

BRIEF COMMUNICATION OPEN



Early pandemic COVID-19 case growth rates increase with city size

Andrew J. Stier¹✉, Marc G. Berman^{1,2} and Luís M. A. Bettencourt^{3,4}

The current outbreak of COVID-19 poses an unprecedented global health and economic threat to interconnected human societies. Strategies for controlling the outbreak rely on social distancing and face covering measures which largely disconnect the social network fabric of cities. We demonstrate that early in the US outbreak, COVID-19 spread faster on average in larger cities and discuss the implications of these observations, emphasizing the need for faster responses to novel infectious diseases in larger cities.

npj Urban Sustainability (2021)1:31; <https://doi.org/10.1038/s42949-021-00030-0>

INTRODUCTION

The coronavirus pandemic of 2019–2021 (COVID-19) is an unprecedented worldwide event. Its speed of propagation, its international reach, and the unprecedented coordinated measures for its mitigation are only possible in a world that is more connected and more urbanized than at any other time in history.

As a novel infectious disease in human populations, COVID-19 has a number of quantitative signatures to its pattern of spread. These signatures make its dynamics more difficult to contain but also easier to understand, compared to well established contagions such as influenza, for the following reasons.

First, because there is no history of previous exposure, all human populations in contact with the virus are (presumably) susceptible. Second, because COVID-19 is a respiratory disease, it is easily transmissible resulting in high reproductive numbers, $R = 2.2$ – 6.5 (refs. ^{1–3}), though considerable uncertainty remains about these estimates. Third, COVID-19 appears to be characterized by reproductive numbers above the epidemic threshold ($R_0 > 1$) everywhere around the world. This, despite slightly reduced transmission associated with higher humidity and temperature^{4–7}. So, while heat and humidity may slow the spread of Covid-19, they alone do not bring the reproductive number below 1 (i.e., the epidemic threshold). In addition, these reproductive numbers are considerably higher than seasonal influenza⁸. Bringing the disease reproductive number below the epidemic threshold ($R \rightarrow R < 1$) is the main goal of all public health interventions; once this is achieved the disease's transmission chain reaction will shut down.

The reproductive number is the product of two factors $R = \beta/\gamma$, the infectious period $1/\gamma$ (a physiological property) and the contact rate β , which is a property of the population, essentially measuring the number of social contacts that can transmit the disease per unit time. Of these, only the contact rate can be changed via public health interventions.

In the absence of a vaccine, social distancing and face coverings remain the only options to slow down the spread of the disease and arrest potential mortality. Governments around the world have enacted aggressive policies, including "shelter in place" and emergency closures of all non-essential services, which carry severe economic and social consequences. However, there is still a

great deal of uncertainty as to how strong social distancing recommendations must be or how long they must last^{9,10}. In addition, regional and local variation in the severity of and compliance with restrictions are likely to impact individual cities differently.

Cities are predicated on extensive and intense face-to-face socioeconomic interactions. These interactions make cities sources of social and economic opportunity but also increase the possibility of infectious disease transmission. Many measurable properties of cities—from the size of their economies, to their crime rates, to their miles of road length—are mediated by socioeconomic interactions and are subject to well known scaling effects with population size predicted by urban scaling theory¹¹. All of these relationships are tied to socioeconomic networks with average degree (number of social connections per capita, k) that increases approximately as a power law of city size $k(N) = k_0 N^\delta$, with $\delta \approx 1/6$, and city size, N , as measured by population size^{11,12}. This is thought to then cause scaling of crime rates, infectious disease spread, number of patents, etc., with city size (i.e. population). We will refer to population/city size as the number of people living in a Metropolitan Statistical Area (MSA).

