters that do not have direct analogues in the literature, and that we therefore need to calibrate with data: (1) the transmission constant in POIs, ψ (equation (9)); (2) the base transmission rate, β_{base} (equation (11)); and (3) the initial proportion of exposed individuals at time $t = 0$, p_0 (equation (16)).

In this section, we describe how we fitted these parameters to published numbers of confirmed cases, as reported by *The New York Times*. We fitted models for each metro area separately.

Selecting parameter ranges for ψ, β_{base} and p_0. We select parameter ranges for the transmission rate factors ψ and β_{base} by checking whether the model outputs match plausible ranges of the basic reproduction number R_0 before lockdown, as R_0 has been the study of substantial previous work on SARS-CoV-2[59]. Under our model, we can decompose $R_0 = R_{\text{base}} + R_{\text{POI}}$, where R_{POI} describes transmission due to POIs and R_{base} describes the remaining transmission (as in equation (12)). We first establish plausible ranges for R_{base} and R_{POI} before translating these into plausible ranges for β_{base} and ψ.

We assume that R_{base} ranges from 0.1 to 2. R_{base} models transmission that is not correlated with activity at POIs in the SafeGraph dataset, including within-household transmission and transmission at POI categories that are not well-captured in the SafeGraph dataset. We chose the lower limit of 0.1 because beyond that point, base transmission would only contribute minimally to overall R, whereas previous studies have suggested that within-household transmission is a substantial contributor to overall transmission[60–62]. Household transmission alone is not estimated to be sufficient to tip the overall R_0 above 1; for example, a single infected individual has been estimated to cause an average of 0.32 (0.22–0.42) secondary within-household infections[60]. However, because R_{base} may also capture transmission at POIs that are not captured in the SafeGraph dataset, to be conservative, we chose an upper limit of $R_{\text{base}} = 2$; as we describe below, the best-fit models for all 10 metro areas have $R_{\text{base}} < 2$, and 9 out of 10 have $R_{\text{base}} < 1$. We allow R_{POI} to range from 1 to 3, which corresponds to allowing $R_0 = R_{\text{POI}} + R_{\text{base}}$ to range from 1.1 to 5. This is a conservatively wide range, as a previous study[59] estimated a pre-lockdown R_0 of 2–3.

To determine the values of R_{base} and R_{POI} that a given pair of β_{base} and ψ imply, we seeded a fraction of index cases and then ran the model on looped mobility data from the first week of March to capture pre-lockdown conditions. We initialized the model by setting p_0, the initial proportion of exposed individuals at time $t = 0$, to $p_0 = 10^{-4}$, and then sampling in accordance with equation (16). Let N_0 be the number of initial exposed individuals sampled. We computed the number of individuals that these N_0 index cases went on to infect through base transmission, N_{base}, and POI transmission, N_{POI}, which gives

$$R_{\text{POI}} = \frac{N_{\text{POI}}}{N_0} \tag{22}$$

$$R_{\text{base}} = \frac{N_{\text{base}}}{N_0}. \tag{23}$$

We averaged these quantities over stochastic realizations for each metro area. Supplementary Figure 6 shows that, as expected, R_{base} is linear in β_{base} and R_{POI} is linear in ψ. R_{base} lies in the plausible range when β_{base} ranges from 0.0012 to 0.024, and R_{POI} lies in the plausible range (for at least one metro area) when ψ ranges from 515 to 4,886; we therefore consider these parameter ranges when fitting the model.

The extent to which SARS-CoV-2 infections had spread in the USA by the start of our simulation (1 March 2020) is currently unclear[63]. To account for this uncertainty, we allow p_0 to vary across a large range between 10^{-5} and 10^{-2}. As described in the next section, we verified that case count data for all metro areas can be fit using parameter settings for β_{base}, ψ and p_0 within this range.

Fitting to the number of confirmed cases. Using the parameter ranges described above, we grid-searched over ψ, β_{base} and p_0 to find the models that best fit the number of confirmed cases reported by *The New York Times*[32]. For each metro area, we tested 1,500 different combinations of ψ, β_{base} and p_0 in the parameter ranges specified above, with parameters linearly spaced for ψ and β_{base} and logarithmically spread for p_0.

In the 'Model dynamics' section, we directly model the number of infections but not the number of confirmed cases. To estimate the number of confirmed cases, we assume that an $r_c = 0.1$ proportion[20,58,64–66] of infections will be confirmed, and moreover that they will confirmed exactly $\delta_c = 168$ h (7 days)[20,66] after becoming infectious. From these assumptions, we can calculate the predicted number of newly confirmed cases across all CBGs in the metro area on day d,

$$N_{\text{cases}}^{(\text{day } d)} = r_c \sum_{i=1}^{m} \sum_{\tau=24(d-1)+1-\delta_c}^{24d-\delta_c} N_{E_{c_i}\to I_{c_i}}^{(\tau)}, \tag{24}$$

where m indicates the total number of CBGs in the metro area and for convenience we define $N_{E_{c_i}\to I_{c_i}}^{(\tau)}$, the number of newly infectious people at hour τ, to be 0 when $\tau < 1$.

From the dataset of *The New York Times*, we have the reported number of new cases $\hat{N}_{\text{cases}}^{(\text{day } d)}$ for each day d, summed over each county in the metro area. We compare the reported number of cases and the number of cases that our model predicts by computing the r.m.s.e. between each of the $D = \lfloor T/24 \rfloor$ days of our simulations,

$$\text{r.m.s.e.} = \sqrt{\frac{1}{D}\sum_{d=1}^{D}\left(N_{\text{cases}}^{(\text{day } d)} - \hat{N}_{\text{cases}}^{(\text{day } d)}\right)^2}. \tag{25}$$

For each combination of model parameters and for each metro area, we quantify the model fit with the data from *The New York Times* by running 30 stochastic realizations and averaging their r.m.s.e. Note that we measure model fit based on the daily number of new reported cases (as opposed to the cumulative number of reported cases)[67].

Our simulation spans 1 March to 2 May 2020, and we use mobility data from that period. However, because we assume that cases will be confirmed $\delta_c = 7$ days after individuals become infectious, we predict the number of cases with a 7-day offset, from 8 March to 9 May 2020.

Parameter selection and uncertainty quantification. Throughout this paper, we report aggregate predictions from different parameter sets of ψ, β_{base} and p_0, and multiple stochastic realizations. For each metro area, we: (1) find the best-fit parameter set, that is, with the lowest average r.m.s.e. on daily incident cases over stochastic realizations; (2) select all parameter sets that achieve an r.m.s.e. (averaged over stochastic realizations) within 20% of the r.m.s.e. of the best-fit parameter set; and (3) pool together all predictions across those parameter sets and all of their stochastic realizations, and report their mean and 2.5–97.5th percentiles.

On average, each metro area has 9.7 parameter sets that achieve an r.m.s.e. within 20% of the best-fitting parameter set (Supplementary Table 6). For each parameter set, we have results for 30 stochastic realizations.

This procedure corresponds to rejection sampling in an approximate Bayesian computation framework[15], for which we assume an error model that is Gaussian with constant variance; we pick an acceptance threshold based on what the best-fit model achieves; and we use a uniform parameter grid instead of sampling from a uniform prior. It quantifies uncertainty from two sources. First, the multiple realizations capture stochastic variability between model runs with the same parameters. Second, simulating with all parameter sets that are within 20% of the r.m.s.e. of the best fit captures uncertainty in the model parameters ψ, β_{base} and p_0. This procedure is equivalent to assuming that the posterior probability over the true parameters is uniformly spread among all parameter sets within the 20% threshold.

Model validation on out-of-sample cases. We validate our models by showing that they predict the number of confirmed cases on out-of-sample data when we have access to corresponding mobility

data. For each metro area, we split the available dataset from the *The New York Times* into a training set (spanning from 8 March 2020 to 14 April 2020) and a test set (spanning from 15 April 2020 to 9 May 2020). We fit the model parameters ψ, β_{base} and p_0, as described in the 'Mobility network' section, but using only the training set. We then evaluate the predictive accuracy of the resulting model on the test set. When running our models on the test set, we still use mobility data from the test period. Thus, this is an evaluation of whether the models can accurately predict the number of cases, given mobility data, in a time period that was not used for model calibration. Extended Data Figure 1 shows that our network model fits the out-of-sample case data fairly well, and that our model substantially outperforms alternative models that use aggregated mobility data (without a network) or do not use mobility data at all (see 'Aggregate mobility and no-mobility baseline models'). Note that we only use this train/test split to evaluate out-of-sample model accuracy. All other results are generated using parameter sets that best fit the entire dataset, as described above.

