COVID-19 Superspreading Suggests Mitigation by Social Network Modulation

Nielsen, Bjarke Frost; Simonsen, Lone; Sneppen, Kim

Published in:
Physical Review Letters

DOI:
10.1103/PhysRevLett.126.118301

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Nielsen, B. F., Simonsen, L., & Sneppen, K. (2021). COVID-19 Superspreading Suggests Mitigation by Social Network Modulation. *Physical Review Letters*, *126*(11), Article 118301. https://doi.org/10.1103/PhysRevLett.126.118301

Editors' Suggestion | Featured in Physics

COVID-19 Superspreading Suggests Mitigation by Social Network Modulation

Bjarke Frost Nielsen,[1,*] Lone Simonsen,[2,†] and Kim Sneppen[1,‡]

[1]Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark
[2]Department of Science and Environment, Roskilde University, Universitetsvej 1, 4000 Roskilde, Denmark

(Received 1 October 2020; revised 6 January 2021; accepted 22 January 2021; published 15 March 2021)

Although COVID-19 has caused severe suffering globally, the efficacy of nonpharmaceutical interventions has been greater than typical models have predicted. Meanwhile, evidence is mounting that the pandemic is characterized by superspreading. Capturing this phenomenon theoretically requires modeling at the scale of individuals. Using a mathematical model, we show that superspreading drastically enhances mitigations which reduce the overall personal contact number and that social clustering increases this effect.

DOI: 10.1103/PhysRevLett.126.118301

During the ongoing COVID-19 pandemic, news stories have frequently appeared detailing spectacular events where single individuals—so-called *superspreaders*—have infected a large number of people within a short time frame [1–3]. By now, there is substantial evidence that these are not just singular events but that they reflect a marked transmission heterogeneity [4–6], a signature feature of the disease. In a well-mixed population, such heterogeneity has little bearing on the trajectory of an epidemic, but, when public sphere contacts are restricted, heterogeneity takes on a decisive role, as shown in Ref. [7]. In this Letter, we investigate the effects of transmission heterogeneity—i.e., superspreading—on mitigation strategies which rely on a general reduction in social network size and probe the influence of social clustering on such interventions.

The origins of superspreading can be diverse, depending on the characteristics of the pathogen in question. Superspreading events may occur due to circumstances and behavior as well as biology. Even medical procedures, such as intubation and bronchoscopy, which facilitate the production of aerosols [8], can lead to superspreading events in respiratory diseases. However, the most straightforward model of superspreading is that some individuals simply shed the virus to a much greater extent than the average infected person. For COVID-19, this *"biological superspreader"* phenomenon has some traction and is supported by the observation that household transmission is limited, despite the relatively high *average* infectiousness of COVID-19 [9–11].

Superspreading is not a phenomenon which is particular to SARS-CoV-2 but has been observed in connection with several other pathogens, including coronaviruses such as SARS [12,13] and MERS [14], as well as in diseases such as measles [15] and Ebola virus disease [16,17]. Pandemic influenzas such as the 1918 Spanish flu, on the other hand, are believed to be far more "democratic" [18]. The heterogeneity of transmission is usually quantified using the Gamma distribution [15]. This is the origin of the *dispersion parameter* or k *value*, which determines the fraction of infectious individuals who account for the majority of infections (Fig. 1). Smaller k means greater heterogeneity—in fact, when k is small ($|k| \ll 1$), it approximates the fraction of infected individuals who give rise to 80% of infections. For COVID-19, which is believed to have a k value of perhaps 0.1 [4–6], the most infectious 10% of individuals thus cause approximately 80% of infections.

The fundamental difference between a homogeneously spreading disease and a highly heterogeneous one is reflected in the infection networks they give rise to, as visualized in Fig. 1. When only a small fraction of individuals cause the bulk of infections, a reduction in social network connectivity amounts to decreasing the likelihood that a superspreader infects another superspreader and thus propagates the disease. Consequently, in a network characterized by superspreading [Fig. 1(a)], the outbreak can be stopped by cutting only a few select edges. Not so for the network in Fig. 1(c).

