Mobility network modeling explains higher SARS-CoV-2 infection rates among disadvantaged groups and informs reopening strategies

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

[1] Department of Computer Science, Stanford University, Stanford, CA 94305, USA
[2] Department of Preventive Medicine, Northwestern University, Chicago, IL 60611, USA
[3] Department of Sociology, Northwestern University, Evanston, IL, 60208, USA
[4] Institute for Policy Research, Northwestern University, Evanston, IL, 60208, USA
[5] Department of Sociology, Stanford University, Stanford, CA 94305, USA
[6] Center on Poverty and Inequality, Stanford University, Stanford, CA 94305, USA
[7] Chan Zuckerberg Biohub, San Francisco, CA 94158, USA
[*] These authors contributed equally to this work
[†] Corresponding author. Email: jure@cs.stanford.edu

Fine-grained epidemiological modeling of the spread of SARS-CoV-2—capturing who is infected at which locations—can aid the development of policy responses that account for heterogeneous risks of different locations as well as the disparities in infections among different demographic groups. Here, we develop a metapopulation SEIR disease model that uses dynamic mobility networks, derived from US cell phone data, to capture the hourly movements of millions of people from local neighborhoods (census block groups, or CBGs) to points of interest (POIs) such as restaurants, grocery stores, or religious establishments. We simulate the spread of SARS-CoV-2 from March 1–May 2, 2020 among a population of 105 million people in 10 of the largest US metropolitan statistical areas. We show that by integrating these mobility networks, which connect 60k CBGs to 565k POIs with a total of 5.4 billion hourly edges, even a relatively simple epidemiological model can accurately capture the case trajectory despite dramatic changes in population behavior due to the virus. Furthermore, by modeling detailed information about each POI, like visitor density and visit length, we can estimate the impacts of fine-grained reopening plans: we predict that a small minority of "superspreader" POIs account for a large majority of infections, that reopening some POI categories (like full-service restaurants) poses especially large risks, and that strategies restricting maximum occupancy at each POI are more effective than uniformly reducing mobility. Our models also predict higher infection rates among disadvantaged racial and socioeconomic groups solely from differences in mobility: disadvantaged groups have not been able to reduce mobility as sharply, and the POIs they visit (even within the same category) tend to be smaller, more crowded, and therefore more dangerous. By modeling who is infected at which locations, our model supports fine-grained analyses that can inform more effective and equitable policy responses to SARS-CoV-2.

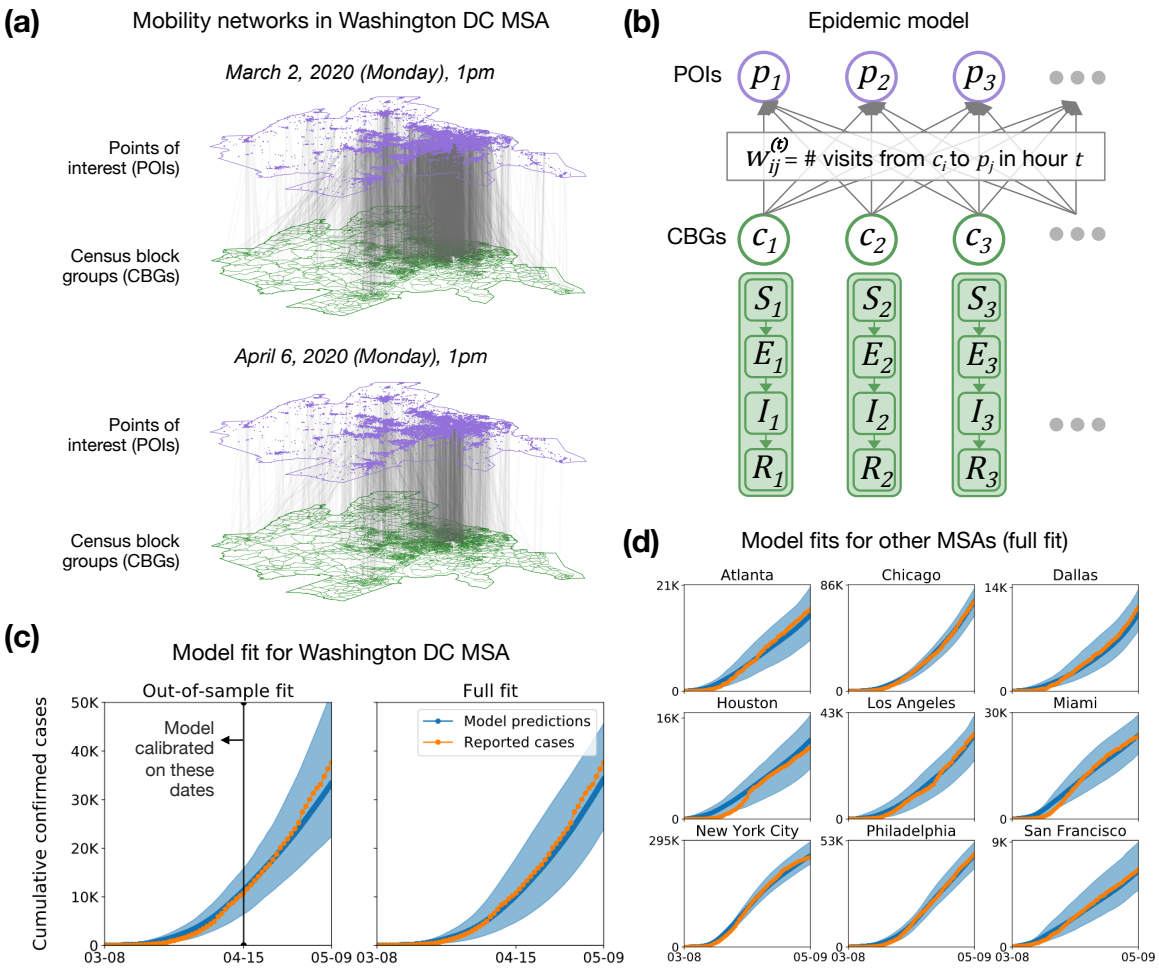


Figure 1: Model description and fit. **(a)** The mobility network captures hourly visits from each census block group (CBG) to each point of interest (POI). The vertical lines indicate that most visits are between nearby POIs and CBGs. Visits dropped dramatically from March (top) to April (bottom), as indicated by the lower density of grey lines. **(b)** We overlaid an SEIR disease model on the mobility network, with each CBG having its own set of SEIR compartments. New infections occur at both POIs and CBGs. The model has just three free parameters, which remain fixed over time, scaling transmission rates at POIs; transmission rates at CBGs; and the initial fraction of infected individuals. To determine the transmission rate at a given time at each POI we use the mobility network, which captures population movements as well as visit duration and the POI physical area, to estimate the density of visitors at each POI. **(c)** Left: To test out-of-sample prediction, we calibrated the model on data before April 15, 2020 (vertical black line). Even though its parameters remain fixed over time, the model accurately predicts the case trajectory after April 15 by using mobility data. Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. Right: Model fit further improved when we calibrated the model on the full range of data. **(d)** We fit separate models to 10 of the largest US metropolitan statistical areas (MSAs), modeling a total population of 105 million people; here, we show full model fits, as in (c)-Right. While we use Washington DC as a running example throughout the paper, we include results for all other MSAs in the SI.

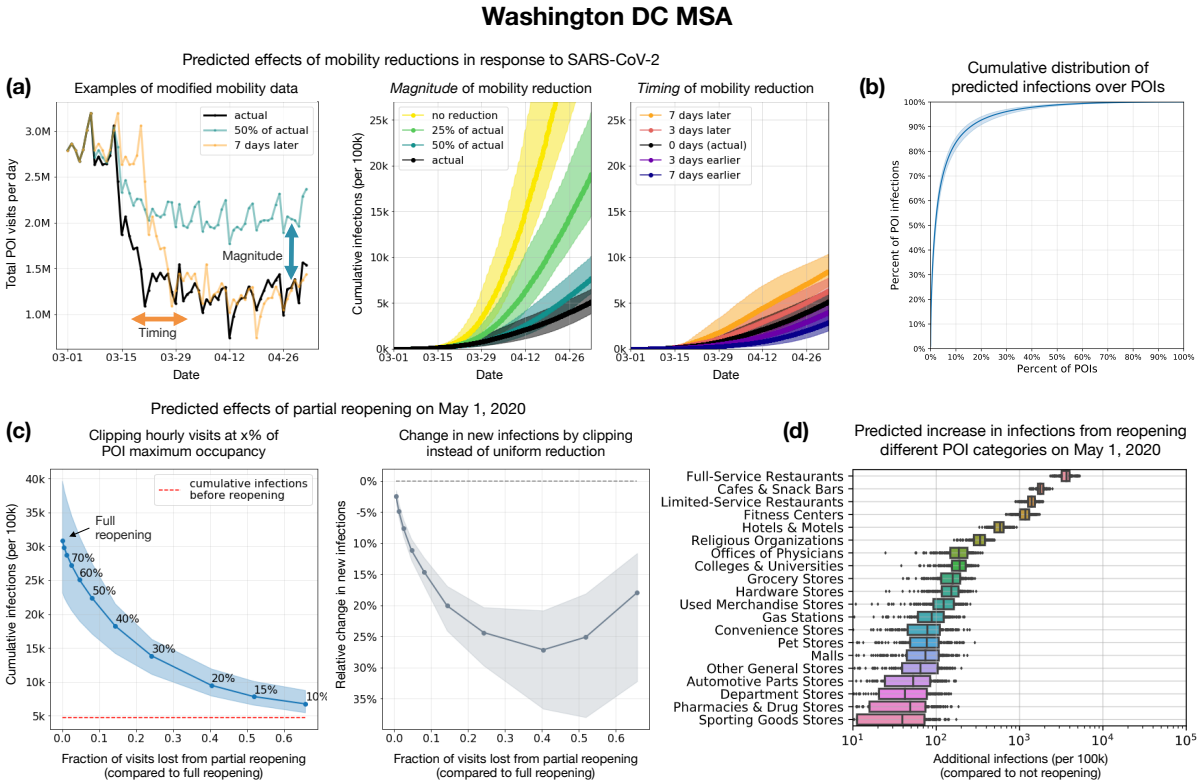


Figure 2: Assessing mobility reduction and reopening policies. **(a)** Counterfactual simulations (left) of the mobility reduction in March 2020—scaling its magnitude down, or shifting the timeline earlier or later—illustrate that the magnitude of mobility reduction (middle) was at least as important as its timing (right). Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. **(b)** Most infections at POIs occur at a small fraction of "super-spreader" POIs: 10% of POIs account for more than 80% of the total infections that occurred at POIs in the Washington DC MSA (results for other MSAs in Extended Data Figure 3). **(c)** Left: We simulated partial reopening by clipping hourly visits if they exceeded a fraction of each POI's maximum occupancy. We plot cumulative infections at the end of one month of reopening against the fraction of visits lost by partial instead of full reopening; the annotations within the plot show the fraction of maximum occupancy used for clipping. Reopening leads to an additional 26% of the population becoming infected by the end of the month, but clipping at 20% maximum occupancy cuts down new infections by more than 80%, while only losing 40% of overall visits. Right: Compared to partially reopening by uniformly reducing visits, the clipping strategy—which disproportionately targets high-risk POIs with sustained high occupancy—always results in a smaller increase in infections for the same number of visits. The y-axis plots the relative difference between the increase in cumulative infections (from May 1 to May 31) under the clipping strategy as compared to the uniform reduction strategy. **(d)** We simulated reopening each POI category while keeping reduced mobility levels at all other POIs. Boxes indicate the interquartile range across parameter sets and stochastic realizations. 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.

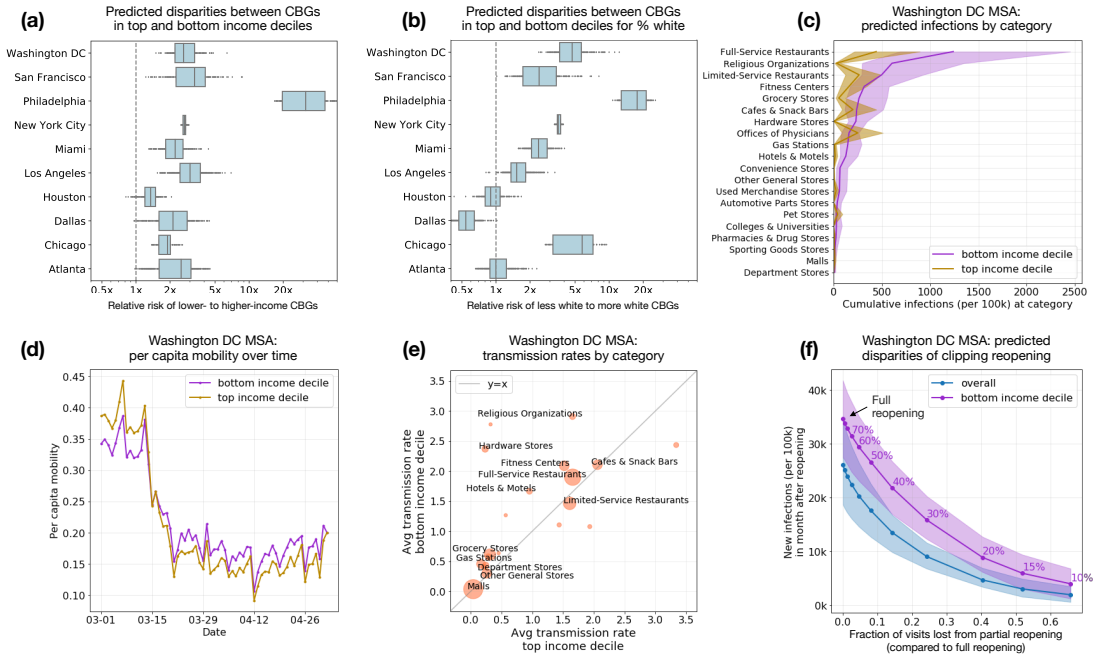


Figure 3: Mobility patterns give rise to socioeconomic and racial disparities in infections. **(a)** Across all MSAs, our model predicts that people in lower-income census block groups (CBGs) are more likely to be infected, even though they start with equal probabilities of being infected. Disparities are especially prominent in Philadelphia, which we discuss in SI Section S2. Boxes indicate the interquartile range across parameter sets and stochastic realizations. **(b)** Racial disparities are similar: people in non-white CBGs are typically more likely to be infected, although results are more variable. (c-e) illustrate how mobility patterns give rise to socioeconomic disparities; similar mechanisms underlie racial disparities (Extended Data Figure 6, Table S4). **(c)** The overall disparity is driven by a few POI categories like full-service restaurants. Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. **(d)** One reason for the disparities is that higher-income CBGs were able to reduce their overall mobility levels below those of lower-income CBGs. **(e)** Within each category, the POIs that people from lower-income CBGs visit also tend to have higher transmission rates because they are smaller and more crowded. Thus, even if a lower-income and a higher-income person went out equally often and went to the same types of places, the lower-income person would still have a greater risk of infection. The size of each dot indicates the total number of visits to that category. **(f)** We predict the effect of reopening (at different levels of clipping maximum occupancy) on different demographic groups. Reopening leads to more infections in lower-income CBGs (purple) than the overall population (blue), underscoring the need to account for disadvantaged subpopulations when assessing reopening plans.

Introduction

In response to the SARS-CoV-2 crisis, numerous stay-at-home orders were enacted across the United States in order to reduce contact between individuals and slow the spread of the virus.[1] As of May 2020, these orders are being relaxed, businesses are beginning to reopen, and mobility is increasing, causing concern among public officials about the potential resurgence of cases.[2] Epidemiological models that can capture the effects of changes in mobility on virus spread are a powerful tool for evaluating the effectiveness and equity of various strategies for reopening or responding to a resurgence. In particular, findings of SARS-CoV-2 "super-spreader" events[3–7] motivate models that can reflect the heterogeneous risks of visiting different locations, while well-reported racial and socioeconomic disparities in infection rates[8–14] require models that can explain the disproportionate impact of the virus on disadvantaged demographic groups.

To address these needs, we construct a mobility network using US cell phone data from March 1–May 2, 2020 that captures the hourly movements of millions of people from census block groups (CBGs), which are geographical units that typically contain 600–3,000 people, to points of interest (POIs) such as restaurants, grocery stores, or religious establishments. On top of this dynamic bipartite network, we overlay a metapopulation SEIR disease model that tracks the infection trajectories of each CBG over time as well as the POIs at which these infections are likely to have occurred. The key idea is that combining even a relatively simple epidemiological model with our fine-grained, dynamic mobility network allows us to not only accurately model the case trajectory, but also identify the most risky POIs; the most at-risk populations; and the impacts of different reopening policies. This builds upon prior work that models disease spread using mobility data, which has used aggregate[15–21], historical[22–24], or synthetic[25–27] mobility data; separately, other work has directly analyzed mobility data and the effects of mobility reductions in the context of SARS-CoV-2, but without an underlying epidemiological model of disease spread.[28–33]

We use our model to simulate the spread of SARS-CoV-2 within 10 of the largest metropolitan statistical areas (MSAs) in the US, starting from a low, homogeneous prevalence of SARS-CoV-2 across CBGs. For each MSA, we examine the infection risks at individual POIs, the effects of past stay-at-home policies, and the effects of reopening strategies that target specific types of POIs. We also analyze disparities in infection rates across racial and socioeconomic groups, identify mobility-related mechanisms driving these disparities, and assess the disparate impacts of reopening policies on disadvantaged groups.

Results

Mobility network modeling

Mobility network. We study mobility patterns from March 1–May 2, 2020 among a population of 105 million people in 10 of the largest US metropolitan statistical areas (MSAs). For each MSA, we represent the movement of individuals between census block groups (CBGs) and points of interest (POIs) as a bipartite network with time-varying edges, where the weight of an edge between a CBG and POI is the number of visitors from that CBG to that POI at a given hour (Figure 1a). We use iterative proportional fitting[34] to derive these networks from geolocation data from SafeGraph, a data company that aggregates anonymized location data from mobile applications. We validate the SafeGraph data by comparing to Google mobility data (SI Section S1). Overall, these networks comprise 5.4 billion hourly edges between 59,519 CBGs and 565,286 POIs (Extended Data Table 1).

Model. We overlay a SEIR disease model on each mobility network,[15,22] where each CBG maintains its own susceptible (S), exposed (E), infectious (I), and removed (R) states (Figure 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 inferred density of infectious individuals at each POI to determine its transmission rate. The model has only three free parameters, which scale (1) transmission rates at POIs, (2) transmission rates at CBGs, and (3) the initial proportion of infected individuals. All three parameters remain constant over time. We calibrate a separate model for each MSA using confirmed case counts from the *The New York Times*.[35]

Model validation. We validated our models by showing that they can predict out-of-sample case and death counts, i.e., on a held-out time period not used for model calibration. Specifically, we calibrated models for each MSA on case counts from March 8–April 14, 2020 and evaluated them on case and death counts from April 15–May 9, 2020 (these dates are offset by a week from the mobility data to account for the delay between infection and case confirmation). Our key technical result is that even with a relatively simple SEIR model with three free parameters, the mobility networks allow us to accurately model out-of-sample cases (Figure 1c and Extended Data Figure 1a) and deaths (Extended Data Figure 2a) without needing to directly incorporate information about the case trajectory or social distancing measures. In contrast, a baseline SEIR

model that does not use the mobility network has considerably worse out-of-sample fit (Extended Data Figures 1b and 2b). All subsequent results were generated using the models calibrated on the entire range of case counts from March 8–May 9, 2020.

Evaluating mobility reduction and reopening policies

We can estimate the impact of a wide range of mobility reduction and reopening policies by applying our model to a modified mobility network that reflects the expected effects of a hypothetical policy. We start by studying the effect of the magnitude and timing of mobility reduction policies from March 2020. We then assess several fine-grained reopening plans, such as placing a maximum occupancy cap or only reopening certain categories of POIs, by leveraging the detailed information that the mobility network contains on each POI, like its average visit length and visitor density at each hour.

The magnitude of mobility reduction is as important as its timing. US population mobility dropped sharply in March 2020 in response to SARS-CoV-2; for example, overall mobility in the Washington DC MSA fell by 58.5% 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 of this mobility reduction earlier and later (Figure 2a), and used our model to simulate the resulting infection trajectories. As expected, shifting the onset of mobility reduction earlier decreased the predicted number of infections incurred, and shifting it later or reducing the magnitude of reduction both increased predicted infections. What was notable was that reducing the magnitude of reduction resulted in far larger increases in predicted infections than shifting the timeline later (Figure 2a). For example, if only a quarter of mobility reduction had occurred in the DC MSA, the predicted number of infections would have increased by $3\times$, compared to a less than $2\times$ increase had people begun reducing their mobility one full week later. We observe similar trends across other MSAs (Tables S1 and S2).