Analysis details

In this section, we include additional details about the experiments that underlie the figures in the paper. We do not include explanations for figures that are completely described in the main text.

Counterfactuals of mobility reduction. Associated with Fig. 2a and Supplementary Tables 4, 5. To simulate what would have happened if we changed the magnitude or timing of mobility reduction, we modified the real mobility networks from 1 March to 2 May 2020, and then ran our models on the hypothetical data. In Fig. 2a, we report the total number of people per 100,000 of the population ever infected (that is, in the exposed, infectious and removed states) by the end of the simulation.

To simulate a smaller magnitude of mobility reduction, we interpolate between the mobility network from the first week of simulation (1–7 March 2020), which we use to represent typical mobility levels, and the actual observed mobility network for each week. Let $W^{(t)}$ represent the observed visit matrix at the tth hour of simulation, and let $f(t) = t \bmod 168$ map t to its corresponding hour in the first week of simulation, since there are 168 h in a week. To represent the scenario in which people had committed to $\alpha \in [0, 1]$ times the actual observed reduction in mobility, we construct a visit matrix $\widetilde{W}_{\alpha}^{(t)}$ that is an α-convex combination of $W^{(t)}$ and $W^{f(t)}$,

$$\widetilde{W}_{\alpha}^{(t)} = \alpha W^{(t)} + (1-\alpha) W^{f(t)}. \tag{26}$$

If α is 1, then $\widetilde{W}_{\alpha}^{(t)} = W^{(t)}$, and we use the actual observed mobility network for the simulation. On the other hand, if $\alpha = 0$, then $\widetilde{W}_{\alpha}^{(t)} = W^{f(t)}$, and we assume that people did not reduce their mobility levels at all by looping the visit matrix for the first week of March throughout the simulation. Any other $\alpha \in [0, 1]$ interpolates between these two extremes.

To simulate changing the timing of mobility reduction, we shift the mobility network by $d \in [-7, 7]$ days. Let T represent the last hour in our simulation (2 May 2020, 23:00), let $f(t) = t \bmod 168$ map t to its corresponding hour in the first week of simulation as above, and similarly let $g(t)$ map t to its corresponding hour in the last week of simulation (27 April–2 May 2020). We construct the time-shifted visit matrix $\widetilde{W}_{d}^{(t)}$

$$\widetilde{W}_{d}^{(t)} = \begin{cases} W^{(t-24d)} & \text{if } 0 \le t - 24d \le T, \\ W^{f(t-24d)} & \text{if } t - 24d < 0, \\ W^{g(t-24d)} & \text{otherwise.} \end{cases} \tag{27}$$

If d is positive, this corresponds to starting mobility reduction d days later; if we imagine time on a horizontal line, this shifts the time series to the right by $24d$ hours. However, doing so leaves the first $24d$ hours without visit data, so we fill it in by reusing visit data from the first week of simulation. Likewise, if d is negative, this corresponds to starting mobility reduction d days earlier, and we fill in the last $24d$ hours with visit data from the last week of simulation.

Distribution of predicted infections across POIs. Associated with Fig. 2b, Extended Data Fig. 2 and Supplementary Fig. 10. We run our models on the observed mobility data from 1 March–2 May 2020 and record the number of predicted infections that occur at each POI. Specifically, for each hour t, we compute the number of expected infections that occur at each POI p_j by taking the number of susceptible people who visit p_j in that hour multiplied by the POI infection rate $\lambda_{p_j}^{(t)}$ (equation (9)). In Fig. 2b and Supplementary Fig. 10, we sort the POIs by their total predicted number of infections (summed over hours) and plot the cumulative distribution of infections over this ordering of POIs. In Extended Data Fig. 2, we select the POI categories that contribute the most to predicted infections and plot the daily proportion of POI infections each category accounted for (summed over POIs within the category) over time.

Reducing mobility by capping maximum occupancy. Associated with Figs. 2c and Extended Data Fig. 3. We implemented two partial reopening strategies: one that uniformly reduced visits at POIs to a fraction of full activity, and the other that 'capped' the number of hourly visits at each POI to a fraction of the maximum occupancy of that the POI. For each reopening strategy, we started the simulation on 1 March 2020 and ran it until 30 May 2020, using the observed mobility network from 1 March to 30 April 2020, and then using a hypothetical post-reopening mobility network from 1 May to 30 May 2020, corresponding to the projected impact of that reopening strategy. Because we only have observed mobility data from 1 March to 2 May 2020, we impute the missing mobility data up to 30 May 2020 by looping mobility data from the first week of March, as in the above analysis on the effect of past reductions in mobility. Let T represent the last hour for which we have observed mobility data (2 May 2020, 23:00). To simplify the notation, we define

$$h(t) = \begin{cases} t & \text{if } t < T, \\ f(t) & \text{otherwise,} \end{cases} \tag{28}$$

where, as above, $f(t) = t \bmod 168$. This function leaves t unchanged if there is observed mobility data at time t, and otherwise maps t to the corresponding hour in the first week of our simulation.

To simulate a reopening strategy that uniformly reduced visits to an γ fraction of their original level, where $\gamma \in [0, 1]$, we constructed the visit matrix

$$\widetilde{W}_{\gamma}^{(t)} = \begin{cases} W^{h(t)} & \text{if } t < \tau, \\ \gamma W^{h(t)} & \text{otherwise,} \end{cases} \tag{29}$$

where τ represents the first hour of reopening (1 May 2020, 00:00). In other words, we use the actual observed mobility network up until hour τ, and then subsequently simulate an γ fraction of full mobility levels.

To simulate the reduced occupancy strategy, we first estimated the maximum occupancy M_{pj} of each POI p_j as the maximum number of visits that it ever had in one hour, across all of 1 March–2 May 2020. As in previous sections, let $w_{ij}^{(t)}$ represent the ijth entry in the observed visit matrix $W^{(t)}$, that is, the number of people from CBG c_i who visited p_j in hour t, and let $V_{p_j}^{(t)}$ represent the total number of visitors to p_j in that hour, that is, $\sum_i w_{ij}^{(t)}$. We simulated capping at a β fraction of maximum occupancy, where $\beta \in [0, 1]$, by constructing the visit matrix $\widetilde{W}_{\beta}^{(t)}$ for which the ijth entry is

$$\widetilde{w}_{ij\beta}^{(t)} = \begin{cases} w_{ij}^{h(t)} & \text{if } t < \tau \text{ or } V_{p_j}^{(t)} \le \beta M_{p_j}, \\ \dfrac{\beta M_{p_j}}{V_{p_j}^{(t)}} w_{ij}^{h(t)} & \text{otherwise.} \end{cases} \tag{30}$$

This corresponds to the following procedure: for each POI p_j and time t, we first check whether $t < \tau$ (reopening has not started) or whether $V_{p_j}^{(t)} \le \beta M_{p_j}$ (the total number of visits to p_j at time t is below the allowed maximum βM_{p_j}). If so, we leave $w_{ij}^{h(t)}$ unchanged. Otherwise, we compute the scaling factor $\frac{\beta M_{p_j}}{V_{p_j}^{(t)}}$ that would reduce the total visits to p_j at time t down to the allowed maximum βM_{p_j}, and then scale down all visits from each CBG c_i to p_j proportionately. For both reopening strategies, we calculate the predicted increase in cumulative incidence at the end of the reopening period (30 May 2020) compared to the start of the reopening period (1 May 2020).

Relative risk of reopening different categories of POIs. Associated with Fig. 2d, Extended Data Fig. 5 and Supplementary Figs. 11, 15–24. We study separately the reopening of the 20 POI categories with the most visits in SafeGraph data. In this analysis, we follow previous studies[28] and do not study four categories: 'child day-care services' and 'elementary and secondary schools' (because children under 13 are not well-tracked by SafeGraph); 'drinking places (alcoholic beverages)' (because SafeGraph seems to undercount these locations[28]) and 'nature parks and other similar institutions' (because boundaries and therefore areas are not well-defined by SafeGraph). We also exclude 'general medical and surgical hospitals' and 'other airport operations' (because hospitals and air travel both involve many additional risk factors that our model is not designed to capture). We do not filter out these POIs during model fitting (that is, we assume that people visit these POIs, and that transmissions occur there) because including them still increases the proportion of overall mobility that our dataset captures; we simply do not analyse these categories, because we wish to be conservative and only focus on categories for which we are most confident that we are capturing transmission faithfully.

