Supplementary Information

Individual preventive social distancing during an epidemic may have negative population-level outcomes

Ka Yin Leung[a,1], Frank Ball[2], David Sirl[2], and Tom Britton[1]

10th July 2018

Contents

This supplementary information accompanies the manuscript 'Individual preventive social distancing during an epidemic may have negative population-level outcomes'. It contains details on mathematical results, additional figures, and summary statistics for the two real-world networks.

In Section S1 analytical results and additional figures are presented for the social distancing model on the configuration network model. In Section S1.3 we provide a proof that the final size can increase with increasing social distancing rate ω by considering an asymptotic lower bound for the final size. In Section S1.4 we provide some additional simulation studies on the

[a]kayin.leung@math.su.se
[1]Department of Mathematics, Stockholm University, 106 91 Stockholm, Sweden
[2]School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

number of index cases. In Section S2 we derive the epidemic threshold parameter R_* for the social distancing model on the clique network, the so-called clique reproduction number. We show that social distancing can move R_* from below to above the threshold value of one, thereby leading to a positive probability of a major outbreak to occur. In Section S3 some properties of the real-world networks are presented. In Section S4 we show that negative population level effects can occur for a range of α values large enough, where 'large enough' depends on the network structure under consideration and other parameter value choices. Finally in Section S5 we describe the technicalities of the simulation studies.

Section S1 Configuration network

Section S1.1 Notation and construction

The (Newman-Strogatz-Watts version of the) configuration model is defined in the main text in Section 2.2.1. The degree D of an individual in the network is distributed as the pre-defined distribution $\{p_d\}$. The mean and variance of D are denoted by μ_D and σ_D^2, respectively. In the configuration network construction a given half-edge is d times as likely to be paired with a given individual with degree d than with a given individual with degree 1. Therefore, a neighbour has size-biased degree $\tilde{D}$, where $P(\tilde{D} = d) = dp_d/\mu_D$, $d = 1, 2, \ldots$ The configuration model is well studied, e.g. [7]. It is for instance known that, as the population size $n \to \infty$, there will be zero connected components that have size of exact order n if $\mu_D + \sigma_D^2/\mu_D - 1 \leq 1$, and precisely one connected component of exact order n if $\mu_D + \sigma_D^2/\mu_D - 1 > 1$. Since there is at most one component that has size of exact order n this component is often called the giant component. Further, if the degrees are not small, then the giant often makes up close to all individuals. If for example all individuals have the same degree $d > 2$, then all nodes are part of the giant component with a probability tending to 1 as $n \to \infty$.

Section S1.2 Epidemic threshold parameter R_0

The basic reproduction number R_0 for the configuration network model with rewiring and dropping has been derived previously in [8]:

$$R_0 = \frac{\beta}{\beta + \omega + \gamma} \left(\mu_D + \frac{\sigma_D^2}{\mu_D} - 1 \right). \tag{S.1}$$

From (S.1) we see that R_0 is a monotonically decreasing function of ω, i.e. any social distancing always reduces R_0. In particular, if R_0 is larger than one in the baseline setting without any social distancing ($\omega = 0$), then social distancing reduces R_0 to below its epidemic threshold of one if and only if $\omega \geq \beta(\mu_D + \sigma_D^2/\mu_D - 1) - \gamma$. Therefore, social distancing is always beneficial at the beginning of an epidemic in the configuration network.

Section S1.3 Asymptotic lower bound for the final size

Assume that the degree distribution of the population only takes on values 0 and d, so $p_0 = 1 - p_d$ with $d > 2$. For ease of argument we consider the asymptotic lower bound first for the simpler setting of our model where $\gamma = 0$ and $\alpha = 1$, i.e. an SI infection (without recovery) and social distancing is always through rewiring to new individuals. We indicate later how the argument may be extended to the case when $\gamma = 0$ and $\alpha \in (0, 1]$. By a continuity argument it follows that, for any $\alpha \in (0, 1]$, the conclusions also hold for all sufficiently small $\gamma > 0$. The degree of an individual may change over the course of an epidemic owing to rewiring. We say that an

individual has 'original degree' 0 or d if that was the degree of the individual in the configuration-network construction before the epidemic.

We consider R_0 as a function of ω. In the setting of this section, $R_0 = R_0(\omega) = \beta(d-1)/(\beta+\omega)$ (in the early stages of the epidemic, all newly infected individuals have degree d). Note that $R_0(0) = d-1$, i.e. in an SI epidemic without rewiring, a newly infected individual transmits infection to all its susceptible neighbours. Note that $R_0(0) > 1$ since we have assumed that $d > 2$. Let $\bar{Z}_n(\omega)$ denote the final fraction of the population that gets infected in the model with population size n and rewiring rate ω. If $\omega = 0$, then only transmission can occur (recall that we consider $\gamma = 0$), so an index case will generate an outbreak of the size equal to the size of the connected component this index case is part of. Asymptotically, as population size $n \to \infty$, as mentioned earlier, all individuals with degree d will be part of the giant component of the network so the giant component consists of a fraction p_d of the entire population. Therefore, as n tends to infinity, $\bar{Z}_n(0)$ tends to a two-point distribution $\bar{Z}(0)$ with $P(\bar{Z}(0) = 0) = p_0 = 1 - P(\bar{Z}(0) = p_d)$. That is, in the setting without recovery, the asymptotic final fraction infected is either zero (if the index case has degree 0) or p_d (if the index case has degree d).

Note that for large enough ω, $R_0(\omega) < 1$. We consider $\omega > 0$ sufficiently small such that $R_0(\omega) > 1$. Furthermore, we assume that a major outbreak occurs. Therefore we assume that the index case has degree d (if the index case has degree 0, then the final fraction $\bar{Z}_n(\omega)$ tends to zero as n tends to infinity). Then the final fraction $\bar{Z}_n(\omega)$ tends in probability to $\tau(\omega)$ as $n \to \infty$. We show that the relative final size $\tau(\omega)$ of a major outbreak can increase for ω compared to the baseline $\omega = 0$ model. We do so by giving a lower bound $\tilde{\tau}(\omega)$ for $\tau(\omega)$. This lower bound is obtained by considering a model with slightly different social distancing rules. Suppose that a susceptible individual distances him/herself from its infectious neighbour. Then he/she tries to rewire to a randomly chosen individual v in the population. If v is an individual with original degree d, then no connection is made and the existing connection is dropped instead. On the other hand, if v has original degree 0, then (i) a connection is made as in the original model and (ii) v is not allowed to transmit infection to an individual with original degree d. Note that all these modifications make infection less likely, so the final size of this modified model is less than that of the original social distancing model. Furthermore, note that in the baseline model without rewiring ($\omega = 0$) the final size of the original model is equal to the lower bound obtained from the modified model.

Mean final size

The $n \to \infty$ limiting final fraction of the population with original degree d that becomes infected in the epidemic in the event of a major outbreak, which we denote by ρ, can be determined using a susceptibility set argument (cf. [1, Section 2.1.2]). Let $\theta = \omega/(\beta+\omega)$ be the probability that social distancing occurs before transmission for any given connection between a susceptible and an infectious individual. An individual, i^* say, having original degree d has d neighbours, each of whom, if they were to become infected, infects i^* independently with probability $1-\theta$. Further, any such neighbour of i^*, say j_*, has $d-1$ neighbours excluding i^*, who would infect j_* independently with probability $1-\theta$ if they were to become infected. It follows that the susceptibility set of i^* can be approximated by a branching process having offspring distribution $\mathrm{Bin}(d, 1-\theta)$ in the initial generation and offspring distribution $\mathrm{Bin}(d-1, 1-\theta)$ in all subsequent generations. The above limiting final fraction ρ is given by the probability that this branching process does not go extinct. Thus, by standard branching process theory,

