

EPIDEMICS WITH TWO LEVELS OF MIXING

BY FRANK BALL, DENIS MOLLISON AND GIANPAOLO SCALIA-TOMBA

*University of Nottingham, Heriot-Watt University
and Universita La Sapienza*

We consider epidemics with removal (SIR epidemics) in populations that mix at two levels: *global* and *local*. We develop a general modelling framework for such processes, which allows us to analyze the conditions under which a large outbreak is possible, the size of such outbreaks when they can occur and the implications for vaccination strategies, in each case comparing our results with the simpler homogeneous mixing case.

More precisely, we consider models in which each infectious individual i has a global probability p_G for infecting each other individual in the population and a local probability p_L , typically much larger, of infecting each other individual among a set of neighbors $\mathcal{N}(i)$. Our main concern is the case where the population is partitioned into local groups or households, but our approach also applies to cases where neighborhoods do not form a partition, for instance, to spatial models with a mixture of local (e.g., nearest-neighbor) and global contacts.

We use a variety of theoretical approaches: a random graph framework for the initial exposition of the simple case where an individual's contacts are independent; branching process approximations for the general threshold result; and an embedding representation for rigorous results on the final size of outbreaks.

From the applied viewpoint the key result is that, compared with the homogeneous mixing model in which individuals make contacts simply with probability p_G , the local infectious contacts have an “amplification” effect. The basic reproductive ratio of the epidemic is increased from its individual-to-individual value R_G in the absence of local infections to a group-to-group value $R_* = \mu R_G$, where μ is the mean size of an outbreak, started by a randomly chosen individual, in which only local infections count. Where the groups are large and the within-group epidemics are above threshold, this amplification can permit an outbreak in the whole population at very low levels of p_G , for instance, for $p_G = O(1/Nn)$ in a population of N divided into groups of size n .

The implication of these results for control strategies is that vaccination should be directed preferentially toward reducing μ ; we discuss the conditions under which the *equalizing strategy*, aimed at leaving unvaccinated sets of neighbors of equal sizes, is optimal. We also discuss the estimation of our threshold parameter R_* from data on epidemics among households.

1. Introduction.

1.1. Mixing at two levels. In the spread of infectious disease, heterogeneities in population behavior often play a key role in determining whether a major epidemic outbreak occurs and, if it does, its rate of spread and the

Received September 1995; revised July 1996.

AMS 1991 subject classifications. Primary 92D30, 60K35; secondary 05C80, 60J80.

Key words and phrases. Epidemic, household, vaccination, estimation, final outcome, random graphs, Reed–Frost, branching processes, Gontcharoff polynomials, threshold parameters, asymptotic distribution, embedding, R_0 .

final size of the epidemic. Here we shall analyze in depth one of the simplest and most basic kinds of heterogeneity, in which each infectious individual i has not only a global probability p_G for infecting each other individual in the population, but also a local probability p_L , typically much larger, of infecting each other individual among a set of neighbors $\mathcal{N}(i)$. We shall call these *global* and *local* infections, respectively.

This kind of model is of application to a wide variety of epidemic situations, especially in infectious diseases of humans, where local mixing in social groups such as households and schools can play a crucial role in facilitating the spread of infection. Such models are also of considerable interest in ecology: see, for instance, the reviews by Kareiva (1990) of “patch dynamics,” especially the references to “island” and “meta-population” models, and by Hanski and Gilpin (1991).

Perhaps the simplest such model is that of a population partitioned into equal sized groups. That is to say, we have m groups, each of n individuals, giving a total population of size $N = mn$; two individuals are neighbors if and only if they belong to the same group. We shall analyze two cases, that of *households* where n takes a fixed, typically fairly small, value, and the case of *large groups*, where we consider what happens as $n \rightarrow \infty$. In either case, we can generalize the model to allow for unequal group sizes $\{n_i: 1 \leq i \leq m\}$, with $\sum_{i=1}^m n_i = N$; this level of generality is of course vital in applications (see Section 5).

Our basic model is a generalization of the standard SIR epidemic: we consider individuals of just one type, who once infected make contacts in Poisson processes during an infectious period T_I , where the T_I 's for different individuals are independent and identically distributed (see Section 3.1). There is no difficulty of principle, however, to extending the same treatment to models with different within-group contact structures, for instance, the models of Gertsbakh (1977) (see Section 5.2.2) or de Koeijer, Diekmann and Reijnders (1995), or with several types of individual with different contact probabilities, which might for instance represent children and adults (see Section 3.6).

Our basic model assumes that global contacts are equally likely to be with any other individual. Because of this, it is natural to describe the variability of group size through its *size-biased* distribution $\{\pi_k\}$, with π_k being defined as the probability that a randomly chosen *individual* lives in a group of size k (see Section 2.1). Much of our analysis would apply equally well if global contacts were chosen on some other basis, for instance, according to the ordinary group size distribution $\{h_k\}$, where h_k is defined as the probability that a randomly chosen *group* is of size k ; notable exceptions to this are our results on optimal vaccination strategies, which do depend on how global contacts are chosen (see Section 5.2.1).

Returning to simple models, another basic case is where individuals are arranged in space (e.g., equally spaced around a circle) and the neighbors of an individual are defined as those within a certain distance, in the simplest case as just an individual's nearest neighbors. Note that in this case the sets of neighbors will overlap, rather than partitioning the population. From the

theoretical point of view, this model can be regarded as a kind of limit of a dispersal model with local and long distance interactions, in which the distribution of the latter degenerates into the uniform distribution [for dispersal distributions with extreme behavior, such as “great leaps forward,” or with infinite velocity, see Mollison (1972) and Mollison and Levin 1995)]. Possible applications include the spread of infection between pigs in a line of stalls (M. de Jong, personal communication) and the recent North Sea seal epidemic [Bolker et al. (1995)]. Note that the latter requires a model with three levels—long distance and local contacts between seal colonies, perhaps following the great circle model (see Section 2.3), plus within-colony contacts, perhaps following the model of de Koeijer, Diekmann and Reijnders (1995)—but there is no difficulty in principle in extending our analysis to deal with mixing at three levels.

1.2. Contents. We shall restrict attention to the SIR model, that is, where there are just three possible states for an individual, *susceptible* (S), *infected* and *infectious* (I) and *removed* (R), and the only possible transitions are $S \rightarrow I$ and $I \rightarrow R$. We shall also assume that the sets of contacts made by different individuals are independent of each other. In general, infections made by the same individual will be correlated, if only because of the dependence induced by the variability of the length of the infectious period.

In Section 2 we describe how a random graph framework can be used to analyze the non-time-dependent aspects of epidemics. This is especially helpful in the “independent links” (or Reed–Frost) case where infections made by the same individual are independent, since we can then use an undirected graph. Further, the technique of first considering only local infections allows us to partition the population, whether or not it was in separate groups initially, so that global contacts can be considered using a “clumped Reed–Frost” model: essentially simple homogeneous mixing, but with individuals replaced by the clumps formed by local contacts.

We are thus able to derive explicit expressions for thresholds and the final size of epidemics. Recall that in an epidemic with a single level of mixing, the condition for a large outbreak to be possible is that the reproductive ratio R_0 should be greater than 1, where R_0 can be loosely defined as the expected number of potentially infectious contacts of a single infectious individual [see, e.g., Diekmann, Heesterbeek and Metz (1990) and Dietz (1993)].