A minority of POIs account for a majority of infections. Since overall mobility reduction reduces infections, we next investigated if *how* we reduce mobility—i.e., to which POIs—matters. Using the observed mobility networks to simulate the infection trajectory from March 1–May 2, 2020, we found that a majority of predicted infections occurred at a small fraction of "super-spreader" POIs; e.g., in the DC MSA, 10% of POIs account for more than 80% of the predicted

infections at POIs (Figure 2b; Extended Data Figure 3 shows similar results across MSAs). Note that infections at POIs represent a majority, but not all, of the total infections, since we also model infections within CBGs; across MSAs, the median proportion of total infections that occur at POIs is 73%. These "superspreader" POIs are smaller and more densely occupied, and their occupants stay longer, suggesting that it is especially important to reduce mobility at these high-risk POIs. In the DC MSA, the median number of hourly visitors per square foot was 4.6× higher for the riskiest 10% of POIs than for the remaining POIs; the median dwell time was 2.3× higher.

Reducing mobility by clipping maximum occupancy. We simulated the effects of two reopening strategies, implemented beginning on May 1, on the increase in infections by the end of May. First, we evaluated a "clipping" reopening strategy, in which hourly visits to each POI return to those in the first week of March (prior to widespread adoption of stay-at-home measures), but are capped if they exceed a fraction of the POI's maximum occupancy,[36] which we estimated as the maximum hourly number of visitors ever recorded at that POI. A full return to early March mobility levels without clipping produces a spike in predicted infections: in the DC MSA, we project that an additional 26% of the population will be infected within a month (Figure 2c). However, clipping substantially reduces risk without sharply reducing overall mobility: clipping at 20% maximum occupancy in the DC MSA cuts down new infections by more than 80% but only loses 40% of overall visits, and we observe similar trends across other MSAs (Extended Data Figure 4). This highlights the non-linearity of infections as a function of visits: one can achieve a disproportionately large reduction in infections with a small reduction in visits.

We also compared the clipping strategy to a baseline that uniformly reduces visits to each POI from their levels in early March. Clipping always results in fewer infections for the same total number of visits: e.g., clipping at 20% maximum occupancy reduces new infections by more than 25% compared to the uniform baseline for the same total number of visits in the Washington DC MSA (Figure 2c). This is because clipping takes advantage of the heterogeneous risks across POIs, disproportionately reducing visits at high-risk POIs with sustained high occupancy, but allowing lower-risk POIs to return fully to prior mobility levels.

Relative risk of reopening different categories of POIs. We assessed the relative risk of reopening different categories of POIs by reopening each category in turn on May 1 (and returning its mobility patterns to early March levels) while keeping mobility patterns at all other POIs at

their reduced, stay-at-home levels (Figure 2d). We find a large variation in reopening risks: on average across the 10 evaluated MSAs (Extended Data Figure 5), full-service restaurants, cafes, gyms, limited-service restaurants, and religious establishments produce the largest increases in infections when reopened. Reopening full-service restaurants is particularly risky: in the Washington DC MSA, we predict an additional 296k infections by the end of May, more than double the next riskiest POI category. These risks are the total risks summed over all POIs in the category, but the relative risks after normalizing by the number of POIs are broadly similar, with restaurants, gyms, cafes, and religious establishments predicted to be the most dangerous on average per individual POI. These categories are more dangerous because their POIs tend to have higher visit densities and/or visitors stay there longer (Figures S4–S13).

Infection disparities between socioeconomic and racial groups

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 how infection disparities arise across groups. We use this approach to study the mobility mechanisms behind disparities and to quantify how different reopening strategies impact disadvantaged groups.

Mobility patterns contribute to disparities in infection rates. Despite only having access to mobility data and no other demographic information, our models correctly predicted higher risks of infection among disadvantaged racial and socioeconomic groups.[8–14] Across all MSAs, individuals from CBGs in the bottom decile for income were substantially likelier to have been infected by the end of the simulation, even though all individuals began with equal likelihoods of infection in our simulation (Figure 3a). This overall disparity was driven primarily by a few POI categories (e.g., full-service restaurants), which infected far larger proportions of lower-income CBGs than higher-income CBGs (Figure 3c; similar trends hold across all MSAs in Figure S1). We similarly found that CBGs with fewer white residents had higher relative risks of infection, although results were more variable (Figure 3b). Our models also recapitulated known associations between population density and infection risk[37] (median Spearman correlation between CBG density and cumulative incidence proportion, 0.42 across MSAs), despite not being given any information on population density. In SI Section S2, we confirm that the magnitude of the disparities our model predicts are generally consistent with real-world disparities and explore the large predicted disparities in Philadelphia, which stem from substantial differences in density that correlate with income and

race. In the analysis below, we focus on the mechanisms producing higher relative risks of infection among lower-income CBGs, and we show in Extended Data Figure 6 and Table S4 that similar results hold for racial disparities as well.

Lower-income CBGs saw smaller reductions in mobility. Across all MSAs, we found that lower-income CBGs were not able to reduce their mobility as sharply in the first few weeks of March 2020, and had higher mobility than higher-income CBGs for most of March through May (Figure 3d, Extended Data Figure 6). For example, over the month of April, lower-income CBGs in the Washington DC MSA had 17% more visits per capita than higher-income CBGs. Differences in mobility patterns within categories partially explained the within-category infection disparities: e.g., lower-income CBGs made substantially more visits per capita to full-service restaurants than did higher-income CBGs, and consequently experienced more infections at that category (Extended Data Figure 7).

POIs visited by lower-income CBGs tend to be more dangerous. Differences in the number of visits per capita between lower- and higher-income CBGs do not fully explain the infection disparities: for example, in the DC MSA, grocery stores were visited more often by higher-income CBGs but still caused more predicted infections among lower-income CBGs. We found that even within a POI category, the transmission rate at POIs frequented by people from lower-income CBGs tended to be higher than the corresponding rate for higher-income CBGs (Figure 3e; Table S3), because these POIs tended to be smaller and more crowded. It follows that, even if a lower-income and higher-income person had the same mobility patterns and went to the same types of places, the lower-income person would still have a greater risk of infection.

As a case study, we examine grocery stores in further detail. Across all MSAs but Dallas, visitors from lower-income CBGs encountered more dangerous grocery stores than those from higher-income CBGs (median transmission rate ratio of 2.11, Table S3). Why was one visit to the grocery store twice as dangerous for a lower-income individual? Taking medians across MSAs, we found that the average grocery store visited by lower-income individuals had 45% more hourly visitors per square foot, and their visitors stayed 27% longer on average. These findings highlight how fine-grained differences in mobility patterns—how often people go out, which categories of places they go to, which POIs they choose within those categories—can ultimately contribute to dramatic disparities in infection outcomes.

Reopening plans must account for disparate impact. Because disadvantaged groups suffer a larger burden of infection, it is critical to not just consider the overall impact of reopening plans but also their disparate impact on disadvantaged groups specifically. For example, our model predicted that full reopening in the Washington DC MSA would result in an additional 35% of the population of CBGs in the bottom income decile being infected within a month, compared to 26% of the overall population (Figure 3f; results for all MSAs in Extended Data Figure 4). Similarly, Extended Data Figure 8 illustrates that reopening individual POI categories tends to have a larger impact on the bottom income decile. More conservative reopening plans produce smaller absolute disparities in infections—e.g., we predict that clipping visits at 20% occupancy would result in infections among an additional 4% of the overall population and 9% of CBGs in the bottom income decile (Figure 3f)—though the relative disparity remains.

Discussion

We model the spread of SARS-CoV-2 using a dynamic mobility network that encodes the hourly movements of millions of people between 60k neighborhoods (census block groups, or CBGs) and 565k points of interest (POIs). Because our data contains detailed information on each POI, like visit length and visitor density, we can estimate the impacts of fine-grained reopening plans—predicting that a small minority of "superspreader" POIs account for a large majority of infections, that reopening some POI categories (like full-service restaurants) poses especially large risks, and that strategies that restrict the maximum occupancy at each POI are more effective than uniformly reducing mobility. Because we model infections in each CBG, we can infer the approximate demographics of the infected population, and thereby assess the disparate socioeconomic and racial impacts of SARS-CoV-2. Our model correctly predicts that disadvantaged groups are more likely to become infected, and also illuminates two mechanisms that drive these disparities: (1) disadvantaged groups have not been able to reduce their mobility as dramatically (consistent with previously-reported data, and likely in part because lower-income individuals are more likely to have to leave their homes to work[10]) and (2) when they do go out, they visit POIs which, even within the same category, are smaller, more crowded, and therefore more dangerous.

The cell phone mobility dataset we use has limitations: it does not cover all populations (e.g., prisoners, young children), does not contain all POIs (e.g., nursing homes), and cannot capture sub-CBG heterogeneity in demographics. These limitations notwithstanding, cell phone mobility data in general and SafeGraph data in particular have been instrumental and widely used in

modeling SARS-CoV-2 spread.[15–17,28–32,38] Our model itself is parsimonious, and does not include such relevant features as asymptomatic transmission, variation in household size, travel between MSAs, differentials in susceptibility (due to pre-existing conditions or access to care), various transmission-reducing behaviors (e.g., hand-washing, mask-wearing), as well as POI-specific risk factors (e.g., ventilation). Although our model recovers case trajectories and known infection disparities even without incorporating these processes, we caution that this predictive accuracy does not mean that our predictions should be interpreted in a narrow causal sense, and that it is important to recognize that certain types of POIs or subpopulations may disproportionately select for certain types of omitted processes. 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—e.g., that lower-income CBGs have higher infection rates in part because they have not been able to reduce mobility by as much, and because they tend to visit smaller, denser POIs—to hold robustly.

Our results can guide policymakers seeking to assess competing approaches to reopening and tamping down post-reopening resurgence. Despite growing concern about racial and socioeconomic disparities in infections and deaths, it has been difficult for policymakers to act on those concerns; they are currently operating without much evidence on the disparate impacts of reopening policies, prompting calls for research which both identifies the causes of observed disparities and suggests policy approaches to mitigate them.[11,14,39,40] Our fine-grained mobility modeling 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, like disparities in preexisting conditions; on the contrary, short-term policy decisions substantially affect infection disparities by altering the overall amount of mobility allowed, the types of POIs reopened, and the extent to which POI occupancies are clipped. Considering the disparate impact of reopening plans may lead policymakers to, e.g., (1) favor more conservative reopening plans, (2) increase testing in disadvantaged neighborhoods predicted to be high risk (especially given known disparities in access to tests[8]), and (3) prioritize distributing masks and other personal protective equipment to disadvantaged populations that cannot reduce their mobility as much and must frequent riskier POIs.

As society reopens and we face the possibility of a resurgence in cases, it is critical to build models which allow for fine-grained assessments of the effects of reopening policies. We hope that our approach, by capturing heterogeneity across POIs, demographic groups, and cities, helps address this need.

References

1. Wu, J. W., Smith, S., Khurana, M., Siemaszko, C. & DeJesus-Banos, B. Stay-at-home orders across the country. *NBC News* (2020). Available at https://www.nbcnews.com/health/health-news/here-are-stay-home-orders-across-country-n1168736.

2. Rojas, R. & Delkic, M. As states reopen, governors balance existing risks with new ones. *The New York Times* (2020). Available at https://www.nytimes.com/2020/05/17/us/coronavirus-states-reopen.html.

3. Endo, A., Abbott, S., Kucharski, A. J., Funk, S. *et al.* Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Research* **5**, 67 (2020).

4. Adam, D. *et al.* Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong. Available at https://europepmc.org/article/ppr/ppr165671.

5. Miller, D. *et al.* Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel. *medRxiv* (2020). Available at https://www.medrxiv.org/content/10.1101/2020.05.21.20104521v1.full.pdf.

6. Park, S. Y. *et al.* Coronavirus Disease Outbreak in Call Center, South Korea. *Emerging Infectious Diseases* **26** (2020).

7. 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. *The Lancet Infectious Diseases* (2020).

8. Wilson, C. These Graphs Show How COVID-19 Is Ravaging New York City's Low-Income Neighborhoods. *Time* (2020). Available at https://time.com/5821212/coronavirus-low-income-communities/.

9. Garg *et al.* Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 — COVID-NET, 14 States, March 1—30, 2020 (CDC Morbidity and Mortality Weekly Report (MMWR), 2020). Available at https://www.cdc.gov/mmwr/volumes/69/wr/mm6915e3.htm.

10. Reeves, R. V. & Rothwell, J. Class and COVID: How the less affluent face double risks. *The Brookings Institution* (2020). Available at https://www.brookings.edu/blog/up-front/2020/03/27/class-and-covid-how-the-less-affluent-face-double-risks/.

11. Pareek, M. *et al.* Ethnicity and COVID-19: an urgent public health research priority. *The Lancet* **395**, 1421–1422 (2020).

12. van Dorn, A., Cooney, R. E. & Sabin, M. L. COVID-19 exacerbating inequalities in the US. *The Lancet* **395**, 1243–1244 (2020).

13. Yancy, C. W. COVID-19 and African Americans. *JAMA* **323**, 1891–1892 (2020).

14. Chowkwanyun, M. & Reed Jr, A. L. Racial Health Disparities and Covid-19—Caution and Context. *New England Journal of Medicine* (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. Woody, S. *et al.* Projections for first-wave COVID-19 deaths across the US using social-distancing measures derived from mobile phones. *medRxiv* (2020). Available at doi.org/10.1101/2020.04.16.20068163.

17. Fenichel, E. P., Berry, K., Bayham, J. & Gonsalves, G. A cell phone data driven time use analysis of the COVID-19 epidemic. *medRxiv* (2020). Available at doi.org/10.1101/2020.04.20.20073098.

18. Unwin, H. *et al.* Report 23: State-level tracking of COVID-19 in the United States (2020). Available at spiral.imperial.ac.uk/handle/10044/1/79231.

19. Jia, J. S. *et al.* Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* (2020).

20. Pei, S., Kandula, S. & Shaman, J. Differential Effects of Intervention Timing on COVID-19 Spread in the United States. *medRxiv* (2020). Available at doi.org/10.1101/2020.05.15.20103655.

21. Lai, S. *et al.* Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* (2020).

22. Li, R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* **368**, 489–493 (2020).

23. Pei, S. & Shaman, J. Initial Simulation of SARS-CoV2 Spread and Intervention Effects in the Continental US. *medRxiv* (2020). Available at doi.org/10.1101/2020.03.21.20040303.

24. Aleta, A. *et al.* Modeling the impact of social distancing, testing, contact tracing and household quarantine on second-wave scenarios of the COVID-19 epidemic. *medRxiv* (2020). Available at doi.org/10.1101/2020.05.06.20092841.

25. Duque, D. *et al.* COVID-19: How to Relax Social Distancing If You Must. *medRxiv* (2020). Available at doi.org/10.1101/2020.04.29.20085134.

26. Block, P. *et al.* Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour* (2020).

27. Karin, O. *et al.* Adaptive cyclic exit strategies from lockdown to suppress COVID-19 and allow economic activity. *medRxiv* (2020). Available at doi.org/10.1101/2020.04.04.20053579.

28. Gao, S., Rao, J., Kang, Y., Liang, Y. & Kruse, J. Mapping county-level mobility pattern changes in the united states in response to covid-19. *SIGSPATIAL Special* **12**, 16–26 (2020).

29. Klein, B. *et al.* Assessing changes in commuting and individual mobility in major metropolitan areas in the United States during the COVID-19 outbreak (2020). Available at networkscienceinstitute.org/publications/assessing-changes-in-commuting-and-individual-mobility-in-major-metropolitan-areas-in-the-united-states-during-the-covid-19-outbreak.

30. Benzell, S. G., Collis, A. & Nicolaides, C. Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations. *Proceedings of the National Academy of Sciences* (2020).

31. Galeazzi, A. *et al.* Human Mobility in Response to COVID-19 in France, Italy and UK (2020). Available at arxiv.org/abs/2005.06341.

32. Baicker, K., Dube, O., Mullainathan, S., Devin, P. & Wezerek, G. Is It Safer to Visit a Coffee Shop or a Gym? *The New York Times* (2020). Available at https://nytimes.com/interactive/2020/05/06/opinion/coronavirus-us-reopen.html.

33. Hsiang, S. *et al.* The effect of large-scale anti-contagion policies on the coronavirus (COVID-19) pandemic. *Nature* (2020).

34. Deming, W. E. & Stephan, F. F. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics* **11**, 427–444 (1940).

35. The New York Times. Coronavirus (Covid-19) Data in the United States (2020). Available at https://github.com/nytimes/covid-19-data.

36. California Department of Public Health. COVID-19 Industry Guidance: Retail (2020). Available at https://covid19.ca.gov/pdf/guidance-retail.pdf.

37. Rosenthal, B. M. Density Is New York City's Big 'Enemy' in the Coronavirus Fight. *The New York Times* (2020). Available at https://nytimes.com/2020/03/23/nyregion/coronavirus-nyc-crowds-density.html.

38. Buckee, C. O. *et al.* Aggregated mobility data could help fight COVID-19. *Science* **368**, 145 (2020).

39. Webb Hooper, M., Nápoles, A. M. & Pérez-Stable, E. J. COVID-19 and Racial/Ethnic Disparities. *JAMA* (2020).

40. Laurencin, C. T. & McClinton, A. The COVID-19 Pandemic: a Call to Action to Identify and Address Racial and Ethnic Disparities. *Journal of Racial and Ethnic Health Disparities* **7**, 398–402 (2020).

41. Hu, H., Nigmatulina, K. & Eckhoff, P. The scaling of contact rates with population density for the infectious disease models. *Mathematical biosciences* **244**, 125–134 (2013).

42. Kucharski, A. J. *et al.* Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases* **20**, 553 – 558 (2020).

Methods

The methods section is structured as follows. We describe the datasets we use in Methods M1 and the mobility network that we derive from these datasets in Methods M2. In Methods M3, we discuss the SEIR model we overlay on the mobility network, and in Methods M4, we describe how we calibrate this model and quantify uncertainty in its predictions. In Methods M5, we provide details on the experimental procedures used for our analysis of physical distancing, reopening, and demographic disparities. Finally, in Methods M6, we elaborate on how we estimate the mobility network from the raw mobility data.

M1 Datasets

SafeGraph. We use geolocation data provided by SafeGraph, a data company that aggregates anonymized location data from numerous applications. SafeGraph data captures the movement of people between census block groups (CBGs), which are geographical units that typically contain a population of between 600 and 3,000 people, and points of interest (POIs) like restaurants, grocery stores, or religious establishments. Specifically, we use the following SafeGraph datasets:

1. Places Patterns[43] and Weekly Patterns (v1)[44], which 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 visitors' home CBGs. For privacy reasons, SafeGraph excludes a home CBG if too few devices were recorded at the POI from that CBG. For each POI, SafeGraph also provides their North American Industry Classification System (NAICS) category, and an estimate of their physical area in square feet. We analyze Places Patterns data from January 1, 2019 to February 29, 2020 and Weekly Patterns data from March 1, 2020 to May 2, 2020.

2. Social Distancing Metrics,[45] which contains hourly estimates of the proportion of people staying home in each CBG. We analyze Social Distancing Metrics data from March 1, 2020 to May 2, 2020.

We focus on 10 of the largest metropolitan statistical areas (MSAs) in the US (Extended Data Table 1). We chose these MSAs by taking a random subset of the SafeGraph Patterns data and picking the 10 MSAs with the most POIs in the data. Our methods in this paper can be straightforwardly applied, in principle, to the other MSAs in the original SafeGraph data. For each MSA, we include all POIs that meet all of the following requirements: (1) the POI is located in the MSA; (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. We then include all CBGs that have at least 1 recorded visit to at least 10 of these POIs; this means that CBGs from outside the MSA may be included if they visit this MSA frequently enough.

As described in Methods M3.1, our model necessarily makes parametric assumptions about the relationship between POI characteristics (area, hourly visitors, and dwell time) and transmission rate at the POI; these assumptions may fail to hold for POIs which are outliers, particularly if SafeGraph data has errors. We mitigate this concern by truncating extreme values for POI characteristics to prevent data errors from unduly influencing our conclusions. Specifically, we truncate each POI's area to the 1st and 99th percentile of areas in the POI's category. Similarly, for every hour, we truncate each POI's visit count to its category's 99th percentile of visit counts in that hour, and for every time period, we truncate each POI's median dwell time to its category's 99th percentile of median dwell times in that period. Summary statistics of the post-processed data are in Extended Data Table 1. Overall, we analyze over 59,000 CBGs from the 10 MSAs, and over 250M visits from these CBGs to over 565,000 POIs.