$$\rho = 1 - [\theta + (1-\theta)\xi]^d,$$

where $\xi \in (0,1)$ is the unique solution to

$$\xi = [\theta + (1-\theta)\xi]^{d-1}. \tag{S.2}$$

Let $z = \theta + (1-\theta)\xi$. Then $\rho = 1 - z^d$. Further, $\xi = (z-\theta)/(1-\theta)$ and substitution into (S.2) shows that z is the unique solution in $(0,1)$ of

$$(\beta+\omega)z - \omega - \beta z^{d-1} = 0 \tag{S.3}$$

Any original degree d individual that is exposed to at least one infectious individual will rewire at least once with probability θ, so the fraction of the original degree d population that rewires at least once is at least $\rho\theta$. These will include individuals that are not subsequently infected. However, since the fraction of the original degree d population that avoid infection is $1-\rho$, the fraction of that population which rewire at least once and are subsequently infected is at least $\rho\theta - (1-\rho)$. The probability that such a rewiring is to an original degree 0 individual is p_0 and, given it is to an original degree 0 individual, the probability that individual itself rewires before he/she is infected is $1-\theta$. Thus the fraction of the original degree d population that rewires and transmits infection to an original degree 0 individual is at least

$$\phi = \big(\rho\theta - (1-\rho)\big)p_0(1-\theta). \tag{S.4}$$

Asymptotically, the probability that an individual of degree 0 escapes infection is at most $(1-1/(np_0)^{\phi n p_d}) \to e^{-\phi p_d/p_0}$ as $n \to \infty$, since approximately at least $n\phi p_d$ individuals rewire and transmit infection to an original degree 0 individual, and for each such rewiring event the individual rewired to is chosen uniformly at random from the np_0 original degree 0 individuals. Therefore, a lower bound for the fraction of the population that originally had degree 0 and get infected is $p_0(1-e^{-\phi p_d/p_0})$. Hence, the total fraction of the population that eventually get infected is, in the limit as $n \to \infty$, at least

$$\tilde{\tau}(\omega) = p_d\rho + p_0(1-e^{-\phi p_d/p_0}). \tag{S.5}$$

The latter yields an asymptotic lower bound for the final fraction infected in the SI epidemic on the configuration network with rewiring rate ω (i.e. where $\gamma = 0$, $\alpha = 1$). We consider the first order approximation of (S.5) in ω to show that the lower bound for the relative final size is exact when $\omega = 0$ and increasing for small ω, thus the final size of the original model also increases for small ω. We approximate ρ by considering (S.3). Denote the left-hand side of (S.3) by $F(z(\omega),\omega)$, so $F(z(\omega),\omega) = 0$. Then $\partial F/\partial z \cdot z'(\omega) + \partial F/\partial\omega = 0$, which yields $z'(\omega) = -(z-1)/(\beta+\omega-\beta(d-1)z^{d-2})$. Hence $z(\omega) = z(0) + z'(0)\omega + o(\omega) = \omega/\beta + o(\omega)$, and we can approximate $\rho = 1-(\omega/\beta)^d + o(\omega^d)$ for $\omega > 0$ small enough. Next, recalling that $\theta = \omega/(\beta+\omega)$ we make the approximation

$$\begin{aligned}\phi &= \left[\left(1-(\omega/\beta)^d\right)\theta - (\omega/\beta)^d + o(\omega^d)\right]p_0(1-\theta)\\ &= p_0\omega/\beta + o(\omega)\end{aligned}$$

as $\omega \downarrow 0$. The first order approximation of (S.5) is therefore

$$\tilde{\tau}(\omega) = p_d(1 + p_0\omega/\beta) + o(\omega).$$

Recall that the limiting mean fraction infected in the event of a large outbreak in the model with no rewiring ($\omega = 0$) is p_d, i.e. this approximation is exact when $\omega = 0$. Thus, since $p_d\,(1+p_0\omega/\beta) \geq p_d$ for all $p_0, \omega > 0$, we find that the final fraction infected in case of a major outbreak is increased for all sufficiently small but strictly positive rewiring rates $\omega > 0$ for the modified model. In other words, the asymptotic lower bound for the relative final size given a major outbreak is increasing in ω for small $\omega > 0$.

Mean final size of a major outbreak

For the expected relative final size we have

$$E(\bar{Z}(\omega)) = \tau(\omega)P_\omega(\text{major outbreak}) \geq p_d(1+\frac{p_0}{\beta}\omega)P_\omega(\text{major outbreak}) + o(\omega).$$

since the expected relative final size of a minor epidemic is 0. In order to show that also $E(\bar{Z}(\omega))$ has a lower bound that is increasing in ω for sufficiently small ω, it suffices to show that $P_\omega(\text{major outbreak})$ does not decrease too fast. We show that $1 - P_\omega(\text{major outbreak}) = P_\omega(\text{minor outbreak}) \leq p_0 + p_d\omega^{d-1} + o(\omega^{d-1})$ or equivalently that

$$P_\theta(\text{minor outbreak}) \leq p_0 + p_d\theta^{d-1} + o(\theta^{d-1}). \tag{S.6}$$

Note that there is a minor outbreak if either the index case has degree 0 (with probability p_0), or the index case has degree d (with probability $p_d = 1 - p_0$) but the outbreak that takes place amongst the individuals with degree d is a minor outbreak. Consider the configuration model without the degree 0 individuals. We modify the model by letting the index case have degree $d-1$ and all other individuals have degree d. Consequently, *all* individuals, including the index case, have $d-1$ susceptible neighbours when newly infected in the beginning of the epidemic. This modified model has a probability of a minor outbreak that is larger than the original model (and is exact in the case $\theta = 0$), i.e. $P_\theta(\text{minor outbreak}) \leq P_\theta(\text{minor outbreak of the modified model})$. Let $\pi(\theta)$ denote the probability of a minor outbreak for the modified model, given that the index case is of degree $d-1$ (with probability $p_d = 1-p_0$). By conditioning on the number of neighbours that avoid infection in the first generation, i.e. the number of neighbours that rewire away before infection, we find the consistency equation

$$\pi(\theta) = \sum_{j=0}^{d-1} \pi(\theta)^{d-1-j}\binom{d-1}{j}\theta^j(1-\theta)^{d-1-j} = \big(\theta + \pi(\theta)(1-\theta)\big)^{d-1}.$$

Note that $\pi(0) = 0$ since a major outbreak will occur without any rewiring. By the Taylor expansion of $\pi(\theta)$ about $\theta = 0$, we find that $\pi(\theta) = \theta^{d-1} + o(\theta^{d-1})$, proving (S.6). Since $E(\bar{Z}(0)) = p_d$, we have a lower bound for the expected relative final size that is increasing for sufficiently small $\omega > 0$, and is exact for the baseline model $\omega > 0$ without rewiring.

Allowing dropping

Suppose that $\alpha \in (0, 1)$. Then since in the model with modified social distancing all attempted rewirings to an original degree d individual results in the edge being dropped, the only change in the argument leading to (S.5) is that the expression for ϕ in (S.4) becomes

$$\phi = \big(\rho\theta - (1-\rho)\big)p_0(1-\theta)\alpha.$$

It follows that $\phi = p_0\alpha\omega/\beta + o(\omega)$ as $\omega \downarrow 0$ and, since $\alpha > 0$, the asymptotic lower bound for the relative final size given a major outbreak is again increasing in ω for small $\omega > 0$. Finally, note that the probability of a major outbreak is independent of α, so the above argument shows that the expected relative final size is also increasing in ω for small $\omega > 0$. These results are illustrated with simulation studies in Fig. S4

Section S1.4 The number of index cases

We consider the social distancing model on the configuration network with degree distribution $\{p_d\}_{d=0}^{10}$ where $p_d = c/(d+1)$, $d = 0, 1, \ldots, 10$, with $c = 0.331$ the normalization constant. In

the main text, each epidemic is initiated with ten initially infected individuals that are randomly chosen from the population. The rationale behind this choice is to ensure sufficiently many epidemics result in major outbreaks. Owing to the degree distribution of the population, a large fraction of the population has degree 0 or degree 1. An individual with degree 0 has no connections and is therefore unable to transmit infection in the population. Also an individual with degree 1 is unlikely to generate a major outbreak. In Fig. S1 we compare the final epidemic size for the case that each epidemic starts with 10 randomly chosen infectives to the case that each epidemic starts with only 1 randomly chosen infective (all other parameters are the same as in Fig. 1 of the main text).