For epidemics with two levels of mixing, we find that the condition for a large outbreak to be possible is given by $R_* > 1$, where $R_* = R_G \mu$, the product of the reproductive ratio R_G for global contacts and the mean size μ of outbreaks utilizing only local contacts.

The threshold parameter R_* thus provides a natural generalization of R_0 to two levels of mixing, but we must emphasize that R_* is a group-to-group—or more precisely clump-to-clump—reproductive ratio: it is the expected number of clumps contacted by all individuals in the clump of a random individual. While it reduces to R_0 when only local contacts matter (and the clump size is therefore equal to 1), it can differ substantially when group sizes are large (see example of Section 2.5).

The case of large subgroups (Section 2.4) is of particular interest, because it is then meaningful to ask whether individual subgroups are above their local threshold. If so, then the “amplification factor” μ is $O(n)$, so that a global epidemic can occur when p_G is only $O(1/(Nn))$.

An important practical implication of this threshold result is the difference it makes for a vaccination strategy according to whether it succeeds in reducing individual groups below their local threshold. This is explored informally in Section 2.5 in an example showing how dramatic the difference can be; more generally, optimal vaccination strategies are considered later in Section 5.2.

In Section 3 we consider a more general model for the spread of an epidemic among a population consisting of m groups or households. We start with the case where the households are of equal size n . Infectious individuals make local contacts at rate $n\lambda_L$ and global contacts at rate λ_G during an infectious period T_I that follows any arbitrary but specified distribution. (The independent links case of Section 2 corresponds to the special case when T_I is constant.) The individual contacted by a local (global) contact is chosen uniformly at random from the n ($N = mn$) individuals in the infective’s group (whole population). Thus the individual-to-individual local and global infection rates are λ_L and λ_G/N , respectively.

In order to derive explicit expressions relating to the threshold behavior of our model, a number of properties of single population SIR stochastic epidemics are required. For convenience, these are collected together in Section 3.2. In Section 3.3.1 we show that the early stages of our epidemic can be approximated by a branching process, whose individuals are single group epidemic processes. Moreover, this approximation can be made precise by considering a sequence of epidemics in which the number of groups $m \rightarrow \infty$. This enables us to determine a threshold parameter R_* for our epidemic, such that, in the limit as $m \rightarrow \infty$, global epidemics occur with nonzero probability if and only if $R_* > 1$. Here, a global epidemic is one which affects infinitely many groups as $m \rightarrow \infty$. We also determine the probability that a global epidemic occurs and various properties of nonglobal epidemics. In Section 3.3.2 we discuss the threshold parameter R_* . In particular, we compare it with the classical basic reproductive ratio R_G that applies when only global contacts are considered, and we show that our model displays a similar amplification effect to that described in Section 2.3 for the independent links case.

In Section 3.4 we use a heuristic argument to determine the distribution of the total size within a typical group in the event of a global epidemic occurring. In Section 3.5, we extend our results to the situation in which the group sizes are not all equal. In Section 3.6 we derive a threshold parameter for the proliferation of infectious individuals for our model and discuss the relationship of our methodology and results to recent papers of Becker and Dietz (1995) and Becker and Hall (1996).

In Section 4 we show that the embedding approach of Scalia-Tomba (1985, 1990) can be extended to epidemics with two levels of mixing and we make use of this to provide a formal derivation of the asymptotic distribution for the final size of a global epidemic.

In Section 5 we consider applications of our results. In Section 5.1 we describe a method of estimating the threshold parameter R_* from household final size data, using data on the spread of an influenza epidemic as an illustration.

Finally, in Section 5.2 we consider vaccination programs, whose aim must be to reduce R_* to below unity, and for a variety of models we examine conditions under which the *equalizing strategy*, which leaves the numbers of susceptibles in each group as nearly equal as possible, is optimal.

1.3. Related work. Papers on epidemics with two levels of mixing go back at least as far as Rushton and Mautner (1955) on deterministic simple epidemics and Bartlett (1957) and Daley (1967) on the stochastic side, but the earliest works relevant to the present approach are two independent papers dating from 1972. Watson (1972) considered a model with a fixed number of large groups, giving deterministic results and some heuristic stochastic approximations supported by simulations. [A more detailed analysis of the threshold behavior of Watson's deterministic model is given in Daley and Gani (1994).]

A more abstract paper, but perhaps representing a greater advance, was that of Bartoszyński (1972), who considered a group epidemic model which corresponds to the limiting branching process that we use in Section 3.1 and derived a threshold condition for it which is essentially the same as our $R_G\mu > 1$. However, his model was described in rather general terms and hence his results are not so explicit as ours. Indeed, he does not in general say how the probability of choosing a contact in a group of size k is to be specified, leaving it open that it might be either $\{\pi_k\}$ or $\{h_k\}$ (in the notation of Section 1.2), though in one specific example he makes it clear that he means the latter.

May and Anderson (1984), who acknowledged earlier work by Hethcote (1978) and Post, DeAngelis and Travis (1983), who considered the vaccination problem for a deterministic epidemic among a finite number of (large) groups, showing that the equalizing strategy is optimal (see Section 5.2.3).

Work on outbreaks within households in the presence of community infection—but without considering the dynamics of the latter—has a long history: see, for example, the discussion in Longini and Koopman (1982) and, for more recent work, Becker (1989) and Addy, Longini and Haber (1991).

The present work is the elaboration of a “back-of-envelope” answer to a question raised by Klaus Dietz at the Newton Institute in 1993, as to whether whole-population models could be found that would justify such applied work on household models and provide a framework for their development and extension. We are grateful to Klaus Dietz and Niels Becker for exchanging preprints as our work has proceeded in parallel: we comment on their papers [Becker and Dietz (1995) and Becker and Hall (1996)] in Section 3.6.

2. Random graphs and independent contacts.

2.1. Introduction: the random graph framework. We will often not be interested in the time course of the epidemic, but only in which individuals

become infected (indeed, we may only be interested in their total number, the *final size* of the epidemic). In that case we can make good use of the representation of the spread of the epidemic by a directed graph, in which the individuals are represented by the nodes of the graph, and we draw an arrow from one individual to another to indicate that the first, if infected, will make an infectious contact with the second [see, e.g., Barbour and Mollison (1989)].

As already mentioned (Section 1.2), in general infections made by the same individual will be correlated, because of the dependence induced by the variability of the length of the infectious period. We shall later prove results on thresholds and final outbreak size for such more general models. However, the special case where they are independent—the generalization to two levels of mixing of the well-known Reed–Frost model—is well worth considering first, as then we can use an *undirected graph*, with links rather than arrows [see Barbour and Mollison (1989); also von Bahr and Martin-Löf (1980), and Ball (1983a)], and analysis is much clearer and simpler. The key idea here is that if an individual's contacts are independent of each other and if the probability that i infects j is the same as the probability that j infects i , then we can represent both the latter events by the *same*, undirected link in the contact graph.