SafeGraph data has been used to study consumer preferences[46] and political polarization.[47] More recently, it has been used as one of the primary sources of mobility data in the US for tracking the effects of the SARS-CoV-2 pandemic.[28,30,48–50] In SI Section S1, we show that aggregate trends in SafeGraph mobility data broadly match up to aggregate trends in Google mobility data in the US,[51] 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 over-represent individuals from higher-income areas.[52,53]

US Census. Our data on the demographics of census block groups (CBGs) comes from the US Census Bureau's American Community Survey (ACS).[54] We use the 5-year ACS (2013-2017) to extract the median household income, proportion of white residents, and 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 wish to minimize systematic downward bias in our total population counts (due to population growth) by making them as recent as possible.

New York Times. We calibrate our models using the SARS-CoV-2 dataset published by the *The New York Times*.[35] Their dataset consists of cumulative counts of cases and deaths in the United States over time, at the state and county level. For each MSA that we model, we sum over the county-level counts to produce overall counts for the entire MSA.

M2 Mobility network

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 census block groups (CBGs), and $\mathcal{P} = \{p_1, \ldots, p_n\}$, representing n points of interest (POIs). 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 t-th hour of simulation. We record the number of edges (with non-zero weights) in each MSA and over all hours from March 1, 2020 to May 2, 2020 in Extended Data Table 1. Across all 10 MSAs, we study 5.4 billion edges between 59,519 CBGs and 565,286 POIs.

From US Census data, each CBG c_i is labeled with its population N_{c_i}, income distribution, and racial and age demographics. From SafeGraph data, each POI p_j is similarly labeled with its category (e.g., 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 central technical challenge in constructing this network is estimating the network weights $W^{(t)} = \{w_{ij}^{(t)}\}$ from SafeGraph data, since this visit matrix is not directly available from the data. Because the estimation procedure is involved, we defer describing it in detail until Methods M6; in Methods M3–M5, we will assume that we already have the network weights.

M3 Model dynamics

To model the spread of SARS-CoV-2, we overlay a metapopulation disease transmission model on the mobility network defined in Methods M2. The transmission model structure follows prior work on epidemiological models of SARS-CoV-2[15,22] but incorporates a fine-grained mobility network into the calculations of the transmission rate (Methods M3.1). We construct separate mobility networks and models for each metropolitan statistical area (MSA).

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 cannot infect others, 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:

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

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

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

where $\lambda_{p_j}^{(t)}$ is the rate of infection at POI p_j at time t; $w_{ij}^{(t)}$, the ij-th entry of the visit matrix from the mobility network (Methods M2), 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,

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

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

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

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

M3.1 The number of new exposures $N_{S_{c_i} \to E_{c_i}}^{(t)}$

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 $\text{Pois}\left(\sum_{j=1}^{n} \lambda_{p_j}^{(t)} w_{ij}^{(t)}\right)$, and other cases not captured by visiting POIs, which are sampled from $\text{Binom}\left(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)}\right)$.

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. Since 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 $\text{Binom}(w_{ij}^{(t)} S_{c_i}^{(t)}/N_{c_i}, \lambda_{p_j}^{(t)}) \approx \text{Pois}(\lambda_{p_j}^{(t)} w_{ij}^{(t)} S_{c_i}^{(t)}/N_{c_i})$. 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, $\text{Pois}\big((S_{c_i}^{(t)}/N_{c_i}) \sum_{j=1}^{n} \lambda_{p_j}^{(t)} w_{ij}^{(t)}\big)$.

We model the infection rate at POI p_j at time t, $\lambda_{p_j}^{(t)} := \beta_{p_j}^{(t)} \cdot I_{p_j}^{(t)}/V_{p_j}^{(t)}$, 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

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

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.[41] 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 (Methods M6.2); it has a quadratic effect on the POI transmission rate $\beta_{p_j}^{(t)}$ because it reduces both (1) the time that a susceptible visitor spends at p_j and (2) 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

$$\begin{aligned} \lambda_{p_j}^{(t)} &= \beta_{p_j}^{(t)} \cdot \frac{I_{p_j}^{(t)}}{V_{p_j}^{(t)}} \\ &= \psi \cdot d_{p_j}^2 \cdot \frac{I_{p_j}^{(t)}}{a_{p_j}}. \end{aligned} \tag{9}$$

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

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

Base rate of new exposures not captured by visiting POIs. 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, e.g., 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 $\lambda_{c_i}^{(t)}$ probability of becoming infected and transitioning to the exposed state, where

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

is proportional to the infection density at CBG c_i, and β_{base} is a constant that we fit to data.

Overall number of new exposures. Putting all of the above together yields the expression for the distribution of new exposures in CBG c_i at time t,

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

M3.2 The number of new infectious and removed cases

We model exposed individuals as becoming infectious at a rate 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 inversely proportional to the mean infectious period δ_I, with

$$N^{(t)}_{I_{c_i} \to R_{c_i}} \sim \text{Binom}\Big(I^{(t)}_{c_i}, 1/\delta_I\Big), \tag{14}$$

We estimate both δ_E and δ_I from prior literature; see Methods M4.

M3.3 Model initialization

In our experiments, $t = 0$ is the first hour of March 1, 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 the same expected initial prevalence p_0 in every CBG c_i. At $t = 0$, every individual in the MSA has the same independent probability p_0 of being exposed E instead of susceptible S. We thus initialize the model state by setting

$$S^{(0)}_{c_i} = N_{c_i} - E^{(0)}_{c_i} \tag{15}$$

$$E^{(0)}_{c_i} \sim \text{Binom}(N_{c_i}, p_0) \tag{16}$$

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

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

M4 Model calibration and validation

Most of our model parameters can either be estimated from SafeGraph and US Census data, or taken from prior work (see Extended Data Table 2 for a summary). This leaves 3 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))
3. The initial proportion of exposed individuals at time $t = 0$, p_0 (Equation (16)).

In this section, we describe how we fit these parameters to published numbers of confirmed cases, as reported by *The New York Times*. We fit models for each MSA separately. In Methods M4.4, we show that the resulting models can accurately predict the number of confirmed cases in out-of-sample data that was not used for model fitting.

M4.1 Selecting parameter ranges

Transmission rate factors ψ and β_{base}. We select parameter ranges for the transmission rate factors ψ and β_{base} by checking if the model outputs match plausible ranges of the basic reproduction number R_0 pre-lockdown, since R_0 has been the study of substantial prior work on SARS-CoV-2.[55] 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 approximately 0.1–1. R_{base} models transmission that is not correlated with POI activity, which includes within-household transmission. We chose the lower limit of 0.1 because beyond that point, base transmission would only contribute minimally to overall R, whereas previous work suggests that within-household transmission is a substantial contributor to overall transmission.[56,57] However, household transmission alone is not estimated to be sufficient to tip 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.[56] We therefore chose an upper limit of 1, corresponding to the assumption that $R_0 < 1$ when there is no POI activity whatsoever (i.e., $R_{\text{POI}} = 0$).

The plausible range for R_{POI} is then determined by combining $R_{\text{POI}} = R_0 - R_{\text{base}}$ with an overall range, estimated from prior work, for pre-lockdown R_0 of 2–3.[55] Thus, R_{POI} pre-lockdown plausibly ranges from roughly 1–3.

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

$$R_{\text{POI}} = \frac{N_{\text{POI}}}{N_0} \tag{19}$$

$$R_{\text{base}} = \frac{N_{\text{base}}}{N_0}. \tag{20}$$

We averaged these quantities over 20 stochastic realizations per MSA. Figure S2 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 approximately 0.001–0.012, and R_{POI} lies in the plausible range (for at least one MSA) when ψ ranges from approximately 1,000–10,000, so these are the parameter ranges we consider when fitting the model. As described in Methods M4.2, we verified that case count data for all MSAs can be fit using parameter settings for β_{base} and ψ within these ranges.

Initial prevalence of exposures, p_0. The extent to which SARS-CoV-2 infections had spread in the U.S. by the start of our simulation (March 1, 2020) is currently unclear.[58] To account for this uncertainty, we allow p_0 to vary across a large range between 10^{-5} and 10^{-2}. As described in Methods M4.2, we verified that case count data for all MSAs can be fit using parameter settings for p_0 within this range.

M4.2 Fitting to the number of confirmed cases

Using the parameter ranges above, we grid searched over $\psi, \beta_{\text{base}}$, and p_0 to find the models that best fit the number of confirmed cases reported by *The New York Times* (NYT).[35] For each of the 10 MSAs studied, we tested 1,260 different combinations of $\psi, \beta_{\text{base}}$, 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 Methods M3, 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 of infections will be confirmed, and moreover that they will confirmed exactly $\delta_c = 168$ hours (7 days) after becoming infectious. We assume that these parameters are time-invariant. As a sensitivity analysis, we alternatively stochastically sampled the number of confirmed cases and the confirmation delay from distributions with mean r_c and δ_c, but found that this did not change predictions noticeably. We estimated these parameters, r_c and δ_c, from prior work (Extended Data Table 2).

From these assumptions, we can calculate the predicted number of newly confirmed cases across all CBGs in the MSA on day d,

$$N_{\text{cases}}^{(\text{day } d)} = r_c \cdot \sum_{i=1}^{m} \sum_{\tau=24(d-1)+1-\delta_c}^{24d-\delta_c} N_{E_{c_i} \to I_{c_i}}^{(\tau)}, \tag{21}$$

where 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 NYT data, 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 MSA. We compare the reported number of cases and the number

of cases that our model predicts by computing the root-mean-squared-error (RMSE)[59] over the $D = \lfloor T/24 \rfloor$ days of our simulations,

$$\text{RMSE} = \sqrt{\frac{1}{D}\sum_{d=1}^{D}\left(N_{\text{cases}}^{(\text{day } d)} - \hat{N}_{\text{cases}}^{(\text{day } d)}\right)^2}. \tag{22}$$

For each combination of model parameters and for each MSA, we quantify model fit with the NYT data by running 20 stochastic realizations and averaging their RMSE.

Our simulation spans March 1 to May 2, 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 (Extended Data Table 2), we predict the number of cases with a 7 day offset, from March 8 to May 9, 2020.

M4.3 Parameter selection and uncertainty quantification

Throughout this paper, we report aggregate predictions from different parameter sets of $\psi, \beta_{\text{base}}$, and p_0 and multiple stochastic realizations. For each MSA, we:

1. Find the best-fit parameter set, i.e., with the lowest average RMSE over stochastic realizations.

2. Select all parameter sets that achieve an RMSE (averaged over stochastic realizations) within 20% of the RMSE of the best-fit parameter set.

3. Pool together all predictions across those parameter sets and all of their stochastic realizations, and report their mean and 2.5th/97.5th percentiles.

On average, each MSA has 10 parameter sets that achieve an RMSE within 20% of the best-fitting parameter set (Table S7). For each parameter set, we have results for 20 stochastic realizations. All uncertainty intervals in our results show the 2.5th/97.5th percentiles across these pooled results.

This procedure 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 RMSE of the best fit captures uncertainty in the model parameters $\psi, \beta_{\text{base}}$, and p_0. The latter is equivalent to assuming that the posterior probability over the true parameters is uniformly spread among all parameter sets within the 20% threshold.

M4.4 Model validation on out-of-sample data

We validate our models by showing that they predict the number of confirmed cases and deaths on out-of-sample data when we have access to corresponding mobility data. We then confirm that the mobility data used as input in the model improves the fit to case and death data by comparing to a model that does not use mobility data.

Out-of-sample prediction of the number of cases (Extended Data Figure 1). For each MSA, we split the available NYT dataset into a training set (spanning March 8, 2020 to April 14, 2020) and a test set (spanning April 15, 2020 to May 9, 2020). We fit the model parameters ψ, β_{base}, and p_0, as described in Methods M4.2, but only using 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 1a shows that the models fit the out-of-sample case data fairly well, demonstrating that they can extrapolate beyond the training set to future time periods.

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 in Methods M4.2.

Out-of-sample prediction of the number of deaths (Extended Data Figure 2). In addition to the number of confirmed cases, the NYT data also contains the daily reported number of deaths due to COVID-19 by county. We use this death data as an additional source of validation. To estimate the number of deaths N_{deaths}, we use a similar process as for the number of cases N_{cases}, except that we replace r_c with $r_d = 0.66\%$, the infection fatality rate for COVID-19, and δ_c with $\delta_d = 432$ hours (18 days), the number of days between becoming infectious and passing away (Extended Data Table 2). This gives

$$N_{\text{deaths}}^{(d)} = r_d \cdot \sum_{i=1}^{m} \sum_{\tau=24(d-1)+1-\delta_d}^{24d-\delta_d} N_{E_{c_i} \to I_{c_i}}^{(\tau)}. \tag{23}$$

Because we assume that deaths occur $\delta_d = 18$ days after individuals become infectious, we compare with NYT death data starting on March 19, 2020.

Extended Data Figure 2a demonstrates that the calibrated models also fit death counts surprisingly well, even though their parameters are selected to minimize RMSE in predicting cases, not deaths. In some MSAs, the model fits the death data less well; this is unsurprising, because our case and death count predictions assume constant case detection rates and fatality rates across MSAs.

Comparison to baseline that does not use mobility data. To determine whether mobility data aids in modeling case and death counts, we compare to a baseline SLIR model that does not use mobility data and simply assumes that all individuals within an MSA mix uniformly. In this baseline, an individual's risk of being infected and transitioning to the exposed state at time t is

$$\lambda^{(t)} := \beta_{\text{base}} \cdot \frac{I^{(t)}}{N}, \tag{24}$$

where $I^{(t)}$ is the total number of infectious individuals at time t, and N is the total population size of the MSA. As above, we performed a grid search over β_{base} and p_0, and calibrated the models on the training set. Extended Data Figure 1b shows that this model fits case counts less well than the model that uses mobility data: while it fits the training time period fairly well, it has poor generalization performance. Results are similar for deaths (Extended Data Figure 2b). The baseline model has a higher RMSE in predicting daily case counts during both the training and testing time periods in all 10 MSAs. As expected, using mobility data allows us to more accurately predict the number of cases.

M5 Analysis details

In this section, we include additional details about the experiments underlying the figures in the paper. We omit explanations for figures that are completely described in the main text.

Comparing the magnitude vs. timing of mobility reduction (Figure 2a). To simulate what would have happened if we changed the magnitude or timing of mobility reduction, we modify the real mobility networks from March 1–May 2, 2020, and then run our models on the hypothetical data. In Figure 2a, we report the cumulative incidence proportion at the end of the simulation (May 2, 2020), i.e., the total fraction of people in the exposed, infectious, and removed states at that time.

To simulate a smaller magnitude of mobility reduction, we interpolate between the mobility

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

$$\tilde{W}_\alpha^{(t)} := \alpha W^{(t)} + (1 - \alpha) W^{f(t)}. \tag{25}$$

If α is 1, then $\tilde{W}_\alpha^{(t)} = W^{(t)}$, and we use the actual observed mobility network for the simulation. On the other hand, if $\alpha = 0$, then $\tilde{W}_\alpha^{(t)} = W^{f(t)}$, 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 (May 2, 2020, 11PM), let $f(t) = t \mod 168$ 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 (April 27–May 2, 2020). We construct the time-shifted visit matrix $\tilde{W}_d^{(t)}$

$$\tilde{W}_d^{(t)} := \begin{cases} W^{(t-24d)} & \text{if } 0 \le t - 24d \le T, \\ W^{f(t-24d)} & \text{if } t - 24d < 0, \\ W^{g(t-24d)} & \text{otherwise.} \end{cases} \tag{26}$$

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.

A minority of POIs account for a majority of infections (Figure 2b and Extended Data Figure 3). To evaluate the distribution of infections over POIs, we run our models on the observed

mobility data from March 1–May 2, 2020 and record the number of 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)). Then, we count the total expected number of infections per POI by summing over hours. In Figure 2b, we sort the POIs by their expected number of infections and report the proportion of all infections caused by the top $x\%$ of POIs.

Reducing mobility by clipping maximum occupancy (Figure 2c, Extended Data Figure 4). We implemented two partial reopening strategies: one that uniformly reduced visits at POIs to a fraction of full activity, and the other that "clipped" each POI's hourly visits to a fraction of the POI's maximum occupancy. For each reopening strategy, we started the simulation at March 1, 2020 and ran it until May 31, 2020, using the observed mobility network from March 1–April 30, 2020, and then using a hypothetical post-reopening mobility network from May 1–31, 2020, corresponding to the projected impact of that reopening strategy. Because we only have observed mobility data from March 1–May 2, 2020, we impute the missing mobility data up to May 31, 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 (May 2, 2020, 11PM). To simplify notation, we define

$$h(t) := \begin{cases} t & \text{if } t < T, \\ f(t) & \text{otherwise,} \end{cases} \tag{27}$$

where, as above, $f(t) = t \mod 168$. 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

$$\tilde{W}_{\gamma}^{(t)} := \begin{cases} W^{h(t)} & \text{if } t < R, \\ \alpha W^{h(t)} & \text{otherwise,} \end{cases} \tag{28}$$

where R represents the first hour of reopening (May 1, 2020, 12AM). In other words, we use the

actual observed mobility network up until hour R, and then subsequently simulate an γ-fraction of full mobility levels.

To simulate the clipping strategy, we first estimated the maximum occupancy M_{p_j} of each POI p_j as the maximum number of visits that it ever had in one hour, across all of March 1 to May 2, 2020. As in previous sections, let $w_{ij}^{(t)}$ represent the i, j-th entry in the observed visit matrix $W^{(t)}$, i.e., 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, i.e., $\sum_i w_{ij}^{(t)}$. We simulated clipping at a β-fraction of maximum occupancy, where $\beta \in [0, 1]$, by constructing the visit matrix $\tilde{W}_{\beta}^{(t)}$ whose i, j-th entry is

$$\tilde{w}_{ij\beta}^{(t)} := \begin{cases} w_{ij}^{h(t)} & \text{if } t < R \text{ or } V_{p_j}^{(t)} \leq \beta M_{p_j}, \\ \frac{\beta M_{p_j}}{V_{p_j}^{(t)}} w_{ij}^{h(t)} & \text{otherwise.} \end{cases} \tag{29}$$

This corresponds to the following procedure: for each POI p_j and time t, we first check if $t < R$ (reopening has not started) or if $V_{p_j}^{(t)} \leq \beta M_{p_j}$ (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 increase in cumulative incidence at the end of the reopening period (May 31, 2020), compared to the start of the reopening period (May 1, 2020).

Relative risk of reopening different categories of POIs (Figure 2d, Extended Data Figures 5 and 8, Figures S4-S13). We study separately reopening the 20 POI categories with the most visits in SafeGraph data. We exclude four categories due to data quality concerns from prior work[30]: "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) 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 our model is not designed to capture).

This reopening analysis is similar to the above analysis on clipping vs. uniform reopening. As above, we set the reopening time R to May 1, 2020, 12AM. 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, i.e., we simply repeat the activity levels of the last week of our data (April 27–May 2, 2020): This gives us the visit matrix $\tilde{W}^{(t)}$ with entries

$$\tilde{w}_{ij}^{(t)} := \begin{cases} w_{ij}^{(t)} & \text{if } t < R, \\ w_{ij}^{f(t)} & \text{if } t \geq R, p_j \in \mathcal{V} \\ w_{ij}^{g(t)} & \text{if } t \geq R, p_j \notin \mathcal{V}. \end{cases} \tag{30}$$

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 difference between (1) the cumulative fraction of people who have been infected by the end of the reopening period (May 31, 2020) and (2) the cumulative fraction of people infected by May 31 had we not reopened the POI category (i.e., if we simply repeated the activity levels of the last week of our data). This seeks to model the increase in cumulative incidence by end of May from reopening the POI category. In Extended Data Figure 5 and Figures S4-S13, the bottom right panel shows the increase for the category as a whole, and the bottom left panel shows the increase *per POI* (i.e., the total increase divided by the number of POIs in the category).

Per-capita mobility (Figure 3d, Extended Data Figures 6 and 7). Each group of CBGs (e.g., the bottom income decile) comprises a set $\mathcal{U}$ of CBGs that fit the corresponding criteria. In Extended Data 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 Extended Data Figure 7, we show the total number of visits made by each group to each POI category, accumulated over the entire data period (March 1–May 2, 2020) and then divided by the total population of the group.