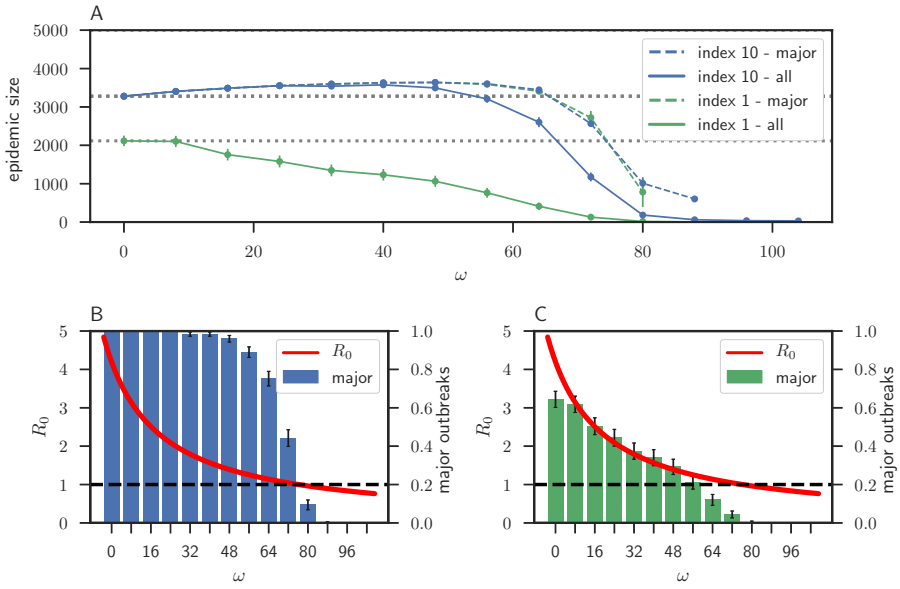

Figure S1: The effect of social distancing on the final size of epidemics with different numbers of initially infected individuals. (A) Average final size compared to the average final size in the baseline model (with 95% CI) over all outbreaks (solid line) and restricted to major outbreaks (dashed line) for 1 initial infective (green) and 10 initial infectives (blue); dotted horizontal lines are at the size of the epidemic when $\omega = 0$, for reference. (B) R_0 and the fraction of outbreaks resulting in a major outbreak (with 95% CI) with 10 initial infectives. (C) R_0 and the fraction of outbreaks resulting in a major outbreak (with 95% CI) with 1 initial infective. In (B) and (C) the black dashed line at $R_0 = 1$ indicates the threshold value. We consider the configuration network model with degree distribution $\{p_d\}_{d=0}^{10}$ where $p_d = c/(d+1)$, $d = 0, 1, \ldots, 10$, with $c = 0.331$ the normalization constant. The population size is 5000 and 500 simulation runs are performed for each value of ω. Parameter values are $\beta = 20, \gamma = 0.2, \alpha = 0.9$. The index cases are randomly chosen in the entire population. The plots for 10 initial infectives (in blue) are the plots given in Fig. 1 of the main text.

Note that the effects of the number of initial infectives as shown in Fig. S1 are not surprising. The average final size over all outbreaks is much smaller in the setting where epidemics are initiated by one initial infective individual. Since this individual is randomly chosen over the entire population, it is quite likely that he/she has degree 0 or 1. In such cases, epidemics will (most likely) die out (Fig. S1C). By considering 10 initial infectives instead, it becomes more likely that an epidemic outbreak occurs (Fig. S1B). This is both because there are simply more infectives to start with, making it less likely that the epidemic dies out owing to stochastic

effects, and it being more likely that some initial infectives have higher degree, enabling further transmission in the population. On the other hand, the average final size in the event of a major outbreak does not depend on the initial number of infectives provided that number is $o(n)$ as $n \to \infty$ (cf. Fig. S1A, blue and green dashed line).

Section S2 The clique-network model

In the main text as well as in this supplement we consider the clique-network model where all cliques consist of exactly three individuals. The reason for this choice is that it allows us to illustrate our point that social distancing can have a negative effect on the initial stages of an outbreak by increasing the epidemic threshold parameter R_*. At the same time, the choice of cliques of size three simplifies the bookkeeping in the calculations below. Moreover, it yields an explicit expression for the clique reproduction number R_*. However, in principle similar calculations can be done also for clique sizes larger than three, or having variable clique sizes.

The clique network with variable clique sizes can be constructed as follows (see also [3, 4] where the same model is referred to as household-network models). A clique-size distribution $\{\pi_h\}$ is predefined that describes the sizes of cliques in the population. The population can be divided into cliques with this distribution by drawing random sizes $h_1, h_2, \ldots$ independently from $\{\pi_h\}$, then labelling all individuals 1 up to n, and letting the first h_1 individuals make up clique 1, the next h_2 individuals make up clique 2, and so on. In the final network, individual 1 is then connected to all individuals he/she is connected to from the construction of the configuration model and individuals 2 up to h_1 from the clique construction, and so on.

We denote the degree distribution for the global network connections by $\{p_d\}$, and the mean and variance by μ_D and σ_D^2, respectively. We also use the size-biased degree distribution $\{\tilde{p}_d\}$ with $\tilde{p}_d = dp_d/\mu_D$. We let $\tilde{\mu}_D$ denote the mean of the size-biased degree distribution, where $\tilde{\mu}_D = \sum_{d=1}^{\infty} d\tilde{p}_d = \sum_{d=1}^{\infty} d^2 p_d/\mu_D = \mu_D + \sigma_D^2/\mu_D$.

Section S2.1 Derivation of the epidemic threshold parameter R_*

We consider the clique reproduction number R_* that can be interpreted as the expected number of secondary cliques generated by one typical newly infected clique at the beginning of an epidemic [6, 2, 5]. So, rather than considering individuals as the units of interest we consider cliques. The clique reproduction number R_* satisfies the desired threshold behaviour that, as the number of cliques tends to infinity, there is a strictly positive probability of a major outbreak if $R_* > 1$ and only minor outbreaks occur if $R_* \leq 1$.

At the start of an epidemic there are only few infected individuals and most individuals are susceptible. Therefore, a newly infected clique at the beginning of an epidemic has most likely only one newly infected individual that got infected by a neighbour outside his/her own clique. In the SIR epidemic on the clique-network with social distancing, there are exactly two types of newly infected clique: the first infected individual in the clique was infected through (1) a global network neighbour or (2) a rewired network neighbour. Note that in case (1) the individual has $d-1$ susceptible global connections with probability $\tilde{p}_d$ while in case (2) the individual has d susceptible global connections with probability p_d (note that the probability that an individual is rewired to more than once is negligible in the early stages). The other two clique members have d susceptible global connections with probability p_d.

The clique reproduction number R_* is given by the dominant eigenvalue of the 2×2 matrix $K = (K_{ij})_{i,j=1,2}$, where K_{ij} is the expected number of cliques of type j generated by one newly infected clique of type i. In the remainder of this section we derive expressions for the K_{ij} by considering all possible events that can occur within the clique. We denote the clique member

that was infected from outside of the clique by u_*. The other two clique members are denoted by u_1 and u_2. We let

$$\theta_1 = \frac{\beta}{\beta+\omega+\gamma}, \quad \theta_2 = \frac{\omega}{\beta+\omega+\gamma}, \quad \theta_3 = 1-\theta_1-\theta_2 = \frac{\gamma}{\beta+\omega+\gamma},$$

so that θ_1 is the probability that transmission occurs from an infectious individual to a given susceptible connection before recovery or social distancing, θ_2 is the probability that social distancing occurs before transmission or recovery, and θ_3 is the remaining probability (of recovery before either event).