In this Letter, we present a model of superspreading phenomena which assumes that the driving force is a biological heterogeneity in infectiousness. We implement this as an agent-based model with contact networks and are also able to capture much of the phenomenology in analytical formulas. In the model, N agents are placed as the nodes in a contact network. We investigate different types of network, but our base case is the Erdös-Renyi

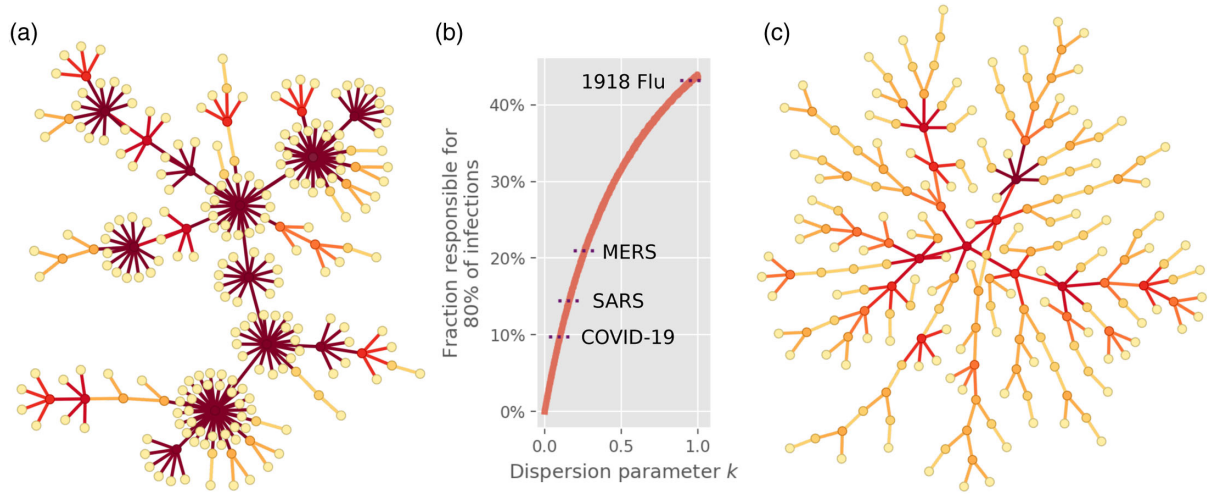

FIG. 1. The characteristics of superspreading. (a) Simulated infection network characterized by superspreading, with a dispersion parameter $k = 0.1$, within what has been observed for COVID-19 [4,5]. Superspreaders appear as hubs, while most individuals are "dead ends," meaning that they do not transmit the disease. The epidemic mainly grows by spreading from one superspreader to the next. (b) The dispersion parameter k provides a measure of superspreading, with lower k values corresponding to a greater heterogeneity. With a k value of 0.1 for COVID-19, approximately 10% of the population has the infectiousness to cause 80% of transmission. SARS and MERS are also characterized by a significant heterogeneity [14,15], while pandemic influenza is believed to be more homogeneous [18]. (c) Simulated infection network without superspreading (all individuals have equal infectiousness). Here, most individuals spread the disease to a few others, leading to a branched structure.

network, which is characterized by a Poisson degree distribution and an absence of clustering.

At initialization, the infectiousness of each individual is drawn from a Gamma distribution [15]. As such, it is an innate property of each individual. The possible states of each individual are susceptible, exposed, infected, and recovered (for details, see Supplemental Material [19], which includes Refs. [20–27]). At each time step, each individual randomly selects one of its contacts to interact with, meaning that only a subset of the network is active at any given time. While a link between an infectious and a susceptible individual is active, there is a constant probability of infection per unit of time, as determined by the individual infectiousness.

