Infectious disease dynamics and restrictions on social gathering size

Christopher Boyer[1,*], Eva Rumpler[1,*], Stephen Kissler[*], and Marc Lipsitch[*]

[*]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

Draft Version: December 31, 2021

Abstract

Social gatherings can be an important locus of transmission for many pathogens including SARS-CoV-2. During an outbreak, restricting the size of these gatherings is one of several non-pharmaceutical interventions available to policy-makers to reduce transmission. Often these restrictions take the form of prohibitions on gatherings above a certain size. While it is generally agreed that such restrictions reduce contacts, the specific size threshold separating "allowed" from "prohibited" gatherings often does not have a clear scientific basis, which leads to dramatic differences in guidance across location and time. Building on the observation that gathering size distributions are often heavy-tailed, we develop a theoretical model of transmission during gatherings and their contribution to general disease dynamics. We find that a key, but often overlooked, determinant of the optimal threshold is the distribution of gathering sizes. Using data on pre-pandemic contact patterns from several sources as well as empirical estimates of transmission parameters for SARS-CoV-2, we apply our model to better understand relationship between restriction threshold and reduction in cases. We find that, under reasonable transmission parameter ranges, restrictions may have to be set quite low to have any demonstrable effect on cases due to relative frequency of smaller gatherings. We compare our conceptual model with observed changes in reported contacts during lockdown in March of 2020.

Keywords: epidemiology, gatherings, COVID-19, SARS-CoV-2, non-pharmaceutical intervention

[1] C. B. and E R. are co-first authors. Corresponding emails: cboyer@g.harvard.edu and erumpler@g.harvard.edu

1 Introduction

Social gatherings in which people meet and interact provide a conducive environment for the spread of pathogens. During an outbreak, restricting the size of such gatherings is one of several nonpharmaceutical interventions (NPIs) available to policymakers. An advantage of these restrictions is that they are simple to articulate and easy for the public to understand and, in some circumstances, for authorities to enforce. Indeed, during the COVID-19 outbreak, gathering size restrictions were among the most commonly used NPIs globally [1]. Some have claimed that these restrictions were among the most effective at reducing transmission [2–5]; however, given the rapid and often haphazard nature of their rollout and the methodological challenges of proper indentification, estimates of the causal effects of specific NPIs may be severely biased [6].

Theory and intuition suggest that, when properly followed, gathering size restrictions should lower transmission by limiting the number of contacts between individuals and thus reducing the opportunity for the pathogen to spread. Yet, it's often unclear precisely how low restrictions should be set to achieve a certain disease control target, be it a stable number of cases or the elimination of the pathogen from the population. Indeed, evidence suggests, policymakers have taken a variety of approaches in practice to set gathering size thresholds, which may reflect different goals or local disease dynamics, but also might reflect ambiguity in the optimal strategy. As a case in point, in the UK the government first banned gatherings above 500 in March 2020 before initiating a lockdown on March 23. Then after restrictions eased in the late summer a ban on gatherings above 30 was declared, but then this was famously revised down to the "rule of six" in September to prevent gatherings with more than six people. In this paper, we use epidemiological theory to better understand the relationship between gathering size and general disease dynamics. We also attempt to enumerate the necessary elements to quantify or predict the impact of a given threshold on the incidence of new cases.

We emphasize restrictions on gathering sizes for several reasons. Firstly, we note that a significant proportion of the superspreading events in the literature, including some of the most spectacular accounts, have occurred during social gatherings. Given the outsized role these events seem to play at the start of outbreaks, some have hypothesized that control and/or suppression of an emerging pathogen could largely be achieved via targeted reduction in mass gatherings. Secondly,

several retrospective reports comparing confirmed COVID-19 cases and test-negative controls [7, 8] have found an association between attending family and friends gatherings and infection with SARS-CoV-2, suggesting that gatherings may be important source of new cases. Thirdly, social gathering restrictions seem to be among the first and most frequent measures to be implemented, which perhaps can be explained by a perception that social gatherings have less social value than other gatherings that occur in venues such a schools and hospitals. Finally, while both the United States and European Centers for Disease Control recommend limiting the size and duration of gatherings [9, 10], the specific timing of when to implement restrictions and the numeric threshold separating "allowable" from "prohibited" gatherings generally do not have a clear scientific basis. This can lead to dramatic differences in guidance across location and time. For instance, while most countries implemented limitations in the spring of 2020, the intensity and duration of these restrictions varied extensively from country to country [11] with maximum gathering sizes permitted ranging from 2 to 5000 and subject to frequent and somewhat erratic changes as the epidemic progressed.

2 Theory

As in the standard compartmental model, we consider the epidemic spread of a pathogen in a population which can be divided into three disjoint sets of individuals: susceptible and not yet infected (S), infected and infectious (I), or recovered, no longer infectious, and immune (R). As time passes, individuals in the population come into contact with one another and the pathogen spreads through contacts between susceptible and infectious individuals. Gathering size restrictions limit the number of contacts that individuals have, but apply only to a subset of contacts that occur during social gatherings. Therefore, we categorize all contacts between individuals as either occurring during "gatherings", i.e. non-household settings that are presumably affected by gathering size restrictions, or at home or other settings not affected by gathering size restrictions, and we focus on the former as the source of the contribution of gatherings to disease dynamics.

At each time point, individuals attend M gatherings of size K, where K is a random variable defined by some distribution $f(k)$. To simplify matters here we include the possibility that an individual does not attend a gathering in $f(k)$ through defining it as a gathering size of 1 so that

the same distribution applies to everyone . We assume for now that policies target the expected number of new infections that can be attributed to gatherings, $E(X_t)$, which, using the law of total probability, can be written as:

$$E(X_t) = \sum_k E(X_t \mid K = k) f(k) \tag{1}$$

where $E(X_t \mid K = k)$ is the expected incidence of new infections at a given gathering size of $K = k$. This expression suggests that we could write the expected rate under an idealized restriction, that is a restriction that is strictly enforced such that no one attends a gathering with $k > k_{max}$ as:

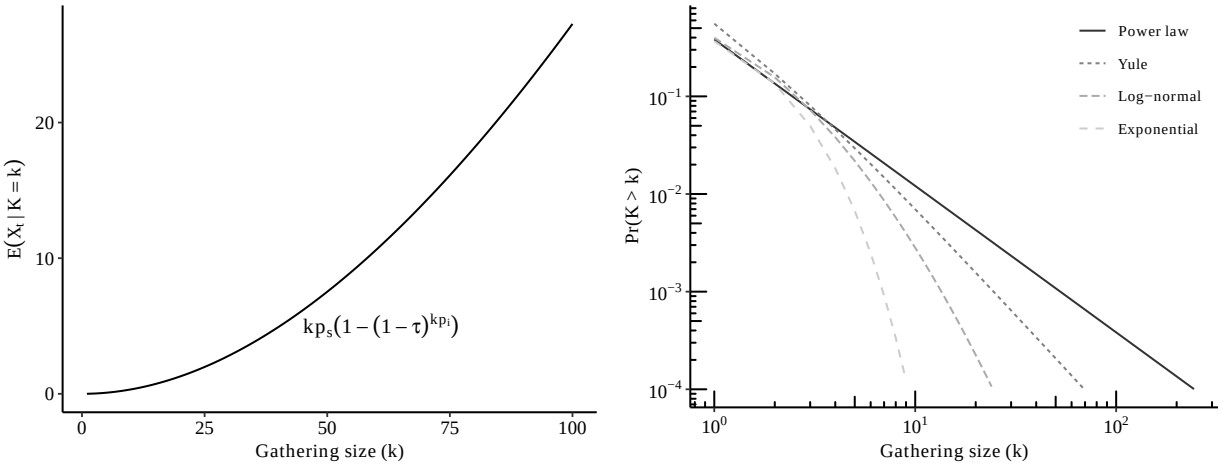
$$E(X_t^{k_{max}}) = \sum_{k=1}^{k_{max}} E(X_t \mid K = k) f_{k_{max}}(k) \tag{2}$$

where the sum is now over the restricted range of gathering sizes and $f_{k_{max}}(k)$ is the distribution of gathering sizes after the restriction has been applied recognizing that it could differ from simply a truncated $f(k)$ as people may respond to the restriction in different ways. Therefore, in order to estimate the potential impact of a gathering size restriction, we need two essential inputs: (1) the distribution of gathering sizes and (2) the relationship between gathering size and expected number of infections. Then, given a range of k_{max} values, policymakers could ideally target a specific reduction in new cases X_t^*, and select $k^* = \max(k_{max})$ such that $X_t < X_t^*$, pserhaps weighing them against the cost of imposing the restriction.

Starting with the second input, as we show in section A.1 of the Appendix, the expected number of incident cases X_t that occur at a gathering of size $K = k$ is :

$$E(X_t \mid K = k) = kp_s(1 - (1 - \tau)^{kp_i}) \tag{3}$$

where τ is the probability of transmission given contact, and p_s and p_i are the population proportions of susceptible and infectious individuals respectively. We assume susceptible, infectious, and recovered individuals attend gatherings at rates roughly equivalent to their population proportions and that everyone who attends a gathering comes into contact with all other attendees. Figure 1a plots this expression for example values of τ, p_i, and p_s. As intuition might suggest, it shows that,



(a) expected incident infections given size

(b) gathering size distribution

Figure 1: Plots of necessary components of effect of gathering size restriction.

for a single gathering, larger gatherings produce more secondary infections than smaller gatherings and that this relationship is nonlinear as larger gatherings both increase the number of potential contacts and as well as the expected number of infectious individuals in attendance. Indeed, as shown in section A.2 of the Appendix, for small values of τ and p_i that are typical of an infectious disease outbreak, i.e. $|\tau k p_i| \ll 1$, we can use a Binomial approximation to simplify this to:

$$E(X_t \mid K = k) \approx k^2 p_s p_i \tau \tag{4}$$

which is quadratic in the size of the gathering.

As for the other input, the distribution of gathering sizes, empirical studies suggest that human contact distributions may be subexponential, or even scale-free or heavy-tailed, with considerable probability mass in the extreme tail of the distribution [12–15]. This observation applies equally to distributions of gathering size, i.e. $f(k)$, as most gatherings are small, but gatherings of tens or hundreds of thousands of individuals are possible. Several generative models of human social interaction have been proposed to explain this phenomena based on random walks [16] or attracting sites [12]. Figure 1b shows a few common examples of heavy-tailed distributions. In the extreme case, the limit or asymptotic behavior of these distributions can be characterized by a discrete power law of the form

$$f(k) = \frac{k^{-\alpha}}{\zeta(k_{min}, \alpha)}, \qquad \forall k \geq k_{min} \tag{5}$$

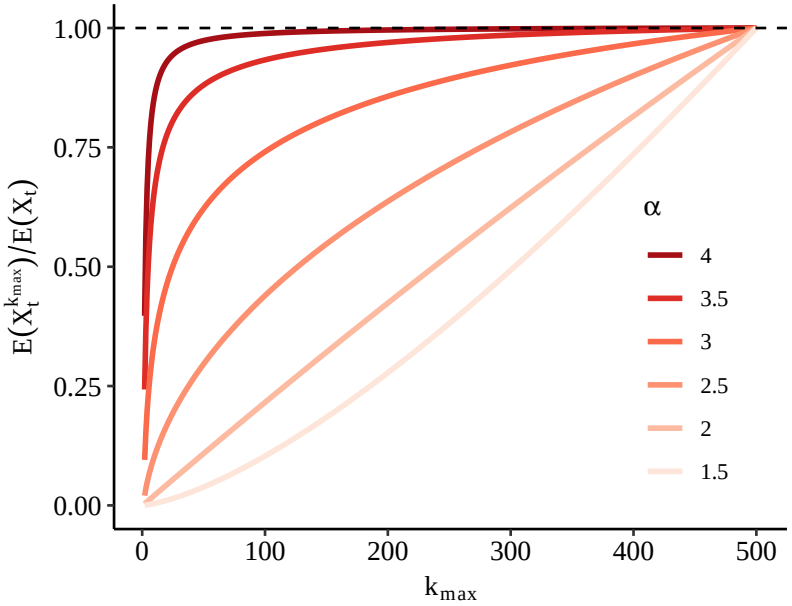
where $\zeta(k_{min}, \alpha) = \sum_{n=0}^{\infty} (n + k_{\min})^{-\alpha}$ is generalized zeta function and k_{min} is the threshold for power-law behavior. This has important implications as moments of power law distributions may not be finite under some parameterizations as the extreme mass in the tail leads to infinite sums or integrals. For instance, it is well known that the number of finite moments of power-law distributions is determined by the value of α, when $\alpha < 3$ the distribution has finite mean but infinite variance and when $\alpha < 2$ the distribution has no finite moments. Many observed phenomena exhibit power-law behavior with $2 \leq \alpha \leq 3$ [17].