This reopening analysis is similar to the previous experiments on reducing maximum occupancy versus uniform reopening (see 'Reducing mobility by capping maximum occupancy'). As above, we set the reopening time τ to 1 May 2020, 00:00. To simulate reopening a POI category, we take the set of POIs in that category, $\mathcal{V}$, and set their activity levels after reopening to that of the first week of March. For POIs not in the category $\mathcal{V}$, we keep their activity levels after reopening the same, that is, we simply repeat the activity levels of the last week of our data (27 April–2 May 2020): This gives us the visit matrix $\widetilde{W}^{(t)}$ with entries

$$\widetilde{w}_{ij}^{(t)} = \begin{cases} w_{ij}^{(t)} & \text{if } t < \tau, \\ w_{ij}^{f(t)} & \text{if } t \ge \tau, p_j \in \mathcal{V} \\ w_{ij}^{g(t)} & \text{if } t \ge \tau, p_j \notin \mathcal{V}. \end{cases} \tag{31}$$

As in the above reopening analysis, $f(t)$ maps t to the corresponding hour in the first week of March, and $g(t)$ maps t to the corresponding hour in the last week of our data. For each category, we calculate the predicted difference between (1) the cumulative fraction of people who have been infected by the end of the reopening period (30 May 2020) and (2) the cumulative fraction of people infected by 30 May 2020 had we not reopened the POI category (that is, if we simply repeated the activity levels of the last week of our data). This seeks to model the increase in cumulative incidence by the end of May from reopening the POI category. In Extended Data Fig. 5 and Supplementary Figs. 15–24, the bottom right panel shows the predicted increase for the category as a whole, and the bottom left panel shows the predicted increase per POI (that is, the total increase divided by the number of POIs in the category).

Per-capita mobility. Associated with Fig. 3d, Extended Data Fig. 6 and Supplementary Fig. 3. Each group of CBGs (for example, the bottom income decile) comprises a set $\mathcal{U}$ of CBGs that fit the corresponding criteria. In Fig. 3d and Extended Data Fig. 6, we show the daily per-capita mobilities of different pairs of groups (broken down by income and by race). To measure the per-capita mobility of a group on day d, we take the total number of visits made from those CBGs to any POI, $\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{P}} \sum_{t=24d}^{24d+23} w_{ij}^{(t)}$, and divide it by the total population of the CBGs in the group, $\sum_{c_i \in \mathcal{U}} N_{c_i}$. In Supplementary Fig. 3, we show the total number of visits made by each group to each POI category, accumulated over the entire data period (1 March–2 May 2020) and then divided by the total population of the group.

Average predicted transmission rate of a POI category. Associated with Fig. 3e and Extended Data Tables 3, 4. We compute the predicted average hourly transmission rate experienced by a group of CBGs $\mathcal{U}$ at a POI category $\mathcal{V}$ as

$$\overline{\beta}_{\mathcal{U}\mathcal{V}} = \frac{\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{V}} \sum_{t=1}^{T} w_{ij}^{(t)} \beta_{p_j}^{(t)}}{\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{V}} \sum_{t=1}^{T} w_{ij}^{(t)}}, \tag{32}$$

where, as above, $\beta_{p_j}^{(t)}$ is the transmission rate at POI p_j in hour t (equation (8)), $w_{ij}^{(t)}$ is the number of visitors from CBG c_i at POI p_j in hour t, and T is the last hour in our simulation. This represents the expected transmission rate encountered during a visit by someone from a CBG in group $\mathcal{U}$ to a POI in category $\mathcal{V}$.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Inferred hourly mobility networks as well as the estimated models are available at the project website (http://covid-mobility.stanford.edu). Raw census data (https://www.census.gov/programs-surveys/acs), case and death counts from *The New York Times* (https://github.com/nytimes/covid-19-data) and Google mobility data (https://www.google.com/covid19/mobility/) are also publicly available. Mobile phone mobility data are freely available to researchers, non-profit organizations and governments through the SafeGraph COVID-19 Data Consortium (https://www.safegraph.com/covid-19-data-consortium).

Code availability

Code is publicly available at the project website (http://covid-mobility.stanford.edu).

39. SafeGraph. *Places Schema*. https://docs.safegraph.com/docs/places-schema (2020).
40. SafeGraph. *Weekly Patterns*. https://docs.safegraph.com/docs/weekly-patterns (2020).
41. SafeGraph. *Using SafeGraph Polygons to Estimate Point-Of-Interest Square Footage*. https://www.safegraph.com/blog/using-safegraph-polygons-to-estimate-point-of-interest-square-footage (2019).
42. SafeGraph. *Guide to Points-of-Interest Data: POI Data FAQ*. https://www.safegraph.com/points-of-interest-poi-data-guide (2020).
43. SafeGraph. *Social Distancing Metrics*. https://docs.safegraph.com/docs/social-distancing-metrics (2020).
44. Athey, S. et al. Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. In *AEA Papers and Proceedings* Vol. 108, 64–67 (2018).
45. Chen, M. K. & Rohla, R. The effect of partisanship and political advertising on close family ties. *Science* **360**, 1020–1024 (2018).
46. Farboodi, M., Jarosch, G. & Shimer, R. Internal and external effects of social distancing in a pandemic. *NBER Working Paper 27059* https://doi.org/10.3386/w27059 (National Bureau Of Economic Research, 2020).
47. Killeen, B. D. et al. A county-level dataset for informing the United States' response to COVID-19. Preprint at https://arxiv.org/abs/2004.00756 (2020).
48. Allcott, H. et al. Polarization and public health: partisan differences in social distancing during the coronavirus pandemic. *NBER Working Paper 26946* https://doi.org/10.3386/w26946 (National Bureau Of Economic Research, 2020).

49. Google. *COVID-19 Community Mobility Reports*. https://google.com/covid19/mobility/ (2020).
50. Athey, S. et al. Experienced Segregation. *Working Paper 3785* https://gsb.stanford.edu/faculty-research/working-papers/experienced-segregation (2019).
51. Squire, R. F. *What about Bias in the SafeGraph Dataset?* https://safegraph.com/blog/what-about-bias-in-the-safegraph-dataset (2019).
52. US Census Bureau. *American Community Survey (ACS)* https://census.gov/programs-surveys/acs (published 30 January 2020).
53. Bishop, Y. M., Fienberg, S. E. & Holland, P. W. *Discrete Multivariate Analysis: Theory and Practice* (MIT Press, 1975).
54. Birkin, M. & Clarke, M. Synthesis—a synthetic spatial information system for urban and regional analysis: methods and examples. *Environ. Plann. A* **20**, 1645–1671 (1988).
55. Wong, D. W. The reliability of using the iterative proportional fitting procedure. *Prof. Geogr*. **44**, 340–348 (1992).
56. Simpson, L. & Tranmer, M. Combining sample and census data in small area estimates: iterative proportional fitting with standard software. *Prof. Geogr*. **57**, 222–234 (2005).
57. Hu, H., Nigmatulina, K. & Eckhoff, P. The scaling of contact rates with population density for the infectious disease models. *Math. Biosci*. **244**, 125–134 (2013).
58. Kucharski, A. J. et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis*. **20**, 553–558 (2020).
59. Park, M., Cook, A. R., Lim, J. T., Sun, Y. & Dickens, B. L. A systematic review of COVID-19 epidemiology based on current evidence. *J. Clin. Med*. **9**, 967 (2020).
60. Curmei, M. et al. Estimating household transmission of SARS-CoV-2. Preprint at https://doi.org/10.1101/2020.05.23.20111559 (2020).
61. Li, W. et al. The characteristics of household transmission of COVID-19. *Clin. Infect. Dis*. **71**, 1943–1946 (2020).
62. Gudbjartsson, D. F. et al. Spread of SARS-CoV-2 in the Icelandic population. *N. Engl. J. Med*. **382**, 2302–2315 (2020).
63. Carey, B. & Glanz, J. Hidden outbreaks spread through U.S. cities far earlier than Americans knew, estimates say. *The New York Times* (23 April 2020).
64. Bommer, C. & Vollmer, S. *Average Detection Rate of SARS-CoV-2 Infections is Estimated Around Nine Percent*. https://www.uni-goettingen.de/en/606540.html (2020).
65. Javan, E., Fox, S. J. & Meyers, L. A. The unseen and pervasive threat of COVID-19 throughout the US. Preprint at https://doi.org/10.1101/2020.04.06.20053561 (2020).
66. Perkins, T. A. et al. Estimating unobserved SARS-CoV-2 infections in the United States. *Proc. Natl Acad. Sci. USA* **117**, 22597–22602 (2020).
67. King, A. A., Domenech de Cellès, M., Magpantay, F. M. & Rohani, P. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proc. R. Soc. Lond. B* **282**, 20150347 (2015).