This basic setup also lends itself to analytic calculations, as long as saturation effects can be ignored. Consider a single infected person who has c contacts, who are all assumed susceptible. First, the infectiousness r of the individual is drawn from a gamma distribution $P_I(r)$ with dispersion parameter k and mean μ. The distribution of the reproductive number R of an individual with a *known* infectiousness r and degree (i.e., connectivity) c is given by

$$P(R;r,c) = \binom{c}{R}(1-e^{-r/c})^R(e^{-r/c})^{(c-R)}. \quad (1)$$

Taking the variability in infectiousness into account, the overall distribution of R becomes

$$P(R;c) = \int_{r=0}^{r=\infty} dr P_I(r) P(R;r,c). \quad (2)$$

In the limit of infinite connectivity, corresponding to a well-mixed population, this becomes a negative binomial distribution. That particular case has been studied in Ref. [15]. Given a contact network and a corresponding degree distribution $P_C(c)$—for example, a Poisson distribution in the case of an Erdös-Renyi network—the connectivities can be summed over to yield a distribution of individual reproductive numbers, $P(R) = \sum_c P_C(c)P(R;c)$.

As reflected in the equations above, the *actual* number of secondary infections depends not only on biological infectiousness. In Figs. 2(a) and 2(b), we use this analytical framework to explore how the number of personal contacts affects the resultant distribution of infections. Without superspreading [Fig. 2(a)], a reduction in the contact number has a very modest effect and the distributions overlap. When the heterogeneity is at a COVID-like level [Fig. 2(b)], it is quite a different story. Here, a decrease in mean connectivity has a considerable effect, and mitigation suddenly looks feasible. Previously, another mitigation strategy which benefits from superspreading was suggested by Ref. [15], with the crucial difference that it relies on prior identification and targeting of superspreaders, in contrast to the broad reduction in mean connectivity explored here.

To quantify the sensitivity of the epidemic to social network size, we consider the basic reproductive number

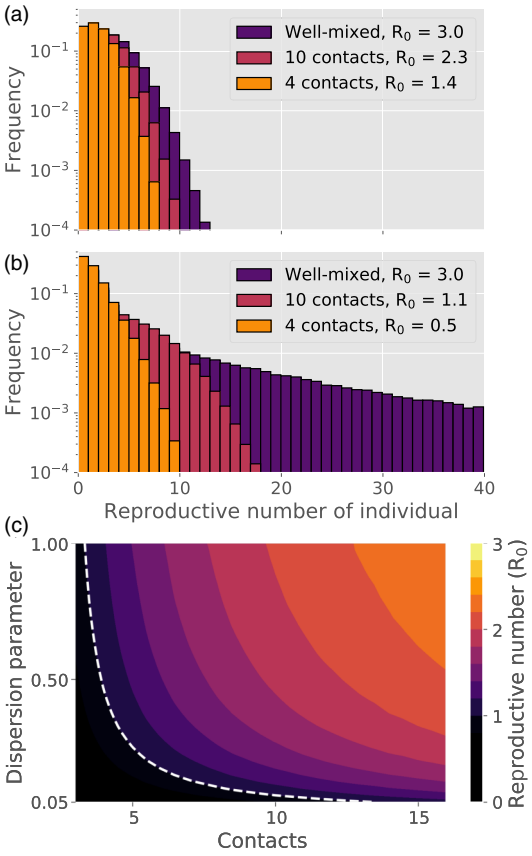

FIG. 2. The reproductive number. Distributions of individual reproductive number R and value of R_0 for different dispersion parameters and number of social contacts during an infectious period. This figure is based on the analytical framework described in the main text. See Supplemental Material [19] for details on the calculation. (a) Distribution of R for a disease where all individuals have equal infectiousness. (b) Distribution of R for a disease characterized by superspreading (dispersion parameter $k = 0.1$). (c) Basic reproductive number R_0 as a function of social connectivity and dispersion. The dashed line represents $R_0 = 1$. These calculations take into account the Poisson distributed contact number and the fact that each infectious person will have one insusceptible person in their network (the individual from whom the infection originated), even when computing the *basic* reproductive number. Details on an analytic computation of R_0 for *fixed* (δ-distributed) contact number can be found in Supplemental Material [19].