An undirected graph can be partitioned into connected components, and the set of those infected during the epidemic will consist precisely of the connected component(s) to which those initially infected belong. In the simple case of homogeneous mixing with contact probability p —the basic Reed–Frost model—the corresponding graph is the simple random graph on a set of N nodes (i.e., a population of size N) $G(N, p)$ [Barbour and Mollison (1989)]. For large N , this graph has a single “giant” component if and only if $R_0 > 1$, where $R_0 = Np$; it then contains a proportion z of the population, where z is the largest root of $z = 1 - \exp(-R_0 z)$ [Bollobas (1985)]. Thus, if there is a large outbreak it will affect approximately this proportion z of the population. Further, if the initial number of infected $I(0) = 1$, the probability of a large outbreak ζ is simply the probability that the initially infected individual lies in the giant component and is therefore equal to z . Note that the equality of ζ to z is quite special to this case. In general (see Section 3.3.1) the probability $\zeta(I(0))$ of a large outbreak depends sensitively both on $I(0)$ [$\zeta(I(0)) = 1 - (1 - \zeta)^{I(0)}$] and on the assumption that the contacts of an individual are independent.

In the remainder of this section, we extend the use of undirected random graphs to find the threshold conditions and final size for epidemics in a large population with two levels of mixing, with particular emphasis on the case of large local groups (Section 2.4) and on the implications for vaccination strategies (Section 2.5).

2.2. Local contacts and the clumped Reed–Frost model. We first describe in detail two models in which each individual has a small number of local contacts, and show how they can both be considered as special cases of a clumped Reed–Frost model.

The first of these is the households model described above, where we fix the size or size distribution of households. Then, if we consider only local contacts, these partition each household into a number of connected components. Once we have done this, these connected components summarize the local interactions—whether two separate components originate from the same household or not is irrelevant when we complete our model by adding the global contacts, since the probability of such contacts is to be the same, independently, for each pair of individuals.

Note that this deconstruction relies on our model allowing individuals in the same group to have both a local probability p_L and (independently) a global probability p_G of contacting each other. Our model is essentially identical to one in which we allow them only a local probability $p'_L = 1 - (1 - p_L)(1 - p_G)$. (We of course require $p_G \leq p'_L$ here, which is no problem as our main interest is in the case $p_G \ll p_L$, when $p'_L \approx p_L$.)

The probability, π_k say, that an individual chosen at random from the whole population belongs to a component of size k can be calculated—in principle at least—from the distribution of household size and standard methods for the Reed–Frost model (see Section 3.2). The component sizes will not be exactly independent of each other, because of dependence of sizes within households, but this effect will be negligible provided the number of households is large.

Our second model, the “great circle,” is one where the population is not partitioned into households. Instead, we have individuals located in one-dimensional space. For simplicity we shall just consider the case where each individual has two neighbors—one on each side; to avoid boundary problems it is convenient to take the space to be the circumference of a circle. We allow infectious local links, of probability p_L , between each pair of neighbors, and global links, as usual, of equal probability p_G for each pair in the population. When we consider first the local contacts, these again partition the whole population into connected components. In this case the probability π_k of belonging to a component of size k is given by the double geometric distribution of parameter p_L [$\pi_k = kp_L^{k-1}(1 - p_L)^2$, $k = 1, 2, \dots$]. Again the component sizes are not exactly independent, but we can neglect their dependence in what follows provided that the population is large relative to the mean component size.

We note that this “neighbors plus global links” model could be generalized to other isotropic spatial structures, such as a regular toroidal lattice or a tessellation of a Poisson process on a sphere, and that we could allow further than nearest-neighbor contacts, provided that the local contacts give only relatively small connected components.

2.3. Threshold and final size for the clumped Reed–Frost model. Both the models introduced above are special cases of the following, which we shall call the clumped Reed–Frost model. In this, the population consists of clumps, the i th clump having weight w_i . We then run a Reed–Frost type epidemic (that is, with independent and symmetric contacts) in this population, with probability $1 - \exp(-cw_iw_j)$ for a contact between clumps i and j . We relate

this to our original models by taking a locally connected component containing k individuals to be a clump of weight k and by taking $p_G = 1 - \exp(-c)$.

We shall use π_k to denote the probability that a randomly chosen individual belongs to a clump of size k and use μ to denote the mean clump size $\sum_k k\pi_k$. We shall assume that the mean clump size is finite. For both the household and nearest-neighbor models, in the limit of large total population the clump weights will be chosen independently from the distribution $\{\pi_k\}$. Note that $\{\pi_k\}$ is the *size-biased* distribution for clump size, as distinct from the formulation in which we define the probability h_k that a randomly chosen clump is of size k . The distributions $\{\pi_k\}$ and $\{h_k\}$ are simply related, with $\pi_k = kh_k / \sum_j j h_j$. Note also that if μ_h and σ_h^2 are, respectively, the mean and variance of the distribution $\{h_k\}$, then $\mu = \mu_h + \sigma_h^2 / \mu_h$.

Assuming that the number of clumps is large, the probability of a large outbreak can be found by considering the branching process which approximates its early stages (see Section 3.3.1), in which individuals correspond to clumps in the epidemic process and the offspring of a given clump are the clumps that it directly tries to infect in the clumped Reed–Frost epidemic. The approximation (which can be made fully rigorous; see Section 3.3.1) assumes that each new clump contacted in the epidemic process is still susceptible. For large N , the number of clumps contacted by an individual in the epidemic process is Poisson($N p_G$), with probability generating function (p.g.f.) $\exp(N p_G(s - 1))$, and the clump size distribution is $\{\pi_k\}$, with p.g.f. $G_\pi(s)$ say. Thus the number of clumps contacted by a given clump, that is, the offspring distribution for the approximating branching process, has p.g.f. $G_\pi(\exp(N p_G(s - 1)))$. It follows that the probability of a large outbreak is the largest solution $\zeta (\leq 1)$ of $1 - \zeta = G_\pi(\exp(-N p_G \zeta))$, that is, of $1 - \zeta = \sum_{k=1}^{\infty} \pi_k \exp(-k N p_G \zeta)$. Further, ζ will be greater than 0 if and only if the mean number of offspring from a clump $N p_G \mu$ is greater than 1.

Thus the basic reproductive ratio for the epidemic among clumps is $R_* = \mu N p_G = \mu R_G$, where $R_G = N p_G$ is the basic reproductive ratio for the ordinary Reed–Frost epidemic—that is, where we only have global contacts so that all clumps are of size 1. [More strictly, we should perhaps use $R_G = (N - 1)p_G$, but if we are interested in values of N sufficiently small that this matters, we should be worrying about the correct definition of thresholds in finite populations; see Nåsell (1995).]

The probability ζ here is that of a large outbreak started by a random individual. If we know that the initial infection(s) is (are) in a clump of size k , consideration of the first generation of contacts shows that the probability of a large outbreak ζ_k is related to ζ by $\zeta_k = 1 - \exp(-k N p_G \zeta)$ (see Section 3.3.1 for more detail).