Acknowledgements We thank Y.-Y. Ahn, R. Appel, C. Chen, J. Feng, N. Fishman, S. Fullerton, T. Hashimoto, M. Kraemer, P. Liang, M. Lipsitch, K. Loh, D. Ouyang, R. Rosenfeld, S. Sagawa, J. Steinhardt, R. Tibshirani, J. Ugander, D. Vrabac, seminar participants and Stanford's Computer Science and Civil Society for support and comments; and N. Singh, R. F. Squire, J. Williams-Holt, J. Wolf, R. Yang and others at SafeGraph for mobile phone mobility data and feedback. This research was supported by US National Science Foundation under OAC-1835598 (CINES), OAC-1934578 (HDR), CCF-1918940 (Expeditions), IIS-2030477 (RAPID), Stanford Data Science Initiative, Wu Tsai Neurosciences Institute and Chan Zuckerberg Biohub. S.C. was supported by an NSF Fellowship. E.P. was supported by a Hertz Fellowship. P.W.K. was supported by the Facebook Fellowship Program. J.L. is a Chan Zuckerberg Biohub investigator.

Author contributions S.C., E.P. and P.W.K. performed computational analysis. All authors jointly analysed the results and wrote the paper.

Competing interests The authors declare no competing interests.

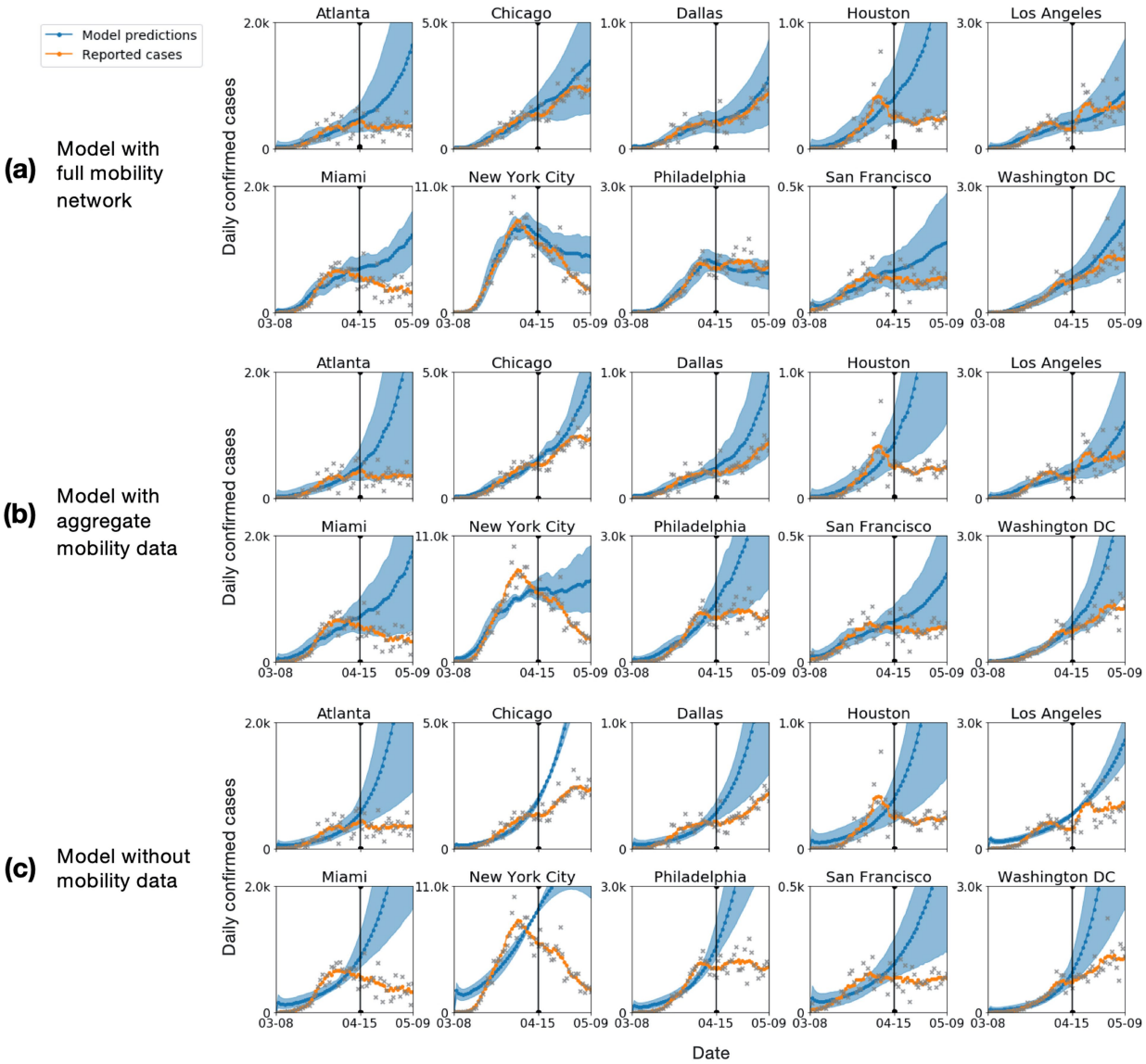
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2923-3.

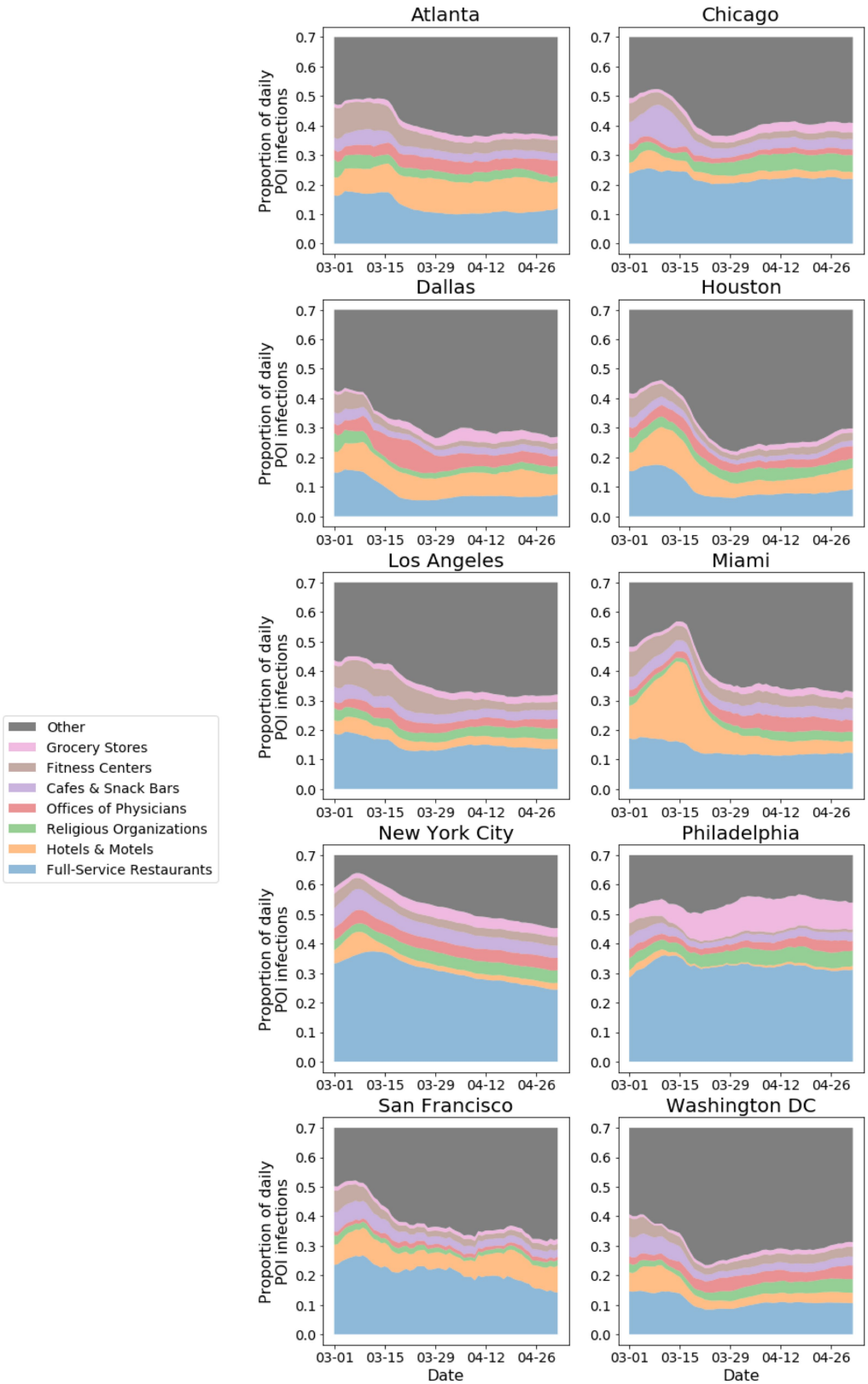
Correspondence and requests for materials should be addressed to J.L.