R_0, meaning the average number of infections that each infected person causes in a situation where all contacts are still susceptible. In Fig. 2(c), the R_0 is given as a function of the dispersion parameter k and the average contact number. The epidemic is evidently much more sensitive to reductions in contact numbers when the transmission heterogeneity is high. A mitigation in which the average number of contacts goes from being unrestricted, down to about 10, causes a reduction in R_0 which lowers both the *peak* and *total* number of persons infected during the course of the epidemic (the *attack rate*). The overall trajectory of a

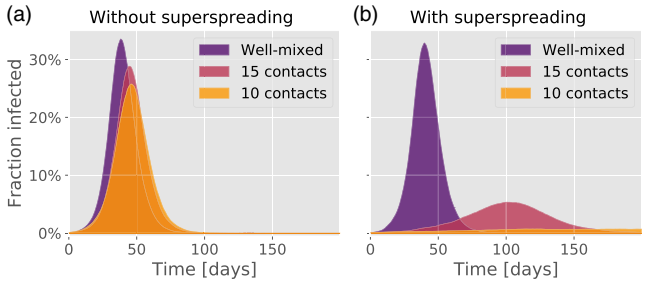

FIG. 3. The epidemic trajectory of a heterogeneous disease is highly sensitive to mitigation. Epidemic trajectories as a function of the number of people that each person interacts with during an infectious period. (a) Time evolution in the absence of any infection heterogeneity. (b) Time evolution for a disease with dispersion parameter $k = 0.1$, roughly representative of COVID-19.

homogeneous disease is largely unaffected by social connectivity [Fig. 3(a)], whereas a heterogeneous epidemic is very sensitive [Fig. 3(b)]. We find a particularly large sensitivity to a reduction of contact number from 15 down to 10 [Fig. 3(b)], indicating a critical threshold for disease spreading, in line with the threshold indicated by the dashed curve in Fig. 2(c).

Crucially, a reduction in contact *time* is not necessary when the disease is characterized by superspreading. What counts is rather a reduction in contact *diversity*, meaning the number of different persons with whom you come into contact during the time you are infectious [7]. This differs fundamentally from SIR models, where contact time and diversity are not differentiated between [28]. In our model, a reduction in the size of an individual's social circle is not accompanied by a reduction in contact time, since the same number of contact events is maintained, with each remaining person being contacted more often. Thus, a mildly infectious individual will not experience appreciable saturation by a reduction in contact diversity, whereas a superspreader will be highly limited by the resultant local saturation.

So far, our analysis has been based on the Erdös-Renyi network, which is largely devoid of clusters. This was chosen as a clean setting in which to probe how social connectivity affects superspreading. However, any realistic social network will involve clusters of people who know each other [29–31]—after all, your colleagues know each other as well as knowing you. It is thus natural to ask whether such *cliquishness* impacts superspreading. In Fig. 4, we compare a cluster-free network to one characterized by a high degree of clustering [32]. See Supplemental Material [19] for the algorithm used to generate this network.

The attack rate of the disease is clearly lowered by clustering, in general (Fig. 4), but the effect is especially significant when heterogeneity is high. The mechanism behind this is that of *local saturation*. If a superspreader

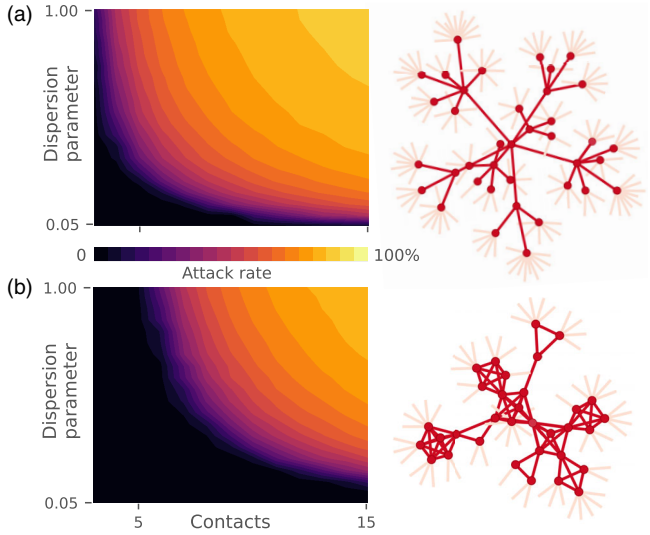

FIG. 4. Final attack rate (total fraction of the population infected) as a function of network connectivity and transmission heterogeneity. In (a), we investigate an Erdös-Renyi network, with the same degree distribution as in Fig. 2. (b) explores a network where each person is assigned to two groups of people, leading to a highly clustered network. The black regions indicate conditions where the disease cannot spread in the population. On the right-hand side, small fragments of the networks in question are shown. Each of the two contour plots in this figure are based on 1500 runs of the model. A detailed description of the algorithm used to generate the clustered network is included in Supplemental Material [19].