Average transmission rate of a POI category (Figure 3e). We compute the average transmission rate experienced by a group of CBGs $\mathcal{U}$ at a POI category $\mathcal{V}$ as

$$\bar{\beta}_{XY} := \frac{\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{V}} \sum_{t=1}^{T} w_{ij}^{(t)} \beta_{p_j}^{(t)}}{\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{V}} \sum_{t} w_{ij}^{(t)}}, \tag{31}$$

where $\beta_{p_j}^{(t)}$ is the POI transmission rate (Equation (8)). This represents the expected transmission rate encountered during a visit by someone from a CBG in group $\mathcal{U}$ to a POI in category $\mathcal{V}$.

M6 Estimating the mobility network from SafeGraph data

Finally, we describe how we estimate the visit matrix $W^{(t)}$ (Methods M6.1) and dwell time d_{p_j} (Methods M6.2) from SafeGraph data.

Notation. We use a hat to denote quantities that we read directly from SafeGraph data, and r instead of t to denote time periods longer than an hour.

M6.1 Estimating the visit matrix $W^{(t)}$

Overview. We estimate the visit matrix $W^{(t)} = \{w_{ij}^{(t)}\}$, which captures the number of visitors from CBG c_i to POI p_j at each hour t from March 1, 2020 to May 2, 2020, through the iterative proportional fitting procedure (IPFP).[34] The idea is as follows:

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

2. However, visit patterns differ substantially from hour to hour (e.g., day versus night) and day to day (e.g., pre- versus post-lockdown). To capture these variations, we use current SafeGraph data to estimate the CBG marginals $U^{(t)}$, i.e., the total number of visitors leaving each CBG at each time t, as well as the POI marginals $V^{(t)}$, i.e., the total number of visitors present at each POI p_j at time t.

3. We then use 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 $\bar{W}$. Here, similarity is defined in terms of Kullback-Leibler divergence; we provide a precise definition below.

Quantities from SafeGraph data. To estimate the visit matrix, we read the following quantities from SafeGraph data:

- The estimated visit matrix $\hat{W}^{(r)}$ aggregated for the month r. This is taken from the Patterns dataset, and is aggregated at a monthly level. To account for non-uniform sampling from

different CBGs, we weight the number of SafeGraph visitors from each CBG by the ratio of the CBG population and the number of SafeGraph devices with homes in that CBG.[60]

- $\hat{V}_{p_j}^{(t)}$: The number of visitors recorded in POI p_j at hour t. This is taken from the Weekly Patterns v1 dataset.

- $\hat{h}_{c_i}^{(t)}$: The estimated fraction of people in CBG c_i that did not leave their home in day $\lfloor t/24 \rfloor$. This is derived by dividing `completely_home_device_count` by `device_count`, which are daily (instead of hourly) metrics in the Social Distancing Metrics dataset.

Estimating the aggregate visit matrix $\bar{W}$. The estimated monthly visit matrices $\hat{W}^{(r)}$ are typically noisy and sparse: SafeGraph only matches a subset of visitors to POIs to their home CBGs, either for privacy reasons (if there are too few visitors from the given CBG) or because they are unable to link the visitor to a home CBG.[61] To mitigate this issue, we aggregate these visit matrices, which are available at the monthly level, over the $R = 14$ months from January 2019 to February 2020:

$$\bar{W} := \frac{1}{R} \sum_{r} \hat{W}^{(r)}. \tag{32}$$

Each entry $\bar{w}_{ij}$ of $\bar{W}$ represents the estimated average number of visitors from CBG c_i to POI p_j per month from January 2019 to February 2020. After March 2020, SafeGraph reports this matrix on a weekly level in the Weekly Patterns v1 dataset. However, due to inconsistencies in the way SafeGraph processes the weekly vs. monthly matrices, we only use the monthly matrices up until February 2020.

Estimating the POI marginals $V^{(t)}$. We estimate the POI marginals $V^{(t)} \in \mathbb{R}^n$, whose j-th element $V_{p_j}^{(t)}$ represents our estimate of the number of visitors at POI p_j (from any CBG) at time t. The number of visitors recorded at POI p_j at hour t in the SafeGraph data, $\hat{V}_{p_j}^{(t)}$, is an underestimate because the SafeGraph data only covers on a fraction of the overall population. To correct for this, we follow Benzell et al.[30] and compute our final estimate of the visitors at POI p_j in time t as

$$V_{p_j}^{(t)} = \frac{\text{US population}}{\text{total number of SafeGraph devices}} \cdot \hat{V}_{p_j}^{(t)}. \tag{33}$$

This correction factor is approximately 7, using population data from the most recent 1-year ACS (2018).

Estimating the CBG marginals $U^{(t)}$. Next, we estimate the CBG marginals $U^{(t)} \in \mathbb{R}^m$. Here, the i-th element $U_{c_i}^{(t)}$ represents our estimate of the number of visitors leaving CBG c_i (to visit any POI) at time t. We will also use N_{c_i}; recall that N_{c_i} is the total population of c_i, which is independent of t.

We first use the POI marginals $V^{(t)}$ to calculate the total number of people who are out visiting any POI from any CBG at time t,

$$N_{\text{POIs}}^{(t)} := \sum_j V_{p_j}^{(t)}. \tag{34}$$

Since the total number of people leaving any CBG to visit a POI must equal the total number of people at all the POIs, we have that $N_{\text{POIs}}^{(t)} = \sum_i U_{c_i}^{(t)}$.

Next, we estimate the number of people from each CBG c_i who are not at home at time t as $\hat{h}_{c_i}^{(t)} N_{c_i}$. In general, the total number of people who are not at home in their CBGs, $\sum_i \hat{h}_{c_i}^{(t)} N_{c_i}$, will not be equal to $N_{\text{POIs}}^{(t)}$, the number of people who are out visiting any POI. This discrepancy occurs for several reasons: for example, some people might have left their homes to travel to places that SafeGraph does not track, SafeGraph might not have been able to determine the home CBG of a POI visitor, etc.

To correct for this discrepancy, we assume that the relative proportions of POI visitors coming from each CBG follows the relative proportions of people who are not at home in each CBG. We thus estimate $U_{c_i}^{(t)}$ by apportioning the $N_{\text{POIs}}^{(t)}$ total POI visitors at time t according to the proportion of people who are not at home in each CBG c_i at time t:

$$U_{c_i}^{(t)} := N_{\text{POIs}}^{(t)} \cdot \frac{\hat{h}_{c_i}^{(t)} N_{c_i}}{\sum_k \hat{h}_{c_k}^{(t)} N_{c_k}}, \tag{35}$$

where N_{c_i} is the total population of CBG i, as derived from US Census data. This construction ensures that the POI and CBG marginals match, i.e., $N_{\text{POIs}}^{(t)} = \sum_j V_{p_j}^{(t)} = \sum_i U_{c_i}^{(t)}$.

Iterative proportional fitting procedure (IPFP). IPFP is a classic statistical method[34] for adjusting joint distributions to match pre-specified marginal distributions, and it is also known in the

literature as biproportional fitting, the RAS algorithm, or raking.[62] In the social sciences, it has been widely used to infer the characteristics of local subpopulations (e.g., within each CBG) from aggregate data.[63–65]

We estimate the visit matrix $W^{(t)}$ by running IPFP on the aggregate visit matrix $\bar{W}$, the CBG marginals $U^{(t)}$, and the POI marginals $V^{(t)}$ constructed above. Our goal is to construct a non-negative matrix $W^{(t)} \in \mathbb{R}^{m \times n}$ whose rows sum up to the CBG marginals $U^{(t)}$,

$$U_{c_i}^{(t)} = \sum_{j=1}^{n} w_{ij}^{(t)}, \tag{36}$$

and whose columns sum up to the POI marginals $V_{p_j}^{(t)}$,

$$V_{p_j}^{(t)} = \sum_{i=1}^{m} w_{ij}^{(t)}, \tag{37}$$

but whose distribution is otherwise "as similar as possible", in the sense of Kullback-Leibler divergence, to the distribution over visits induced by the aggregate visit matrix $\bar{W}$.

Algorithm 1: Iterative proportional fitting procedure to estimate visit matrix $W^{(t)}$

Input: Aggregate visits $\bar{W} \in \mathbb{R}^{m \times n}$
CBG marginals $U^{(t)} \in \mathbb{R}^{m}$; POI marginals $V^{(t)} \in \mathbb{R}^{n}$
Number of iterations τ_{max}

Initialize $W^{(t,0)} = \bar{W}$
for $\tau = 1, \ldots, \tau_{max}$ **do**
  if τ *is odd* **then**
    for $i = 1, \ldots, m$ **do**
      $\alpha_i \leftarrow U_{c_i}^{(t)} / \sum_j w_{ij}^{(t)}$ `// Compute scaling factor for row` i
      $W_{i,:}^{(t,\tau)} \leftarrow \alpha_i * W_{i,:}^{(t,\tau-1)}$ `// Rescale row` i
    end
  else if τ *is even* **then**
    for $j = 1, \ldots, n$ **do**
      $\beta_j \leftarrow V_{p_j}^{(t)} / \sum_i w_{ij}^{(t)}$ `// Compute scaling factor for col` j
      $W_{:,j}^{(t,\tau)} \leftarrow \alpha_i * W_{:,j}^{(t,\tau-1)}$ `// Rescale col` j
    end
  end
end
$W^{(t)} \leftarrow W^{(t,\tau_{\text{max}})}$

IPFP is an iterative algorithm that alternates between scaling each row to match the row (CBG) marginals $U^{(t)}$ and scaling each column to match the column (POI) marginals $V^{(t)}$. We provide pseudocode in Algorithm 1. For each value of t used in our simulation, we run IPFP separately for $\tau_{\text{max}} = 100$ iterations. Note that IPFP is invariant to scaling the absolute magnitude of the entries in $\bar{W}$, since the total number of visits it returns is fixed by the sum of the marginals; instead, its output depends only on the distribution over visits in $\bar{W}$.

The notion of similarity invoked above has a maximum likelihood interpretation: if IPFP converges, then it returns a visit matrix $W^{(t)}$ whose induced distribution minimizes the Kullback-Leibler divergence to the distribution induced by $\bar{W}$.[66] We further discuss the convergence of IPFP in our setting in SI Section S3.

M6.2 Dwell times

Here, we describe how we estimate the dwell time d_{p_j}, which we use to estimate the hourly mean occupancy at each POI p_j. For simplicity, we treat d_{p_j} as independent of t, i.e., we average across all times t in our data for each POI p_j.

Quantities from SafeGraph data. To compute the average time $\hat{\delta}_{p_j}$ spent at each POI p_j, we average the values in the `median_dwell` field in the Patterns datasets from 2020. $\hat{\delta}_{p_j}$ is measured to minute-level resolution and expressed in units of hours, e.g., $\hat{\delta}_{p_j}^{(r)} = 1.5$ means a median visit time of 1.5 hours = 90 minutes.

If a visit straddles multiple hour boundaries at a POI, SafeGraph treats it as multiple visits in each of those hours for the purposes of computing the visit counts that we use in Methods M6.1. As a hypothetical example, consider a POI p_1 which has 1 new visitor come at the start of every hour and stay for exactly 1 hour, and another POI p_2 which also has 1 new visitor come at the start of every hour, except that visitors to p_2 stay for exactly 2 hours. SafeGraph data will reflect this difference in visit times, with $\bar{\delta}_{p_1} = 1$ and $\bar{\delta}_{p_2} = 2$. However, SafeGraph will also record that p_2 has twice as many visitors at every hour than p_1—in the notation of Methods M6.1, we would have that $V_{p_2}^{(t)} = 2V_{p_1}^{(t)}$ for all t—even though the total number of unique visitors to p_1 and p_2 are the same. This overcounting is the central technical issue that we address in this section.

Note that visits are still counted even if a visitor does not stay for the entire hour. For example, a visitor that arrives at 9:30am and leaves at 10:10am will be counted as two visits, once in the 9-10am hour and another time in the 10-11am hour.

Estimating the dwell time fraction d_{p_j}. Our goal is to estimate a correction factor for each POI p_j that corrects for the average visit duration of visitors to p_j. As the example above shows, setting the correction factor to be directly proportional to the average visit time $\bar{\delta}_{p_j}$ would overcount visitors who stay across multiple hour boundaries. Instead, we define the dwell time fraction $d_{p_j} \in [0, 1]$ as the average fraction of an hour that a visitor to POI p_j at any hour will spend there. In other words, conditioned on a visitor being at p_j at some time within an hour t, d_{p_j} is the expected fraction of the hour t that the visitor physically spends at p_j.

To estimate d_{p_j}, we make two assumptions: first, that every visitor to p_j stays for exactly $\bar{\delta}_{p_j}$ hours, and second, that the arrival times of visitors are uniformly distributed over all possible arrival times. Concretely, if a visitor to p_j stays for $\bar{\delta}_{p_j}$ hours and is recorded as present at hour t, then we assume that they are equally likely to have arrived at any time from $[t - \bar{\delta}_{p_j}, t + 1)$. From these assumptions, we can calculate

$$d_{p_j} := \frac{1}{1 + \bar{\delta}_{p_j}} \int_{t-\bar{\delta}_{p_j}}^{t+1} \underbrace{\min(\tau + \bar{\delta}_{p_j}, t + 1)}_{\text{departure time}} - \underbrace{\max(\tau, t)}_{\text{arrival time}} \; d\tau. \tag{38}$$

We truncate the departure time at $t + 1$ because any time spent after $t + 1$ does not count towards the hour t. Similarly, we truncate the arrival time at t. This expression for d_{p_j} simplifies into

$$d_{p_j} = \frac{\bar{\delta}_{p_j}}{1 + \bar{\delta}_{p_j}}. \tag{39}$$

See SI Section S4 for the derivation.

Methods references

43. SafeGraph. Places Schema (2020). Available at https://docs.safegraph.com/docs/places-schema.

44. SafeGraph. Weekly Patterns (2020). Available at https://docs.safegraph.com/docs/weekly-patterns.

45. SafeGraph. Social Distancing Metrics (2020). Available at https://docs.safegraph.com/docs/social-distancing-metrics.

46. Athey, S., Blei, D., Donnelly, R., Ruiz, F. & Schmidt, T. Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. In *AEA Papers and Proceedings*, vol. 108, 64–67 (2018).

47. Chen, M. K. & Rohla, R. The effect of partisanship and political advertising on close family ties. *Science* **360**, 1020–1024 (2018).

48. Farboodi, M., Jarosch, G. & Shimer, R. Internal and external effects of social distancing in a pandemic (2020). Available at https://nber.org/papers/w27059.

49. Killeen, B. D. *et al.* A County-level Dataset for Informing the United States' Response to COVID-19 (2020). Available at https://arxiv.org/abs/2004.00756.

50. Allcott, H. *et al.* Polarization and public health: Partisan differences in social distancing during the Coronavirus pandemic (2020). Available at nber.org/papers/w26946.

51. Google. COVID-19 community mobility reports (2020). Available at https://google.com/covid19/mobility/.

52. Athey, S., Ferguson, B., Gentzkow, M. & Schmidt, T. Experienced Segregation (2019). Available at https://gsb.stanford.edu/faculty-research/working-papers/experienced-segregation.

53. Squire, R. F. What about bias in the SafeGraph dataset? (2019). Available at https://safegraph.com/blog/what-about-bias-in-the-safegraph-dataset.

54. US Census. American Community Survey. Available at https://census.gov/programs-surveys/acs.

55. Park, M., Cook, A. R., Lim, J. T., Sun, Y. & Dickens, B. L. A systematic review of COVID-19 epidemiology based on current evidence. *Journal of Clinical Medicine* **9**, 967 (2020).

56. Curmei, M., Ilyas, A., Evans, O. & Steinhardt, J. Estimating household transmission of sars-cov-2. *medRxiv* (2020). Available at https://doi.org/10.1101/2020.05.23.20111559.

57. Li, W. *et al.* The characteristics of household transmission of COVID-19. *Clinical Infectious Diseases* .

58. Carey, B. & Glanz, J. Hidden outbreaks spread through u.s. cities far earlier than americans knew, estimates say. *The New York Times* (2020). Available at https://nytimes.com/2020/04/23/us/coronavirus-early-outbreaks-cities.html.

59. Chowell, G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling* 379–398 (2017).

60. SafeGraph. Measuring and Correcting Sampling Bias in Safegraph Patterns for More Accurate Demographic Analysis (2020). Available at https://safegraph.com/blog/measuring-and-correcting-sampling-bias-for-accurate-demographic-analysis.

61. SafeGraph. Places Manual (2020). Available at https://docs.safegraph.com/docs/places-manual#section-visitor-home-cbgs.

62. Bishop, Y. M., Fienberg, S. E. & Holland, P. W. Discrete multivariate analysis (1975).

63. Birkin, M. & Clarke, M. Synthesis—a synthetic spatial information system for urban and regional analysis: methods and examples. *Environment and planning A* **20**, 1645–1671 (1988).

64. Wong, D. W. The reliability of using the iterative proportional fitting procedure. *The Professional Geographer* **44**, 340–348 (1992).

65. Simpson, L. & Tranmer, M. Combining sample and census data in small area estimates: Iterative proportional fitting with standard software. *The Professional Geographer* **57**, 222–234 (2005).

66. Csiszár, I. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability* 146–158 (1975).

67. Perkins, A. *et al.* Estimating unobserved SARS-CoV-2 infections in the United States. *medRxiv* (2020). Available at https://doi.org/10.1101/2020.03.15.20036582.

68. Bommer, C. & Vollmer, S. Average detection rate of SARS-CoV-2 infections has improved since our last estimates but is still as low as nine percent on March 30th (2020). Available at https://www.uni-goettingen.de/en/606540.html.

69. Javan, E., Fox, S. J. & Meyers, L. A. The unseen and pervasive threat of COVID-19 throughout the US. *medRxiv* (2020). Available at https://doi.org/10.1101/2020.04.06.20053561.

70. Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet* **20**, 669–677 (2020).

71. APM Research Lab. The color of coronavirus: COVID-19 deaths by race and ethnicity in the U.S. (2020). Available at https://apmresearchlab.org/covid/deaths-by-race.

72. Pukelsheim, F. Biproportional scaling of matrices and the iterative proportional fitting procedure. *Annals of Operations Research* **215**, 269–283 (2014).

73. Gietl, C. & Reffel, F. P. Accumulation points of the iterative proportional fitting procedure. *Metrika* **76**, 783–798 (2013).

Data Availability. Census data, case and death counts from *The New York Times*, and Google mobility data are publicly available. Cell phone mobility data is freely available to researchers, non-profits, and governments through the SafeGraph COVID-19 Data Consortium.

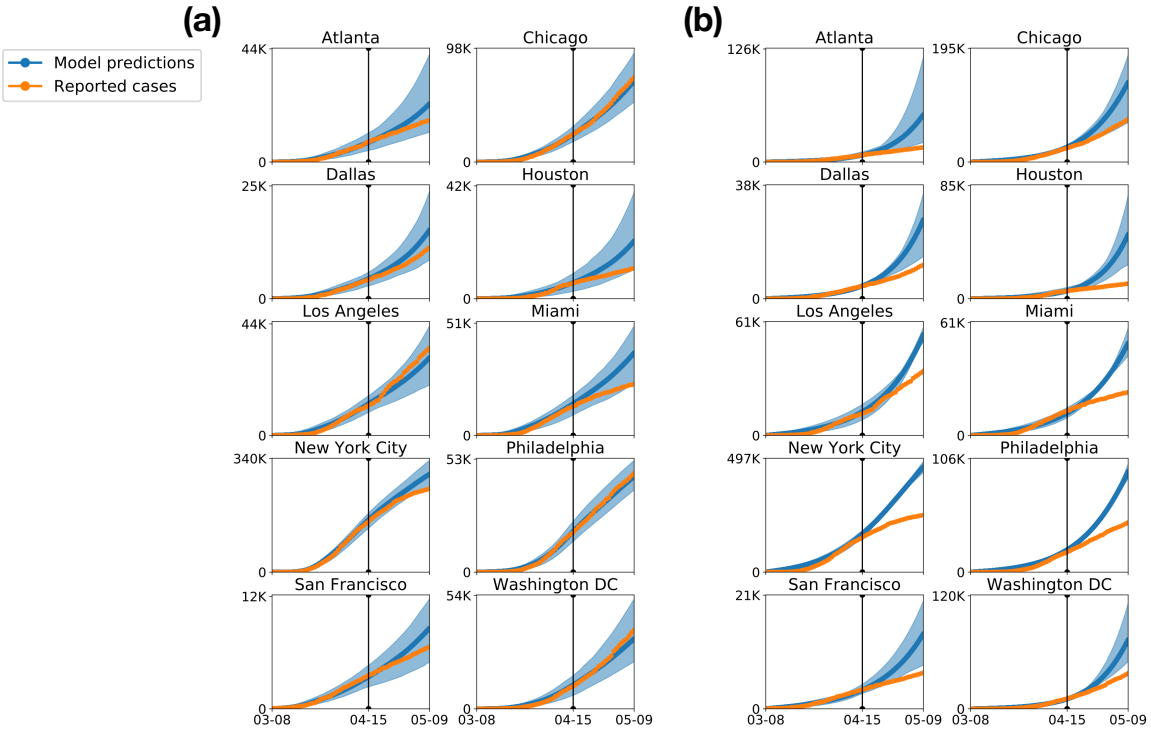
Code Availability. Code is in preparation and will be made publicly available at http://snap.stanford.edu/covid-mobility/.