Peer review information *Nature* thanks Moritz Kraemer, Marc Lipsitch and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

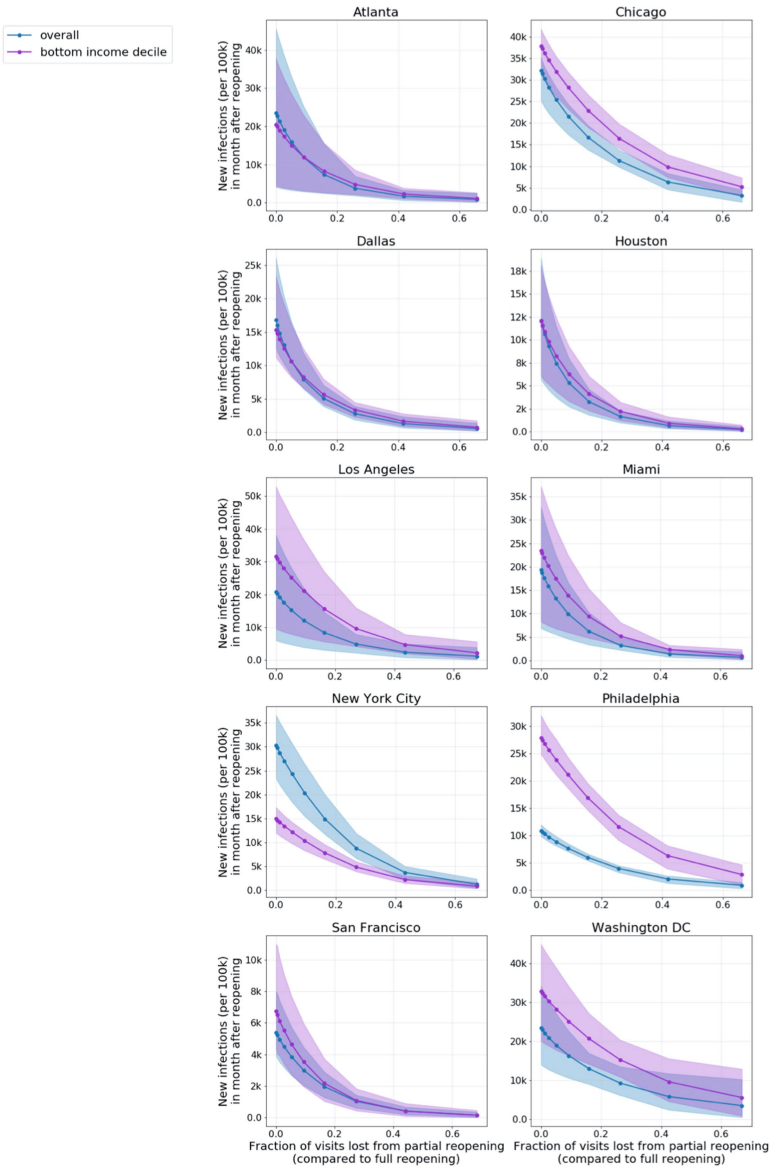
Reprints and permissions information is available at http://www.nature.com/reprints.



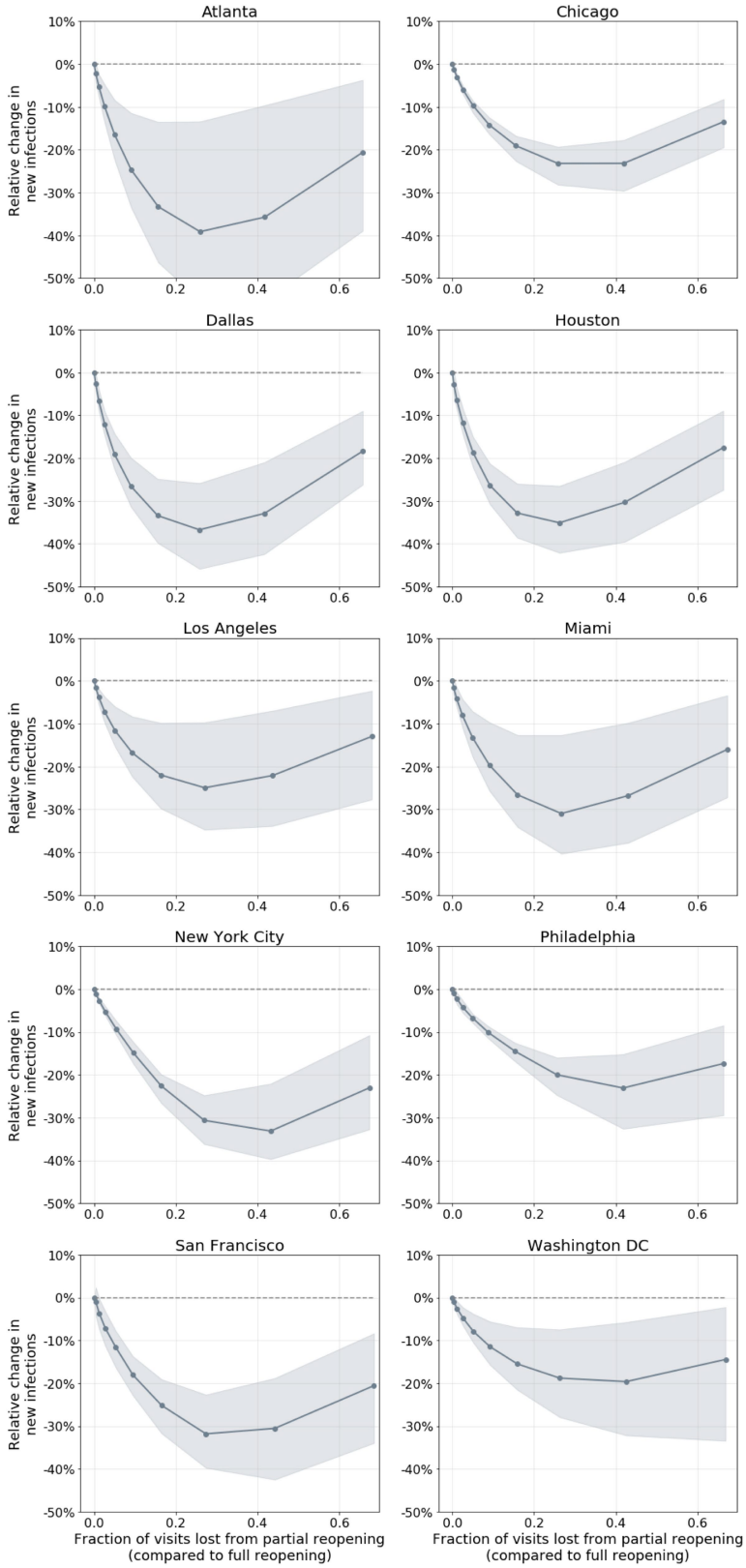
Extended Data Fig. 1 | Mobility-based epidemiological model and its predictions. a–**c**, Predicted (blue) and true (orange) daily case counts for our model (**a**), which uses hourly mobility networks, an SEIR model (**b**) that uses hourly aggregated mobility data and a baseline SEIR model (**c**) that does not use mobility data (see Methods, 'Aggregate mobility and no-mobility baseline models' for details). Incorporating mobility information improves out-of-sample fit and using a mobility network, instead of an aggregate measure of mobility, further improves fit: on average across metro areas, the out-of-sample error (r.m.s.e.) of our best-fit network model was only 58% that of the best-fit aggregate mobility model. All three models are calibrated on observed case counts before 15 April 2020 (vertical black line). The grey crosses represent the daily reported cases; as they tend to have great variability, we also show the smoothed weekly average (orange line). Shaded regions denote the 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations.