For functional effects, including epidemics, occurring against the ecological backdrop of cities, the underlying concept of a city is that of daily networks of socioeconomic interactions. These are most readily defined as integrated labor markets (commuting zones), which in the US are defined by the Census since the 1960s as MSAs¹³. These are the "city" definitions we use in our empirical analyses. For example, the Chicago MSA represents an integrated labor market, which includes 14 counties, across three states (<https://www.census.gov/programs-surveys/metro-micro/about/delineation-files.html>). The MSA definitions are revised annually by the US Census to reflect highly socioeconomically linked urban places, and are the result of theory and analysis at the US Census. Though commuting ties have been used to link places, they also encompass many other activities such as shared real estate and labor markets, transportation networks, and environmental protections. For these reasons, there is an effort to (re)create analogous or improved urban functional definitions worldwide, led by the Organization for Economic Co-operation and Development (OECD)¹⁴. While adopting any other definition for the same effect would require extraordinary arguments, other ways of

¹Department of Psychology, University of Chicago, Chicago, IL, USA. ²The University of Chicago Neuroscience Institute, Chicago, IL, USA. ³Mansueto Institute for Urban Innovation, University of Chicago, Chicago, IL, USA. ⁴Department of Ecology & Evolution, University of Chicago, Chicago, IL, USA. ✉email: andrewstier@uchicago.edu

demarcating the boundaries of cities have been proposed. For example, morphological urban areas are defined via satellite imagery by identification of contiguous areas of development that can be distinguished from peripheral rural areas and areas of lower-density development¹⁵. However, it is as-of-yet unclear whether these alternate methods for demarcating city boundaries are superior in any way. In addition, defining city boundaries by integrated labor markets captures a variety of social interactions relevant to disease transmission beyond labor market interactions. For example, the variety of restaurants types has been shown to be greater in larger cities (integrated labor markets)¹⁶. This suggests that consumer patterns of choosing which restaurants to visit (and socialize at) is meaningfully related to city size. Consequently, we expect well principled (see below) analyses that use MSAs as the definition of a city to capture the best approximation to the vast milieu of social interactions that create urban social life and form the basis for inter-individual disease transmission.

Urban scaling theory, and the observed scaling relationships it explains, are based on a number of parameters which are general to cities and more broadly to human settlements across history¹⁷. These parameters include, but are not limited to, transportation costs, the economic benefits of social interactions, the typical distance traveled per person, the total population, the embedding of transportation networks into the physical area of settlements, and the probability of different types of social and economic interactions¹¹. These factors combine into a cost-benefit calculus for individuals embedded in the social and built spaces of cities. In turn, this allows for the calculation of expected values for many different urban quantities as functions of city size (i.e. they all scale with city population size).

It is important to recognize that scaling relationships between urban quantities and city population size give us a way to calculate the "ecological" effects of urban networks—such as disease transmissibility—but that these quantities are not a model for the behavior at smaller scales, such as those of individuals or neighborhoods. Behaviors at those smaller scales may vary from the mean substantially because of contextual effects. This is a fundamental point that we must heed in order to avoid "ecological fallacies"¹⁸, which is why we do not make claims at that level of analysis. A few additional considerations help make these points more explicit.

First, the mathematical models that explain urban scaling explicitly identify social network densification as the primary mechanistic driver of many measurable properties of cities, including innovation, economic output, and crime. Consequently, these models explicitly link average outcomes for entire cities to the local dynamics of social interactions that may naturally transmit a disease, as averages over these local dynamics. Thus, while a particular exposure to COVID-19 is the result of an individual's history of social contacts, the measurements of the reproductive number and case growth rates are always a population average over these histories. These measures account for the average number of transmission events that infected individuals originate on average in that population. For these reasons, epidemics are often treated as the paradigm for collective behavior on large ecological scales.

Second, urban scaling theory has been shown to apply to cities and settlements, across cultures¹¹, and across historical time scales^{19–21}. While the generality of urban scaling theory across time and culture always allows for places and neighborhoods within cities to deviate from the mean and follow more specific patterns of organization, it does suggest that the average densification of social interactions with the population size of cities is a general organizational principle of (relatively) self-contained human cities and settlements¹¹.