Acknowledgements. The authors thank Yong-Yeol Ahn, Nic Fishman, Tatsunori Hashimoto, Roni Rosenfeld, Jacob Steinhardt, and Ryan Tibshirani for helpful comments. We also thank Nick Singh, Ryan Fox Squire, Jessica Williams-Holt, Jonathan Wolf, and others at SafeGraph for cell phone mobility data and helpful feedback. This research was supported by US National Science Foundation under OAC-1835598 (CINES), OAC-1934578 (HDR), CCF-1918940 (Expeditions), Chan Zuckerberg Biohub, Stanford Data Science Initiative, and the Stanford University Dean's Research Fund. 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.

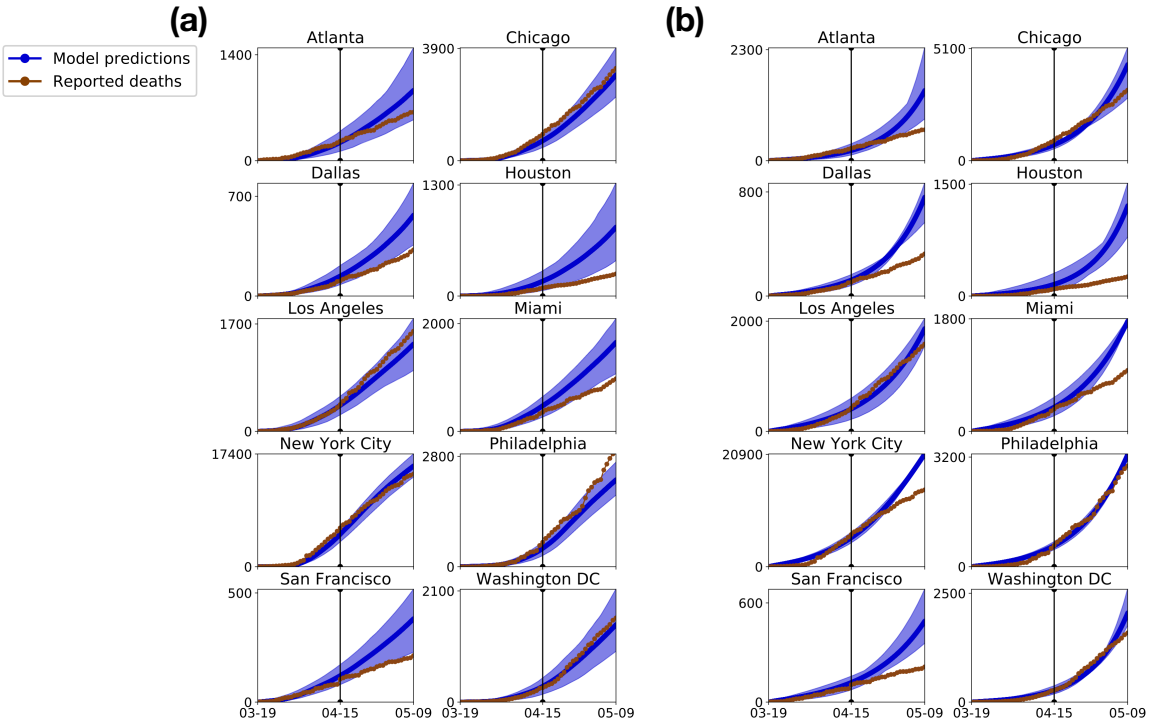
Author Contributions. S.C., E.P., and P.W.K. performed computational analysis. All authors jointly analyzed the results and wrote the paper.

Author Information. The authors declare no conflict of interest. Correspondence should be addressed to jure@cs.stanford.edu.

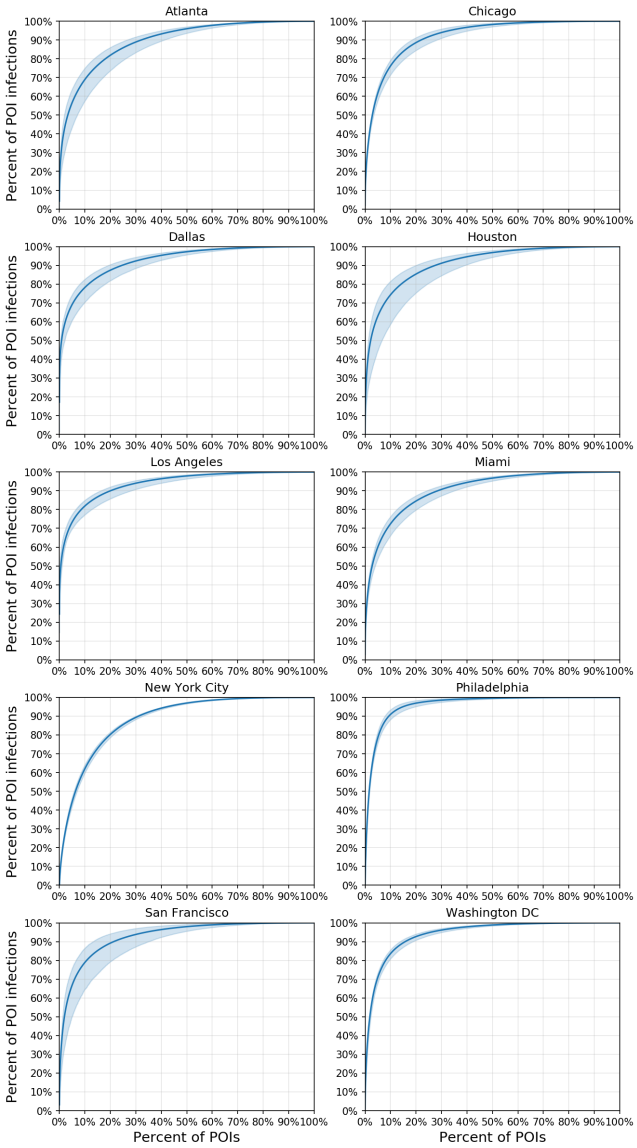
Extended data



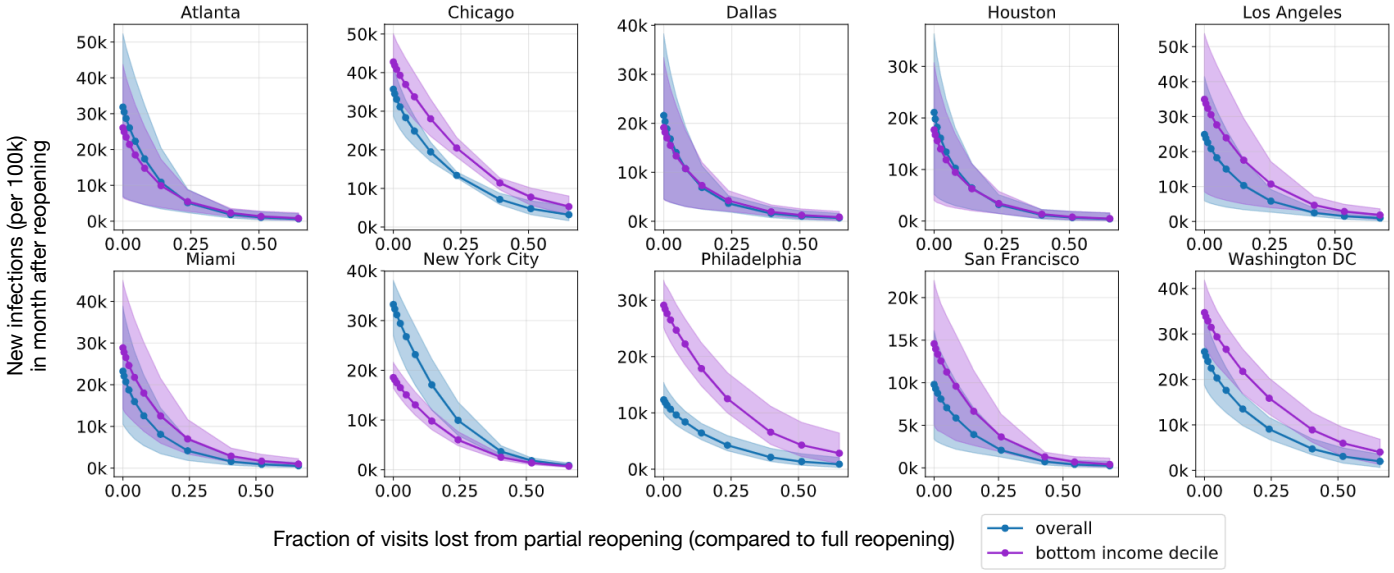
Extended Data Figure 1: Predicted (blue) and true (orange) cumulative case counts for (a) our model, which uses the mobility network, and (b) a baseline SEIR model that does not use mobility data. Incorporating the mobility network improves out-of-sample model fit. Both models are calibrated on observed case counts before April 15, 2020 (vertical black line). Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. See Methods M4.4 for details.



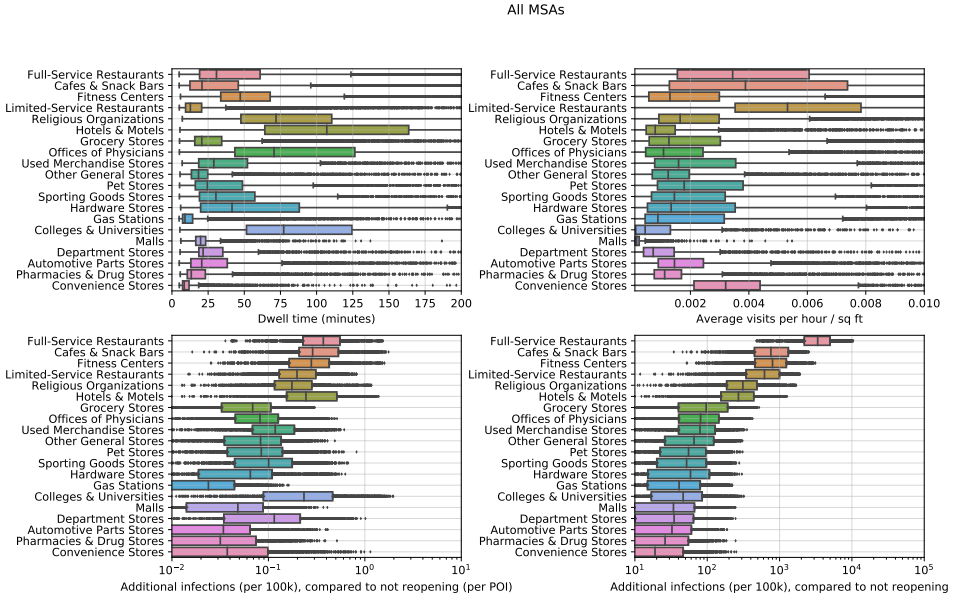
Extended Data Figure 2: Predicted (blue) and true (brown) cumulative death counts for (a) our model, which uses the mobility network, and (b) a baseline SEIR model that does not use mobility data. Incorporating the mobility network improves out-of-sample model fit. Both models are calibrated on observed *case* counts before April 15, 2020 (vertical black line). Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. See Methods M4.4 for details.



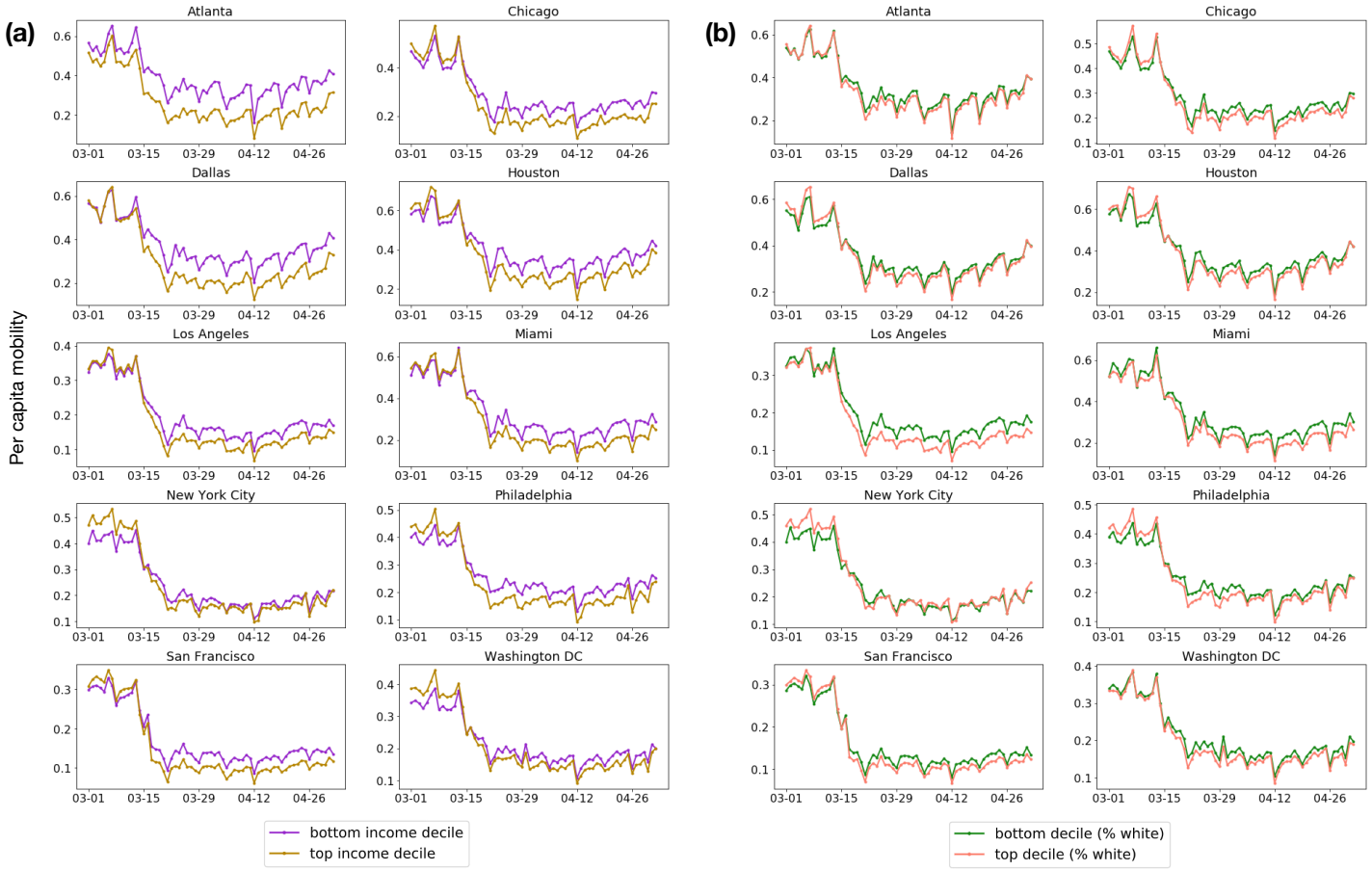
Extended Data Figure 3: A small fraction of POIs account for a large fraction of the predicted infections at POIs. We ran our models on the observed mobility data from March 1–May 2, 2020 and recorded the number of infections that occurred at each POI. Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. See Methods M5 for details.



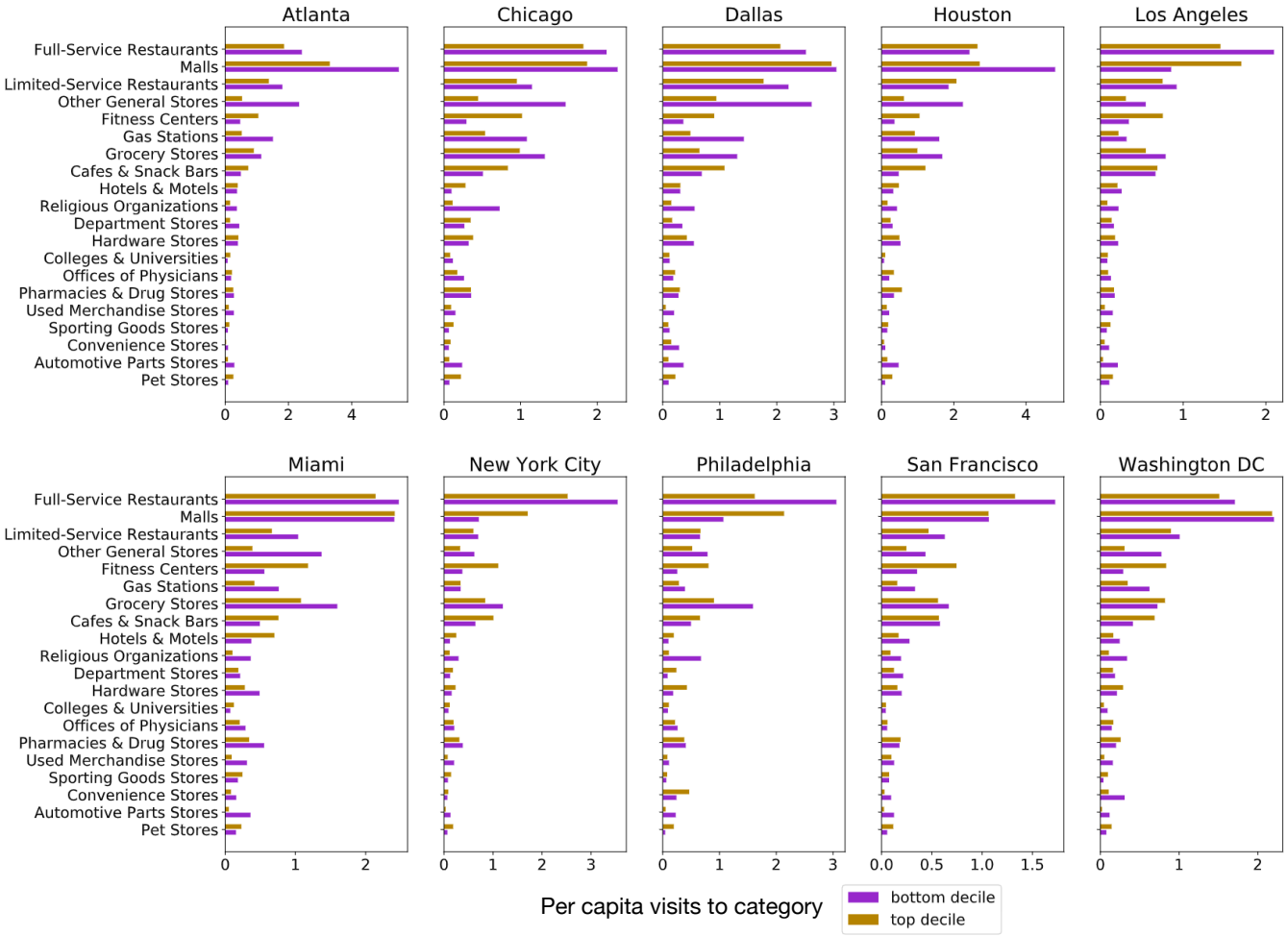
Extended Data Figure 4: The predicted increase in infections with the "clipping" partial reopening strategy. We simulate reopening starting on May 1, 2020 and run the simulation until the end of the month. Each dot represents a degree of clipping, e.g., clipping at 50% maximum occupancy, at 20% maximum occupancy, etc. The y-coordinate of each dot represents the predicted number of new infections incurred after reopening (per 100k population) and its x-coordinate represents the fraction of visits lost from partial reopening compared to full reopening. In 6 MSAs, the lower-income CBGs incur substantially more infections from reopening. Only in New York City (NYC) is this trend reversed; this is because such a high fraction of lower-income CBGs in NYC had been infected before reopening (62%) that after reopening, there was only a minority of the lower-income population that is still susceptible. In comparison, none of the other MSAs saw such a high incidence proportion among lower-income CBGs before reopening; for example, the second highest was 31% for Philadelphia, and the rest ranged from 2–13%. Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. See Methods M5 for reopening details.