Assuming that, in the range of k_{max} restrictions considered, a power law is a good approximation for the distribution of gathering sizes, and combining this with the results in equation 3 and 4, the expected rate of new infections under restriction simplifies to:

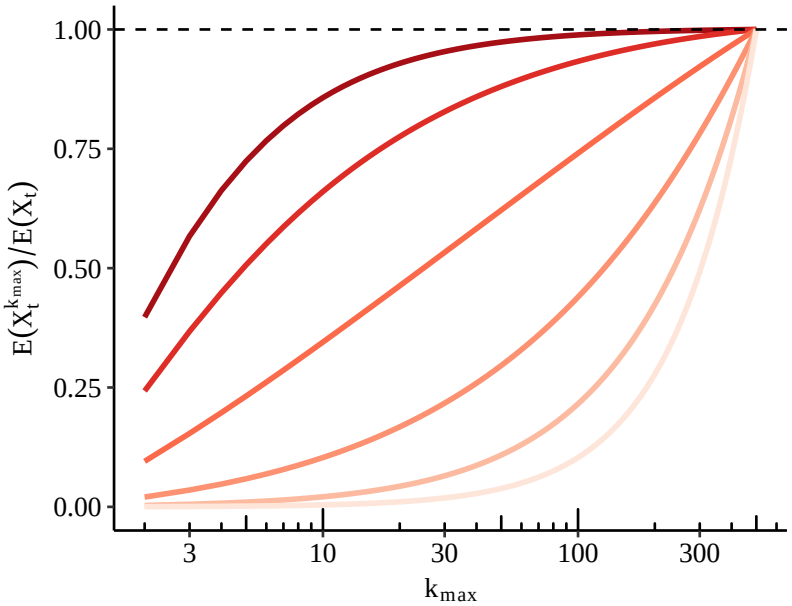
$$\begin{aligned} E(X_t^{k_{max}}) &= \sum_{k=1}^{k_{max}} k p_s (1 - (1-\tau)^{k p_i}) \frac{k^{-\alpha}}{\zeta(k_{min}, \alpha)} \\ &\approx \frac{p_s p_i \tau}{\zeta(k_{min}, \alpha)} \sum_{k=1}^{k_{max}} k^{2-\alpha} \end{aligned} \tag{6}$$

Viewing the $\sum_{k=1}^{k_{max}} k^{2-\alpha}$ as a weighted sum denoting the contribution of gatherings of size between 1 and k_{max} to the rate of new infections yields the following insight: when $\alpha < 2$ the contributions are increasing suggesting that larger gatherings contribute more to the rate of new infections than smaller gatherings and by extension there are diminishing returns to imposing lower restrictions; while, on the other hand, when $\alpha \geq 2$ contributions are flat or decreasing suggesting that smaller gatherings contribute more to the rate of new infections than larger gatherings and by extension there are increasing returns to imposing lower restrictions.

In Figure 2 we plot an example of the relative rate of incident cases under a restriction which prohibits gatherings above size k_{max} for power law distributions of gathering size with different α values. Here, we see that when α is 2 or below, restrictions of larger gatherings quickly leads to a large reduction in cases; however, as α increases vastly more stringent restrictions are required to achieve meaningful reductions. This suggests that the empirical distribution of gathering sizes and the tail-behavior specifically, i.e. the frequency of very large gatherings relative to small ones, are important parameters in determining the optimal threshold for gathering size restrictions.



(a)



(b)

Figure 2: Relative rate of incident cases under restriction which prohibits gatherings above size k_{max} for different power law distributions; shown on both a linear (a) and a log (b) scale.

Notes: Here we fix transmission parameters to following values $p_i = 0.01$, $p_s = 0.99$, $\tau = 0.08$ and assume power law behavior starts at $k_{min} = 1$. We truncate the power law above gatherings of size 500 both to make the sum tractable and given that gathering sizes must at minimum be less than population size. Panel (a) shows the relative rate of incident cases calculated using equation 6 and comparing restrictions with k_{max}-level thresholds to unrestricted rate (e.g. a value of 0.5 implies a 50% fewer per capita incident cases at time t relative to unrestricted rate). Panel (b) shows the same thing but on the log scale.

3 Application

In the previous section, we showed the distribution of gathering sizes, and the tail-behavior more specifically, is an important determinant of the degree to which smaller or larger gatherings contribute to epidemic dynamics. In this section, we use observational data on the size of human gatherings from multiple sources to estimate the empirical power law behavior of gathering size distributions. We use data collected both during "normal" times and during the COVID-19 pandemic as a reference for $f(k)$ an $f_{k_{max}}(k)$ respectively. We estimate the effect of gathering size restrictions during the COVID-19 pandemic using the results from the previous section and empirical estimates of transmission parameters.

Our data on gathering size distributions in the pre-pandemic period are from two primary sources: the BBC Pandemic study [18] and the Copenhagen Networks study of university sources [19]. Both are described in more detail elsewhere. Briefly, the BBC Pandemic study is a citizen science project in which UK citizens self-reported daily contacts using a mobile app in 2018. We extracted the number of contacts made in a day by setting (home, work/school or other) for over 38,000 participants [20]. In the Copenhagen Networks study, the movement and contacts among approximately 1000 university students were intensively tracked and measured via Bluetooth, telecommunication networks, online social media contacts and geolocation over a 5 month period in 2014. In the supplement to the original study the authors report the distribution of 23,231 gatherings observed during the study period. A gathering was defined as groups of individuals in close physical proximity that persists for at least 20 minutes. We extracted the raw data for the probability of observing gatherings of different sizes (Supplementary Figure S9a) using `WebPlotDigitizer`, an online tool that allows the extraction of numerical data from graphs [21].