Consequently, the densification of social networks in larger cities is expected to be relevant to COVID-19 disease transmission,

which is known to spread primarily via social contacts^{10,22}. Specifically, as the average contact rate is proportional to degree $\beta \propto k(N)$ (see "Methods"), we expect that initial growth rates of COVID-19 cases to be higher in larger cities (see "Methods"). For example, based on data of mobile phone social networks in Portugal¹², people living in a city of 500,000 people have, on average, 11 people in their mobile phone social network, while people living in a city of 5,000,000 people have, on average, 15. This means that individuals living in larger cities have more dense social networks. Therefore, we expect the contact rate (and case growth rate) to be larger in a city of five million compared to a city of five hundred thousand. This has been previously demonstrated with early epidemic case growth of AIDS, which spreads through the sexual network subset of social networks²³. However, we do not expect this to hold later in the pandemic, as social distancing and lockdowns disrupt the social networks of cities, specifically to reduce contacts, and this modifies the urban scaling laws. Later in the pandemic, when mitigation measures have been implemented, we expect neighborhood and individual variations in social network size after social distancing [during complete lockdowns this is roughly equivalent to household size, which has been demonstrated to be a significant risk factor for COVID-19 transmission^{24,25}], face covering compliance, and availability of work-from-home²⁶ to influence transmission rates. In addition, neighborhood and individual differences in co-morbidities are likely relevant to understanding the variance in morbidity and mortality outcomes within cities, but are not accounted for here where the focus is on transmission rates. As such, we first focused only on the beginning of the pandemic because that is the time period that was least affected by interventions to curb the spread of COVID-19, which would disrupt the social networks and socioeconomic connections that relate COVID-19 transmission rates to city size.

RESULTS

Scaling of case growth rates

This is exactly what is found empirically. Figure 1a shows the growth rate of COVID-19 cases from March 14 to March 24, the first 11 days in which city level reported case numbers are available, vs. city size (population). While larger cities were faster to adopt preventative COVID-19 measures²⁷, the first stay at home order in the US was effective on March 17. As a result of the incubation period for SARS-COV-2 (ref. ²⁸), this period most likely captures case growth before large-scale behavior modification as instituted by local preventative physical/social distancing and lock-down policies. This demonstrates that, in the US, the early part of the pandemic was characterized by increasing case growth rates with city size, consistent with the expected scaling exponent of $\delta \sim 1/6$. This scaling exponent matches other socioeconomic scaling phenomena such as patents, wealth, and information flow^{11,12,29,30}. In contrast, Fig. 2 shows that this period of higher case growth rates in larger cities was restricted to the first few months of the pandemic. By the end of June, the scaling exponent δ is consistent with 0 (i.e. no changes in case growth rate by city size).

Since these results are based on the temporal growth rate of cases they do not depend on the total number of tests administered. Thus, our results are robust to discrepancies between positive tests and actual cases as well as absolute differences in testing capacity between cities, so long as levels of testing are consistent in each place over time. While the observed scaling of case growth rates with city size may be due to larger cities ramping up testing faster than smaller cities, we are not aware of any county or city level testing data for this early period of the US outbreak that could be used to determine the influence of testing availability on these results.