The K_{ij} can be derived directly from the above interpretation. First of all, note that $K_{12} = K_{22}$, i.e. the expected number of secondary cliques of type 2 generated by a newly infected clique does not depend on the type of that clique as the global network degree of u_* does not play a role. A clique member within a clique can transmit infection to individuals in another clique through a rewired edge if (i) there is at least one clique edge that is rewired away to a new individual in the population, (ii) the individual u_1 that rewires away becomes infected by a clique member, and (iii) u_1 transmits infection along its rewired clique connection. Since we need both that clique connections rewire to new cliques and the individuals that rewire away to become infected, the only possibility for secondary cliques of type 2 to be generated is if there is exactly one rewired clique connection in the index clique.

When u_* is newly infected and the other two clique members u_1 and u_2 are (still) susceptible, the following events can occur: (i) u_* recovers (at rate γ), (ii) u_* transmits to u_1 or u_2 (at rate 2β), or (iii) u_1 or u_2 distances itself from u_* (at rate 2ω). So we find that the probability that u_* transmits to u_1 is

$$\pi = \frac{\beta}{\gamma+2\beta+2\omega}. \tag{S.7}$$

The other probabilities in the derivations below can be derived in similar manner and we leave out the details. We find the following two possibilities.

- u_* transmits to u_1 with probability π, then u_2 rewires one of its clique connections from either u_* or u_1 (with probability $2\alpha\omega/(2\beta+2\omega+2\gamma)$), after which u_2 becomes infected through its clique connection that was not rewired away (with probability θ_1). Finally, u_2 transmits to a global neighbour through its rewired clique edge (with probability θ_1)
- u_1 rewires away from u_* to a new individual (with probability $2\alpha\omega/(2\beta+\gamma+2\omega)$), then u_* transmits to u_2 (with probability θ_1). Next, u_2 transmits infection to u_1 (with probability θ_1), and u_1 transmits to a global neighbour through its rewired clique edge (with probability θ_1)

Therefore,

$$K_{12} = \frac{4\alpha\omega}{(2\beta+\gamma+2\omega)}\theta_1^3 = K_{22}. \tag{S.8}$$

K_{11} and K_{21} are a bit more involved as there are more possibilities to take into account. Note that the only difference between the two K_{i1} is in the degree of u_*. We note that the expected number of newly infected cliques generated by network edges of u_* is the probability θ_1 of transmission through a network connection (before recovery of u_* or rewiring of network edges) times the expected number of susceptible network neighbours of u_*. If u_* is of type 1, then the expected number of susceptible network neighbours is $\tilde{\mu}_D - 1$. If u_* is of type 2, then this expected number is μ_D.

Next, we note that clique members u_1 and u_2 are interchangeable in the sense that they both have network degree d with probability p_d, and all three clique members u_1, u_2, and u_* are connected to each other. If individual u_1 gets infected, then the expected number of susceptible network neighbours it has is μ_D. The expected number of secondary cases an infectious clique member u_1 generates is then $\theta_1\mu_D$. We derive the probability that u_1 becomes infected by taking into account all relevant events and the order in which they occur.

The following events can occur (with corresponding probabilities):

- u_* transmits to u_1 with probability π
- u_* transmits to u_2 with probability π, after which u_* or u_2 transmits to u_1 with probability $\frac{2\beta}{2\gamma+2\beta+2\omega} = \theta_1$.
- u_* transmits to u_2 with probability π, then u_1 distances itself from either u_* or u_2 with probability θ_2, and finally u_1 becomes infected by the clique member it did not distance itself from with probability θ_1
- u_* transmits to u_2 with probability π, then u_* or u_2 recovers with probability θ_3, and finally u_1 becomes infected by the clique member that is still infectious with probability θ_1
- u_2 distances itself from u_* with probability $\frac{\omega}{\gamma+2\beta+2\omega}$ after which u_* transmits to u_1 with probability θ_1
- u_1 distances itself from u_* with probability $\frac{\omega}{\gamma+2\beta+2\omega}$, then u_* transmits to u_2 with probability θ_1, and finally u_2 transmits to u_1 with probability θ_1

The probability that u_2 becomes infected is obtained by interchanging the names u_1 and u_2, so we simply multiply by a factor 2 in the expected number of infected cliques generated by u_1 and u_2.

Putting these pieces together, we find that the expressions for the K_{ij}, $i,j = 1,2$, are as follows:

$$\begin{aligned}
K_{11} &= \theta_1(\tilde{\mu}_D - 1) + 2\theta_1\mu_D \left(\frac{\beta}{\gamma+2\beta+2\omega} + \frac{\beta}{\gamma+2\beta+2\omega}\theta_1 + \frac{\beta}{\gamma+2\beta+2\omega}\theta_1\theta_2 \right. \\
&\qquad \left. + \frac{\beta}{\gamma+2\beta+2\omega}\theta_3\theta_1 + \frac{\omega}{\gamma+2\beta+2\omega}\theta_1 + \frac{\omega}{\gamma+2\beta+2\omega}\theta_1^2 \right) \\
&= \theta_1(\tilde{\mu}_D - 1) + 2\theta_1\mu_D \left(\frac{\beta}{\gamma+2\beta+2\omega}(1 + \theta_1 + \theta_1\theta_2 + \theta_1\theta_3) + \frac{\omega}{\gamma+2\beta+2\omega}\theta_1(1+\theta_1) \right) \\
K_{21} &= \theta_1\mu_D + 2\theta_1\mu_D \left(\frac{\beta}{\gamma+2\beta+2\omega}(1 + \theta_1 + \theta_1\theta_2 + \theta_1\theta_3) + \frac{\omega}{\gamma+2\beta+2\omega}\theta_1(1+\theta_1) \right)
\end{aligned}$$

Note that the sum $2\frac{\omega}{(\gamma+2\beta+2\omega)}\theta_1^2 = \frac{\beta}{\gamma+2\beta+2\omega}\theta_1\theta_2 + \frac{\omega}{\gamma+2\beta+2\omega}\theta_1^2$ of the two terms $\frac{\beta}{\gamma+2\beta+2\omega}\theta_1\theta_2$ and $\frac{\omega}{\gamma+2\beta+2\omega}\theta_1^2$ in K_{11} and K_{21} is the probability of social distancing of an initially susceptible clique member that later becomes infected.

Finally, the clique reproduction number R_* is the dominant eigenvalue of the 2×2 matrix $K = (K_{ij})_{i,j=1,2}$. Therefore, we have an explicit expression of R_* in terms of the model parameters:

$$R_* = \frac{\mathrm{tr} + \sqrt{\mathrm{tr}^2 - 4\mathrm{det}}}{2}, \tag{S.9}$$

where $\mathrm{tr} = K_{11} + K_{22}$ and $\mathrm{det} = K_{11}K_{22} - K_{12}K_{21}$ denote the trace and determinant of the matrix K, respectively.

Section S2.2 Social distancing can increase R_* but this depends on the network structure

We consider two degree distributions for the degree D of the global network, and using (S.9) for R_*, show that R_* *can* increase for sufficiently small distancing rate $\omega > 0$ and recovery rate $\gamma > 0$. We consider a third choice for D to show that this need not always be the case. Assume that $\gamma = 0$ and $\alpha = 1$ and write $R_* = R_*(\omega)$.

First, suppose that $p_1 = p = 1 - p_0$. Then $\mu_D = p$ and $\mu_{\tilde{D}-1} = 0$. Then $R_*(0) = 2p$ and $R'_*(0) = (3 - 2p)/\beta$. In particular, if $p = 1/2$, then $R_*(0) = 1$, and $R'_*(0) > 0$. Therefore, for sufficiently small $\gamma > 0$ and sufficiently large $\alpha < 1$, an increasing rewiring rate $\omega > 0$ can push the epidemic threshold parameter R_* from below to above the threshold value of one as shown numerically in Fig. 2A and Fig. 2C of the main text.