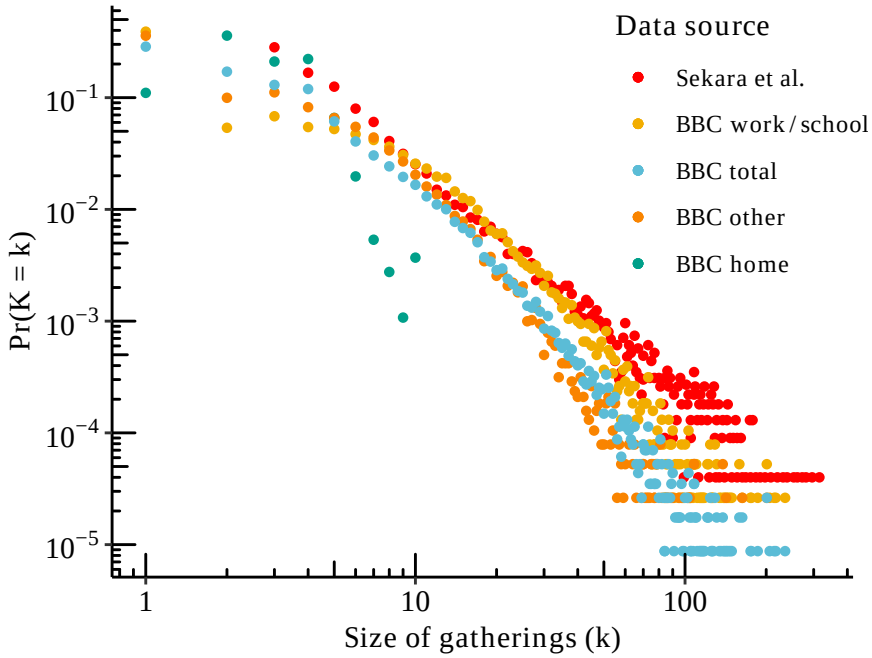
Table 1: Descriptive statistics for empirical distributions of gathering size.

Data source	N	mean	variance	min	max	q90	q99
BBC Pandemic							
Home	38,117	2.9	1.8	1	10	4	7
Work / school	38,131	7.0	117.7	1	235	16	51
Other	38,122	4.8	40.4	1	201	11	28
Total	38,117	4.9	56.1	1	235	11	35
Copenhagen Networks	23,231	9.4	225.9	3	315	15	56

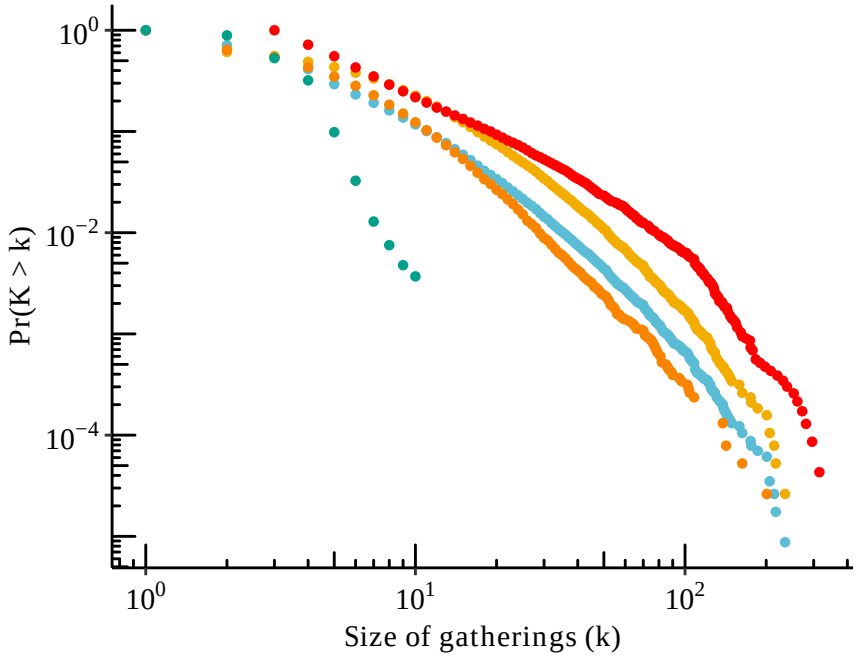
Table 1 provides the descriptive statistics for the empirical distributions of gathering sizes extracted from both sources. In all cases except for household sizes the variance is much greater than the mean which is indicative of overdispersion or a "heavy tail". The maximum gathering sizes outside the household were between 200 and 300. The 90th and 99th empirical quantiles similarly suggest extreme skewness.

Figure 3 plots the full distribution of gathering sizes from both sources using a log-log scale. In all contexts the majority of individuals have very few contacts. For work/school and social gatherings, a very long tail of individuals have very large number of encounters (up to 234 daily contacts at work). We plot both the empirical mass function and the complementary cumulative distribution function (CCDF), also often referred to as Zipf plot, noting that the second is generally preferred for distinguishing power-law type behavior. Typically, a CCDF plot from a power law should be linear on a log-log scale. Here we see that most contexts exhibit approximately linear behavior over significant range; however at the extreme right there may be some nonlinearity which may suggest the presence of an upper bound (for instance the gathering size cannot exceed population size). Interestingly, the distributions of gathering size reported in the Copenhagen Networks study and number of contacts reported in the BBC Pandemic study are similar in range and shape, despite having been measured using completely different methodologies.

Next, we find the best fitting power law for the observed distributions using maximum likelihood. Using the `poweRlaw` [22] package in R, we estimate α as well as k_{min} representing the size beyond which the distribution exhibits power law behavior. For the latter, we use the approach of Clauset



(a) PDF



(b) CCDF

Figure 3: Distribution of gathering sizes from Sekara et al and the BBC Pandemic study by setting.

Notes: Empirical distribution of gathering size from the Sekara et al. study as well as the BBC Pandemic study by setting (home, work/school, other and total). Panel (a) is a log-log plot of the empirical probability that each size is observed. Panel (b) is the empirical complementary cumulative distribution function, i.e. the probability of observing size greater than or equal to k, and is often preferred for understanding tail behavior.