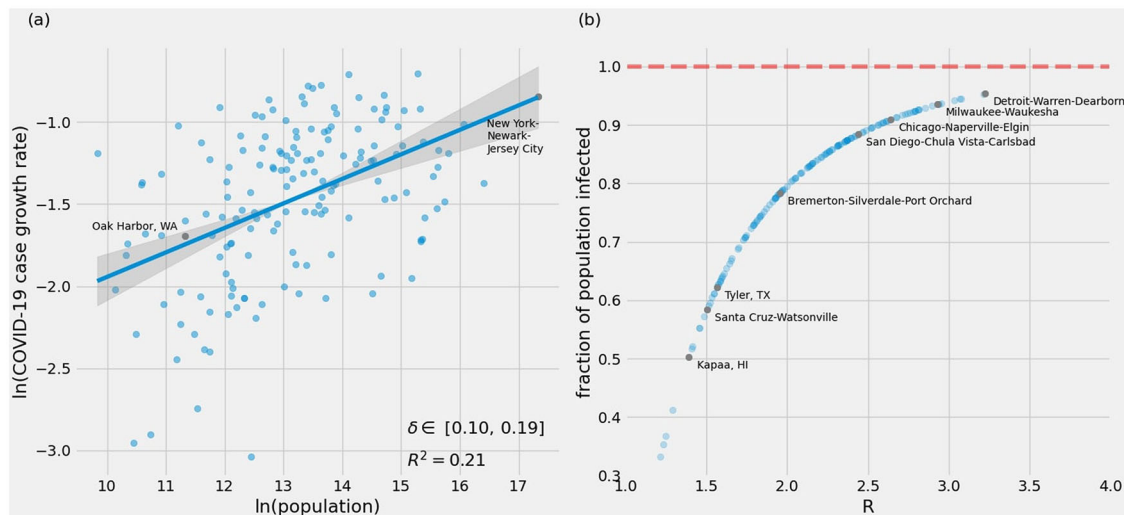


Fig. 1 COVID-19 cases reported by state health agencies grow faster in larger cities. **a** Estimated exponential daily growth rates of COVID-19 in US Metropolitan Areas (MSAs). These estimates were made with the assumption that cities were experiencing exponential growth of cases. The growth rate of COVID-19 cases is approximately 2.4 times faster in New York–Newark–Jersey compared to Oak Harbour, WA. **b** In the absence of effective controls, larger cities are expected to have more extensive epidemics than smaller cities (Eq. (1)). Higher values of R result in a greater percentage of the population eventually infected, unless this effect is curbed by controls that reduce the social contact rate. The translation of growth rates into reproductive numbers was obtained using an infectious period of $1/\gamma = 4.5$ days. These estimated values of R are high in some cases (e.g., New York City) compared to reports in other situations and may in part be the result of the acceleration of testing in larger cities and specific places. Each estimate comes from the 11 previous (inclusive) days of data. Shaded regions represent 95% confidence intervals from OLS scaling fits.

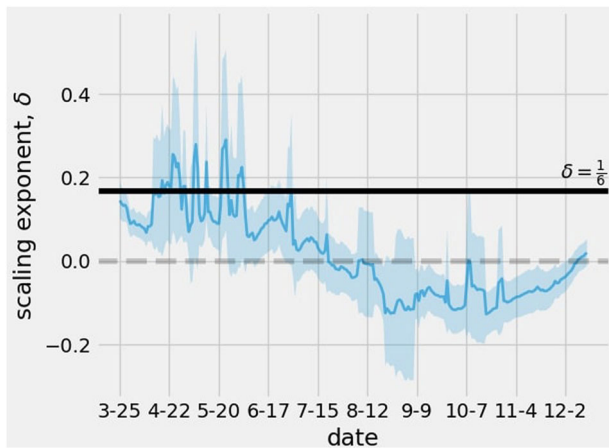


Fig. 2 Estimated scaling exponents (δ) for the scaling of COVID-19 case growth rates with city size in the US over the course of the pandemic. The early pandemic is characterized by faster growth of cases in larger cities, consistent with the predicted scaling exponent $\delta = 1/6$. This is followed by decreasing scaling towards no scaling $\delta = 0$, later in the pandemic as social distancing, lockdowns, and mask wearing disrupt the social/transmission networks of cities. Shaded regions indicate 95% confidence intervals.

Urban scaling theory only explains some (though a significant amount) of the variance in growth rates of COVID-19 cases between cities ($R^2 = 0.21$), i.e. inter-city variation. Besides the expected uncertainty in growth rate estimations, much of this variance is likely due to unique characteristics of each city that influences the speed with which COVID-19 spreads³¹. For example, early pandemic hotspots in New York City were generally in dense working-class and middle-income neighborhoods, while hotspots in Chicago were generally in the city's most vulnerable, low-income neighborhoods³² (though, see ref. ³³ for evidence that

early pandemic impacts were worst in New York City zip codes with the highest social vulnerability). These unique characteristics of individual cities and the neighborhoods within them could include local differences in the ramping up of testing capacity, international connectedness of each city, the type of available public transportation, the distribution of job types across income brackets, and the amount of long distance business travel that each city depends on.

While other research has implicated population density in the spread of COVID-19 (refs. ^{34,35}), urban scaling theory and empirical evidence suggest that population density increases $\sim N^{1-\alpha}$ ($\alpha = 2/3$) with city size³⁶. The observed scaling exponent of COVID-19 case growth of $[0.11, 0.20] \sim \delta = 1/6$ is thus inconsistent with a pure population density effect (i.e., $1/6 = \delta < \alpha = 2/3$, see Supplementary Fig. 6). Instead we associate the scaling relationship observed here with the density of socioeconomic interactions³⁷ which increase slower than population density with city size. These predictions of urban scaling theory have been confirmed in analyses showing that population density at the county level is not a significant predictor of cumulative COVID-19 caseload after accounting for county population²⁴. In summary, a population density account would predict a much more intense and faster spread of COVID-19 in more dense cities than what urban scaling theory predicts based on the growth of socioeconomic interactions as cities increase in population.