In either case, just as for the simple Reed–Frost model, we can argue that the probability of a large outbreak is the same as the probability that an individual belongs to the giant connected component of the contact graph, which in turn is the same as the (proportional) final size of the epidemic conditional on a large outbreak. Thus the final size $\approx \zeta N$, and the probability that an individual in a clump of size k (or equivalently the whole of that clump)

is infected during the epidemic is ζ_k . We may check the consistency of these results: $\zeta = \sum_k \pi_k \zeta_k = \sum \pi_k - \sum \pi_k \exp(-\zeta k N p_G) = 1 - G_\pi(\exp(-\zeta N p_G))$. Also, thinking of $1 - \zeta_k$ as the probability that a clump of size k escapes infection, we note that the number of links from each individual in the clump to the giant component is Poisson($\zeta N p_G$), so that the probability of having no such links should indeed be $\exp(-k N p_G \zeta)$.

Note that for the great circle model with independent links there are only two parameters, p_L and p_G , and it is easy to show that $G_\pi(s) = (1-p_L)^2 s / (1-p_L s)^2$ and $R_* = R_G \mu = N p_G (1 + p_L) / (1 - p_L)$.

2.4. Epidemics among giants. We consider here the simple case of a large number (m) of large households of equal sizes n . For large households the idea of a threshold for local contacts makes sense. When we consider only these local contacts, each household has its own local simple Reed–Frost epidemic, with population size n and basic reproductive ratio $R_L = np_L$. It is well known that the behavior of these single-group models goes through a phase transition at around the value $R_L = 1$ [e.g., Whittle (1955), von Bahr and Martin-Löf (1980), Nåsell (1995) and Ball and Nåsell (1994)], and it is interesting to examine the implications of this for the present two level model.

If $R_L \leq 1$, the epidemics in individual households are below threshold and the contact graph within each household consists of components all small compared with n , in fact of size $O(1)$. Then the analysis of the previous section applies, with R_* being greater than $R_G = N p_G$ by the factor μ equal to the mean size of these components; but μ is only $O(1)$, so we still require $p_G = O(1/N)$ to get a global epidemic—that is, a large outbreak at the interhousehold level.

The situation is more interesting when $R_L > 1$, so that the within-household epidemics are above threshold. Then each has its own giant connected component, of size nz_h say, and it is easy to see that the epidemic among the meta-population of giants has $R_* = nN p_G z_h^2$. Further, it is not too difficult to see that the members of households out with the giants do not significantly affect the probability or size of the overall outbreak.

Since $z_h = O(1)$ in this case, it only requires p_G to be $O(1/Nn)$ for R_* to be greater than 1, and thus make possible a large outbreak among the giants; that is, R_G need only be $O(1/n)$. Then the proportion of giants forming the “meta-giant” component of those involved in this large-scale Reed–Frost epidemic is given by $1 - z_g = \exp(-R_* z_g)$, so that the final proportion of the whole population affected is $z_h z_g$, and, as usual in the independent links case, this is also the probability of a large outbreak arising from an initial infected individual: here z_h is represents the probability that the individual belongs to its local giant, and z_g is the probability that this giant belongs to the meta-giant.

In all cases the local contacts have an amplifying effect on the global epidemic. But, for large households this amplification undergoes a significant change (we might call this a phase transition) from $O(1)$ to $O(n)$ as we reach the local threshold ($R_L = 1$) at which the households go through their individual phase transitions [von Bahr and Martin-Löf (1980)].

Finally here, we note the consistency of these essentially asymptotic results with those of the previous section. If in the clumped Reed–Frost model we let the clump distribution tend to that concentrated on nz_h , with probability z_h , and 0, with probability $1 - z_h$, then $\mu = nz_h^2$ so that both models agree that $R_* = nNp_Gz_h^2$ and $G_\pi(s) = (1 - z_h) + z_h s^{nz_h}$. So that the equation for the final size becomes $1 - z = (1 - z_h)1 + z_h \exp(-Np_Gznz_h)$, which, if we write $z_g = z/z_h$, boils down to $1 - z_g = \exp(-R_*z_g)$ as obtained above for the giant epidemic.

2.5. Vaccination strategies in relation to local thresholds. In a homogeneously mixing population, the minimum proportion v that we need to vaccinate to render the remaining susceptible population sub-threshold is given by $R'_G = (1 - v)R_G = 1$; that is, we require $v \leq 1 - 1/R_G$.

With our two levels of mixing, we have found that the basic reproductive ratio is $R_* = \mu R_G$. For a population divided into large groups, R_* can take large values, since μ will be a significant proportion of group size if groups are above their individual thresholds ($R_L > 1$). (Recall, from Section 1.2, that R_* is a parameter describing group-to-group infection and is therefore not directly comparable with individual-to-individual reproductive ratios such as R_G and R_L .)

Now vaccination of a proportion v of the population will still simply reduce R_G *pro rata* to $R'_G = (1 - v)R_G$, but the effect on μ will depend on the distribution of vaccination among the population. We shall consider the question of optimal vaccination strategies in more detail and generality in Section 5.2; here we simply indicate the practical importance of this question.

For the groups or households model, one strategy is to vaccinate whole groups. Let us assume for simplicity that if they are of different sizes, we choose groups at random, that is, according to the distribution $\{\pi_k\}$. Then μ will be unchanged, so that the overall reproductive ratio will simply become $R'_* = (1 - v)R_*$. However, a strategy in which we vaccinate a proportion of those in each group—for instance, the strategy in which we simply vaccinate members of the overall population chosen at random—can also reduce μ and thus reduce R_* further. In the case where vaccination changes groups from being above to below their local threshold, the difference can be dramatic, as the following simple numerical example illustrates.

Suppose that our population is divided into groups of size $n = 1000$ (perhaps schools or local communities) and that the reproductive ratio R_G for global contacts is 1 (the exact value is not important for what follows). Suppose also that $p_L = 0.003$, so that the reproductive ratio for local contacts is $R_L = (n - 1)p_L \approx 3$. Then $z_h \approx 0.95$, whence (in the notation of the last section) $\mu \approx nz_h^2 \approx 900$, and hence the overall reproductive ratio is $R_* = \mu R_G \approx 900$. (In contrast, the individual-to-individual reproductive ratio here is $R_0 \approx R_L + R_G \approx 4$.)

We now consider two alternative strategies for vaccinating 80% of the population. First note that with any such strategy, R'_G will be $(1 - 0.8)R_G = 0.2$. If we have a “patchy” vaccination program that vaccinates whole groups, we

will have $R'_* = \mu R'_G \approx 900 \times 0.2 = 180$, still far above threshold. However, if we have a uniform vaccination program, in which approximately 80% of each group are vaccinated, the local reproductive ratio will be brought down to $R'_L \approx (1 - 0.8) \times 3 = 0.6$. The groups will thus be below their local thresholds and their new mean clump size is easily calculated (from an approximating branching process, as in Section 3.3.2) to be $\mu' \approx 1/(1 - 0.6) = 2.5$. Thus in this case we will have $R'_* = \mu' R'_G \approx 2.5 \times 0.2 = 0.5$, so that vaccination will succeed in bringing the infection below threshold.

We can go further: from the practical point of view it is interesting to consider a program aimed at uniform coverage, but which is inadequate in some groups, meaning that in them there are still enough susceptibles left for the group to be above its local threshold. We find that where the initial value of R'_* is large, a quite small proportion of groups with inadequate coverage suffices to leave the population as a whole above threshold, that is, $R'_* > 1$. Extending our example of groups of size 1000 with $R_G = 1$, $R_L = 3$, if we have a program which generally vaccinates 80% within each group, the program will fail ($R'_* > 1$) if there is just 1% of the groups in which vaccination coverage is only 50%.