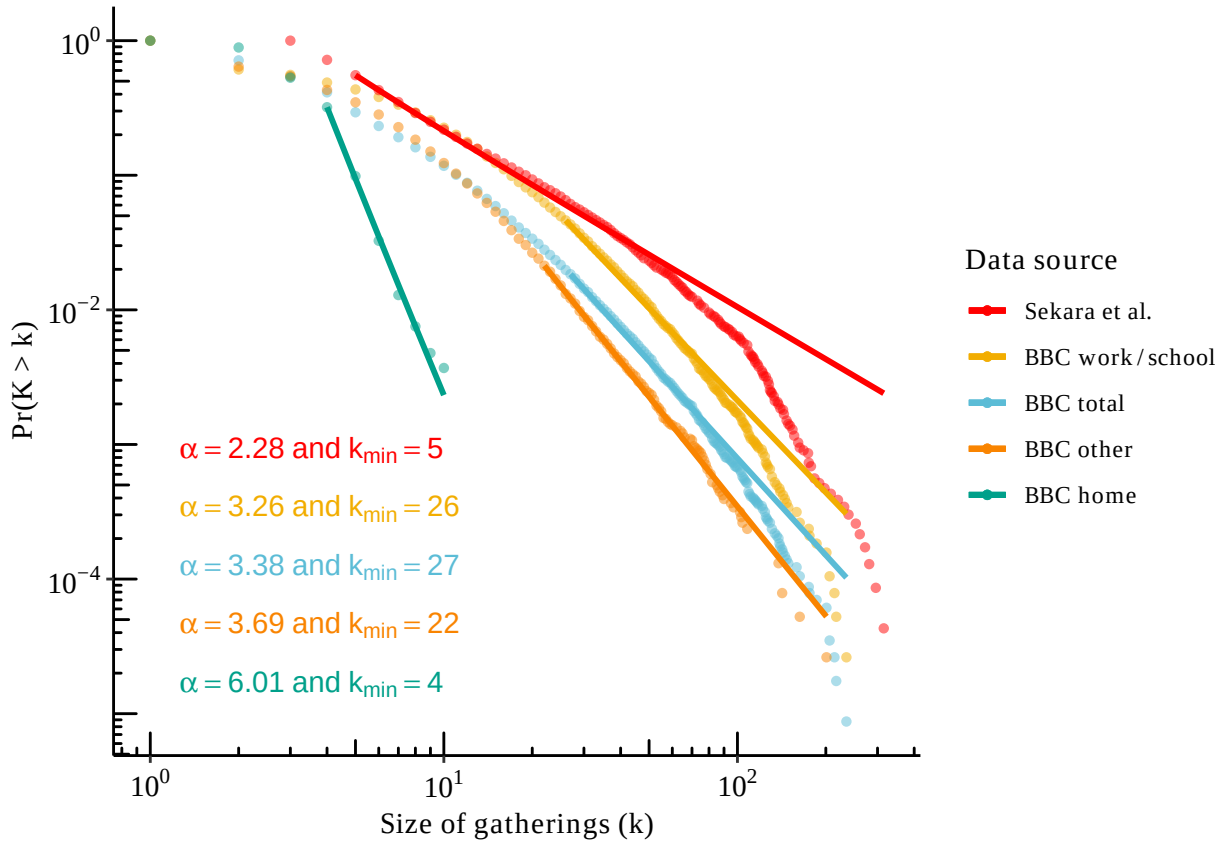


Figure 4: Estimates of power law parameters for Sekara et al. and the BBC Pandemic study by setting

Notes: Plot is complementary cumulative distribution function versus gathering size with lines showing fitted power law distribution. Estimates for α and k_{min} obtained using maximum likelihood for discrete power law using the `poweRlaw` package in R.

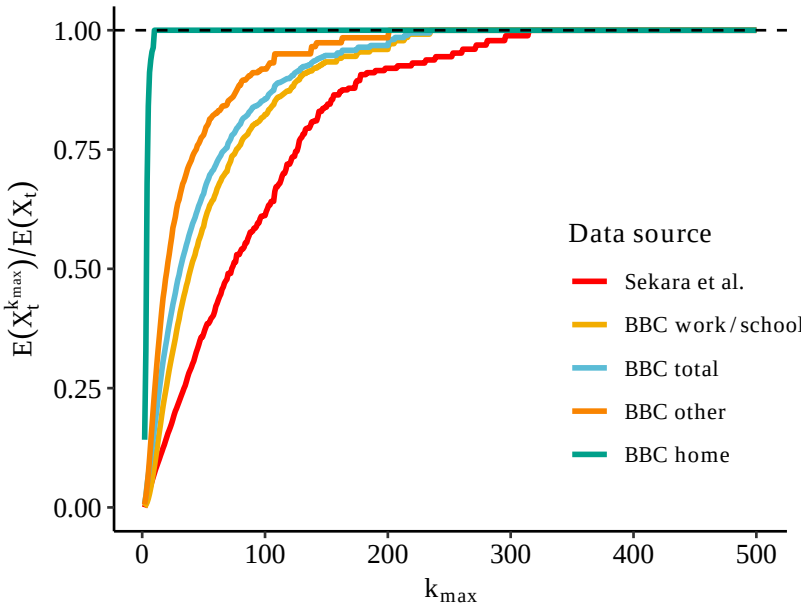
et al. [17] and estimate it by finding the value which minimizes the Kolmogorov-Smirnov statistic. We estimate the standard errors for both using the bootstrap.

Figure 4 shows the resulting power law fits for each of the data sources. The α values estimated range from 2.41 to 6.01, with all settings other than households between 2 and 4. The estimate for the Copenhagen Network study in particular is consistent with infinite second moments (i.e. infinite variance). However, visual inspection suggests that a single power law might not fit well in the extreme tail, with most settings exhibiting considerably lower observed frequencies than suggested by the best-fitting power law. This may be partially due to low cell counts or sampling variability in these extreme quantiles, or as discussed previously may be reflective of the fact that the true distribution is truncated with an upper bound on gathering size.

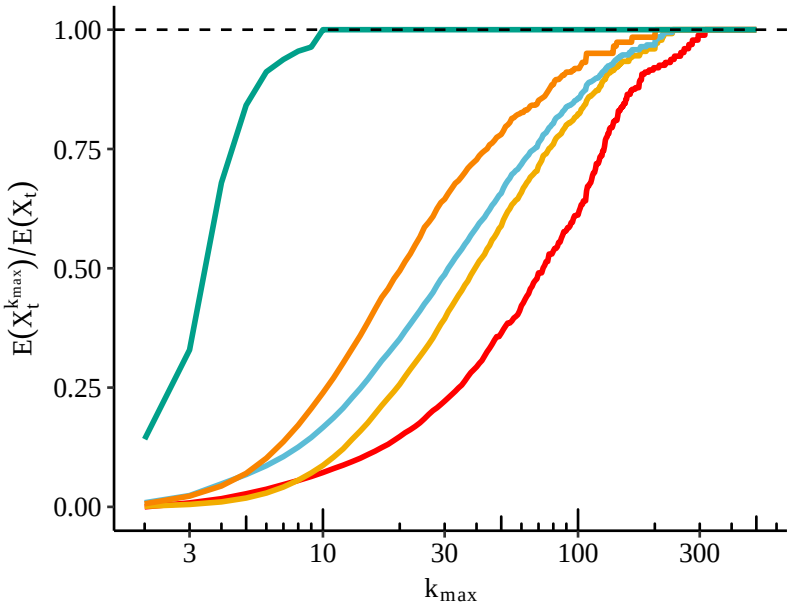
Next, we attempt to estimate the effect of a hypothetical gathering size restriction by replicating the analysis shown in Figure 2 but substituting the empirical gathering size distributions. This would represent an idealized intervention in which everyone followed the restriction by not attending a gathering over the threshold, but their other gathering-seeking behavior is otherwise unaffected. Figure 5 shows the results for the distributions in each of the data sources. Here we see that, to achieve reduction in cases of 50% or more, restrictions must be set below 30 in most settings. Compared with the results in Figure 2, however, we see that the empirical distributions suggest a larger impact of restrictions on medium to large size gatherings, likely because the empirical distributions have slightly less mass in the extreme tail than would be suggested by a true power law.