infects a significant portion of his network, there is a risk that one of these individuals will turn out to be another superspreader. However, *if* there is clustering, a large part of this *second* superspreader's network will already have been exposed, and the second superspreader does comparatively little harm.

In the literature, there exists ample evidence that *social* heterogeneity, implemented either through wide distributions of social activity [28] or through social networks with broad degree distributions [23–26], has a significant effect on the course of an epidemic. Notably, epidemics tend to attain lower final sizes in networks with broad degree distributions [23,24,26,33,34] and in clustered networks [26]. While the effects of varying degree distributions as well as clustering were explored in Ref. [26], mitigation by contact network reduction was not investigated, and no variation in individual infectiousness was assumed. As we have shown, the effects of biological superspreading on mitigation are profound in networks with a representative mean. In Supplemental Material [19], we simulate an epidemic on a much more socially heterogeneous (fat-tailed) network based on data from Ref. [27] and find that our conclusions are robust to alterations in the degree distribution.

Beyond the mitigation strategies discussed here, which rely on broad reductions in contact numbers, more targeted strategies are possible—most prominently, test-trace-isolate (TTI) strategies. While an in-depth treatment of the implications of superspreading for TTI strategies is beyond the scope of this Letter, our simulations do imply that backward contact tracing (see, e.g., [35]) is more effective in the presence of superspreading. When encountering an infected individual, this strategy relies on asking "Who was this person infected by, and who else might *that* person have infected?" rather than simply asking "Who might this person have infected?," as one would in forward tracing schemes. We can estimate the efficacy of backward tracing in our simulations by measuring how many secondary cases each infected person allows one to trace, with and without superspreading. In a well-mixed scenario, we find the answer to be 2.7 without superspreading ($k = \infty$) and 24 with COVID-like superspreading ($k = 0.1$). Of course, such a backward contact tracing scheme may run into practical limitations, especially regarding the temporal constraints arising from a disease with a relatively short generation time. Nevertheless, these results seem to indicate that transmission heterogeneity may profoundly influence TTI mitigation strategies as well.

Superspreading is now a well-established phenomenon for a number of diseases [15], including COVID-19 [4,5]. In spite of this, the extent to which circumstance and person-specific properties contribute to the observed overdispersion in COVID-19 is still not clear. Superspreading can also have a social component, exemplified by highly social individuals, who come into contact with a large number of people in a limited time frame. However, such individuals would also be *superreceivers*, a trait which impacts the epidemic even in the absence of mitigation [36,37]. In any case, ability as well as opportunity is necessary for superspreading to occur. In our model, we have focused on interindividual variation in ability to produce and transmit virus. This simplification is supported by cases of one person infecting many people at different times and locations [38] and by the observation that most infected people do not even infect their spouse [9–11]. However, more complex models could incorporate realistic *social* heterogeneity as well as large temporal variations in viral load [39,40]—effects which we have not probed. Furthermore, studies which address event-driven superspreading as well as contact tracing in the presence of superspreaders are also needed.

Regardless of the origin of superspreading, we emphasize the particular fragility of a disease in which a major part of infections are caused by the minority. If this is the case, the disease is vulnerable to mitigation by reducing the number of *different* people that an individual meets within an infectious period. The significance is clear: Everybody can still be socially active but generally only with relatively few—on the order of ten persons. Importantly, our study further demonstrates that repeated contact with *interconnected* groups (such as at a workplace or in friend groups) is

comparatively less damaging than repeated contacts with independent people.

We thank Robert J. Taylor, Andreas Eilersen, Gorm G. Jensen, and Julius B. Kirkegaard for enlightening discussions. Our research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, Grant Agreement No. 740704, as well as from The Carlsberg Foundation, Grant No. 61114.