A larger reproductive number for COVID-19 in larger cities has two important consequences^{38–41}. First, the reproductive number, R , sets a finite threshold for how an epidemic outbreak propagates in a population, just like the branching rate in a chain reaction^{42,43}. For $R < 1$, an introduced disease will die off because transmission will be dampened. For $R > 1$, disease transmission will be amplified and result in an epidemic where the disease is transmitted quickly to almost everyone (see Fig. 1B). Because we expect $R \sim N^\delta$ to increase with city size, we expect larger cities to be more susceptible to contagious diseases, but also to the spread of information (see below), which we believe is mediated by the same mechanism, i.e. socioeconomic interactions.

Second, the size of an epidemic outbreak, as measured by the percent of the population that becomes infected, is also related to the reproductive number. In complex epidemic models, this needs to be computed numerically, but for a simple Susceptible-Infected-Recovered (SIR) model⁴² we can integrate the dynamics and write the explicit expression (see “Methods”)

$$\frac{S_{\infty}}{S_0} = e^{-R(1-S_{\infty}/N)} \quad (1)$$

where S_0 is the initial susceptible population size (before the outbreak) and S_{∞} is the (smaller) final population of susceptible people. A larger $R \sim N^{\delta}$ leads to more extensive epidemics. The percent of people infected at the end of the outbreak is $1 - S_{\infty}/N$ which is larger in populations with larger R , such as in larger cities. In addition, the vaccination rate, p_R , or the fraction of the population that must be removed from the susceptible class when $R > 1$ to stop the outbreak⁴⁴ is also expected to be dependent on population size. In the SIR model, this is simply $p_R = 1 - 1/R$. As cities get larger, the fraction of individuals who must be vaccinated or, alternatively, follow social distance and face covering policies must also increase (see Supplementary Fig. 2).

DISCUSSION

These observations have a number of implications that can inform evolving national, regional, and local responses to the outbreak of COVID-19 and help to shape response plans for future, novel infectious diseases. First, it is particularly important for larger cities to be able to act more quickly to contain surges in cases. Second, measures taken to contain outbreaks^{9,10} will impact cities differently based on city size. From the perspective of containing the outbreak, larger cities require faster responses, which could consist of a number of proven policies^{9,10}, that can quickly reduce R below the epidemic threshold. These distinctions may help to bring more nuance to ongoing strategies for suppression and control of COVID-19, including gradually restoring socioeconomic activity in context appropriate and safe ways.

Because of their higher network density, insufficient social distancing or face covering compliance in larger cities may lead to bigger outbreaks and establish reservoirs for the disease, which can continue to create introductions elsewhere^{45,46}. These dynamics may also play out within cities, though we caution in ascribing causes to neighborhood level variance in transmission rates from the present data. The results presented here suggest that communities in which people interact more densely from the perspective of disease transmission (e.g., downtowns) may similarly act as contagion reservoirs which may prolong the duration of outbreaks and potentially create secondary reinfection waves. However, the factors that lead to a higher density of social interactions in one community, for example, household size^{24,25} or employment in the service industry or essential worker status²⁶, can often be the result of long term structural inequalities^{47–50}. These factors must be accounted for when attempting to explain causal sources of differences in transmission rates within cities and in planning responses to outbreaks. In summary, our results apply to comparisons across cities, and not to neighborhoods or areas within individual cities. That is an important question, but one which these data and analyses cannot address.

Finally, as strategies for controlling Covid-19 continue to evolve, it is critical to keep in mind that almost everything that we appreciate about urban environments, including their economic prowess, their ability to innovate, and their role in their inhabitants social and mental health, is predicated on network effects mediated by socioeconomic interactions. The ability to succeed against a fast emerging epidemic like COVID-19 depends on preserving person-to-person connectivity (e.g., through technology), while stopping disease transmission. Establishing safe types of socioeconomic

contact is therefore paramount so that we can succeed in controlling COVID-19 while maintaining livelihoods, socioeconomic capacity, and mental health.