Finally, while taking the pre-restriction (and pre-outbreak) distribution such as in the analysis above can help one plan for extreme scenarios, it is clear that humans react to restrictions in complex ways that may not mirror the ideal discussed above. Therefore, we also extracted data from the COMIX study [23], which was designed as a deliberate follow on to the BBC Pandemic during the COVID-19 pandemic. In this study, a representative sample of 1,240 adults in the UK were asked about their contact patterns in the first week of the government-imposed 'lockdown' in March 2020. As before, we extracted the number of contacts made in a day by setting (home, work/school, or other). This additional data provides insight into the distribution of contacts under strong social distancing measures.

Figure 6 shows the full distribution of gathering sizes on a log-log scale comparing COMIX to the pre-pandemic "normal" recorded in the BBC Pandemic study. Although sample sizes were considerably lower in COMIX, several interesting patterns emerge. First, the distribution of household contacts under lockdown is almost identical to its pre-pandemic baseline, which is reassuring given household composition is largely unaffected by lockdown. Next, gatherings at work/school and other settings appear "clipped" relative to their pre-pandemic baseline and there now appears to be a preference for lower gathering sizes with a few outliers. This seems consistent with most people complying with order and a few who can't (for instance because their occupation is among those deemed "essential") or who refuse.



(a)



(b)

Figure 5: Relative rate of incident cases under restriction which prohibits gatherings above size k_{max} for different power law distributions; shown on both a linear (a) and a log (b) scale.

Notes: Again we fix transmission parameters to following values $p_i = 0.01$, $p_s = 0.99$, $\tau = 0.08$ but use draws from empirical distributions in Figure 3. Panel (a) shows the relative rate of incident cases calculated using equation 6 and comparing restrictions with k_{max}-level thresholds to unrestricted rate (e.g. a value of 0.5 implies a 50% fewer per capita incident cases at time t relative to unrestricted rate). Panel (b) shows the same thing but on the log scale.

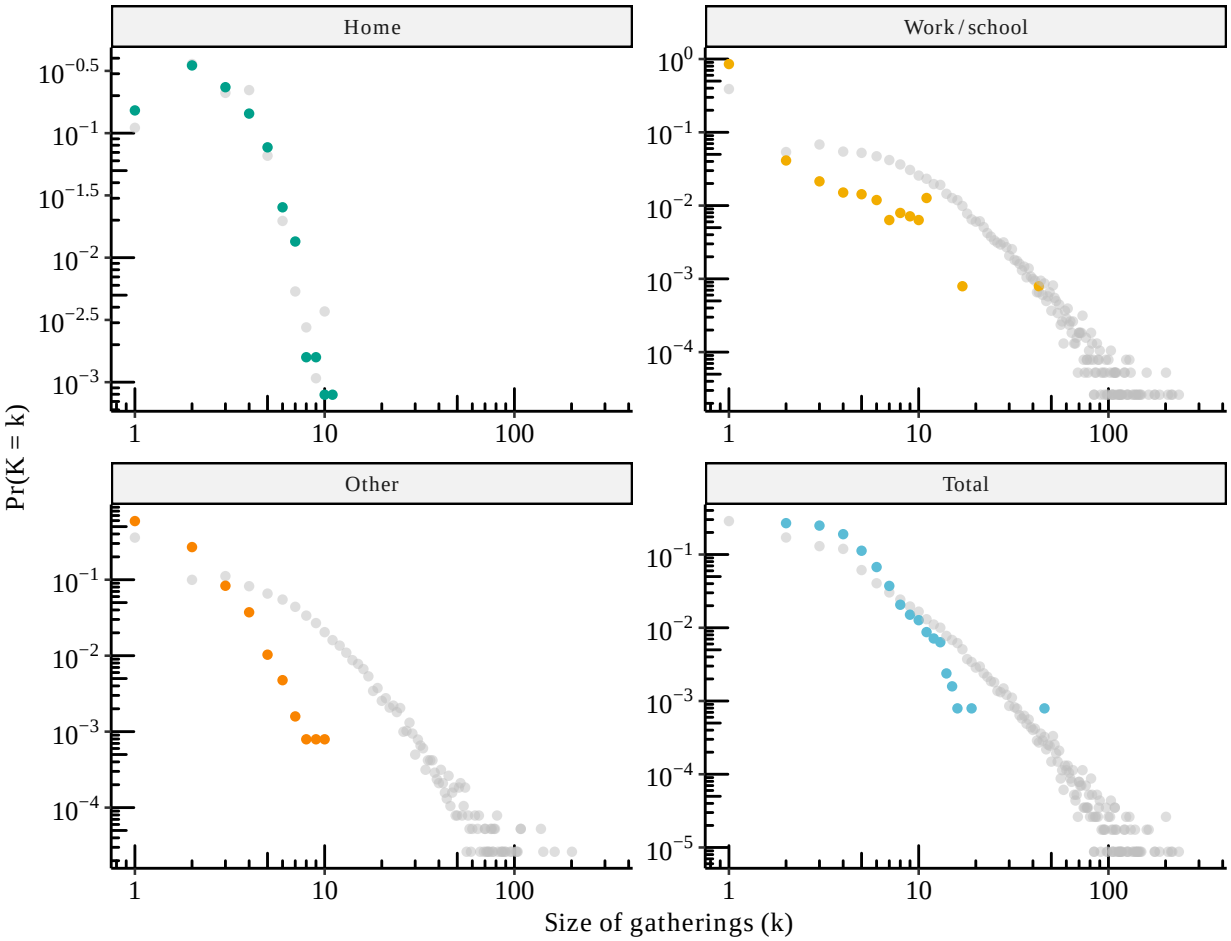


Figure 6: Distribution of gathering sizes before/during UK lockdown from the COMIX study

Notes: Empirical distribution of gathering sizs by setting (home, work/school, other and total) during UK lockdown in March of 2020 are shown in color as measured in the COMIX study. For comparison, the corresponding distributions as measured pre-pandemic in the original BBC Pandemic study are shown in gray for each setting.