Next, suppose that D is Poisson with mean μ_D. Then $\tilde{D} - 1$ is also Poisson with mean μ_D. Then $R_*(0) = 3\mu_D$, and $R'_*(0) = (2 - 3\mu_D)/\beta$. In particular, if $\mu_D = 1/3$, then $R_*(0) = 1$, and $R'_*(0) > 0$. Thus, for sufficiently small $\gamma > 0$ and sufficiently large $\alpha < 1$ and μ_D close to $1/3$, an increasing rewiring rate ω can push the epidemic threshold parameter R_* from below to above the threshold value of one as shown numerically in Fig. S2A.

Note that the derivations above do not rely on the specific choice $\alpha = 1$. Indeed, we have an analytical expression (S.9) for R_* that is explicit in terms of the model parameters so we can very well do the same derivations for some fixed value of α. By doing so, we can also derive a threshold for α. For the model with $p_1 = p = 1 - p_0$, one gets $\text{sign}[R'_*(0)] = \text{sign}[3\alpha - 2p]$ so $R'_*(0) > 0$ if $\alpha > 2p/3$. In particular, if $p = 1/2$ then $R'_*(0) > 0$ if $\alpha > 1/3$. For the Poisson model, one gets $\text{sign}[R'_*(0)] = \text{sign}[2\alpha - 3\mu_D]$, so $R'_*(0) > 0$ if $\alpha > 3\mu_D/2$. In particular, if $\mu_D = 1/3$ then $R'_*(0) > 0$ if $\alpha > 1/2$. Note that this implies that the same conclusions hold for R_* for a much larger range of probabilities α than illustrated in Fig. 2 of the main manuscript. Concretely, for the model with $p_1 = p = 1 - p_0$, R_* can cross the threshold value of one for the range $1/3 < \alpha \leq 1$ and for the model with Poisson degree the range is $1/2 < \alpha \leq 1$. We illustrate these results in Fig. S5 for the model with degree distribution $p_1 = p = 1 - p_0$.

Finally, to show that the result for R_* does not hold in general, we consider the following distribution. Suppose that $p_d = 1$, for some fixed $d > 1$, so all individuals have degree d. Then $\mu_D = d$ and $\mu_{\tilde{D}-1} = d-1$, whence $R_*(0) = 3\mu_D - 1$ and $R'_*(0) = -(1-12\mu_D+9\mu_D^2)/(\beta(3\mu_D-1)$. Note that $1 - 12\mu_D + 9\mu_D^2 = (3\mu_D - 2)^2 > 0$ as $d > 1$. We find that $R'_*(0) < 0$ for $d > 1$. This shows that rewiring decreases the epidemic threshold parameter R_* for sufficiently small $\omega > 0$ (and $\alpha < 1$, $\gamma > 0$) as illustrated in Fig. S2B.

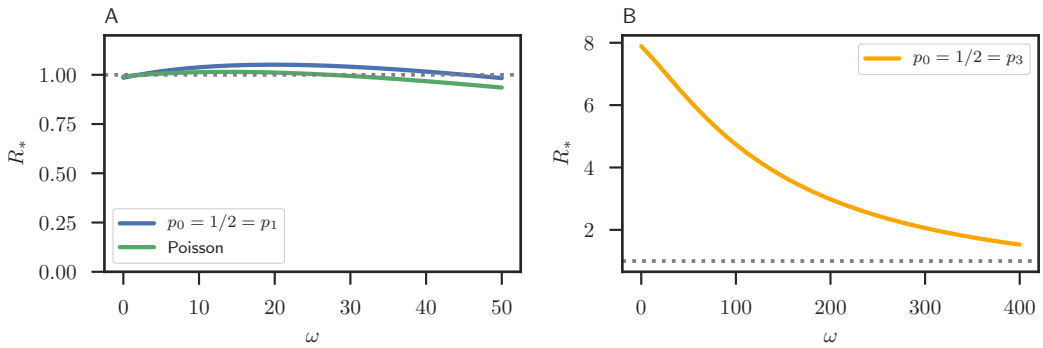

Figure S2: R_* for the clique-network model with cliques of size 3 with mean infectious period $1/\gamma = 5$ days, $\beta = 200$, $\alpha = 0.9$ (A) Social distancing can increase the epidemic threshold parameter R_* from below to above its threshold value of one for clique-network models. Illustration with two-point degree distribution with $p_0 = 1/2 = p_1$ (blue line) and Poisson degree distribution with mean $\mu_D = 0.335$ (green line) (B) Social distancing does not have to increase the epidemic threshold parameter R_*. Illustration with degree distribution $p_3 = 1$. Results for each value of ω are based on 500 simulations.

Section S3 Real-world networks

Section S3.1 Properties of the real-world networks

We summarize the most important characteristics of the real-world networks for our purpose, other network summary statistics and more details are found in [12, 9, 11].

The 'arXiv General Relativity' collaboration network describes scientific collaborations between authors that submitted papers to the arXiv in the General Relativity and Quantum Cosmology category. Edges between nodes represent two co-authors that have written a paper together. There are 5241 nodes and 14484 edges in the network. In total there are 354 connected components that make up the network and the largest connected component covers a fraction of 0.793 nodes and 0.926 edges. The minimum and maximum degree in the population are 1 and 81, with a mean of 5.53 and a variance of 62.7. The median degree is 3. There are in total 777 nodes with degree 3, and out of these nodes, 676 are part of the largest connected component. We choose the index case at random from these 676 nodes.

In the 'Facebook social circles' network, nodes are survey participants of the social network website Facebook that were using a specific app. Edges between nodes represent the 'circles' or 'friends lists' of those participants. There are 4039 nodes and 88234 edges. The largest network component is precisely the network itself. The minimum and maximum degree in the population are 1 and 1045, with a mean of 43.70 and a variance of 2748.44. The median degree in the population is 25, and there are 55 individuals with this median degree of 25. We choose an index case at random from this set of individuals.

Finally, we consider a set of networks that has been collected over a period of 69 days, available at [9]. In the 'Science Gallery' (SG) networks, nodes represent visitors of an exhibition in the Science Gallery in Dublin. In principle, each day there is a unique set of visitors, and therefore network. Edges represent close-range face-to-face proximity between nodes (details on data collection are found in [9, 11]). In principle the data takes into account the duration of contacts through weights on edges but since this is not accounted for in our social distancing model, we neglect the weights and temporal aspects of the network. Instead we use the available (daily)

aggregated networks of [9]. We consider two SG networks that are supposed to be representative for the collected networks over the 69 days, namely 20 May 2009 and 14 July 2009 [11] (which we simply refer to as SG1 and SG2, respectively).

SG1 contains 194 nodes and 958 edges. There are 5 isolated nodes that we ignore in the network. The other 189 nodes make up a network through 5 connected components of sizes 2, 3, 5, 89, and 90. The two largest connected components of sizes 89 and 90 cover a fraction of 0.471 and 0.476 of all nodes and 0.460 and 0.531 of all edges, respectively. The minimum and maximum degree in the population are 1 and 35, with a mean of 10.14 and a variance of 51.5. The median degree is 9. There are in total 9 nodes with degree 9, five of which are in the component of size 90 and the other four are in the component of size 89. We choose the index case at random from these 9 nodes.

SG2 contains 139 nodes and 433 edges. There are 2 isolated nodes that we ignore in the network. The other 137 nodes are all in one connected component. The minimum and maximum degree in the population are 1 and 20, with a mean of 6.32 and a variance of 16.44. The median degree is 6. There are in total 18 nodes with degree 6. We choose the index case at random from these 18 nodes.

Section S4 Effects of dropping and/or rewiring

As previously noted, the setting that $\alpha = 0$, i.e. all social distancing is done through dropping of edges, is always beneficial for the population level. Therefore, only for α large enough, we will find negative population level effects. However, whether α is large enough depends much on the network structure at hand and the choice of other parameter values. In this section we illustrate the analytical results of Section S1.3 and Section S2.2 that negative population level effects can occur for a range of α-values larger than the $\alpha = 0.9$ taken in the studies in the main manuscript. At the same time we show that these negative effects do not always occur.