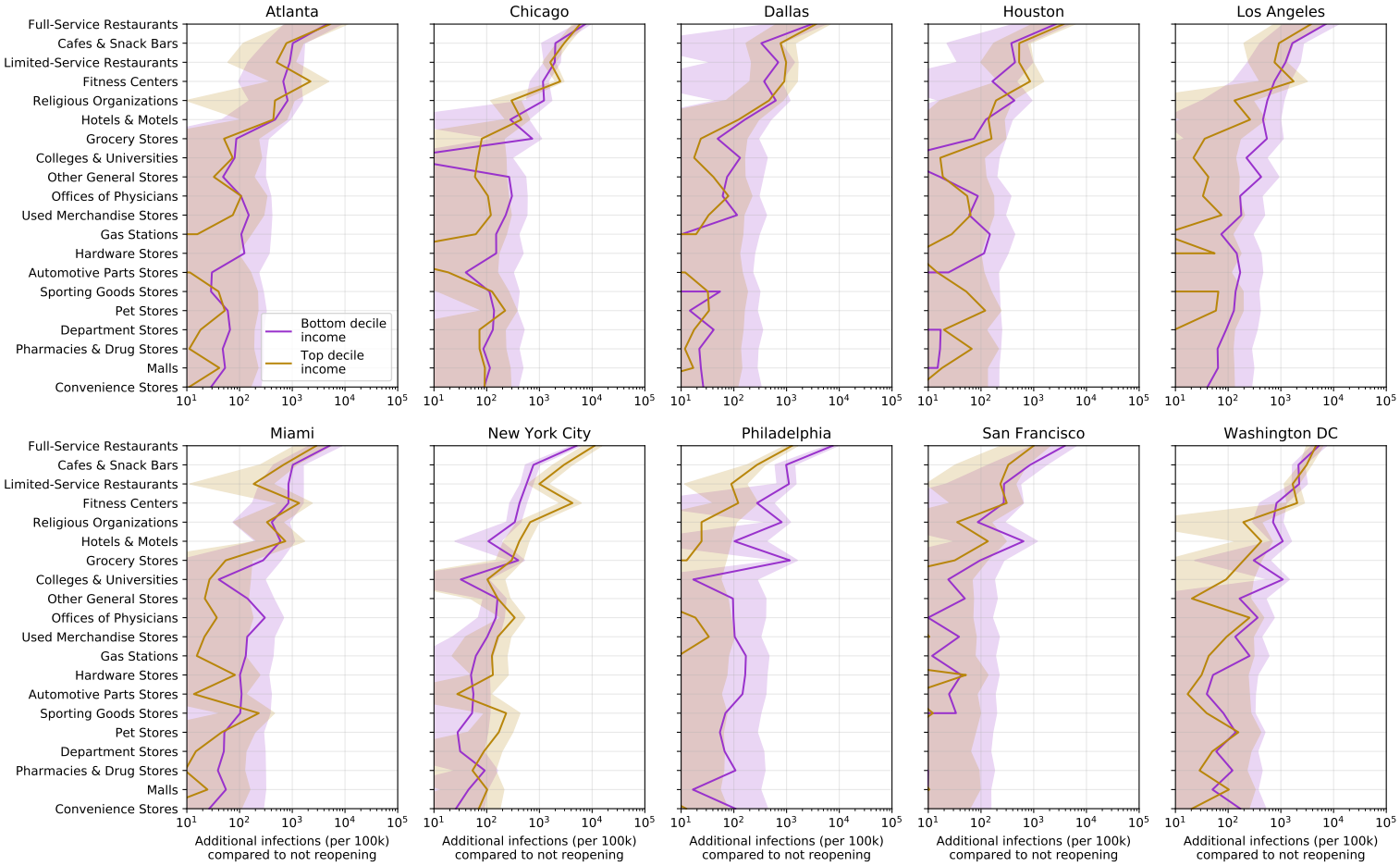
Extended Data Figure 5: POI attributes in all 10 MSAs combined. The top two plots pool POIs from all MSAs, and show the distribution of dwell time and the average number of hourly visitors divided by the area of the POI in square feet. Each point represents one POI; boxes depict the interquartile range across POIs. The bottom two plots pool across models from all MSAs, and show predictions for the increase in infections (per 100k population) from reopening a POI category: per POI (left bottom plot) and for the category as a whole (right bottom plot). Each point represents one model realization; boxes depict the interquartile range across sampled parameters and stochastic realizations. See Methods M5 for reopening details.



Extended Data Figure 6: Daily per-capita mobility over time, (a) comparing lower-income to higher-income CBGs and (b) comparing less white to more white CBGs. See Methods M5 for details.



Extended Data Figure 7: Visits per capita from the bottom- (purple) and top- (gold) income deciles to each POI category, accumulated from March 2–May 1, 2020. See Methods M5 for details.



Extended Data Figure 8: Predicted additional infections (per 100k population) from reopening each POI category, for CBGs in the top- (gold) and bottom- (purple) income deciles. Reopening impacts are generally worse for lower-income CBGs. See Methods M5 for reopening details.

MSA	CBGs	POIs	Hourly edges	Total modeled pop	Total visits
Atlanta	3,418	40,790	564,833,067	8,412,703	25,781,129
Chicago	6,967	63,783	520,620,178	10,630,027	26,317,828
Dallas	5,310	54,425	765,375,701	10,585,511	32,372,502
Houston	3,525	50,974	588,830,794	8,172,419	28,374,914
Los Angeles	9,418	85,498	663,105,494	17,639,622	29,932,105
Miami	3,755	42,124	521,159,453	7,527,368	21,784,546
New York City	14,997	124,123	981,646,415	21,491,228	47,552,731
Philadelphia	4,725	38,938	304,524,970	7,096,928	15,628,173
San Francisco	3,179	29,385	169,668,043	5,758,729	8,300,633
Washington DC	4,225	35,246	326,167,894	8,171,909	14,904,627
All MSAs combined	**59,519**	**565,286**	**5,405,932,009**	**105,486,444**	**250,949,188**

Extended Data Table 1: Dataset summary statistics from March 1–May 2, 2020.

Param.	Description	Value (Source)
δ_E	mean latency period	96 hours[22,42]
δ_I	mean infectious period	84 hours[22]
δ_c	period from infectious to confirmed	7 days[22,67]
r_c	rate of confirmed cases	10%[22,42,67–69]
δ_d	period from infectious to death	18 days[70]
r_d	infection fatality rate	0.66%[70]
β_{base}	base CBG transmission rate	Variable (Estimated)
N_{c_i}	population size of CBG c_i	Variable (2018 US Census[54])
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 latent 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 latent population in CBG c_i	$p_0 N_{c_i}$
$I_{c_i}^{(0)}$	initial infected population in CBG c_i	0
$R_{c_i}^{(0)}$	initial removed population in CBG c_i	0

Extended Data Table 2: Model parameters. If the parameter has a fixed value, we specify it under **Value**; otherwise, we write "Variable" to indicate that it varies across CBG / POI / MSA.

Supplementary methods

S1 Comparison of Google and SafeGraph mobility data

To assess the reliability of the SafeGraph datasets, we measured the correlation between mobility trends according to SafeGraph versus Google.[51] Google provides a high-level picture of mobility changes around the world for several categories of places, such as grocery stores or restaurants. We analyzed four of the six categories defined by Google: *Retail & recreation* (e.g., restaurants, shopping centers, movie theaters), *Grocery & pharmacy* (e.g., grocery stores, farmers markets, pharmacies), *Parks* (e.g., local parks, national parks, public beaches), and *Residential* (i.e. places of residence). We omitted *Transit stations* because they are not well-covered by SafeGraph POIs, and *Workplaces* because we do not model whether people are at work. To account for the first three categories, we used POI visits in the SafeGraph Patterns datasets, identifying POIs in each category based on their 6-digit NAICS codes (Table S5). For the *Residential* category, we used SafeGraph Social Distancing Metrics, which provides daily counts of the number of people in each CBG who stayed at home for the entire day.

For each US region and category, Google tracks how the number of visits to the category has changed over the last few months, compared to typical levels of activity before SARS-CoV-2. To set this baseline, they compute the median number of visits to the category for each day of the week, over a 5-week span from January 3–February 6, 2020. For a given day of interest, they then compute the relative change in number of visits seen on this day compared to the baseline for the corresponding day of week. We replicated this procedure on SafeGraph data, and compared the results to Google's trends for Washington DC and 14 states that appear in the MSAs that we model. For each region and category, we measured the Pearson correlation between the relative change in number of visits according to Google versus Safegraph, from March 1–May 2, 2020. Across the 15 regions, we found that the median Pearson correlation was 0.96 for *Retail & recreation*, 0.79 for *Grocery & pharmacy*, 0.44 for *Parks*, and 0.88 for *Residential*. As an illustrative example, we visualize the results for New York in Figure S3, and provide a full table of results for every state in Table S6. The Pearson correlations are high for all categories aside from *Parks*. Since POIs in *Parks* account for a small fraction of the overall visits in the SafeGraph dataset, we believe that this discrepancy should not significantly impact our models. Specifically, we evaluated the percentage of total SafeGraph POI visits to each category during the baseline period, and we found that *Retail & recreation* accounts for 34.7% of visits, *Grocery & pharmacy* covers 5.4%, and *Parks*

only covers 2.5%. Overall, these results indicate that SafeGraph mobility trends recapitulate those in Google data, providing a validation of the reliability of SafeGraph data.

S2 Plausibility of predicted racial/socioeconomic disparities

To assess the plausibility of the predicted disparities in infection rates in Figure 3, we compared the model's predicted racial disparities to observed racial disparities in mortality rates. (Data on socioeconomic disparities in mortality was not systematically available on a national level.) The racial disparities in Figure 3 are generally of the same magnitude as reported racial disparities in mortality rates—for example, the overall reported black mortality rate is $2.4\times$ higher than the white mortality rate,[71] which is approximately the same as the median racial disparity across MSAs of $2.4\times$ that our model predicts (Figure 3b). However, we note that this is an imperfect comparison because many factors besides mobility contribute to racial disparities in death rates.

In addition, we observed that our model predicted unusually large socioeconomic and racial disparities in infection rates in the Philadelphia MSA. To understand why the model predicted such large disparities, we inspected the mobility factors discussed in the main text; namely, how much each group was able to reduce their mobility, and whether disadvantaged groups encountered higher transmission rates at POIs.

First, we find in Philadelphia that higher-income CBGs were able to reduce their mobility substantially more than lower-income CBGs (Extended Data Figure 6 left). The CBGs with the greatest percentage of white residents were also able to reduce their mobility more than the CBGs with the lowest percentage of white residents (Extended Data Figure 6 right). These gaps are noticeable, but not obviously larger than those in other MSAs. The key to Philadelphia's outlier status seems to lie in the comparison of transmission rates. Within the same category of POI—e.g., full-service restaurants—individuals from lower-income CBGs tend to visit POIs with higher transmission rates than individuals from high-income CBGs (Table S3). This is particularly true for Philadelphia; in every one of the 20 categories, individuals from lower-income CBGs in Philadelphia encounter higher transmission rates than individuals from high-income CBGs, which does not happen for any other MSA. (Results for race are similar: when we compare Philadelphia CBGs in the top and bottom deciles for percentage of white residents, the CBGs in the bottom decile encounter higher transmission rates in 19 out of 20 of the categories, as shown in Table S4.) The transmission rates encountered by individuals from lower-income CBGs in Philadelphia are often dramatically higher than those encountered by higher-income CBGs; for example, up to $11.8\times$

higher for grocery stores. Digging deeper, this is because the average grocery store visited by lower-income CBGs has $4\times$ the number of hourly visitors per square foot, and visitors tend to stay twice as long. Furthermore, Philadelphia's large discrepancy in density between lower-income and higher-income POIs in SafeGraph data is consistent with Census data, which shows that the correlation between population density and median household income is larger in Philadelphia than in any of the other MSAs that we examine (Spearman correlation 0.55 in Philadelphia, as compared to a median of 0.31 across MSAs).

Since there are many other factors of disparity that we do not model, we do not place too much weight on our model's prediction that Philadelphia's disparities will be larger than those of other cities. However, we consider this a valuable finding in terms of Philadelphia's mobility patterns, suggesting that mobility may play an especially strong role in driving socioeconomic and racial infection disparities in this MSA, and we encourage policymakers to be aware of how differences in mobility patterns may exacerbate the disproportionate impact of SARS-CoV-2 on disadvantaged groups.

S3 Convergence of iterative proportional fitting

For completeness, we briefly review the convergence properties of the iterative proportional fitting procedure (IPFP) used to infer our mobility networks. Consider the L_1-error function

$$E^{(t,\tau)} := \underbrace{\sum_i \left| U_{c_i}^{(t)} - \sum_j w_{ij}^{(t)} \right|}_{\text{Error in row marginals}} + \underbrace{\sum_j \left| V_{p_j}^{(t)} - \sum_i w_{ij}^{(t)} \right|}_{\text{Error in column marginals}}, \tag{40}$$

which sums up the errors in the row (CBG) and column (POI) marginals of the visit matrix $W^{(t,\tau)}$ from the τ-th iteration of IPFP. Each iteration of IPFP monotonically reduces this L_1-error $E^{(t,\tau)}$, i.e., $E^{(t,\tau)} \geq E^{(t,\tau+1)}$ for all $\tau \geq 0$.[72] In other words, the row and column sums of $W^{(t,\tau)}$ (which is initialized as $W^{(t,0)} = \bar{W}$) progressively get closer to (or technically, no further from) the target marginals as the iteration number τ increases. Moreover, IPFP maintains the cross-product ratios of the aggregate matrix $\bar{W}$, i.e.,

$$\frac{w_{ij}^{(t,\tau)} w_{k\ell}^{(t,\tau)}}{w_{i\ell}^{(t,\tau)} w_{kj}^{(t,\tau)}} = \frac{\bar{w}_{ij}\bar{w}_{k\ell}}{\bar{w}_{i\ell}\bar{w}_{kj}} \tag{41}$$

for all matrix entries indexed by i, j, k, ℓ, for all t, and for all iterations τ.

IPFP converges to a unique solution, in the sense that $W^{(t)} = \lim_{\tau\to\infty} W^{(t,\tau)}$, if there exists a matrix $W^{(t)}$ that fits the row and column marginals while maintaining the sparsity pattern (i.e., location of zeroes) of $\bar{W}$.[72] If IPFP converges, then the L_1-error also converges to 0 as $\tau \to \infty$,[72] and $W^{(t)}$ is the maximum likelihood solution in the following sense. For a visit matrix $W = \{w_{ij}\}$, let P_W represent a multinomial distribution over the mn entries of W with probability proportional to w_{ij}, and define $\mathcal{U}^{(t)} \subseteq \mathbb{R}_+^{m\times n}$ and $\mathcal{V}^{(t)} \subseteq \mathbb{R}^{m\times n}$ as the set of non-negative matrices whose row and column marginals match $U^{(t)}$ and $V^{(t)}$ respectively. Then, if IPFP converges,

$$W^{(t)} = \underset{W\in\mathcal{U}^{(t)}\cap\mathcal{V}^{(t)}}{\arg\min} \quad \mathrm{KL}\left(P_W \| P_{\bar{W}}\right), \tag{42}$$

where $\mathrm{KL}\left(p\|q\right)$ is the Kullback-Leibler divergence $\mathrm{KL}\left(p\|q\right) = \mathbb{E}_p\left[\log \frac{p(x)}{q(x)}\right]$. In other words, IPFP returns a visit matrix $W^{(t)}$ whose induced distribution $P_{W^{(t)}}$ is the I-projection of the aggregate visit distribution $P_{\bar{W}}$ on the set of distributions with compatible row and column marginals.[66] In fact, IPFP can be viewed as an alternating sequence of I-projections onto the row marginals and I-projections onto the column marginals.[66,73]

However, in our setting, IPFP typically does not return a unique solution and instead oscillates between two accumulation points, one that fits the row marginals and another that fits the column marginals.[73] This is because $\bar{W}$ is highly sparse (there is no recorded interaction between most CBGs and POIs), so the marginals are sometimes impossible to reconcile. For example, suppose there is some CBG c_i and POI p_j such that $\bar{w}_{ij}$ is the only non-zero entry in the i-th row and j-th column of $\bar{W}$, i.e., visitors from c_i only travel to p_j and conversely visitors from p_j are all from c_i. Then, if $U_{c_i}^{(t)} \neq V_{p_j}^{(t)}$, there does not exist any solution $W^{(t)}$ such that $U_{c_i}^{(t)} = V_{p_j}^{(t)} = w_{ij}^{(t)}$. Note that in this scenario, IPFP still monotonically decreases the L_1-error.[72]

In our implementation (Algorithm 1), we take $\tau_{\text{max}} = 100$, so IPFP ends by fitting the column (POI) marginals. This ensures that our visit matrix $W^{(t)}$ is fully compatible with the POI marginals $V^{(t)}$, i.e.,

$$V_{p_j}^{(t)} = \sum_{i=1}^{m} w_{ij}^{(t)}, \tag{43}$$

while still minimizing the L_1-error $E^{(t,\tau)}$ with respect to the CBG marginals $U^{(t)}$. Empirically, we find that $\tau_{\text{max}} = 100$ iterations of IPFP are sufficient to converge to this oscillatory regime.

S4 Derivation of the dwell time fraction d_{p_j}

Recall from Methods M6.2 that we have

$$d_{p_j} := \frac{1}{1+\bar{\delta}_{p_j}} \int_{t-\bar{\delta}_{p_j}}^{t+1} \underbrace{\min(\tau+\bar{\delta}_{p_j}, t+1)}_{\text{departure time}} - \underbrace{\max(\tau, t)}_{\text{arrival time}} \; d\tau. \tag{44}$$

This expression for d_{p_j} simplifies into

$$d_{p_j} = \frac{\bar{\delta}_{p_j}}{1+\bar{\delta}_{p_j}}. \tag{45}$$

To see this, first consider the case where $\bar{\delta}_{p_j} \leq 1$. To keep notation simple, let $t = 0$ mark the start of the hour being considered. Then,

- Visitors who arrive at $-\bar{\delta}_{p_j}$ will spend 0 time during the hour being considered. This time increases linearly from 0 to $\bar{\delta}_{p_j}$ as the arrival time increases from $-\bar{\delta}_{p_j}$ to 0.
- Visitors who arrive during $[0, 1-\bar{\delta}_{p_j}]$ will spend the full $\bar{\delta}_{p_j}$ time in the hour being considered.
- This time decreases linearly from $\bar{\delta}_{p_j}$ to 0 as the arrival time increases from $1-\bar{\delta}_{p_j}$ to 1.

In total, we thus have

$$d_{p_j} = \frac{1}{1+\bar{\delta}_{p_j}} \left(\frac{\bar{\delta}_{p_j}^2}{2} + (1-\bar{\delta}_{p_j})\bar{\delta}_{p_j} + \frac{\bar{\delta}_{p_j}^2}{2} \right) = \frac{\bar{\delta}_{p_j}}{1+\bar{\delta}_{p_j}}. \tag{46}$$

Similarly, consider the case where $\bar{\delta}_{p_j} > 1$. Then,

- Visitors who arrive at $-\bar{\delta}_{p_j}$ will spend 0 time during the hour being considered. This time increases linearly from 0 to 1 as the arrival time increases from $-\bar{\delta}_{p_j}$ to $1-\bar{\delta}_{p_j}$.
- Visitors who arrive during $[1-\bar{\delta}_{p_j}, 0]$ will spend the full hour.
- This time decreases linearly from 1 to 0 as the arrival time increases from 0 to 1.