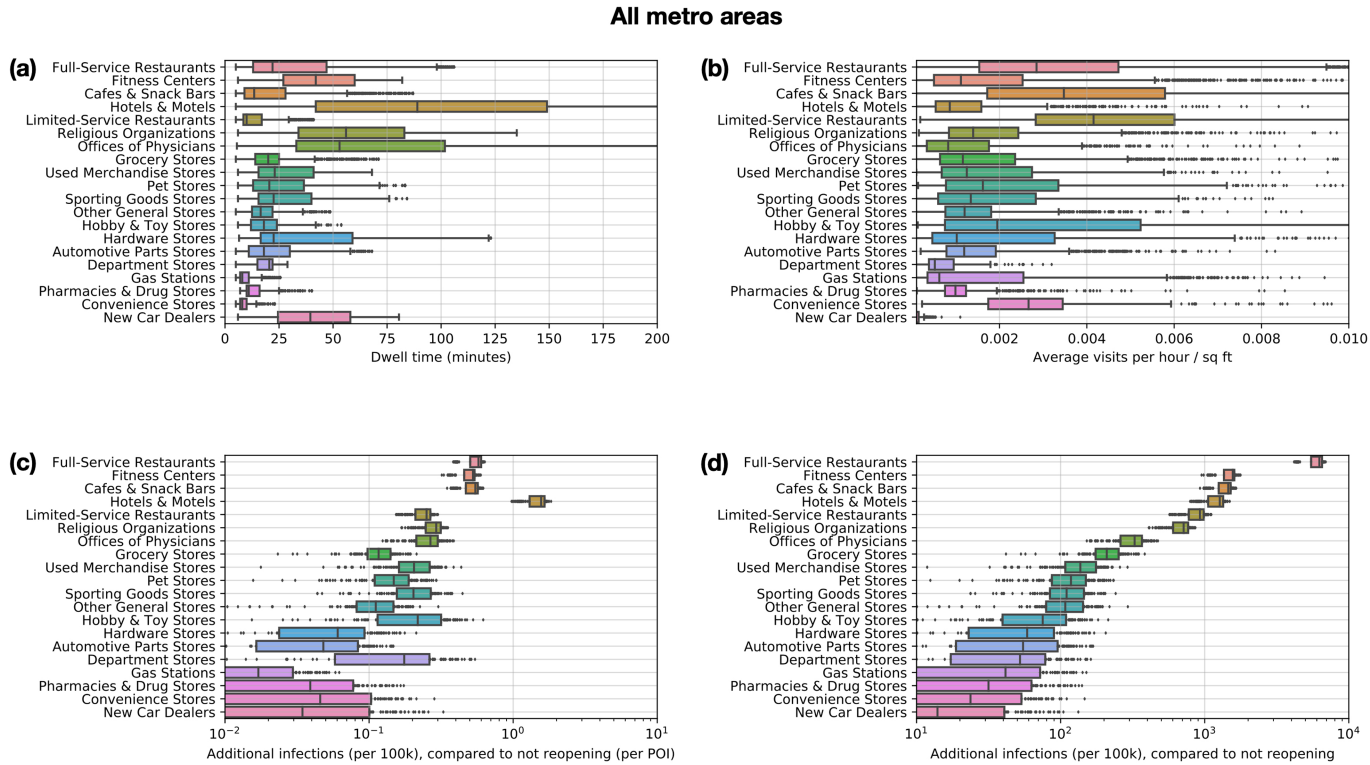
Extended Data Fig. 2 | Distribution of POI infections over time. We selected the POI categories that our models predicted to contribute the most to infections, and plotted the predicted proportion of POI infections that each category accounted for over time. Our model predicts time-dependent variation of where transmissions may have occurred. For example, full-service restaurants (blue) and fitness centres (brown) contributed less to predicted infections over time, probably due to lockdown orders closing these POIs, whereas grocery stores remained steady or even grew in their predicted contribution, probably because they remained open as essential businesses. Hotels and motels (yellow) also feature in these plots; most notably, the model predicts a peak in their contributed infections in Miami around mid-March, which aligns with college spring break, as Miami is a popular vacation spot for students. The proportions are stacked in these plots, and the *y* axes are truncated at 0.7 because every plot would only show 'other' from 0.7 to 1.0.



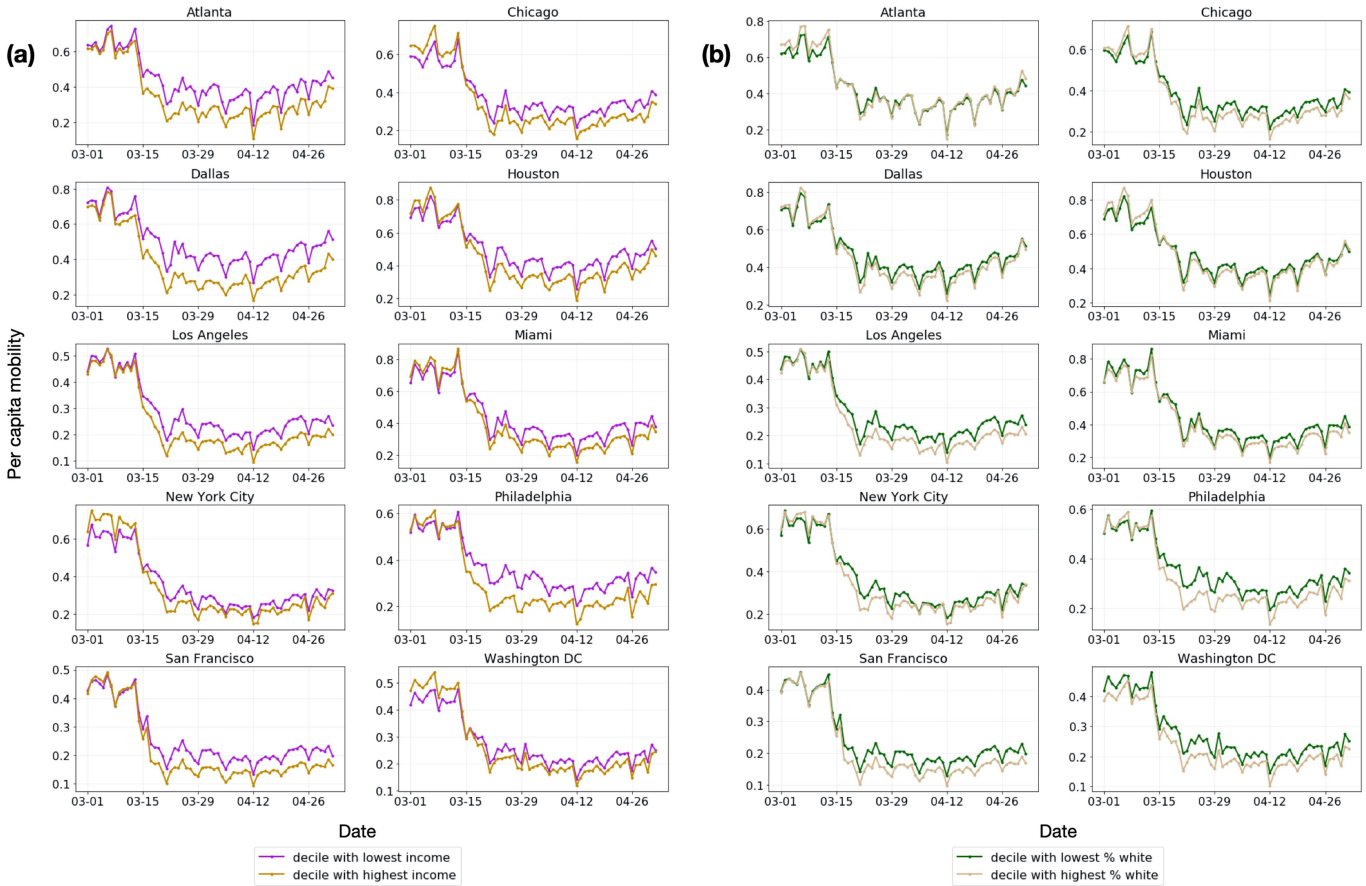
Extended Data Fig. 3 | Trade-off between new infections and visits lost from reopening. We simulate reduced maximum occupancy reopening starting on 1 May 2020 and run the simulation until the end of the month. Each dot represents the level of occupancy reduction: for example, capping visits at 50% of the maximum occupancy. The *y* coordinate represents the predicted number of new infections incurred after reopening (per 100,000 population) and the *x* coordinate represents the fraction of visits lost from partial reopening compared to full reopening. Shaded regions denote the 2.5th and 97.5th percentiles across parameter sets and stochastic realizations. In four metro areas, the predicted cost of new infections from reopening is roughly similar for lower-income CBGs and the overall population, but in five metro areas, the lower-income CBGs incur more predicted infections from reopening. Notably, New York City (NYC) is the only metro area in which this trend is reversed; this is because the model predicts that such a high fraction—65% (95% confidence interval, 62–68%)—of lower-income CBGs in NYC had been infected before reopening that after reopening, only a minority of the lower-income population is still susceptible (in comparison, the second highest fraction infected before reopening was 31% (95% confidence interval, 28–35%) for Philadelphia, and the rest ranged from 1 to 14%).