*bjarkenielsen@nbi.ku.dk
†lonesimo@ruc.dk
‡sneppen@nbi.dk

[1] L. Hamner, High SARS-CoV-2 attack rate following exposure at a choir practice—Skagit County, Washington, March 2020, Morb. Mortal. Wkly. Rep. **69**, 606 (2020).
[2] S. Kumar, S. Jha, and S. K. Rai, Significance of super spreader events in COVID-19, Indian Journal of public health : quarterly Journal of the Indian Public Health Association **64**, 139 (2020).
[3] Y. Liu, R. M. Eggo, and A. J. Kucharski, Secondary attack rate and superspreading events for SARS-CoV-2, Lancet **395**, e47 (2020).
[4] D. Miller, M. A. Martin, N. Harel, T. Kustin, O. Tirosh, M. Meir, N. Sorek, S. Gefen-Halevi, S. Amit, O. Vorontsov *et al.*, Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel, Nat Commun. **11**, 5518 (2020).
[5] A. Endo, S. Abbott, A. J. Kucharski, S. Funk *et al.*, Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China, Wellcome Open Res. **5**, 67 (2020).
[6] M. S. Lau, B. Grenfell, M. Thomas, M. Bryan, K. Nelson, and B. Lopman, Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA, Proc. Natl. Acad. Sci. U.S.A. **117**, 22430 (2020).
[7] K. Sneppen, R. J. Taylor, and L. Simonsen, Impact of superspreaders on dissemination and mitigation of COVID-19, medRxiv https://doi.org/10.1101/2020.05.17.20104745 (2020).
[8] T. R. Frieden and C. T. Lee, Identifying and interrupting superspreading events—implications for control of severe acute respiratory syndrome coronavirus 2, Emerging Infect. Dis. **26**, 1059 (2020).
[9] Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S. A. Truelove, T. Zhang *et al.*, Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study, Lancet Infect. Dis. **20**, 911 (2020).
[10] S. Y. Park, Y.-M. Kim, S. Yi, S. Lee, B.-J. Na, C. B. Kim, J.-i. Kim, H. S. Kim, Y. B. Kim, Y. Park *et al.*, Coronavirus disease outbreak in call center, South Korea, Emerging Infect. Dis. **26**, 1666 (2020).
[11] Q.-L. Jing, M.-J. Liu, Z.-B. Zhang, L.-Q. Fang, J. Yuan, A.-R. Zhang, N. E. Dean, L. Luo, M.-M. Ma, I. Longini *et al.*, Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: A retrospective cohort study, Lancet Infect. Dis. **20**, 1141 (2020).
[12] J. Raboud, A. Shigayeva, A. McGeer, E. Bontovics, M. Chapman, D. Gravel, B. Henry, S. Lapinsky, M. Loeb, L. C. McDonald *et al.*, Risk factors for SARA transmission from patients requiring intubation: A multicentre investigation in Toronto, Canada, PLoS One **5**, e10717 (2010).
[13] Z. Shen, F. Ning, W. Zhou, X. He, C. Lin, D. P. Chin, Z. Zhu, and A. Schuchat, Superspreading SARS events, Beijing, 2003, Emerging Infect. Dis. **10**, 256 (2004).
[14] A. J. Kucharski and C. L. Althaus, The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission, Eurosurveillance **20**, 21167 (2015).
[15] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz, Superspreading and the effect of individual variation on disease emergence, Nature (London) **438**, 355 (2005).
[16] C. L. Althaus, Ebola superspreading, Lancet Infect. Dis. **15**, 507 (2015).
[17] M. S. Lau, B. D. Dalziel, S. Funk, A. McClelland, A. Tiffany, S. Riley, C. J. E. Metcalf, and B. T. Grenfell, Spatial and temporal dynamics of superspreading events in the 2014–2015 West Africa Ebola epidemic, Proc. Natl. Acad. Sci. U.S.A. **114**, 2337 (2017).
[18] C. Fraser, D. A. Cummings, D. Klinkenberg, D. S. Burke, and N. M. Ferguson, Influenza transmission in households during the 1918 pandemic, American Journal of Epidemiology **174**, 505 (2011).
[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.126.118301 for model description and sensitivity analyses.
[20] A. Remuzzi and G. Remuzzi, COVID-19 and Italy: What next?, Lancet **395**, 1225 (2020).
[21] Our World In Data and European Centre for Disease Prevention and Control, COVID-19-Data (Deaths), https://github.com/owid/covid-19-data/ (2020).
[22] M. A. Billah, M. M. Miah, and M. N. Khan, Reproductive number of coronavirus: A systematic review and meta-analysis based on global level evidence, PLoS One **15**, e0242128 (2020).
[23] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Dynamical patterns of epidemic outbreaks in complex heterogeneous networks, J. Theor. Biol. **235**, 275 (2005).
[24] J. Hindes and I. B. Schwartz, Epidemic Extinction and Control in Heterogeneous Networks, Phys. Rev. Lett. **117**, 028302 (2016).
[25] S. Bansal, B. T. Grenfell, and L. A. Meyers, When individual behaviour matters: Homogeneous and network models in epidemiology, J. R. Soc. Interface **4**, 879 (2007).
[26] E. M. Volz, J. C. Miller, A. Galvani, and L. A. Meyers, Effects of heterogeneous and clustered contact patterns on infectious disease dynamics, PLoS Comput. Biol. **7**, e1002042 (2011).
[27] L. Danon, J. M. Read, T. A. House, M. C. Vernon, and M. J. Keeling, Social encounter networks: Characterizing Great Britain, Proc. R. Soc. B **280**, 20131037 (2013).
[28] T. Britton, F. Ball, and P. Trapman, A mathematical model reveals the influence of population heterogeneity on herd immunity to SARA-CoV-2, Science **369**, 846 (2020).