First, in Fig. S3 we consider the scenario of Fig. 1 in the main manuscript, and compare the effect of $\alpha = 0.9$ to $\alpha = 0.5$. We see that when $\alpha = 0.5$, i.e. half of social distancing is through dropping of edges, then this has a positive population level effect in the setting presented in Fig. S3.

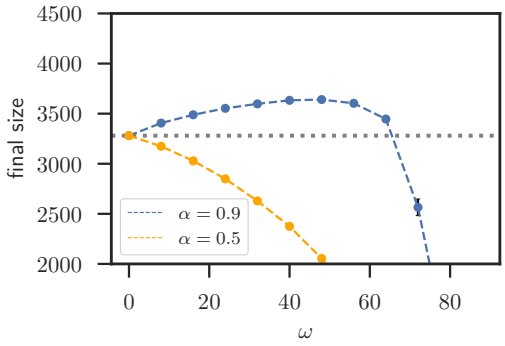

Figure S3: The effect of the probability α that social distancing is through rewiring rather than dropping on the final size of the epidemic given a major outbreak on the configuration network with degree distribution $\{p_d\}_{d=0}^{10}$ where $p_d = c/(d+1)$, $d = 0, 1, \ldots, 10$, with $c = 0.331$ the normalization constant. Dotted horizontal line is for comparison with the size at $\omega = 0$. Model parameters are $1/\gamma = 5$ days, $\beta = 20$ days^{-1}, $\alpha = 0.9$ and 0.5. Compare with Fig. 1 of the main manuscript (note that the y-axis is scaled differently for better comparison with α values). Results for each value of ω are based on 500 simulations. The epidemic is initiated with 10 index case that are chosen uniformly at random.

In Section S1.3 we proved that social distancing can have negative population level effects on the final size of the epidemic for $\alpha \in (0, 1]$. We illustrate this by considering an epidemic on a configuration network with degree distribution $p_0 = 1/2 = p_5$ for different values of α in Fig. S4. Note that generally the increase in the final size is smaller and occurs for a smaller range of social distancing rates $\omega > 0$ for smaller values of α.

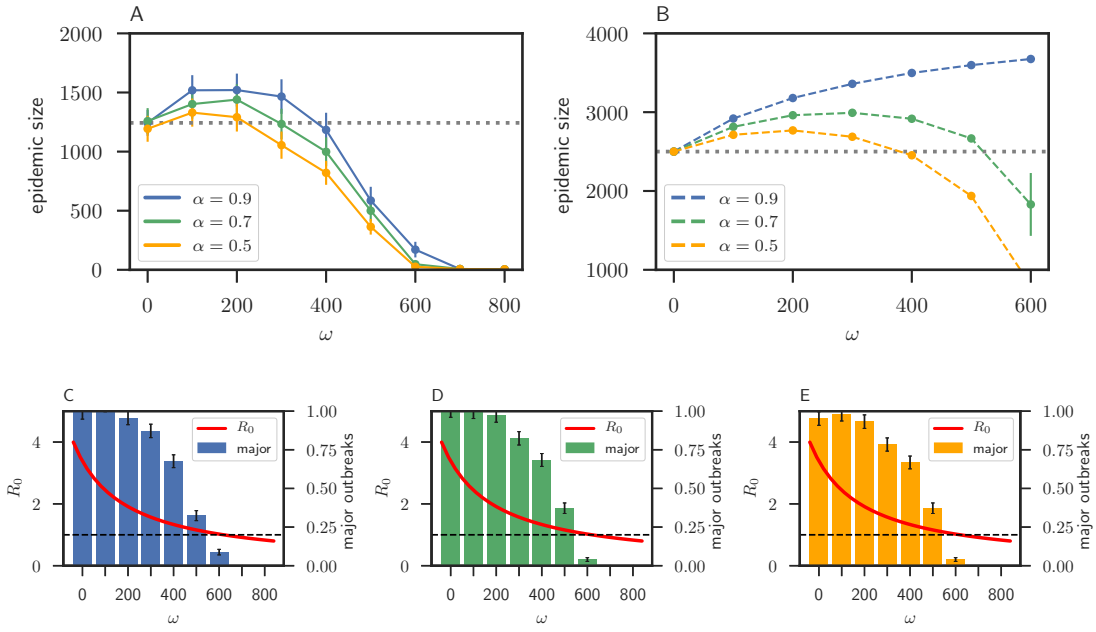

Figure S4: The effect of the probability α that social distancing is through rewiring rather than dropping on the final size of the epidemic over all epidemic outbreaks (A) and given a major outbreak (B) on the configuration network with degree distribution $p_0 = 1/2 = p_5$; dotted horizontal line is for comparison with the size at $\omega = 0$. Model parameters are $1/\gamma = 5$ days, $\beta = 200$ days^{-1}, and $\alpha = 0.9$ (C), $\alpha = 0.7$ (D), $\alpha = 0.5$ (E). (Note that the scaling of the y-axis is different in (A) and (B).) Results for each value of ω are based on 500 simulations. The index case is chosen uniformly at random from degree 5 individuals.

Analysis in Section S2.2 shows that social distancing can push the epidemic threshold parameter R_* across the threshold value of one in a clique network for a range of α values. We illustrate this by considering the clique network with degree distribution $p_0 = 1/2 = p_1$ in Fig. S5. We see in Figs. S5A and B that increasing ω can lead to R_* crossing the threshold value of one for both $\alpha = 0.9$ and $\alpha = 0.5$. Note however that parameter values are such that for $\alpha = 0.5$, R_* remains very close to one, and therefore few epidemics result in major outbreaks. Furthermore, the negative population level effect is generally smaller for smaller values of α.

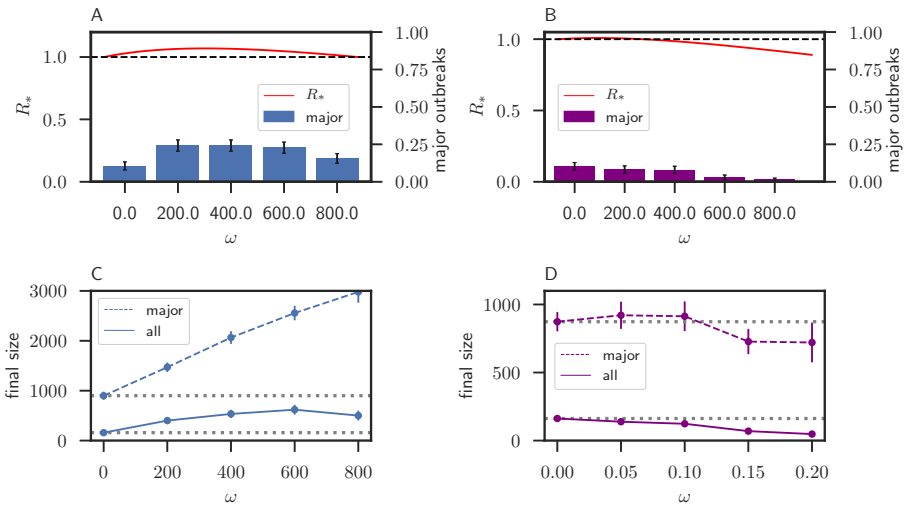

Figure S5: The effect of the probability α that social distancing is through rewiring rather than dropping on the final size of the epidemic given a major outbreak on the clique network with degree distribution $p_0 = 1/2 = p_1$; dotted hoirzontal lines are for comparison with the size at $\omega = 0$. Model parameters are $1/\gamma = 5$ days, $\beta = 4000$ days^{-1}, and (A, C) $\alpha = 0.9$, (B, D) $\alpha = 0.5$ (note that the y-axis is scaled differently in Figs. C and D. Results for each value of ω are based on 500 simulations. The index case is chosen uniformly at random.