The higher socioeconomic connectivity of larger cities in a fast urbanizing world makes containing emergent epidemics harder. But the density of socioeconomic connections in cities can also facilitate the faster spread of information, social coordination, and innovations necessary to stop the spread of COVID-19. This information and associated actions can easily spread more rapidly than the biological viral contagion. To fight an exponential, we need to create an even faster exponential!

METHODS

Calculation of case growth rates

County level daily data from March 14 to December 14 (inclusive) were aggregated to the city level (MSAs, which are integrated labor markets) using delineation files from the US Office of Budget and Management (<https://www.census.gov/programs-surveys/metro-micro/about/delineation-files.html>). MSAs are groups of counties (in their entirety, i.e., not parts of counties), which make up an interconnected labor market. To our knowledge, the data covering March 14 through March 24 represent the first 10 days in which aggregated county level case statistics were available for the entire country⁵¹. Previously the few states that published county level case data did so on state health department websites and there was no central repository for COVID case data in the US. After preprocessing (see Supplementary Materials) we subtracted deaths from cases in each city and found the exponential growth rate r for each time-series of active cases (see Supplementary Material). Finally we plotted the natural logarithm of r and the natural logarithm of city population from 2018 census estimates⁵¹, and performed an ordinary least-squares linear regression to determine the slope of the scaling line.

DATA AVAILABILITY

The data generated and analyzed during this study are described in the following data record: <https://doi.org/10.6084/m9.figshare.14199629> (ref. ⁵¹). All data are openly available online at the following places. The testing data are contained in all csv files in GitHub at <https://github.com/tomquise/covid19-data/tree/master/data/csv>. Population data and Delineation data can be downloaded from The United States Census Bureau. The Population data are contained in the file “co-est2018-alldata.csv”, available at <https://www2.census.gov/programs-surveys/popest/datasets/2010-2018/counties/totals/>, and the Delineation data are contained in the file “list1_Sep_2018.csv”, available at <https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html>.

Received: 10 September 2020; Accepted: 24 April 2021;

Published online: 28 June 2021

REFERENCES

- Liu, Y., Gayle, A. A., Wilder-Smith, A. & Rocklöv, J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J. Travel Med.* **27**, <https://doi.org/10.1093/jtm/taaa021> (2020).
- Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* **395**, 689–697 (2020).
- Yang, S. et al. on behalf of COVID-19 evidence and recommendations working group. Early estimation of the case fatality rate of COVID-19 in mainland China: a data-driven analysis. *Ann. Transl. Med.* **8**, 128. <https://doi.org/10.21037/atm.2020.02.66> (2020).
- Awasthi, R. et al. Temperature and humidity do not influence global COVID-19 incidence as inferred from causal models. Preprint at <https://www.medrxiv.org/content/early/2020/06/29/2020.06.29.20142307> (2020).
- Auler, A., Cássaro, F., da Silva, V. & Pires, L. Evidence that high temperatures and intermediate relative humidity might favor the spread of COVID-19 in tropical climate: a case study for the most affected Brazilian cities. *Sci. Total Environ.* **729**, 139090 (2020).
- Wang, J., Tang, K., Feng, K. & Lv, W. High temperature and high humidity reduce the transmission of COVID-19. *SSRN Electron. J.* <https://www.ssrn.com/abstract=3551767> (2020).