Extended Data Fig. 4 | Reduced maximum occupancy versus uniform reduction reopening. In comparison to partially reopening by uniformly reducing visits, the reduced maximum occupancy strategy—which disproportionately targets POIs during their most risky high-density periods—always results in a smaller predicted increase in infections for the same number of visits. The *y* axis plots the relative difference between the predicted increase in cumulative infections (from 1 May to 30 May 2020) under the reduced occupancy strategy compared to the uniform reduction strategy. The shaded regions denote the 2.5th and 97.5th percentiles across the sampled parameters and stochastic realizations.



Extended Data Fig. 5 | POI attributes in all 10 metro areas combined. a, **b**, The POIs from all metro areas are pooled and the quantities from the mobility data are shown. **a**, The distribution of dwell time. **b**, The average number of hourly visitors divided by the area of the POI in square feet. Each point represents one POI; boxes depict the interquartile range across POIs; data points outside the range are shown as individual dots. **c**, **d**, The data are pooled across model realizations from all metro areas and model predictions are shown for the increase in infections (per 100,000 population) because of reopening a POI category. **c**, Data per POI. **d**, Data for the category as a whole. Each point represents a model realization; boxes depict the interquartile range across realizations; data points outside the range are shown as individual dots. Across MSAs, we model 552,758 POIs in total, and we sample 97 parameters and 30 stochastic realizations ($n = 2{,}910$); see Supplementary Table 6 for the number of sets per metro area. Colours are used to distinguish the different POI categories, but do not have any additional meaning.



Extended Data Fig. 6 | Daily per-capita mobility over time. a, **b**, We compare mobility in the lowest and highest deciles of CBGs based on median household income (**a**) and the percentage of white residents (**b**). See Methods, 'Analysis details' for details.

Extended Data Table 1 | Dataset summary statistics from 1 March to 2 May 2020

Metro area	CBGs	POIs	Hourly edges	Total modeled pop	Total visits
Atlanta	3,130	39,411	540,166,727	7,455,619	27,669,692
Chicago	6,812	62,420	540,112,026	10,169,539	33,785,702
Dallas	4,877	52,999	752,998,455	9,353,561	37,298,053
Houston	3,345	49,622	609,766,288	7,621,541	32,943,613
Los Angeles	8,904	83,954	643,758,979	16,101,274	38,101,674
Miami	3,555	40,964	487,544,190	6,833,129	26,347,947
New York City	14,763	122,428	1,057,789,207	20,729,481	66,581,080
Philadelphia	4,565	37,951	304,697,220	6,759,058	19,551,138
San Francisco	2,943	28,713	161,575,167	5,137,800	10,728,090
Washington DC	4,051	34,296	312,620,619	7,740,276	17,898,324
All metro areas combined	**56,945**	**552,758**	**5,411,028,878**	**97,901,278**	**310,905,313**

Extended Data Table 2 | Model parameters

Param.	Description	Value (Source)
δ_E	mean latency period	96 hours[24,63]
δ_I	mean infectious period	84 hours[24]
δ_c	period from infectious to confirmed	7 days[24,71]
r_c	percentage of cases which are detected	10%[24,63,69–71]
β_{base}	base CBG transmission rate	Variable (Estimated)
N_{c_i}	population size of CBG c_i	Variable (2018 US Census[57])
ψ	scaling factor for POI transmission	Variable (Estimated)
$w_{ij}^{(t)}$	# visitors from CBG c_i to POI p_j at time t	Variable (SafeGraph)
a_{p_j}	area of POI p_j in square feet	Variable (SafeGraph)
p_0	initial proportion of exposed population	Variable (Estimated)
$S_{c_i}^{(0)}$	initial susceptible population in CBG c_i	$(1-p_0)N_{c_i}$
$E_{c_i}^{(0)}$	initial exposed population in CBG c_i	$p_0 N_{c_i}$
$I_{c_i}^{(0)}$	initial infectious population in CBG c_i	0
$R_{c_i}^{(0)}$	initial removed population in CBG c_i	0

If the parameter has a fixed value, we specify it under 'Value'; otherwise, 'Variable' is used to indicate that it varies across CBG, POI or metro area.

Extended Data Table 3 | Predicted transmission rate disparities at each POI category between income groups

Metro area	ATL	CHI	DAL	HOU	LA	MIA	NY	PHL	SF	DC	**Median**
Full-Service Restaurants	0.764	1.204	0.956	1.000	1.445	1.232	2.035	2.883	1.758	1.171	1.218
Limited-Service Restaurants	0.940	0.950	1.002	0.906	1.067	0.872	1.901	1.614	0.994	0.962	0.978
Other General Stores	0.782	1.083	0.957	0.729	0.760	0.894	1.218	1.312	1.045	0.950	0.954
Gas Stations	1.326	1.865	1.310	1.515	2.254	2.195	1.899	6.461	1.357	1.870	1.868
Fitness Centers	0.536	0.907	0.708	0.670	1.461	0.789	1.151	1.516	0.995	1.160	0.951
Grocery Stores	0.948	3.080	0.838	1.333	2.408	1.498	4.984	10.437	2.478	1.977	2.192
Cafes & Snack Bars	1.385	0.919	0.716	1.120	1.327	2.168	1.943	1.757	0.982	0.932	1.224
Hotels & Motels	1.228	1.200	0.814	0.804	1.229	1.134	1.260	1.993	1.199	1.346	1.214
Religious Organizations	1.546	1.763	0.956	0.919	1.746	1.464	1.756	1.736	1.515	1.852	1.641
Hardware Stores	3.938	3.340	1.575	2.111	1.333	0.939	3.553	6.716	4.202	13.560	3.446
Department Stores	1.132	1.230	0.978	0.911	1.083	1.431	1.667	0.976	0.867	1.042	1.062
Offices of Physicians	1.235	0.721	0.667	1.036	1.141	1.687	1.307	1.319	1.193	0.445	1.167
Pharmacies & Drug Stores	1.636	1.389	1.176	0.854	1.718	1.555	2.577	5.624	1.200	1.699	1.596
Sporting Goods Stores	0.936	1.540	1.129	0.812	1.168	0.700	1.253	1.161	0.826	2.777	1.145
Automotive Parts Stores	0.890	1.707	0.862	1.086	1.990	1.414	1.524	2.697	1.753	1.246	1.469
Used Merchandise Stores	0.993	0.931	1.000	1.315	1.017	1.074	1.352	1.668	1.587	0.814	1.046
Convenience Stores	1.208	0.932	1.613	0.647	0.838	0.824	1.736	2.322	1.086	1.428	1.147
Pet Stores	1.260	0.820	1.192	1.487	1.536	0.776	3.558	1.652	2.124	0.905	1.374
New Car Dealers	2.036	1.471	0.741	0.809	1.180	1.377	2.022	1.129	0.395	0.872	1.154
Hobby & Toy Stores	1.168	1.110	1.165	0.853	1.771	1.520	1.525	1.088	0.883	0.926	1.138
Median	1.188	1.202	0.968	0.915	1.330	1.305	1.746	1.702	1.196	1.166	

We report the ratio of the average predicted transmission rate encountered by visitors from CBGs in the bottom income decile to that for the top income decile. A ratio greater than 1 means that visitors from CBGs in the bottom income decile experienced higher (more dangerous) predicted transmission rates. See Methods, 'Analysis details' for details.