4 Discussion

As the COVID-19 pandemic has demonstrated, non-pharmaceutical interventions are an essential tool to limit the spread of infectious diseases, both in the absence of vaccines or effective therapeutics, and when facing surges that test capacity of health systems or the emergence of new variants. We have shown that, when considering limitations on gathering size, decision-makers should consider the distribution of gathering sizes in addition to local conditions when determining the optimal threshold. While a lot of attention has focused on large gatherings, we show that small gatherings, due to their frequency, can be important contributors to transmission dynamics. Using empirical data from previous studies, we find that gathering size distributions are in fact "heavy-tailed" but that meaningful reduction in new cases only occurs once restrictions are set quite low. In theory this conclusion should also apply to future emerging variants of COVID-19 as well as future epidemics other than COVID-19. Our conclusion aligns with that of Brooks-Pollock et al. [24] who have showed that large gatherings of 50+ individuals have relatively small epidemiological impact while small and medium-sized groups of 10 to 50 individuals contribute most to COVID-19 epidemics.

Our work highlights the fact that more detailed data on human gathering sizes dynamics are needed, as datasets on this facet of social dynamics are extremely rare. This should include data on gathering size and duration across contexts and seasons as well as how distributions change during the course of an outbreak. These data would allow for more tailored restrictions and potentially more effective interventions. They would also contribute to better understanding of micro-dynamics of transmission during an outbreak and better parameterization of infectious disease models. Continuously tracked, remotely sensed data from cell phones, with appropriate anonymization and protection of individuals, may be one avenue for collecting this information on a large scale.

Our model relies on multiple simplifying assumptions. Recognizing that violations are not equal and from the point of view of the policymaker the cautious approach is often the most prudent, where possible we have made effort to make conservative assumptions. First, by using a single probability we ignore many important heterogeneities in transmission risk (e.g indoor vs outdoor, use of face coverings, duration, ventilation, etc). However, this would only substantively affect our conclusions if heterogeneity varies with gathering size. For instance, if larger gatherings are

more likely to be outdoor and people perceive them to be more dangerous and therefore adhere more strictly to masking and social distancing guidance then it's possible that the per contact transmission risk may decrease with size of gathering, making restrictions on large gatherings even less effective relative to smaller ones. In this case, it would still be possible to apply the model presented, but specifying the transmission risk for each gathering size, which in practice may be hard to empirically validate.

Second, we assume that the probability of transmission is constant across gathering sizes, which may not be reasonable for very large gatherings (except perhaps in the case of an airborne pathogen in an unventilated and crowded indoor space). Here our model clearly represents a worst case scenario where all individuals have contacts with all other attendees. It thus likely overestimates the contribution of large gatherings to the overall number of new infections.

Third, we assume that susceptible, infected and recovered individuals are exchangeable, mix randomly, exhibit the same behaviour and attend gatherings at the same proportion as their proportion in the underlying population distribution. This may not be the case if, for instance, infectious people self-isolate upon developing symptoms or if there exists significant subsets of susceptibles who avoid gatherings and significant recovereds who believe they are immune and therefore go to gatherings at rates above their population fraction. Again these heterogeneities in behavior will mostly affect our conclusions if they vary with size of the gathering. In particular, they may lead to substantially different conclusions if behavioral dynamics tend to favor transmission at larger gatherings, such as if a core

Similarly, in estimating effect of a certain threshold, we assume that individuals respond to gathering size restrictions uniformly, with perfect compliance and that they do not adapt their social behaviours independently of the regulation, based on, for instance, their knowledge of local epidemic dynamics. This is obviously not true in practice, but most plausible deviations would tend to make our estimates an upper bound on the effect of restrictions above a certain size. However, if announcing any restriction is a sufficient signal that many opt to avoid any gatherings at all, that may lead to a large reduction in cases even at a relatively large threshold. This may be more likely at the start of an outbreak when people are still attempting to ascertain the seriousness of the risk.

Lastly, we assume all new infections to be equivalent, not considering heterogeneity in the impact of secondary infections. This assumption again may not be reasonable at the beginning of

an epidemic when local transmission is not established and we might expect infections at larger gatherings to seed downstream cases in more diverse parts of the population/community. In this case although larger gatherings are less frequent they act as central nodes in the contact graph through which infection reaches sub-communities.

Our work is also subject to several limitations due to the data sources that we used. The three data sources (BBC Pandemic, COMIX and Sekara et al.) had different aims, study designs and limitations. We assume they all provide good estimates of frequency of gathering sizes. In using the BBC Pandemic and COMIX study we approximate the size of gatherings by assuming that all daily contacts in a given context all took place in one gathering. The COMIX study was conducted during the first week of lockdown in March 2020 in the UK and may not be representative of restrictions in other locations or times. Sekara et al. studied university students in Copenhagen, a specific population that may not be representative of other populations. In the case of the BBC Pandemic and COMIX data, a particular threat to our main conclusions might be measurement error that correlates with gathering size, for instance if people get worse at recalling or recording the size of larger gatherings we may underestimate their frequency and therefore their contribution to transmission dynamics. This is a major advantage of the Sekara data which were remotely recorded by cellphone and gps devices.

References

1. Hale, T. *et al.* A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat Hum Behav* **5.** Bandiera_abtest: a Cg_type: Nature Research Journals Number: 4 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Infectious diseases;Politics and international relations;Social policy Subject_term_id: infectious-diseases;politics-and-international-relations;social-policy, 529–538 (Apr. 2021).

2. Haug, N. *et al.* Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat Hum Behav* **4.** Bandiera_abtest: a Cg_type: Nature Research Journals Number: 12 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Epidemiology;Viral infection Subject_term_id: epidemiology;viral-infection, 1303–1312 (Dec. 2020).

3. Sharma, M. *et al.* Understanding the effectiveness of government interventions against the resurgence of COVID-19 in Europe. *Nat Commun* **12.** Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Epidemiology Subject_term_id: epidemiology, 5820 (Oct. 5, 2021).

4. Brauner, J. M. *et al.* Inferring the effectiveness of government interventions against COVID-19. *Science* **371.** Publisher: American Association for the Advancement of Science, eabd9338 (Feb. 19, 2021).