7. Mecenas, P., Bastos, R. T. d. R. M., Vallinoto, A. C. R. & Normando, D. Effects of temperature and humidity on the spread of COVID-19: a systematic review. *PLoS ONE* **15**, 1–21 (2020).
8. Biggerstaff, M., Cauchemez, S., Reed, C., Gambhir, M. & Finelli, L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. *BMC Infect. Dis.* **14**, 480 (2014).
9. Glaeser, E. L., Jin, G. Z., Leyden, B. T. & Luca, M. National Bureau of Economic Research. *Learning from Deregulation: The Asymmetric Impact of Lockdown and Reopening on Risky Behavior during COVID-19*. Tech. Rep. (2020).
10. Gatto, M. et al. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. *Proc. Natl Acad. Sci. USA* **117**, 10484–10491 (2020).
11. Bettencourt, L. M. The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
12. Schläpfer, M. et al. The scaling of human interactions with city size. *J. R. Soc. Interface* **11**, 20130789 (2014).
13. Berry, B. J., Goheen, P. G. & Goldstein, H. *Metropolitan Area Definition: A Re-Evaluation of Concept and Statistical Practice* Vol. 28 (US Bureau of the Census, 1969).
14. Dijkstra, L., Poelman, H. & Veneri, P. *The EU-OECD Definition of a Functional Urban Area*. OECD Regional Development Working Papers, No. 2019/11, OECD Publishing, Paris, <https://doi.org/10.1787/d58cb34d-en>. (2019).
15. Taubenböck, H. et al. A new ranking of the world's largest cities—do administrative units obscure morphological realities? *Remote Sens. Environ.* **232**, 111353 (2019).
16. Schiff, N. Cities and product variety: evidence from restaurants. *J. Econ. Geogr.* **15**, 1085–1123 (2014).
17. Lobo, J., Bettencourt, L. M., Smith, M. E. & Ortman, S. Settlement scaling theory: bridging the study of ancient and contemporary urban systems. *Urban Stud.* **57**, 731–747 (2020).
18. Loney, T. & Nagelkerke, N. J. The individualistic fallacy, ecological studies and instrumental variables: a causal interpretation. *Emerg. Themes Epidemiol.* **11**, 18 (2014).
19. Ortman, S. G., Cabaniss, A. H. F., Sturm, J. O. & Bettencourt, L. M. A. Settlement scaling and increasing returns in an ancient society. *Sci. Adv.* **1** <https://advances.sciencemag.org/content/1/1/e1400066> (2015).
20. Ortman, S. G. et al. Settlement scaling and economic change in the central Andes. *J. Archaeol. Sci.* **73**, 94–106 (2016).
21. Ortman, S. G., Cabaniss, A. H. F., Sturm, J. O. & Bettencourt, L. M. A. The pre-history of urban scaling. *PLoS ONE* **9**, 1–10 (2014).
22. Glaeser, E., Gorbach, C. & Redding, S. *How Much Does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities*. Tech. Rep. w27519 (National Bureau of Economic Research, Cambridge, MA, 2020).
23. Antonio, F. J., de Picoli Jr, S., Teixeira, J. J. V. & dos Santos Mendes, R. Growth patterns and scaling laws governing aids epidemic in Brazilian cities. *PLoS ONE* **9**, e111015 (2014).
24. Hamidi, S., Sabouri, S. & Ewing, R. Does density aggravate the COVID-19 pandemic? Early findings and lessons for planners. *J. Am. Plann. Assoc.* **86**, 495–509 (2020).
25. Nash, D. et al. Household factors and the risk of severe COVID-like illness early in the US pandemic. Preprint at <https://www.medrxiv.org/content/early/2020/12/04/2020.12.03.20243683> (2020).
26. Bartik, A. W., Cullen, Z. B., Glaeser, E. L., Luca, M. & Stanton, C. T. *What Jobs Are Being Done at Home during the COVID-19 Crisis? Evidence from Firm-Level Surveys*. Tech. Rep. (National Bureau of Economic Research, 2020).
27. Brandtner, C., Bettencourt, L. M., Berman, M. G. & Stier, A. J. Creatures of the state? Metropolitan counties compensated for state inaction in initial US response to COVID-19 pandemic. *PLoS ONE* **16**, e0246249 (2021).
28. Lauer, S. A. et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).
29. Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl Acad. Sci. USA* **104**, 7301–7306 (2007).
30. Bettencourt, L. M., Lobo, J. & Strumsky, D. Invention in the city: increasing returns to patenting as a scaling function of metropolitan size. *Res. Policy* **36**, 107–120 (2007).