3. The model with a general infectious period.

3.1. The basic model. We now consider a generalization of the households model of Section 2, in which the infectious period may follow any arbitrary but specified distribution. Let the population consist of N individuals, subdivided into m groups each of size n . (We shall treat the case of unequal group sizes in Section 3.5.) The infectious periods of different infectives are independently and identically distributed according to a random variable T_I . Throughout its infectious period a given infective makes contact with each other susceptible in the population at the points of a homogeneous Poisson process having rate λ_G/N and, additionally, with each susceptible in its own group at the points of a homogeneous Poisson process having rate λ_L . All the Poisson processes describing infectious contacts (whether or not either or both of the individuals involved are the same), as well the random variables describing infectious periods, are assumed to be mutually independent.

Note that we have chosen here to formulate our model so that an individual i can make both local and global contacts with a susceptible in its own group, which may seem slightly unnatural. However, for the groups model of this section it facilitates our analysis by putting all individuals on an equal footing with respect to global contacts, and it is essential for cases such as the great circle model where the population is not partitioned into groups (see Section 2.3).

The alternative would be to treat all within-group contacts as local, at rate $\lambda'_L = \lambda_L + \lambda_G/N$. It follows from the superposition and splitting (or “coloring”) properties of the Poisson process [see, e.g., Kingman (1993)] that for the groups model (provided $\lambda'_L \geq \lambda_G/N$) the two formulations are exactly equivalent; and of course $\lambda'_L \rightarrow \lambda_L$ as $N \rightarrow \infty$.

For ease of exposition we shall assume that there is no latent period. However, all our results can be generalized to a model that incorporates a latent period. In particular, the final outcome of the epidemic is invariant to very general assumptions concerning a latent period. This can be seen by considering the random graph associated with the epidemic, in which for any two nodes, i, j say, a directed arc from i to j is present if and only if i will infect j if i is an infective and j is a susceptible.

The epidemic is initiated by a number of individuals becoming infected at time $t = 0$. We shall consider the spread of the epidemic in the asymptotic situation where the number of groups m tends to infinity while the group size n is held fixed.

If in this model we let T_I take a constant value t_I , then the epidemic has the same final outcome as the Reed–Frost model of Section 2 with $p_L = 1 - \exp(-\lambda_L t_I)$ and $p_G = 1 - \exp(-\lambda_G t_I/N)$. If instead we let T_I follow an exponential distribution, then our model reduces to the “equivalent classes” model of Watson (1972). Watson studied the deterministic version of the equivalent classes model, and also the branching process approximation as the group size n tends to infinity with the number of groups m fixed and finite. This contrasts sharply with our asymptotic regime outlined above.

3.2. Final outcome of a single population SIR stochastic epidemic. Consider a closed homogeneously mixing population consisting initially of n susceptibles and a infectives, who have just been infected. Suppose, as above, that the infectious period is distributed according to a random variable T_I and that throughout its infectious period a given infective infects a given susceptible at rate λ_L . The epidemic ceases as soon as there are no infectives present in the population. Let T be the final size of the epidemic, that is, the total number of initial susceptibles that are ultimately infected by the epidemic. Let T_A be the severity of the epidemic, that is, the sum of the infectious periods of all individuals infected during the course of the epidemic, including the a initial infectives. Note that T_A is equal to the area under the trajectory of infectives; see, for example, Downton (1972). The joint distribution of (T, T_A) is studied in Ball (1986). More recently, a general framework for analyzing the final size and severity of SIR stochastic epidemics has been developed in a series of papers by Lefèvre and Picard; see, for example, Picard and Lefèvre (1990). A key tool in their framework is a nonstandard family of polynomials, first introduced by Gontcharoff (1937), which we now outline.

Let $U = u_0, u_1, \dots$ be a given sequence of real numbers. Then the Gontcharoff polynomials attached to U , $G_0(x|U), G_1(x|U), \dots$, are defined recursively by the triangular system of equations

$$(3.1) \quad \sum_{j=0}^i \frac{u_j^{i-j}}{(i-j)!} G_j(x|U) = \frac{x^i}{i!}, \quad i = 0, 1, \dots$$

For $i = 1, 2, \dots$, the polynomial $G_i(x|U)$ admits the integral representation

$$(3.2) \quad G_i(x|U) = \int_{u_0}^x \int_{u_1}^{\xi_0} \int_{u_2}^{\xi_1} \cdots \int_{u_{i-1}}^{\xi_{i-2}} d\xi_0 d\xi_1 d\xi_2 \cdots d\xi_{i-1};$$

see, for example, Lefèvre and Picard [(1990), (2.5)]. Another property of Gontcharoff polynomials [see (2.7) of Lefèvre and Picard (1990)] that we shall require is

$$(3.3) \quad G_i^{(j)}(x|U) = G_{i-j}(x|E^j U), \quad 0 \leq j \leq i,$$

where $E^j U$ is the sequence u_j, u_{j+1}, \dots and $G_i^{(j)}(x|U)$ is the j th derivative of $G_i(x|U)$. Note that $G_i^{(j)}(x|U) = 0$ if $j > i$.

For the single population epidemic model, let $\phi(\theta) = E[\exp(-\theta T_I)]$, $\theta \geq 0$, be the moment generating function of T_I and let

$$(3.4) \quad \phi_{n,a}(s, \theta) = E[s^{n-T} \exp(-\theta T_A)], \quad \theta \geq 0.$$

Then it follows from Proposition 3.3 of Picard and Lefèvre (1990) [see also Ball and Clancy (1993)] that

$$(3.5) \quad \phi_{n,a}(s, \theta) = \sum_{i=0}^n \frac{n!}{(n-i)!} \phi(\theta + \lambda_L i)^{n+a-i} G_i(s|U),$$

where the sequence U is given by $u_i = \phi(\theta + \lambda_L i)$, $i = 0, 1, \dots$.

Let $\mu_{n,a} = E[T]$ be the mean final size of the above epidemic. Then by differentiating (3.5) with respect to s and setting $s = 1$ and $\theta = 0$, it follows using (3.3) that

$$(3.6) \quad \mu_{n,a} = n - \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n+a-i} \alpha_i,$$

where $q_i = \phi(\lambda_L i)$ and $\alpha_i = G_{i-1}(1|V)$. Here the sequence V is given by $v_i = \phi(\lambda_L(i+1)) = q_{i+1}$ (for $i = 0, 1, \dots$). We may call the q_i 's the escape probabilities, since $q_i = E[\exp(-i\lambda_L T_I)]$ is the probability that a set of i individuals exposed to a single infective in the same group all escape infection by it. From this interpretation it is immediate that the q_i 's, and hence the v_i 's, are monotone nonincreasing:

$$(3.7) \quad q_i \geq q_{i+1} \quad \text{for all } i \geq 0 \text{ (note that } q_0 = 1).$$

Note that it is straightforward to compute $\alpha_1, \alpha_2, \dots$ numerically using the recursive definition of the Gontcharoff family of polynomials given in (3.1).