In Fig. S6 we find that the final size given a major outbreak can increase for small enough social distancing rates ω under different rewiring probabilities α for an epidemic on the arXiv collaboration network. Note however that the range of social distancing rates ω for which this occurs, and the extent to which the final size can increase depend on the value of α. In contrast, in Fig. S7 we find that the probability α has little effect on the final size given a major outbreak for an epidemic on the facebook social circles network.

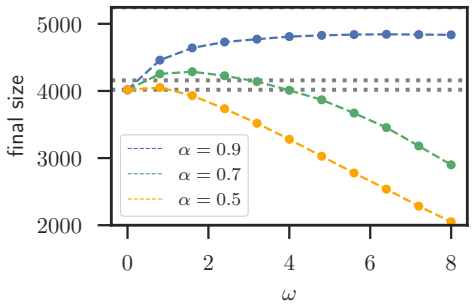

Figure S6: The effect of the probability α that social distancing is through rewiring rather than dropping on the final size of the epidemic given a major outbreak on the arXiv collaboration network; dotted hoirzontal lines are for the size of the giant component (top) and comparison with the size at $\omega = 0$. Model parameters are $1/\gamma = 5$ days, $\beta = 2$ days^{-1}, and $\alpha = 0.9$, 0.7, and 0.5. Results for each value of ω are based on 500 simulations. The index case is chosen uniformly at random from the sub-population of individuals that has median degree and are part of the largest connected component of the network. Compare with Fig. 3 of the main manuscript (note that the y-axis is scaled from 2000 to 5241 for better comparison between the α).

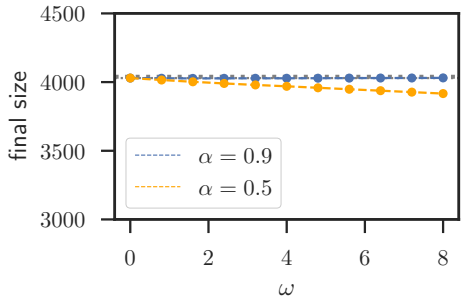

Figure S7: The effect of the probability α that social distancing is through rewiring rather than dropping on the final size of the epidemic given a major outbreak on the facebook social circles network; dotted horizontal lines are for the size of the network (top) and comparison with the size at $\omega = 0$. Model parameters are $1/\gamma = 5$ days, $\beta = 2$ days^{-1}, and $\alpha = 0.9$, and 0.5. Results for each value of ω are based on 500 simulations. The index case is chosen uniformly at random from the sub-population of individuals that has median degree. Compare with Fig. 5 of the main manuscript (note that the y-axis is scaled from 3000 to 4500 for better comparison between the α).

Finally, we consider the effect of the transmission rate β and the probability α on the two Science Gallery networks SG1 ad SG2 in Figs. S8 and S9 (where we adjust for the range of ω-values while keeping the recovery rate γ fixed at 0.2/day). We find that the results are qualitatively the same for different sets of parameter values. Only the results for the average final size given a major outbreak are presented together with the fraction of simulations resulting in major outbreaks. Quantitatively, we find that decreasing β and α result in smaller average final sizes, which is in line with what is expected. Note that SG1 allows for a large range of α-values for which the negative population-level effect on the final size can still be observed (as compared to e.g. arXiv collaboration network in Fig. S6). A major structural difference in

the networks SG1 and the arXiv collaboration network is that the first network has two large connected components while the second has one main connected component. In SG1, rewiring makes it likely that the two connected components become connected through rewiring.

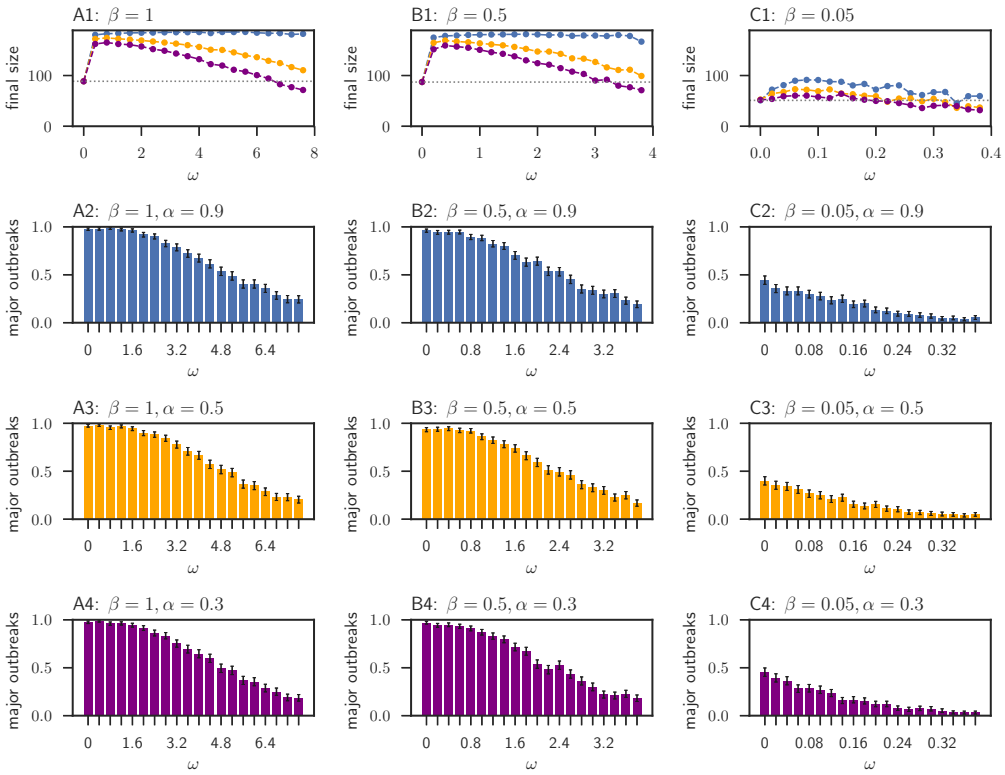

Figure S8: The effect of the transmission rate β and the probability α that social distancing is through rewiring rather than dropping on the final size of the epidemic given a major outbreak on the SG1 network; dotted horizontal line is for comparison with the size at $\omega = 0$. The mean infectious period is $1/\gamma = 5$ days. Note that the range of ω-values is adjusted for each value of β. Results for each value of ω are based on 500 simulations. The index case is chosen uniformly at random from the sub-population of individuals that has median degree.