31. Alves, L. G. A., Ribeiro, H. V., Lenzi, E. K. & Mendes, R. S. Distance to the scaling law: a useful approach for unveiling relationships between crime and urban metrics. *PLoS ONE* **8**, 1–8 (2013).
32. Maroko, A. R., Nash, D. & Pavilonis, B. T. COVID-19 and Inequity: a Comparative Spatial Analysis of New York City and Chicago Hot Spots. *J. Urban Health.* **97**, 461–470. <https://doi.org/10.1007/s11524-020-00468-0> (2020).
33. McPhearson, T., et al. Pandemic Injustice: Spatial and Social Distributions of COVID-19 in the US Epicenter. *J. Extreme Events.* 2150007, (2020).
34. Kang, D., Choi, H., Kim, J.-H. & Choi, J. Spatial epidemic dynamics of the COVID-19 outbreak in China. *Int. J. Infect. Dis.* **94**, 96–102 (2020).
35. Rocklöv, J. & Sjödin, H. High population densities catalyse the spread of COVID-19. *J. Travel Med.* **27**, taaa038 (2020).
36. Bettencourt, L. M. The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
37. Kuchler, T., Russel, D. & Stroebel, J. The Geographic Spread of COVID-19 Correlates with Structure of Social Networks as Measured by Facebook. Tech. Rep. (National Bureau of Economic Research, 2020).
38. Dalziel, B. D. et al. Urbanization and humidity shape the intensity of influenza epidemics in US cities. *Science* **362**, 75–79 (2018).
39. Bettencourt, L. M., Lobo, J. & Strumsky, D. Invention in the city: increasing returns to patenting as a scaling function of metropolitan size. *Res. Policy* **36**, 107–120 (2007).
40. Feldman, M. P. & Audretsch, D. B. Innovation in cities: science-based diversity, specialization and localized competition. *Eur. Econ. Rev.* **43**, 409–429 (1999).
41. Acs, Z. J. in *Urban Dynamics and Growth: Advances in Urban Economics*, (eds Capello, R. & Nijkamp, P.) Vol. 266 of *Contributions to Economic Analysis* 635–658 (Elsevier, 2004).
42. Anderson, R. M., Anderson, B. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Reprinted, 2010).
43. Bettencourt, L. M., Cintrón-Arias, A., Kaiser, D. I. & Castillo-Chávez, C. The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models. *Physica A* **364**, 513–536 (2006).
44. Randolph, H. E. & Barreiro, L. B. Herd immunity: understanding COVID-19. *Immunity* **52**, 737–741 (2020).
45. Miranda, J. G. V. et al. Scaling effect in COVID-19 spreading: the role of heterogeneity in a hybrid ode-network model with restrictions on the inter-cities flow. *Physica D* **415**, 132792 (2020).
46. Pujari, B. S. & Shekatkar, S. M. Multi-city modeling of epidemics using spatial networks: application to 2019-nCoV (COVID-19) coronavirus in India. Preprint at *Medrxiv* <https://doi.org/10.1101/2020.03.13.20035386> (2020).
47. Tai, D. B. G., Shah, A., Doubeni, C. A., Sia, I. G. & Wieland, M. L. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciaa815> (2020).
48. Rumberger, R. W. Education and the reproduction of economic inequality in the United States: an empirical investigation. *Econ. Educ. Rev.* **29**, 246–254 (2010).
49. Mollborn, S., Fomby, P. & Dennis, J. A. Who matters for children's early development? Race/ethnicity and extended household structures in the United States. *Child Indic. Res.* **4**, 389–411 (2011).
50. Ellis, M., Holloway, S. R., Wright, R. & Fowler, C. S. Agents of change: mixed-race households and the dynamics of neighborhood segregation in the United States. *Ann. Assoc. Am. Geogr.* **102**, 549–570 (2012).
51. Stier, A. J., Berman, M. G. & Bettencourt, L. M. A. Metadata record for the article: Early pandemic COVID-19 case growth rates increase with city size. Tech. Rep. <https://doi.org/10.6084/m9.figshare.1419962> (2021).

ACKNOWLEDGEMENTS

This research is partially supported by the Mansueti Institute for Urban Innovation, a Social Science Research grant from the University of Chicago, and NSF S&CC grant #1952050.

AUTHOR CONTRIBUTIONS

A.J.S., M.G.B. and L.M.A.B. contributed equally to this work, including conception of the manuscript, analysis of data, and writing.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42949-021-00030-0>.

Correspondence and requests for materials should be addressed to A.J.S.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021