[29] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature (London) **393**, 440 (1998).
[30] J. Davidsen, H. Ebel, and S. Bornholdt, Emergence of a Small World from Local Interactions: Modeling Acquaintance Networks, Phys. Rev. Lett. **88**, 128701 (2002).
[31] M. Rosvall and C. T. Bergstrom, Maps of random walks on complex networks reveal community structure, Proc. Natl. Acad. Sci. U.S.A. **105**, 1118 (2008).
[32] A. Eilersen and K. Sneppen, Estimating cost-benefit of quarantine length for COVID-19 mitigation, medRxiv https://doi.org/10.1101/2020.04.09.20059790 (2020).
[33] M. E. J. Newman, Spread of epidemic disease on networks, Phys. Rev. E **66**, 016128 (2002).
[34] R. Pastor-Satorras and A. Vespignani, Epidemic Spreading in Scale-Free Networks, Phys. Rev. Lett. **86**, 3200 (2001).
[35] A. Endo *et al.*, Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks, Wellcome Open Res. **5** (2020).
[36] B. F. Nielsen, K. Sneppen, L. Simonsen, and J. Mathiesen, Social network heterogeneity is essential for contact tracing, medRxiv https://doi.org/10.1101/2020.06.05.20123141 (2020).
[37] A. V. Tkachenko, S. Maslov, A. Elbanna, G. N. Wong, Z. J. Weiner, and N. Goldenfeld, Persistent heterogeneity not short-term overdispersion determines herd immunity to COVID-19, arXiv:2008.08142.
[38] Y. Shen, C. Li, H. Dong, Z. Wang, L. Martinez, Z. Sun, A. Handel, Z. Chen, E. Chen, M. H. Ebell *et al.*, Community outbreak investigation of SARS-CoV-2 transmission among bus riders in Eastern China, JAMA Intern. Med. **180**, 1665 (2020).
[39] L. Zou, F. Ruan, M. Huang, L. Liang, H. Huang, Z. Hong, J. Yu, M. Kang, Y. Song, J. Xia *et al.*, SARS-CoV-2 viral load in upper respiratory specimens of infected patients, N. Engl. J. Med. **382**, 1177 (2020).
[40] J. B. Kirkegaard, J. Mathiesen, and K. Sneppen, Airborne pathogens in a heterogeneous world: Superspreading & mitigation, medRxiv (to be published).