5. Li, Y. *et al.* The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *The Lancet Infectious Diseases* **21,** 193–202 (Feb. 1, 2021).

6. Haber, N. A., Clarke-Deelder, E., Salomon, J. A., Feller, A. & Stuart, E. A. Impact Evaluation of Coronavirus Disease 2019 Policy: A Guide to Common Design Issues. *American Journal of Epidemiology* **190,** 2474–2486 (Nov. 2, 2021).

7. Fisher, K. A. *et al.* Community and Close Contact Exposures Associated with COVID-19 Among Symptomatic Adults 18 Years in 11 Outpatient Health Care Facilities — United States, July 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69,** 1258–1264 (Sept. 11, 2020).

8. *COVID-19: Note by the Chief Medical Officer, Chief Nursing Officer and National Clinical Director* (Oct. 7, 2020).

9. Center for Disease Control and Prevention (CDC). *Events and Gatherings: Readiness and Planning Tool* (Stockholm, July 6, 2020).

10. European Center for Disease Control (ECDC). *Guidelines for non-pharmaceutical interventions to reduce the impact of COVID-19 in the EU/EEA and the UK* (Sept. 24, 2020).

11. Roser, M., Ritchie, H., Ortiz-Ospina, E. & Hasell, J. *Coronavirus Pandemic (COVID-19)* (Our World in Data, 2020).

12. Barabási, A.-L. & Albert, R. Emergence of Scaling in Random Networks. *Science* **286.** Publisher: American Association for the Advancement of Science, 509–512 (1999).

13. May, R. M. & Lloyd, A. L. Infection dynamics on scale-free networks. *PHYSICAL REVIEW E,* 4.

14. Pastor-Satorras, R. & Vespignani, A. Epidemic Spreading in Scale-Free Networks. *Phys. Rev. Lett.* **86.** Publisher: American Physical Society, 3200–3203 (Apr. 2, 2001).

15. Bansal, S., Grenfell, B. T. & Meyers, L. A. When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of The Royal Society Interface* **4.** Publisher: Royal Society, 879–891 (Oct. 22, 2007).

16. Kelker, D. A Random Walk Epidemic Simulation. *Journal of the American Statistical Association* **68.** Publisher: [American Statistical Association, Taylor & Francis, Ltd.], 821–823 (1973).

17. Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-Law Distributions in Empirical Data. *SIAM Review* **51.** Publisher: Society for Industrial and Applied Mathematics, 661–703 (2009).

18. Klepac, P., Kissler, S. & Gog, J. Contagion! The BBC Four Pandemic – The model behind the documentary. *Epidemics* **24,** 49–59 (Sept. 1, 2018).

19. Sekara, V., Stopczynski, A. & Lehmann, S. Fundamental structures of dynamic social networks. *Proc Natl Acad Sci USA* **113,** 9977–9982 (Sept. 6, 2016).

20. Kucharski, A. J. *et al.* Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet Infectious Diseases* **20,** 1151–1160 (Oct. 1, 2020).

21. Rohatgi, A. WebPlotDigitizer, Version 4.2. *URL https://automeris. io/WebPlotDigitizer* (2020).

22. Gillespie, C. S. Fitting heavy tailed distributions: the poweRlaw package. *arXiv:1407.3492 [physics, stat]* (July 13, 2014).

23. CMMID COVID-19 working group *et al.* Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med* **18,** 124 (Dec. 2020).

24. Brooks-Pollock, E. *et al.* The population attributable fraction of cases due to gatherings and groups with relevance to COVID-19 mitigation strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376.** Publisher: Royal Society, 20200273 (July 19, 2021).

A Appendix

A.1 Derivation of relationship between expected cases and gathering size

Assuming fixed transmission probability τ for contacts between susceptibles and infectious individuals, the number of secondary cases generated by a single infectious individual with $K_s = k_s$ susceptible contacts is binomially distributed, i.e.

$$X_{it} \mid K_s = k_s \sim \text{Binomial}(k_s, \tau). \tag{7}$$

If susceptibles, infectious, and recovered individuals attend gatherings at rates equivalent to their population proportions then attendance at a gathering of size $K = k$ can be represented by a multinomial sampling model of the form

$$(K_s, K_i, K_r)' \mid K = k \sim \text{Multinomial}(k, (p_s, p_i, p_r)') \tag{8}$$

where $p_s = \frac{S(t)}{N}$, $p_i = \frac{I(t)}{N}$, and $p_r = \frac{R(t)}{N}$. Under the model, the expected number of susceptibles is kp_s, the expected number of infectious is kp_i, and the expected number of recovereds is kp_r. To calculate the expected number of secondary cases, note that, on average, only the K_s susceptibles are at risk of infection and they are exposed to K_i infectious individuals. Then the probability that the K_s susceptibles "escape", i.e. that they are not infected by any of the K_i infectious individuals, is $(1-\tau)^{K_i}$ and thus, by extension, the probability that they are infected by at least one of the K_i infectious individuals in attendance is $1-(1-\tau)^{K_i}$. Therefore, the total number of secondary cases at a gathering of size $K = k$ is

$$\begin{aligned} (K_s, K_i, K_r)' \mid K = k &\sim \text{Multinomial}(k, (p_s, p_i, p_r)') \\ X_t \mid K_s = k_s, K_i = k_i &\sim \text{Binomial}(k_s, 1-(1-\tau)^{k_i}) \end{aligned} \tag{9}$$

and taking iterated expectations, the expected number of secondary cases given a gathering of size $K = k$ is simply

$$E(X_t \mid K = k) = kp_s(1-(1-\tau)^{kp_i}). \tag{10}$$

For a more intuitive way to think about this equation, notice that kp_s is the expected number

of susceptibles and kp_i is the expected number of infectious individuals; the expression $kp_s(1-(1-\tau)^{kp_i})$ is then just the expected number of susceptibles times the probability of being infected by any of the infectious individuals who attend, where the latter is equivalent to the one minus the "escape" probability, i.e. the probability that no susceptible is infected by any of the infectious individuals expected to attend.

A.2 Binomial Approximation

More specifically, when $|\tau kp_i| \ll 1$ a Binomial approximation gives

$$(1-\tau)^{kp_i} \approx 1 - kp_i\tau$$

and thus

$$kp_s(1-(1-\tau)^{kp_i}) \approx kp_s(1-(1-kp_i\tau)) \approx k^2 p_s p_i \tau$$