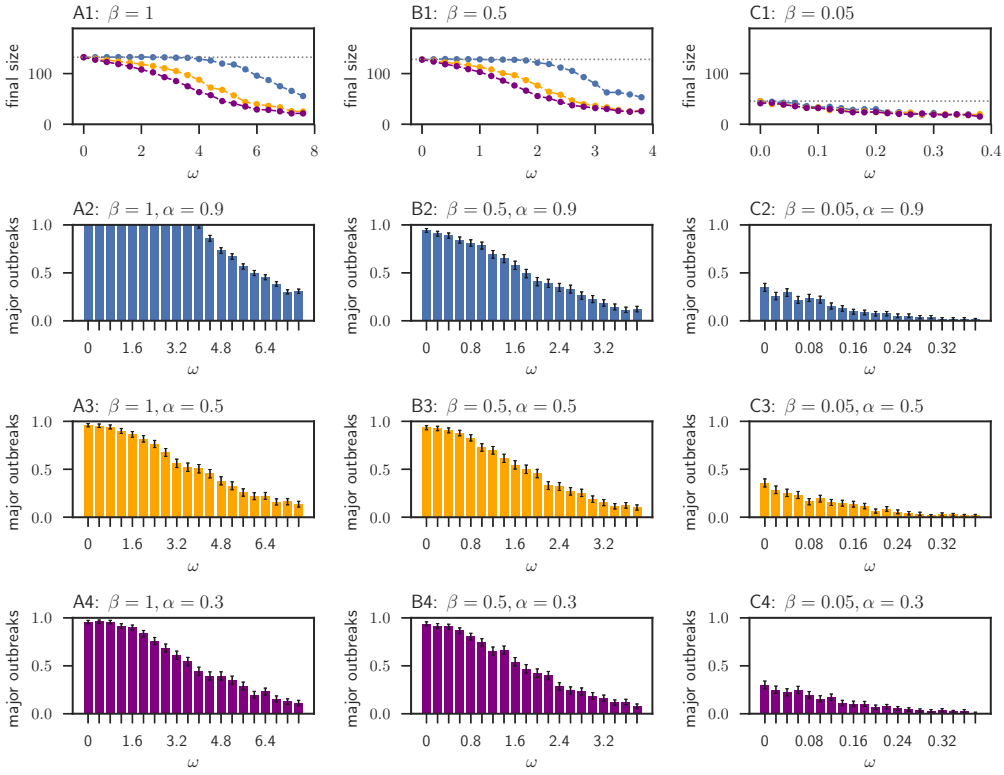

Figure S9: The effect of the transmission rate β and the probability α that social distancing is through rewiring rather than dropping on the final size of the epidemic given a major outbreak on the SG2 network; dotted horizontal line is for comparison with the size at $\omega = 0$. The mean infectious period is $1/\gamma = 5$ days. Note that the range of ω-values is adjusted for each value of β. Results for each value of ω are based on 500 simulations. The index case is chosen uniformly at random from the sub-population of individuals that has median degree.

Section S5 Simulation algorithm

We are interested in the epidemic final size and the probability of a major outbreak for the social distancing model applied to (i) the configuration network model, (ii) the clique network model, and (iii) two real-world networks. The social distancing model is mathematically challenging to analyse. Although we have some analytical results (presented in Sections Section S1 and Section S2), e.g. for the threshold parameter R_0 for the configuration network model and R_* for the clique network model, other quantities of interest, i.e. the final size and the probability of a major outbreak we have not managed to characterize mathematically. Moreover, on large real-world networks it is impractical to exactly compute the quantities we are interested in. The results in this text are obtained through simulation methods.

Given the structure of the networks it is fairly straightforward (since the epidemic process is a continuous-time Markov chain) to write code to simulate realisations of the final size of stochastic SIR epidemic process described in section *Model description* of the main article. By inspecting histograms of these simulated final sizes we find that the cut-off of 10% of population size (as stated in section *Model description* of the main text) is satisfactory for the networks we consider.

The simulation program is set up such that it keeps track of the network structure through (i) a unique labelling of all the nodes in the population and (ii) a list of edges between corresponding nodes at each time step (additionally we track the degree of each node). The epidemic dynamics are tracked through (i) a list of the disease states of all nodes and (ii) a list of edges between susceptible and infectious nodes (additionally we track the number of edges between susceptible and infectious nodes and the number of infectious nodes). The events that may occur are (i) recovery of an infectious node (update the list of disease states and the list of edges between susceptible and infectious nodes), (ii) transmission from an infectious to susceptible neighbour (update the list of disease states and the list of edges between susceptible and infectious nodes), (iii) dropping of an edge from a susceptible to an infectious neighbour (update the list of existing edges in the network and the list of edges between susceptible and infectious nodes), and (iv) rewiring of an edge from a susceptible to an infectious neighbour (update the list of existing edges in the network and the list of edges between susceptible and infectious nodes). The epidemic stops when there are no more infectious nodes in the population.

We use the Gillespie algorithm to determine the next event and the time to the next event. The program additionally keeps track of the event times and number of infectious nodes at each time step, but these quantities are not used in the analysis in the manuscript. A simulation run is performed for a given network, either generated through a model network or a real-world network. In both cases the network information required is the total population size, the degrees of each node, and the list of edges between nodes. An epidemic can be initiated by a random number of index cases or by choosing a specific node satisfying specific properties. The initialization of the epidemic determines the list of diseases states and the list of edges between susceptible and infectious nodes. The final size of an epidemic is the number of recovered nodes at the end of the simulation that is found through the list of disease states of all nodes. The main simulation program is found in an additional SI file to the main manuscript where also more details on the key variables can be found and an explanation of how the real-world networks are extracted to use in the simulations (code for extracting the real-world networks is available as separate SI files). The file rewire_sim.py contains functions to generate random graphs and simulate epidemics upon them, extract_networks.py contains code to import the empirical graphs from a text file to an appropriate format for use with the simulation code, extractPos.py contains helper functions for extracting the networks.

In plots depicting the results of simulations we present point estimates, for example the proportion of simulations that result in a final size greater than 10% of population size as an estimate of major outbreak probability, the mean final size amongst outbreaks with a final size greater than 10% of population size as an estimate of the relative final size of a major outbreak. We also present confidence intervals (CI) around these estimates in cases where those intervals are large enough to be visible on the scales used in the plots.

In simulations that are based on the configuration network construction imperfections of the network such as self-loops and multiple edges may arise. However, there are generally only a small number of such imperfections. The theory (see [10, Thm. 3.1.2]) predicts that the number of self-loops and pairs of parallel edges should be asymptotically independent Poisson variables with mean $\mu = \sum_k k(k-1)p_k/2\mu_D$ and μ^2, respectively. Table 1 shows good agreement, as the population size increases, with this prediction of an average of $\mu + 2\mu^2 \approx 14.4$ edges involved in imperfections for the configuration network with degree distribution $p_d = c/(d+1)$, $d = 0, \ldots, 10$ and $c = 0.331$ the normalization constant.

Population size	Number of edges	Number of imperfections	Fraction of imperfections
100	134	13.51	0.10
500	663	14.73	0.022
1000	1318	14.56	0.011
5000	6612	14.21	0.0021

Table 1: The average number of self-loops and multiple edges for different population sizes for the configuration network with degree distribution $\{p_d\}_{d=0}^{10}$ with $p_d = c/(d+1)$, and $c = 0.331$ the normalization constant. Imperfections are counted as 1 for self-loops and 2 for each pair of multiple edges, so the last column of the table gives the fraction of edges involved in such imperfections.

References

[1] F. Ball and D. Sirl. Acquaintance vaccination in an epidemic on a random graph with specified degree distribution. *J Appl Prob*, 50:1147–1168, 2013.

[2] F. Ball, D. Mollison, and G. Scalia-Tomba. Epidemics with two levels of mixing. *Ann. Appl. Probab.*, 7(1):46–89, 1997.

[3] F. Ball, D. Sirl, and P. Trapman. Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Adv Appl Probab*, 41:765–796, 2009.

[4] F. Ball, D. Sirl, and P. Trapman. Analysis of a stochastic SIR epidemic on a random network incorporating household structure. *Math Biosci*, 224:53–73, 2010.

[5] F. Ball, L. Pellis, and P. Trapman. Reproduction numbers for epidemic models with households and other social structures II: Comparisons and implications for vaccination. *Math. Biosci.*, 274:108–139, 2016.

[6] N. G. Becker and K. Dietz. The effect of household distribution on transmission and control of highly infectious diseases. *Math. Biosci.*, 127:207–219, 1995.

[7] B. Bollobás. *Random Graphs*. Academic Press, New York, 2001.

[8] T. Britton, D. Juher, and J. Saldaña. A network epidemic model with preventive rewiring: comparative analysis of the initial phase. *Bull Math Biol*, 78:2427–2454, 2016.

[9] The SocioPatterns collaboration. Infectious sociopatterns. `http://www.sociopatterns.org/datasets/infectious-sociopatterns/`, 2011.

[10] R. Durrett. *Random graph dynamics*. Cambridge University Press, 2006.

[11] L. Isella, J. Stehle, A. Barrat, C. Cattuto, J.-F. Pinton, and W. Van den Broeck. What's in a crowd? Analysis of face-to-face behavioral networks. *J Theor Biol*, 271:166–180, 2011.

[12] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. `http://snap.stanford.edu/data`, June 2014.