In Sections 3.4 and 3.5 we shall require the fact that $\alpha_i > 0$, $i = 1, 2, \dots$, which we now prove. The integral definition of $G_i(x|U)$ given in (3.2) implies that $G_i(x|U) > 0$ for $i = 0, 1, \dots$, provided that $x > u_0 \geq u_1 \geq \dots \geq 0$. [This gives a new and elegant proof of a result proved in Gani and Shanbhag (1974).] The strict positivity of the α_i 's follows immediately from (3.7) (remembering that $v_i = q_{i+1}$).

We shall need the moment generating function of T_A ,

$$\psi_{n,a}(\theta) = E[\exp(-\theta T_A)]$$

say, which can be obtained by setting $s = 1$ in (3.5).

Consider now an extension of the single population epidemic model, in which susceptibles can also be infected from outside the population. Specifically, suppose that each of the n initial susceptibles has probability π of avoiding infection from outside the population during the course of the epidemic, independently of other susceptibles in the population. This extended model has been considered by Addy, Longini and Haber (1991), who derived recursive expressions for the probability generating function of T and the moment generating function of T_A . The final outcome of the extended model with outside infection has the same distribution as that of the single population model with initial numbers of infectives and susceptibles $a + Y$ and $n - Y$, respectively, where Y is a realization of a binomial random variable with parameters n and $1 - \pi$. (This follows by considering the random graph associated with the epidemic.)

Let $\tilde{\phi}_{n,a}(s, \theta) = E[s^{n-T} \exp(-\theta T_A)]$, $\theta \geq 0$, be the joint generating function of (T, T_A) for the model with outside infection. Then conditioning on the value of Y and using (3.5) yields

$$(3.8) \quad \tilde{\phi}_{n,a}(s, \theta) = \sum_{k=0}^n \binom{n}{k} \pi^k (1 - \pi)^{n-k} \sum_{i=0}^k \frac{k!}{(k-i)!} \phi(\theta + \lambda_L i)^{n+a-i} G_i(s|U),$$

which on changing the order of summation gives, after a little algebra,

$$(3.9) \quad \tilde{\phi}_{n,a}(s, \theta) = \sum_{i=0}^n \frac{n!}{(n-i)!} \phi(\theta + \lambda_L i)^{n+a-i} \pi^i G_i(s|U).$$

Let $\tilde{\mu}_{n,a} = E[T]$ be the mean final size for the epidemic with outside infection. Then arguing as in the derivation of (3.6) yields

$$(3.10) \quad \tilde{\mu}_{n,a} = n - \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n+a-i} \pi^i \alpha_i.$$

We now give expressions for the final size distribution of the single population epidemic model with outside infection. Let $\tilde{P}_k^n = \Pr\{T = k\}$, $k = 0, 1, \dots, n$. Then setting $\theta = 0$ in (3.9), differentiating $n - k$ times with respect to s and using (3.3) yields

$$(3.11) \quad \tilde{P}_k^n = \frac{1}{(n-k)!} \sum_{i=n-k}^n \frac{n!}{(n-i)!} q_i^{n+a-i} \pi^i G_{i-n+k}(0|E^{n-k}U), \quad k = 0, 1, \dots, n,$$

where the sequence U is given by $u_i = q_i = \phi(\lambda_L i)$, $i = 0, 1, \dots$.

Addy, Longini and Haber (1991) gave a similar expression to (3.11), but not using Gontcharoff polynomials. They also showed that the final size probabilities can be determined from the triangular system of linear equations

$$(3.12) \quad \sum_{i=0}^k \binom{n-i}{k-i} \tilde{P}_i^n / \{q_{n-k}^{a+i} \pi^{n-k}\} = \binom{n}{k}, \quad k = 0, 1, \dots, n.$$

Setting $\pi = 1$ in (3.12) yields a set of linear equations governing the final size distribution of the epidemic without outside infection; see Ball (1986).

The above systems of equations are in principle straightforward to solve numerically, because of their triangular structure. Numerical problems due to rounding errors can occur even for moderate values of n , perhaps $n = 50$ or 100. However, in many applications, n will correspond to group or household size and will typically be small, say $n \leq 5$ or 10, permitting the required properties to be calculated accurately.

3.3. Initial stages of a multigroup epidemic.

3.3.1. Branching process approximation. Suppose that the number of groups m , and hence the total population N , is large. Then during the early stages of the epidemic, every time a between group infection occurs the contacted individual is likely to be in a previously uninfected group. Thus the initial stages of the epidemic can be approximated by a branching process, in which the units are single group epidemic processes and the offspring of a given unit are those groups that are directly infected by infectives in that unit.

The approximation can be made precise by considering a sequence of epidemics with fixed group size n , indexed by the number of groups m , and using the coupling argument of Ball (1983b) and Ball and Donnelly (1995). Specifically, the epidemic processes and the approximating branching process can be constructed on the same probability space (Ω, \mathcal{F}, P) such that if $A \subseteq \Omega$ denotes the set on which the branching process goes extinct, then (i) for P -almost all $\omega \in A$ the process of infectives in the epidemic process and the branching process agree over the time interval $[0, \infty)$ for all sufficiently large m and (ii) for P -almost all $\omega \in \Omega \setminus A$ the epidemic process and the branching process agree over $[0, c \log m]$ for all sufficiently large m , for any $c < (2\alpha)^{-1}$, where α is the Malthusian parameter of the branching process. The result in (ii) is the best possible in the sense that if $c > (2\alpha)^{-1}$, then, for all sufficiently large m , the epidemic process and the branching process disagree over part of the interval $[0, c \log m]$, and the maximum difference tends to infinity as $m \rightarrow \infty$. The Malthusian parameter α can be obtained as follows. For $t \geq 0$, let $Y(t)$ denote the number of infectives at time t in the single group epidemic model of Section 3.2, when initially there are one infective and $n - 1$ susceptibles. Then, provided that the branching process is supercritical, α is the unique solution in $(0, \infty)$ of the equation

$$(3.13) \quad \int_0^\infty \lambda_G E[Y(t)] \exp(-at) dt = 1$$

[see Ball (1996) for details].

The final size of the approximating branching process can be obtained by considering its embedded Galton–Watson process, whose offspring distribution can be derived as follows. A typical unit in the branching process commences with one of the susceptibles in the group being infected from outside. That infective will start an epidemic within its own group. Each infective in this single-group epidemic independently makes infections outside the group at

rate λ_G throughout its infectious period. Hence the total number of outside infections emanating from the group under consideration follows a Poisson distribution with random mean $\lambda_G T_A$, where T_A is the severity of the single-group epidemic. Further, in the branching process approximation, all of these outside infections are with susceptibles in distinct groups, so the offspring distribution, R say, of the embedded Galton–Watson process is also Poisson with random mean $\lambda_G T_A$. Let $R_* = \mathbb{E}[R]$. Then, letting T be the final size of the single group epidemic and using the Wald identity for epidemics proved in Ball (1986), we obtain

$$\begin{aligned} R_* &= \lambda_G \mathbb{E}[T_A] \\ (3.14) \quad &= \lambda_G(1 + \mathbb{E}[T])\mathbb{E}[T_I] \\ &= \lambda_G(1 + \mu_{n-1,1})\mathbb{E}[T_I]. \end{aligned}$$

Note that, as in the clumped Reed–Frost model, R_* is of the form $R_* = \mu R_G$, where $R_G = \lambda_G \mathbb{E}[T_I]$ is the basic reproductive ratio for the model in which all the groups are of size 1, that is, $n = 1$, and $\mu = 1 + \mu_{n-1,1}$ is the mean clump size.