In total, we likewise have

$$d_{p_j} = \frac{1}{1+\bar{\delta}_{p_j}} \left(\frac{1}{2} + \bar{\delta}_{p_j} - 1 + \frac{1}{2} \right) = \frac{\bar{\delta}_{p_j}}{1+\bar{\delta}_{p_j}}. \tag{47}$$

Supplementary tables

MSA	7 days earlier	3 days earlier	3 days later	7 days later
Atlanta	0.443 (0.257, 0.738)	0.732 (0.540, 0.924)	1.537 (1.161, 2.021)	2.509 (1.364, 3.409)
Chicago	0.495 (0.362, 0.622)	0.779 (0.641, 0.904)	1.314 (1.120, 1.693)	1.705 (1.508, 1.925)
Dallas	0.522 (0.337, 0.691)	0.807 (0.589, 1.007)	1.360 (1.080, 1.762)	1.859 (1.369, 2.415)
Houston	0.525 (0.339, 0.759)	0.753 (0.554, 0.944)	1.471 (1.120, 1.842)	2.737 (1.465, 3.864)
Los Angeles	0.636 (0.487, 0.787)	0.783 (0.617, 0.923)	1.259 (1.044, 1.637)	2.009 (1.306, 2.619)
Miami	0.480 (0.351, 0.632)	0.733 (0.625, 0.827)	1.371 (1.203, 1.617)	2.042 (1.553, 2.503)
New York City	0.782 (0.745, 0.816)	0.895 (0.878, 0.916)	1.129 (1.110, 1.145)	1.312 (1.292, 1.335)
Philadelphia	0.783 (0.734, 0.853)	0.899 (0.845, 0.970)	1.135 (1.059, 1.220)	1.328 (1.272, 1.392)
San Francisco	0.469 (0.275, 0.794)	0.725 (0.476, 0.949)	1.416 (1.059, 1.931)	2.164 (1.303, 2.945)
Washington DC	0.560 (0.372, 0.737)	0.783 (0.578, 0.993)	1.241 (0.941, 1.602)	1.655 (1.510, 1.859)

Table S1: What if people had begun mobility reductions earlier or later? We report the expected ratio of the number of infections predicted under the counterfactual to the number of infections predicted using observed mobility data; a ratio lower than 1 means that fewer infections occurred under the counterfactual. The numbers in parentheses indicate the 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. See Methods M5 for details.

MSA	0%	25%	50%
Atlanta	22.906 (4.869, 39.730)	11.827 (2.304, 22.741)	2.815 (1.317, 4.404)
Chicago	6.844 (4.967, 8.971)	3.567 (2.626, 4.651)	1.511 (1.289, 1.766)
Dallas	22.484 (5.051, 38.465)	7.709 (2.303, 14.337)	1.925 (1.307, 2.676)
Houston	27.363 (8.704, 46.183)	11.184 (3.412, 18.570)	2.720 (1.419, 3.963)
Los Angeles	15.136 (4.024, 24.037)	6.345 (2.071, 10.758)	1.912 (1.271, 2.524)
Miami	13.847 (7.094, 21.344)	5.650 (2.804, 9.470)	1.525 (1.291, 1.938)
New York City	2.160 (2.042, 2.252)	1.803 (1.754, 1.874)	1.283 (1.229, 1.331)
Philadelphia	3.820 (3.385, 4.386)	2.019 (1.837, 2.248)	0.988 (0.933, 1.066)
San Francisco	16.893 (5.656, 30.163)	6.627 (2.284, 11.907)	1.742 (1.151, 2.439)
Washington DC	8.138 (6.409, 10.525)	3.733 (3.050, 4.794)	1.500 (1.324, 1.710)

Table S2: What if the magnitude of mobility reduction changed? Each column represents a counterfactual scenario where the magnitude of mobility reduction is only a some percentage of the observed mobility reduction, i.e., 0% corresponds to no mobility reduction, and 100% corresponds to the real, observed level of mobility reduction. We report the expected ratio of the number of infections predicted under the counterfactual to the number of infections predicted using observed mobility data; a ratio lower than 1 means that fewer infections occurred under the counterfactual. The numbers in parentheses indicate the 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations. See Methods M5 for details.

	Atlanta	Chicago	Dallas	Houst.	LA	Miami	NY	Phila.	SF	DC	**Median**
Full-Service Restaurants	0.730	1.130	0.874	0.948	1.222	1.148	1.707	2.664	1.609	1.147	1.148
Malls	0.869	0.740	1.458	0.947	1.510	0.829	1.333	1.022	0.405	0.879	0.913
Limited-Service Restaurants	0.872	0.752	0.796	0.762	1.008	0.762	1.408	1.301	0.857	0.911	0.864
Other General Stores	0.945	1.037	0.987	0.665	1.577	0.965	1.296	1.252	1.231	1.157	1.097
Fitness Centers	0.563	1.093	0.776	0.608	1.635	0.741	1.052	1.255	1.018	1.375	1.035
Gas Stations	1.876	1.833	1.258	1.827	2.907	1.977	1.857	6.251	1.251	2.706	1.866
Grocery Stores	1.239	2.880	0.784	1.461	2.493	1.638	4.085	11.834	2.275	1.948	2.112
Cafes & Snack Bars	1.163	0.664	0.680	1.364	1.233	1.240	1.397	1.652	1.056	1.022	1.198
Hotels & Motels	0.957	1.447	0.875	1.015	1.378	1.023	1.100	1.252	1.415	1.744	1.176
Religious Organizations	1.487	1.544	1.062	1.061	1.764	0.875	1.695	1.830	1.916	1.759	1.620
Department Stores	1.300	1.173	0.985	0.929	1.734	1.978	1.507	4.924	1.059	1.898	1.404
Hardware Stores	2.677	2.862	1.459	2.262	1.527	0.949	2.809	4.138	2.787	10.062	2.732
Colleges & Universities	2.045	0.673	3.973	1.503	1.809	1.445	1.509	2.928	4.844	8.590	1.927
Offices of Physicians	1.215	0.745	0.707	1.086	1.063	1.112	1.098	1.432	1.049	0.729	1.074
Pharmacies & Drug Stores	1.939	1.297	1.249	0.616	2.065	1.753	2.771	16.693	1.661	1.771	1.762
Used Merch. Stores	1.023	0.746	0.748	0.798	0.846	0.878	0.988	1.573	1.497	0.557	0.862
Sporting Goods Stores	0.728	2.010	0.981	0.949	1.094	0.682	1.132	1.406	1.072	2.223	1.083
Convenience Stores	1.414	1.472	2.710	1.692	1.341	1.398	2.685	6.042	1.344	2.001	1.582
Automotive Parts Stores	1.228	1.602	0.878	1.240	2.208	1.540	1.219	1.928	2.248	1.399	1.470
Pet Stores	1.362	1.120	1.067	1.463	1.260	1.038	2.597	2.047	1.590	0.769	1.311
Median	1.222	1.152	0.983	1.038	1.518	1.075	1.458	1.879	1.38	1.572	

Table S3: Transmission rate disparities at each POI category between income groups. We report the ratio of the average 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) transmission rates. See Methods M5 for details.

	Atlanta	Chicago	Dallas	Houst.	LA	Miami	NY	Phila.	SF	DC	**Median**
Full-Service Restaurants	0.683	1.304	0.888	0.780	0.957	1.089	1.963	2.259	1.093	0.943	1.023
Malls	0.659	0.736	1.497	1.239	0.834	1.161	1.736	1.160	0.127	0.913	1.036
Limited-Service Restaurants	0.891	0.740	0.837	0.888	0.589	0.924	1.980	1.221	0.864	1.453	0.890
Other General Stores	0.668	1.306	0.849	0.794	1.634	0.974	2.089	0.971	1.155	1.315	1.064
Fitness Centers	0.712	1.036	0.620	0.935	0.749	0.818	1.371	1.535	0.848	1.434	0.892
Gas Stations	1.547	2.160	1.600	1.005	1.317	1.766	1.284	2.692	1.980	1.117	1.574
Grocery Stores	0.517	3.804	0.497	1.428	2.469	1.173	4.644	12.266	2.757	2.137	2.303
Cafes & Snack Bars	1.810	0.973	0.764	0.971	0.772	1.764	1.476	2.056	1.045	0.943	1.009
Hotels & Motels	0.834	1.405	1.329	1.180	1.473	1.792	1.258	2.442	0.631	0.869	1.293
Religious Organizations	0.990	1.349	1.061	0.910	1.730	1.645	1.611	2.067	1.650	1.160	1.480
Department Stores	1.642	1.474	0.643	0.692	1.057	4.022	4.784	5.686	1.340	1.664	1.558
Hardware Stores	0.860	3.487	1.674	1.536	2.007	1.006	3.426	2.811	7.348	6.038	2.409
Colleges & Universities	4.880	1.927	3.187	2.454	0.857	8.080	2.580	14.961	9.119	3.409	3.298
Offices of Physicians	1.014	1.393	0.803	1.187	1.041	1.029	1.207	1.504	0.646	0.626	1.035
Pharmacies & Drug Stores	0.824	1.566	0.937	0.725	1.468	2.141	3.404	3.108	1.272	1.464	1.466
Used Merchandise Stores	0.881	0.930	0.527	0.431	1.150	0.643	1.058	2.249	1.345	0.684	0.906
Sporting Goods Stores	0.927	1.661	1.142	0.505	1.789	0.815	1.199	1.078	0.612	1.292	1.110
Convenience Stores	2.061	16.190	1.856	2.757	0.881	1.023	2.814	3.980	1.309	1.864	1.962
Automotive Parts Stores	1.143	2.024	1.000	1.420	2.649	2.220	1.354	2.601	2.262	0.517	1.722
Pet Stores	0.921	1.307	0.577	1.218	1.735	1.018	1.231	9.558	1.135	2.415	1.224
Median	0.906	1.399	0.913	0.988	1.233	1.125	1.674	2.351	1.214	1.304	

Table S4: Transmission rate disparities at each POI category between racial groups. We report the ratio of the average 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) transmission rates. See Methods M5 for details.

Google category	Google description	NAICS categories
Retail & recreation	Restaurants	Full-Service Restaurants
	Cafes	Limited-Service Restaurants
	Shopping centers	Snack and Nonalcoholic Beverage Bars
	Theme parks	Drinking Places (Alcoholic Beverages)
	Museums	Malls, Amusement and Theme Parks
	Libraries	Museums, Libraries and Archives
	Movie theaters	Motion Picture Theaters (except Drive-Ins)
Grocery & pharmacy	Grocery markets	Supermarkets and Other Grocery (except Convenience) Stores
	Food warehouses	
	Farmers markets	Food (Health) Supplement Stores
	Specialty food shops	Fish and Seafood Markets
	Drug stores	All Other Specialty Food Stores
	Pharmacies	Pharmacies and Drug Stores
Parks	Local parks	Nature Parks and Other Similar Institutions
	National parks	
	Public beaches	
	Marinas	
	Dog parks	
	Plazas	
	Public gardens	

Table S5: Mapping of Google mobility data categories to NAICS categories. Google descriptions taken from `https://www.google.com/covid19/mobility/data_documentation.html`.

State	Retail & recreation	Grocery & pharmacy	Parks	Residential
California	0.947	0.834	0.824	0.876
Delaware	0.957	0.847	0.366	0.856
Florida	0.963	0.814	0.905	0.885
Georgia	0.948	0.682	0.442	0.868
Illinois	0.964	0.710	0.512	0.899
Indiana	0.956	0.741	0.070	0.877
Maryland	0.956	0.825	0.351	0.886
New Jersey	0.951	0.720	0.497	0.935
New York	0.958	0.763	0.569	0.909
Pennsylvania	0.971	0.850	0.339	0.875
Texas	0.965	0.789	0.755	0.886
Virginia	0.967	0.840	0.336	0.877
Washington, DC	0.959	0.889	0.841	0.780
West Virginia	0.960	0.740	0.311	0.814
Wisconsin	0.967	0.783	-0.093	0.886
Median	0.959	0.789	0.442	0.877

Table S6: Pearson correlations between the Google and SafeGraph mobility timeseries. We report correlations over the period of March 1–May 2, 2020 for the 15 states that we model. See SI Section S1 for details.

MSA	**# sets**	β_{base}	ψ	p_0
Atlanta	16	0.005 (0.001, 0.012)	4000 (1000, 6000)	2×10^{-4} (5×10^{-5}, 1×10^{-3})
Chicago	3	0.012 (0.005, 0.012)	2000 (2000, 4500)	5×10^{-4} (5×10^{-5}, 5×10^{-4})
Dallas	12	0.002 (0.001, 0.012)	4500 (1000, 5000)	1×10^{-4} (5×10^{-5}, 5×10^{-4})
Houston	9	0.001 (0.001, 0.012)	3500 (1000, 4000)	2×10^{-4} (1×10^{-4}, 1×10^{-3})
Los Angeles	20	0.010 (0.001, 0.012)	2500 (1000, 6000)	2×10^{-4} (2×10^{-5}, 1×10^{-3})
Miami	11	0.005 (0.001, 0.010)	3000 (1500, 4500)	5×10^{-4} (1×10^{-4}, 1×10^{-3})
New York City	7	0.001 (0.001, 0.005)	5500 (3500, 6000)	2×10^{-4} (1×10^{-4}, 1×10^{-3})
Philadelphia	5	0.005 (0.001, 0.010)	2000 (1500, 3000)	5×10^{-4} (1×10^{-4}, 1×10^{-3})
San Francisco	8	0.001 (0.001, 0.012)	4000 (1000, 4500)	2×10^{-4} (1×10^{-4}, 5×10^{-4})
Washington DC	9	0.008 (0.002, 0.012)	4500 (3000, 7000)	5×10^{-5} (2×10^{-5}, 2×10^{-4})

Table S7: Model parameters used for each MSA. # sets counts the number of parameter sets that are within 20% of the RMSE of the best-fit parameter set, as described in Section M4. For each of β_{base}, ψ, and p_0, we show the best-fit parameter set and, in parentheses, the corresponding minimum and maximum within the 20% threshold.

Supplementary figures

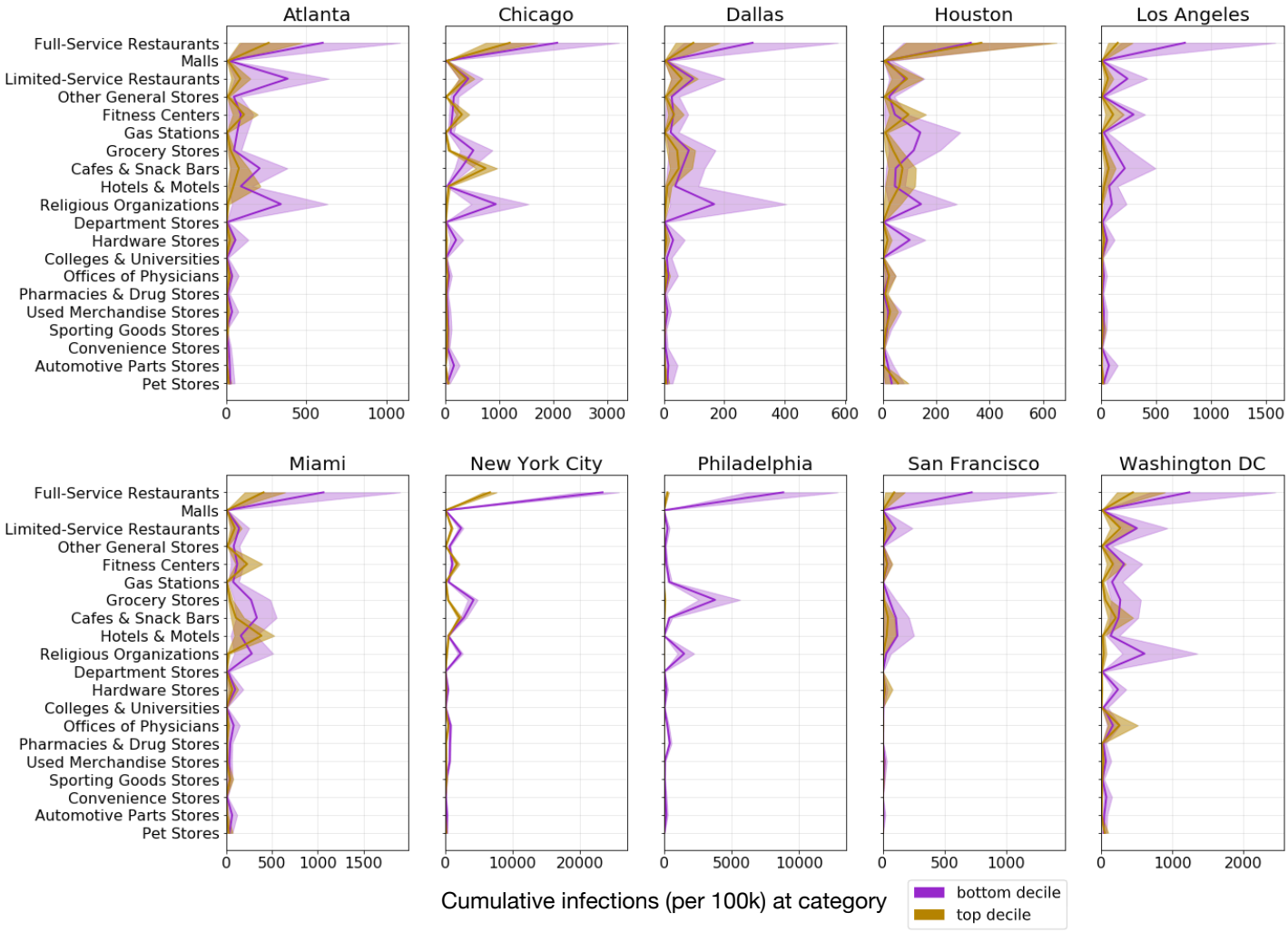


Figure S1: For each POI category, we plot the predicted cumulative number of infections (per 100k population) that occurred at that category for CBGs in the bottom- (purple) and top- (gold) income deciles. Shaded regions denote 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations.

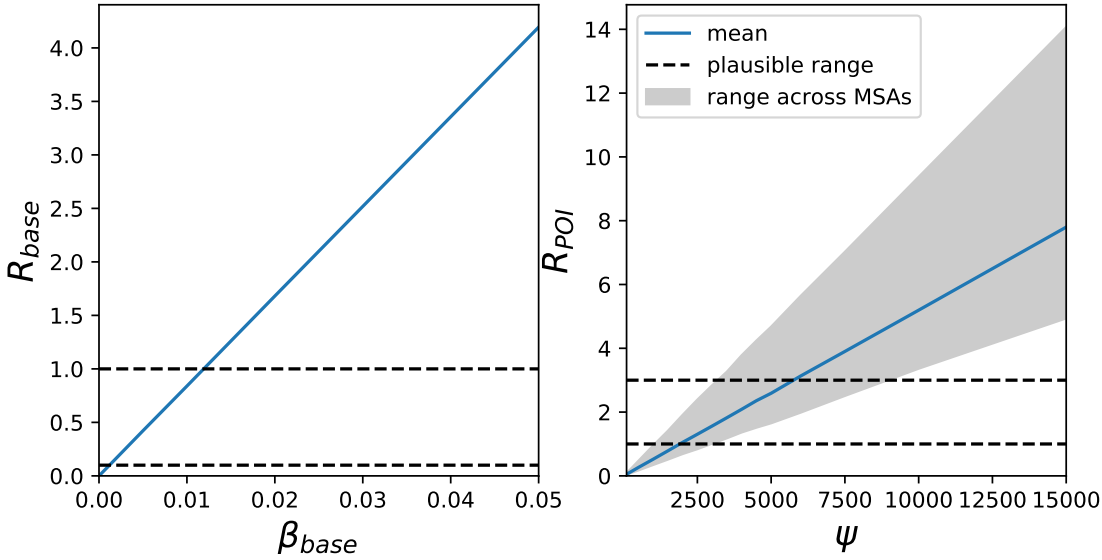


Figure S2: R_{base} and R_{POI} implied by model parameter settings. Dotted black lines denote plausible ranges from prior work, the blue line shows the mean across MSAs, and the grey shaded area indicates the range across MSAs. R_{base} does not vary across MSAs because it does not depend on MSA-specific social activity. See Methods M4.1 for details.

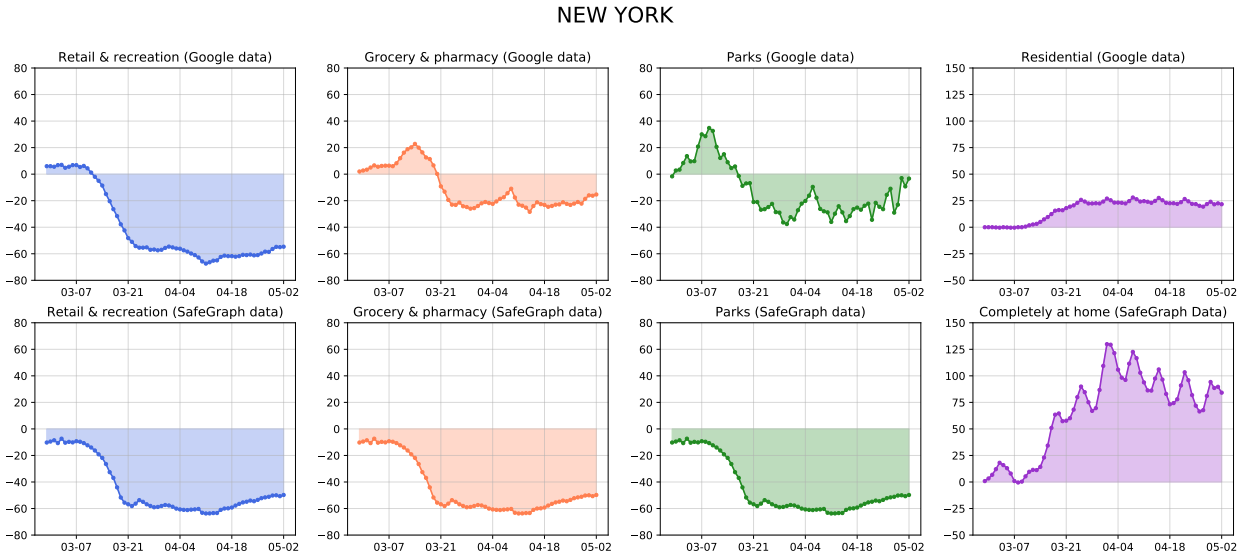


Figure S3: Google versus SafeGraph mobility trends for New York state. The y-axis represents mobility levels compared to baseline activity in January and February 2020. For the categories from left to right, the Pearson correlations between the datasets are 0.96, 0.76, 0.57, and 0.91. See SI Section S1 for details.