Extended Data Table 4 | Predicted transmission rate disparities at each POI category between racial groups

Metro area	ATL	CHI	DAL	HOU	LA	MIA	NY	PHL	SF	DC	**Median**
Full-Service Restaurants	0.802	1.354	0.981	0.965	1.065	1.167	2.418	2.661	1.223	1.013	1.116
Limited-Service Restaurants	0.940	1.144	1.028	0.940	0.820	0.919	2.136	1.523	0.799	1.346	0.984
Other General Stores	0.776	1.277	0.838	0.841	1.527	1.132	2.158	1.313	0.925	1.312	1.204
Gas Stations	1.402	1.891	1.389	1.190	1.336	1.857	1.818	2.286	2.321	1.316	1.610
Fitness Centers	0.607	1.167	0.670	0.831	0.780	1.066	1.447	1.977	1.103	1.205	1.084
Grocery Stores	0.589	3.664	0.613	1.195	2.386	0.950	5.864	13.705	2.243	2.262	2.252
Cafes & Snack Bars	1.308	1.104	0.845	0.840	0.976	2.619	1.767	2.456	1.045	0.867	1.074
Hotels & Motels	0.977	1.007	1.366	0.718	1.112	1.024	1.449	2.494	0.654	0.899	1.015
Religious Organizations	0.938	1.606	1.060	0.953	2.096	1.795	1.933	2.040	1.674	1.188	1.640
Hardware Stores	0.909	3.900	1.523	1.461	1.952	0.586	5.032	3.898	11.103	13.432	2.925
Department Stores	1.081	1.301	0.805	0.777	0.992	2.337	2.479	1.357	1.089	1.402	1.195
Offices of Physicians	0.894	1.323	1.006	1.415	0.898	1.117	1.652	2.073	0.694	1.911	1.220
Pharmacies & Drug Stores	0.888	1.376	0.930	0.732	1.538	1.674	3.315	3.366	1.135	1.715	1.457
Sporting Goods Stores	0.767	0.674	0.650	0.506	1.946	0.818	1.532	2.152	0.880	1.715	0.849
Automotive Parts Stores	1.049	1.479	1.010	1.353	2.998	2.657	1.740	3.387	1.646	0.601	1.562
Used Merchandise Stores	0.858	1.195	0.699	1.060	1.270	0.593	1.500	3.024	1.425	0.799	1.128
Convenience Stores	2.016	5.055	1.272	2.188	0.761	0.902	1.911	2.276	1.239	1.844	1.878
Pet Stores	0.925	1.624	0.724	1.465	1.506	0.881	2.715	10.182	1.568	2.408	1.537
New Car Dealers	1.008	1.398	0.812	0.736	0.942	0.998	1.977	0.866	0.772	0.383	0.904
Hobby & Toy Stores	2.569	0.853	0.628	0.979	1.373	1.388	2.237	0.825	0.864	1.286	1.132
Median	0.932	1.339	0.888	0.959	1.303	1.092	1.955	2.281	1.119	1.314	

We report the ratio of the average predicted transmission rate encountered by visitors from CBGs with the lowest (bottom decile) proportion of white residents versus that for the top decile. A ratio greater than 1 means that visitors from CBGs in the bottom decile experienced higher (more dangerous) predicted transmission rates. See Methods, 'Analysis details' for details.

nature research

Corresponding author(s): Jure Leskovec

Last updated by author(s): Oct 7, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No data collection was performed for this study; all analysis relied on previously collected datasets, as described in the Data section below.
Data analysis	All data analysis was performed using Python with standard libraries. Code is available at http://snap.stanford.edu/covid-mobility and https://github.com/snap-stanford/covid-mobility.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Census data (https://www.census.gov/programs-surveys/acs), case and death counts from The New York Times (https://github.com/nytimes/covid-19-data), and Google mobility data (https://www.google.com/covid19/mobility/) are publicly available. Cell phone mobility data is freely available to researchers, non-profits, and governments through the SafeGraph COVID-19 Data Consortium (https://www.safegraph.com/covid-19-data-consortium).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a quantitative epidemiological modeling study.
Research sample	We study previously-collected data on 10 of the largest American metropolitan statistical areas. The data came from the following: 1) Anonymized mobility data from smartphone users from SafeGraph. SafeGraph data is obtained by "partnering with mobile applications that obtain opt-in consent from its users to collect anonymous location data", per official SafeGraph documentation. 2) US Census 3) Case and death counts from The New York Times 4) Google mobility data
Sampling strategy	We did not perform sampling, but relied on previously collected datasets. All datasets were chosen as they were comprehensive: the case and death counts and Census data were designed to cover the entire population; Google mobility data covers a large subset of the population; and previous analyses have shown that the SafeGraph anonymized cell phone mobility dataset is geographically representative: for example, it does not systematically over-represent individuals from higher-income areas (https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset).
Data collection	We did not perform data collection, but relied on previously collected datasets. Census data is collected as described here: https://www.census.gov/programs-surveys/acs. The cell phone mobility data is collected from cell phone geolocation information, and is anonymized and aggregated. We accessed and downloaded the datasets in May 2020.
Timing	We make use of data in the following ranges: 1) Safegraph mobility data: Jan 1, 2019 - May 2, 2020 2) US Census: 2013-2018 3) Case and death counts from The New York Times: March 8 - May 9, 2020 4) Google mobility data: March 1 - May 2, 2020
Data exclusions	The original SafeGraph dataset contains 5.4 million points of interest (POIs). We retain a POI in our final dataset if it satisfies the following criteria: (1) it lies within one of the 10 American metropolitan areas that we analyze (out of 384 metropolitan statistical areas total); (2) SafeGraph has visit data for this POI for every hour that we model, from 12am on March 1, 2020 to 11pm on May 2, 2020; (3) SafeGraph has recorded the home CBGs of this POI's visitors for at least one month from January 2019 to February 2020; (4) the POI is not a "parent" POI, as defined in the Methods section. After applying these filters, our dataset contains 553k POIs. Most POIs are filtered out because they do not lie within the 10 large metropolitan statistical areas that we study; this filtering decision was made prior to any analysis for computational tractability reasons. In our analysis of POI-specific category risks, we do not analyze 6 categories of POIs because we wish to be conservative and only focus on categories where we are most confident we are fully capturing transmission at the category: Child Day Care Services, Elementary and Secondary Schools, Drinking Places (Alcoholic Beverages), Nature Parks and Other Similar Institutions, General Medical and Surgical Hospitals, and Other Airport Operations. The justifications for these exclusions, which are based on prior work, are given in the Methods section.
Non-participation	Because we relied on previously collected anonymized, aggregated data from cell phone mobility tracking, we did not have access to individual-level data and do not know how many participants dropped out/declined participation.
Randomization	This is not a randomized controlled trial and participants were not randomized into experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above.
Recruitment	See above. Prior work has looked into biases in the SafeGraph dataset (https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset).
Ethics oversight	The dataset from The New York Times consists of aggregated COVID-19 confirmed case and death counts collected by journalists from public news conferences and public data releases. For the mobility data, consent was obtained by the third-party sources collecting the data. SafeGraph aggregates data from mobile applications that obtain opt-in consent from their users to collect anonymous location data. Google's mobility data consists of aggregated, anonymized sets of data from users who have chosen to turn on the Location History setting. Additionally, we obtained IRB exemption for SafeGraph data from the Northwestern University IRB office.

Note that full information on the approval of the study protocol must also be provided in the manuscript.