To obtain a threshold theorem for the multigroup epidemic process, we say that a global epidemic occurs if in the limit as $m \rightarrow \infty$ the epidemic infects infinitely many groups. By standard branching process theory [see, e.g., Jagers (1975)], global epidemics can occur if and only if $R_* > 1$, so R_* may be viewed as the threshold parameter for the multigroup epidemic. Note that for any given set of parameter values, R_* can be computed using (3.6). Indeed, for small values of the group size n , explicit expressions for α_i , and hence for R_* , can easily be obtained.

The probability of a global epidemic depends on the number and configuration of initial infectives. Consider first the case in which the epidemic is initiated by just one of the susceptibles becoming infected. Then, again by standard branching process theory, the probability of a global epidemic is $\zeta = 1 - \tau$, where τ is the smallest root in $[0, 1]$ of the equation $f(s) = s$. Here $f(s)$ is the probability generating function of R , which, conditioning on the value of T_A , is given by

$$\begin{aligned} f(s) &= \mathbb{E}[s^R] \\ (3.15) \quad &= \mathbb{E}[\mathbb{E}[s^R | T_A]] \\ &= \mathbb{E}[\exp(-\lambda_G T_A(1-s))] \\ &= \psi_{n-1,1}(\lambda_G(1-s)), \quad 0 \leq s \leq 1. \end{aligned}$$

For $i = 1, 2, \dots, n$, let τ_i be the probability of a nonglobal epidemic when initially there is one infectious group containing i infectives and $n - i$ susceptibles, so $\tau_1 = \tau$. Let Z be the size of the first generation in the embedded Galton–Watson process; that is, Z is the total number of outside infections emanating from the initial single-group epidemic. Then, again conditioning

on the value of T_A ,

$$\begin{aligned}
 \tau_i &= \mathbb{E}[\tau^Z] \\
 (3.16) \quad &= \mathbb{E}[\mathbb{E}[\tau^Z | T_A]] \\
 &= \mathbb{E}[\exp(-\lambda_G T_A(1 - \tau))] \\
 &= \psi_{n-i, i}(\lambda_G(1 - \tau)).
 \end{aligned}$$

Finally, if initially there are a_i infectious groups with i infectives and $n - i$ susceptibles, $i = 1, 2, \dots, n$, then

$$(3.17) \quad \Pr\{\text{global epidemic}\} = 1 - \prod_{i=1}^n \tau_i^{a_i}.$$

Note that $\psi_{n-i, i}(\theta)$, and hence τ_i , $i = 1, 2, \dots, n$, are straightforward to compute by setting $s = 1$ in (3.5) and using the recursive definition (3.1) for the quantities $G_0(1|U)$, $G_1(1|U), \dots, G_{n-1}(1|U)$.

Other properties of the approximating branching process are straightforward to determine. Suppose that initially there is one infectious group containing just one infective. Let \tilde{N} and \tilde{G} be, respectively, the total number of individuals and total number of groups infected by the epidemic, where now the initial infective and the initial infectious group are included. As before, let T and T_A be, respectively, the final size and severity of the single group epidemic in the initial infectious group. Let $h(s_1, s_2) = \mathbb{E}[s_1^{\tilde{N}} s_2^{\tilde{G}}]$ be the joint probability generating function of (\tilde{N}, \tilde{G}) under the branching process approximation. Then, conditioning on (T, T_A) ,

$$\begin{aligned}
 h(s_1, s_2) &= \mathbb{E}\left[\mathbb{E}\left[s_1^{\tilde{N}} s_2^{\tilde{G}} | T, T_A\right]\right] \\
 (3.18) \quad &= \mathbb{E}\left[\mathbb{E}\left[s_1^{1+T+\sum_{i=1}^Z \tilde{N}_i} s_2^{1+\sum_{i=1}^Z \tilde{G}_i} | T, T_A\right]\right],
 \end{aligned}$$

where, as above, Z is the size of the first generation in the embedded Galton–Watson process and $(\tilde{N}_1, \tilde{G}_1), (\tilde{N}_2, \tilde{G}_2), \dots, (\tilde{N}_Z, \tilde{G}_Z)$ are independent and identically distributed copies of (\tilde{N}, \tilde{G}) . Now Z is Poisson with mean $\lambda_G T_A$, so

$$\begin{aligned}
 h(s_1, s_2) &= s_1 s_2 \mathbb{E}[s_1^T \mathbb{E}[h(s_1, s_2)^Z | T, T_A]] \\
 (3.19) \quad &= s_1 s_2 \mathbb{E}[s_1^T \exp(-\lambda_G T_A(1 - h(s_1, s_2)))] \\
 &= s_1 s_2 \hat{\phi}_{n-1, 1}(s_1, \lambda_G(1 - h(s_1, s_2))),
 \end{aligned}$$

where

$$\begin{aligned}
 \hat{\phi}_{n-1, 1}(s, \theta) &= \mathbb{E}[s^T \exp(-\theta T_A)] \\
 (3.20) \quad &= s^{n-1} \phi_{n-1, 1}(s^{-1}, \theta)
 \end{aligned}$$

is the joint generating function of (T, T_A) . Thus $h(s_1, s_2)$ satisfies the functional equation

$$(3.21) \quad h(s_1, s_2) = s_1^n s_2 \phi_{n-1, 1}(s_1^{-1}, \lambda_G(1 - h(s_1, s_2))).$$

Appropriate differentiation of (3.21) yields expressions for the moments of \tilde{N} and \tilde{G} , such as $E[\tilde{N}]$, $E[\tilde{G}]$, $\text{var}(\tilde{N})$, $\text{var}(\tilde{G})$ and $\text{cov}(\tilde{N}, \tilde{G})$. Note that (3.3) and the recursive definition (3.1) of Gontcharoff polynomials enables the derivatives of $\phi_{n-1, 1}(s, \theta)$ and hence the above moments to be calculated. When $R_* \geq 1$, the above moments are all infinite. However, if $R_* > 1$, then moments conditional upon the occurrence of a nonglobal epidemic can be derived from (3.21).

An alternative approach to determining the limiting properties of (\tilde{N}, \tilde{G}) as the number of groups $m \rightarrow \infty$ is via the two-type branching process described in Section 3.6.

3.3.2. Discussion of the threshold parameter R_* . We now discuss the relationship of our threshold parameter R_* to the classical reproductive ratio R_0 [see, e.g., Diekmann, Heesterbeek and Metz (1990)] for the multigroup epidemic. For definiteness of argument, suppose that the infectious period T_I follows an exponential distribution with mean γ^{-1} , so that our model becomes a multigroup generalization of the general stochastic epidemic [see, e.g., Bailey (1975), Chapter 6]. The deterministic version of our model is then expressed by the differential equations

$$(3.22) \quad \begin{aligned} \frac{dx_i}{dt} &= -\left(\lambda_L y_i + N^{-1} \lambda_G \sum_{j \neq i} y_j\right)x_i, \\ \frac{dy_i}{dt} &= \left(\lambda_L y_i + N^{-1} \lambda_G \sum_{j \neq i} y_j\right)x_i - \gamma y_i, \quad i = 1, 2, \dots, m, \end{aligned}$$

where the groups are labelled $1, 2, \dots, m$ and $x_i(t)$ and $y_i(t)$ are, respectively, the numbers of susceptibles and infectives in the i th group at time t .