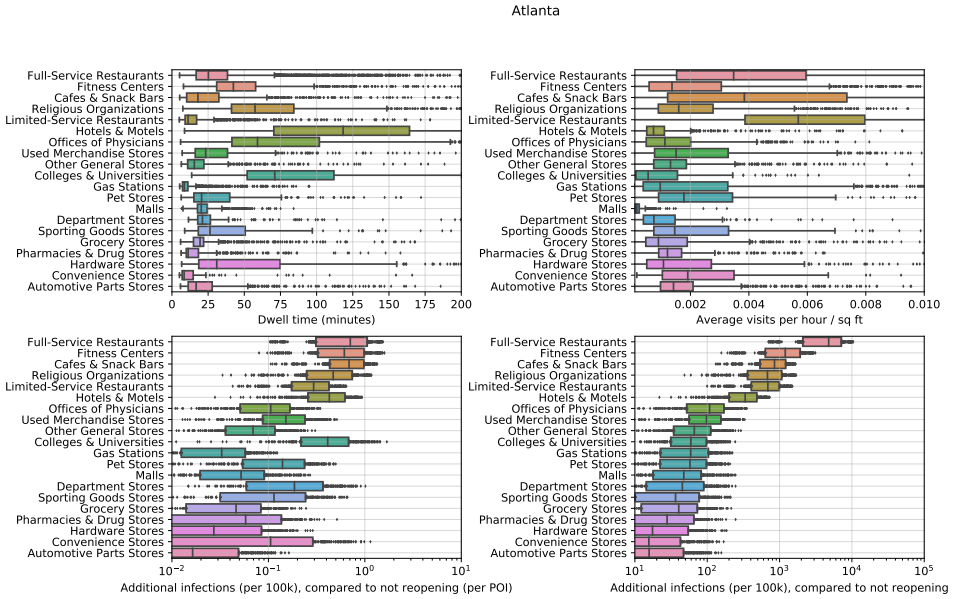


Figure S4: POI attributes in Atlanta. The top two plots show the distribution of dwell time and the average number of hourly visitors divided by the area of the POI in square feet. Each point represents one POI; boxes depict the interquartile range across POIs. The bottom two plots show predictions for the increase in infections (per 100,000 people) from reopening a POI category: per POI (left bottom) and for the category as a whole (right bottom). Each point represents one model realization; boxes depict the interquartile range across sampled parameters and stochastic realizations.

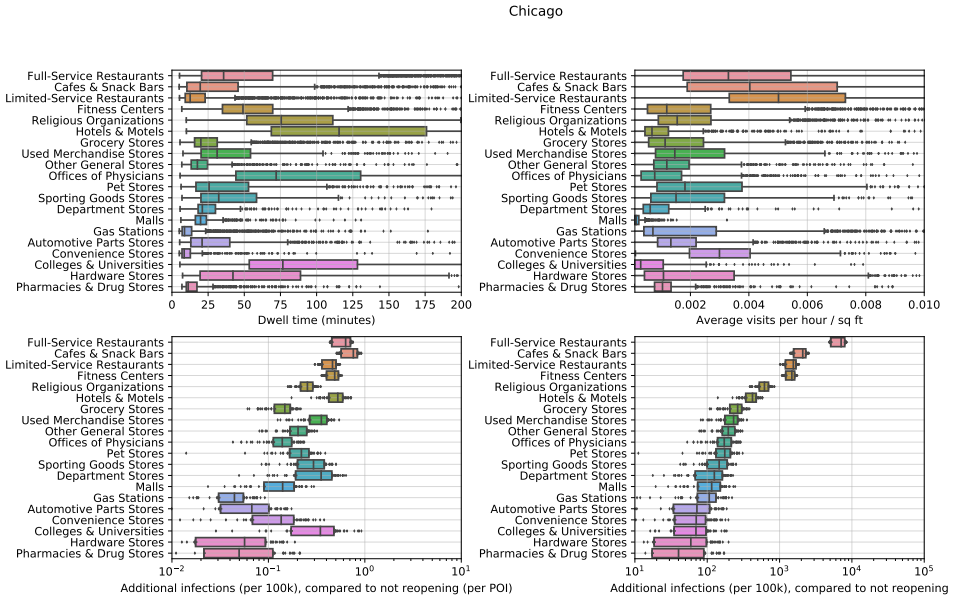


Figure S5: POI attributes in Chicago. See Figure S4 for details.

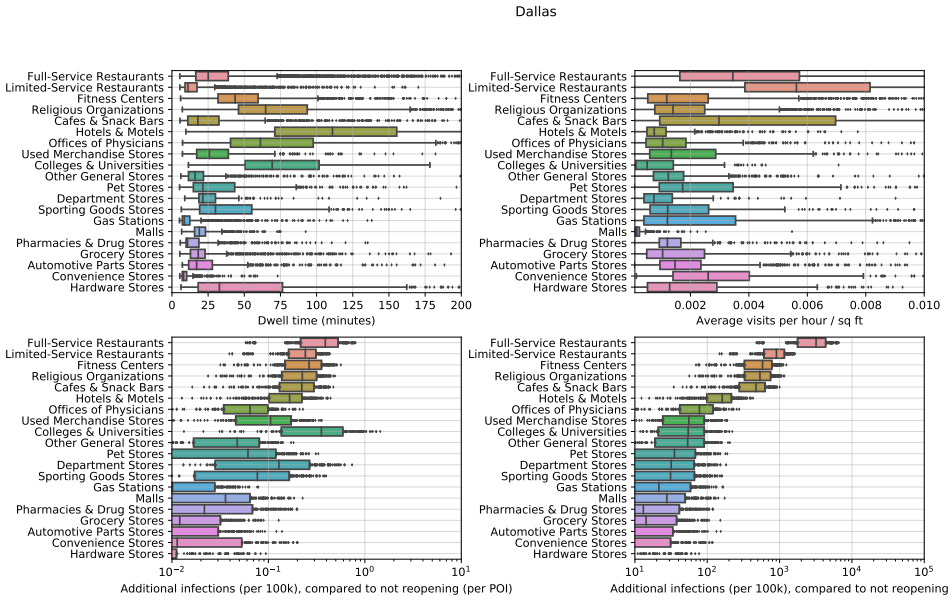


Figure S6: POI attributes in Dallas. See Figure S4 for details.

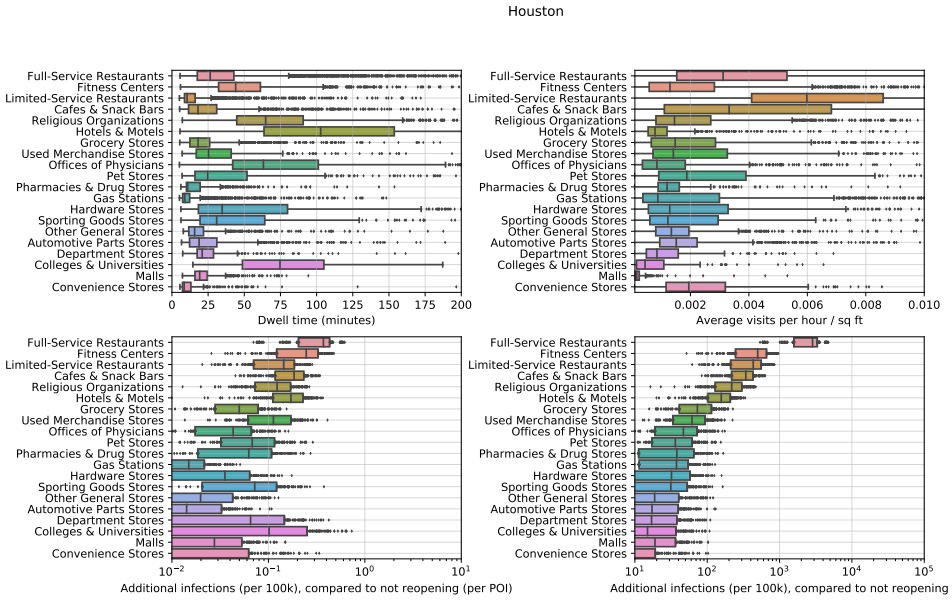


Figure S7: POI attributes in Houston. See Figure S4 for details.

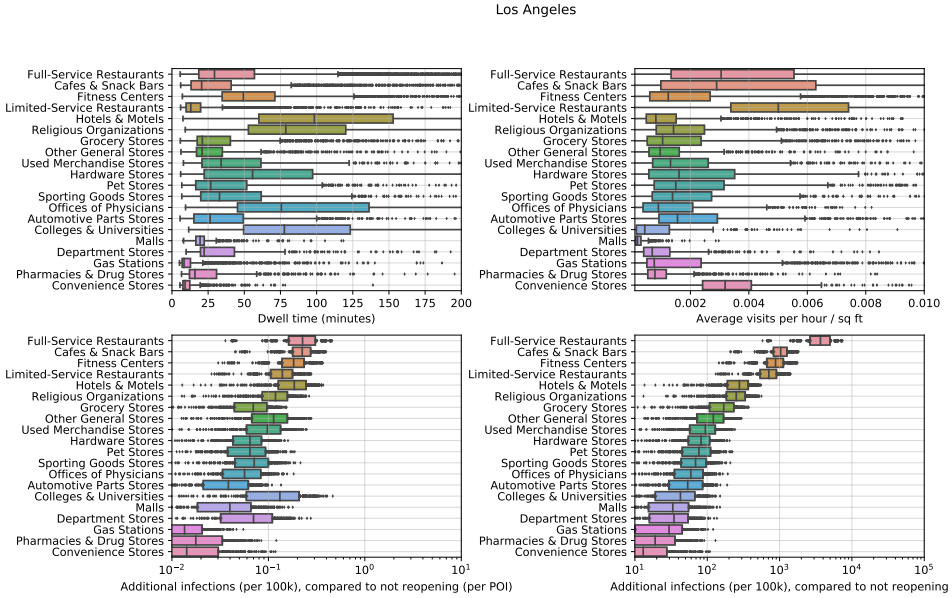


Figure S8: POI attributes in Los Angeles. See Figure S4 for details.

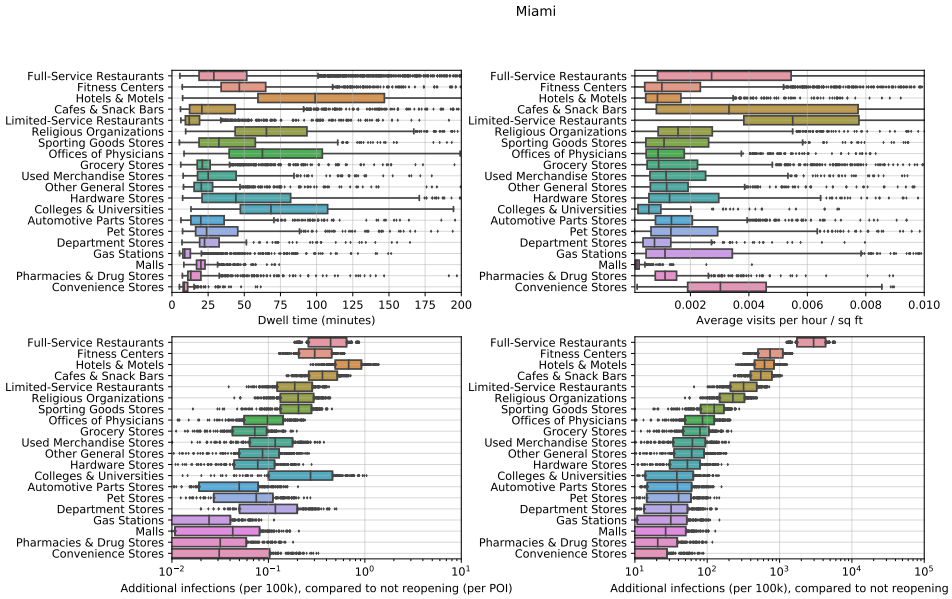


Figure S9: POI attributes in Miami. See Figure S4 for details.

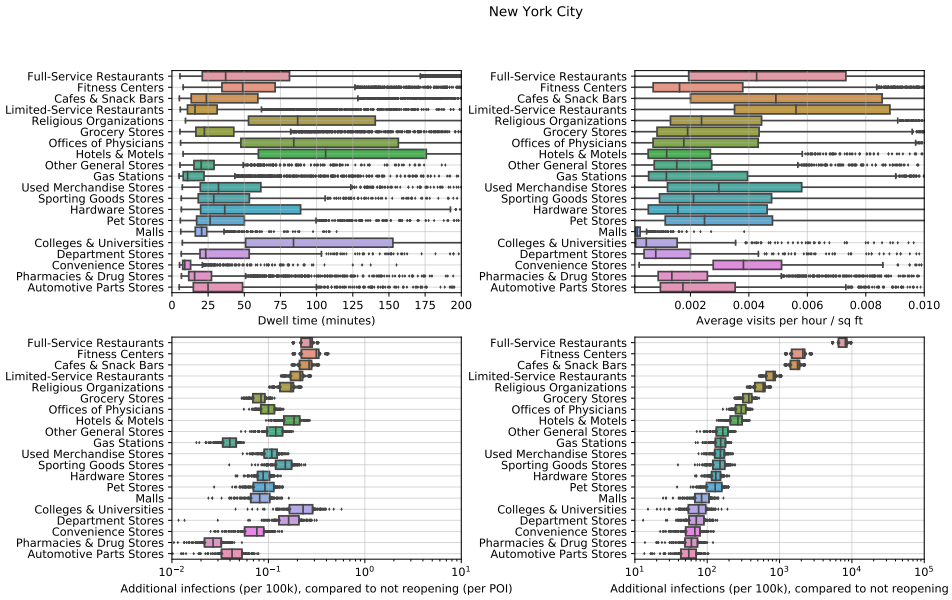


Figure S10: POI attributes in New York. See Figure S4 for details.

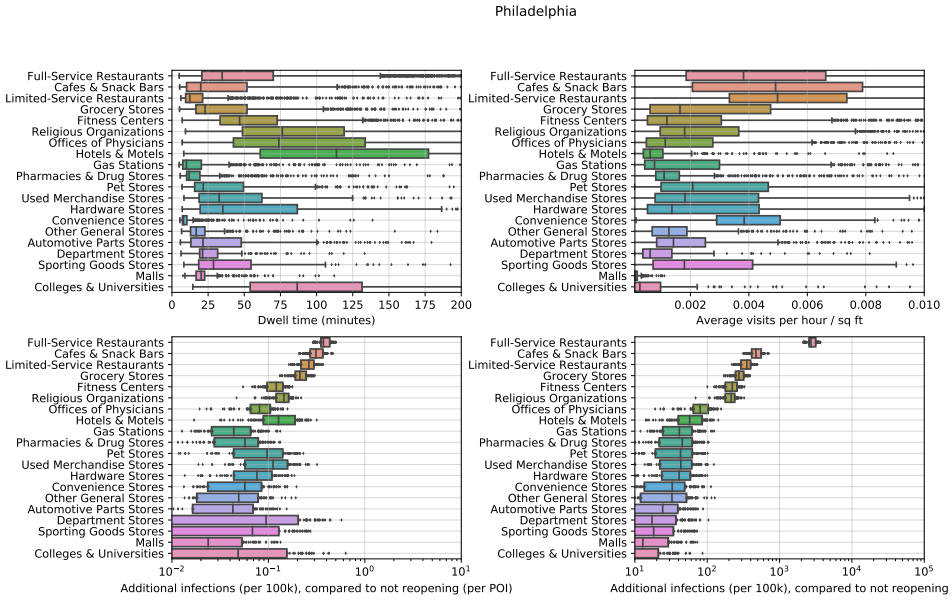


Figure S11: POI attributes in Philadelphia. See Figure S4 for details.

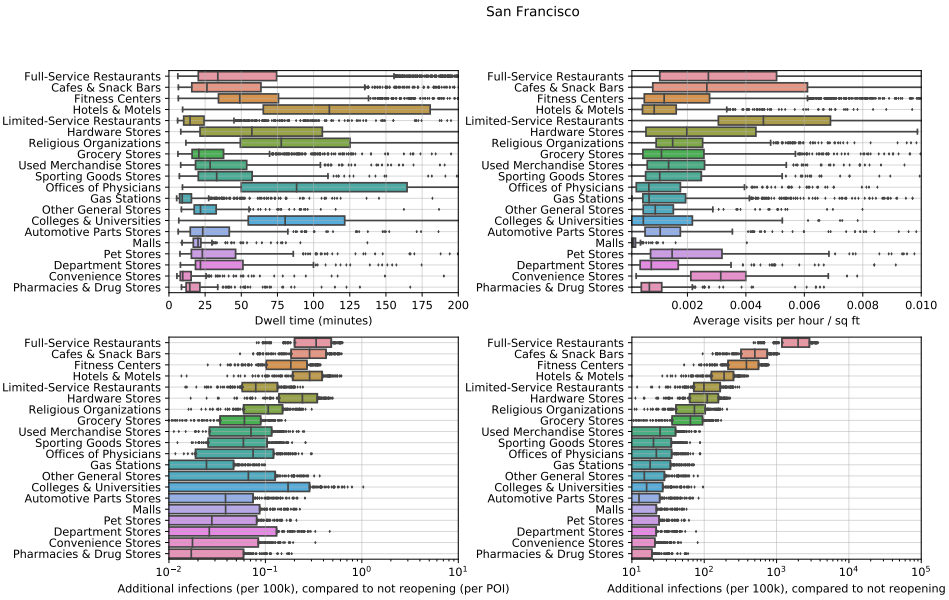


Figure S12: POI attributes in San Francisco. See Figure S4 for details.

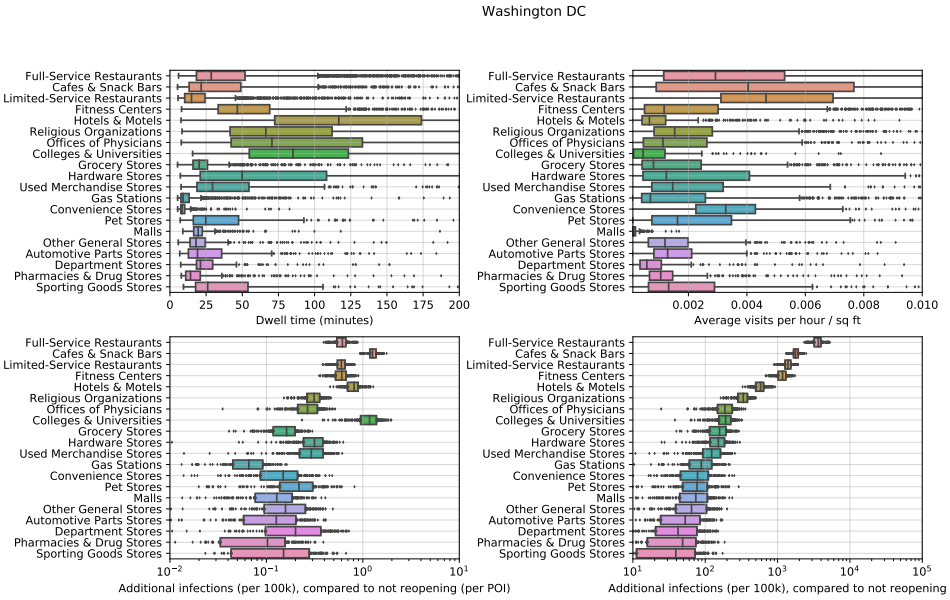


Figure S13: POI attributes in Washington DC. See Figure S4 for details.