The reproductive ratio for the above deterministic model, usually defined informally (in a stochastic sense!) as the expected number of infectious contacts made by a single initial infective in an otherwise susceptible population, is $R_0 = \{(n-1)\lambda_L + \lambda_G\}/\gamma$. In the deterministic setting, a major epidemic occurs if and only if $R_0 > 1$. However, as we shall see soon, $R_0 > 1$ does not generally provide a good indication as to whether a global epidemic can occur in our stochastic model. This is because a deterministic model can only be a good approximation to the more realistic stochastic model if all the population sizes are large [cf. the convergence theorems of Kurtz (1970, 1981)], but in the multigroup epidemic the group size n is often small. Thus the deterministic model (3.22) will not generally provide an adequate description of the multigroup epidemic. Indeed, a more appropriate deterministic model is one described by a system of differential equations for $x_{i,j}(t)$, $0 \leq i, j \leq n$, where $x_{i,j}(t)$ is the number of groups with i susceptibles and j infectives at time t .

It is now convenient to assume that the multigroup epidemic model is parameterized so that the within-group infection rate is $(n-1)^{-1}\lambda_L$ and, for the purpose of illustration, that the time axis is linearly rescaled so that $\gamma = 1$. Under these assumptions, $R_0 = \lambda_L + \lambda_G$ independently of the group size n . The threshold parameter R_* can be calculated using (3.14). Figure 1 shows for various group sizes n the graph of critical values of (λ_L, λ_G) so that $R_* = 1$.

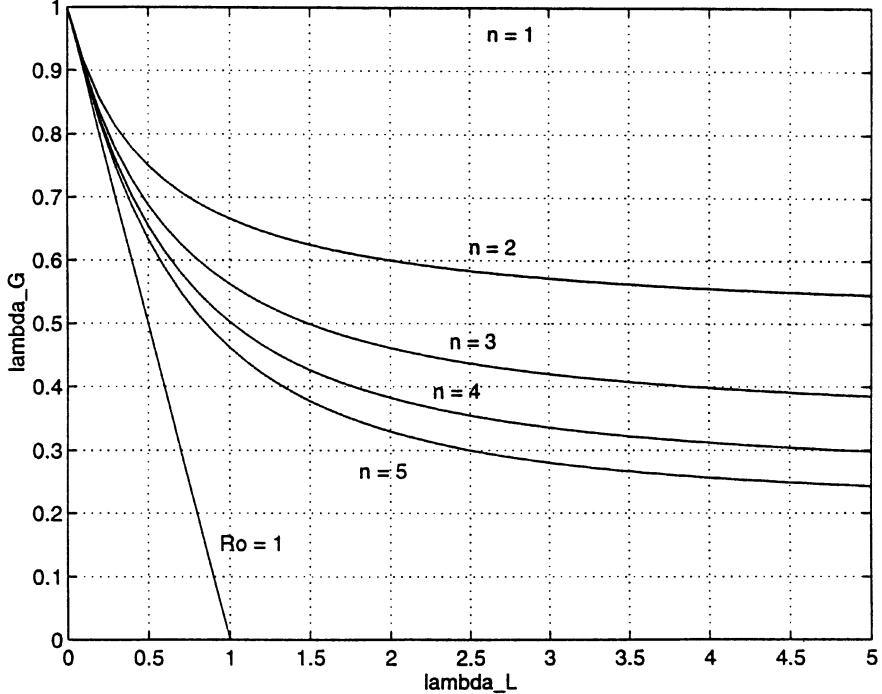


FIG. 1. Critical values of (λ_L, λ_G) so that $R_* = 1$.

The corresponding graph for $R_0 = 1$ is also shown. When the group size $n = 1$, the graph corresponding to $R_* = 1$ is constant at $\lambda_G = 1$, since then there can be no within-group spread of infection and the value of λ_L is irrelevant. For $n = 1, 2, \dots$, the $R_* = 1$ graph for $n + 1$ lies below that for n and it is shown below that the $R_* = 1$ graph converges to the $R_0 = 1$ graph as $n \rightarrow \infty$. As noted above, $R_0 = 1$ does not provide a good indicator as to whether a global epidemic can occur when the group size n is small.

We now return to the model with general T_I and examine the asymptotic behavior of R_* as $n \rightarrow \infty$. We still assume without loss of generality that $E[T_I] = 1$ and that the within-group infection rate is $(n - 1)^{-1}\lambda_L$. Thus from (3.14), $R_* = \lambda_G(1 + \mu_{n-1,1})$, so we are interested in the asymptotic behavior of $\mu_{n-1,1}$ as $n \rightarrow \infty$. For large n , the early stages of the single-group epidemic process can be approximated by a branching process and Ball (1983b) shows how to make the approximation precise in the limit as $n \rightarrow \infty$. Specifically, a sequence of epidemic processes indexed by n and the approximating branching process can be constructed on the same probability space so that, as $n \rightarrow \infty$, if the branching process goes extinct, then the final size of the epidemic process converges almost surely to the final size of the branching process, and if the branching process does not go extinct, then the final size of the epidemic process converges almost surely to ∞ . Moreover, in the latter case von Bahr and Martin-Löf (1980) show that $n^{1/2}(T/n - \rho)$ converges in distribution to

a normal random variable with zero mean. Here T is the final size of the single population epidemic and ρ is the largest root in $(0, 1)$ of the equation $1 - x = \exp(-\lambda_L x)$. Note that ρ is the proportion of initial susceptibles that are ultimately infected in the limiting deterministic epidemic as $n \rightarrow \infty$.

Suppose first that $\lambda_L < 1$, so that the single-group epidemic is below threshold. The mean final size of the approximating branching process is $\lambda_L/(1-\lambda_L)$, so by the dominated convergence theorem $\lim_{n \rightarrow \infty} \mu_{n-1,1} = \lambda_L/(1-\lambda_L)$. Thus R_* converges up to $\lambda_G(1+(1-\lambda_L)^{-1}\lambda_L) = \lambda_G/(1-\lambda_L)$ as n tends to ∞ . Hence, if $\lambda_G < 1 - \lambda_L$ only nonglobal epidemics can occur, however large the group size n is, while if $\lambda_G > 1 - \lambda_L$ global epidemics can occur provided that n is sufficiently large. Note that as $n \rightarrow \infty$ the equation $R_* = 1$ converges to $\lambda_L + \lambda_G = 1$, that is, $R_0 = 1$.

Now suppose that $\lambda_L > 1$ so that the single-group epidemic is above threshold. Let q be the probability that the approximating branching process (to the single-group epidemic) goes extinct. Recall from Waugh (1958) and Daly (1979) that, conditional upon extinction, a supercritical Galton–Watson process with offspring probability generating function $g(s)$ behaves as a (subcritical) Galton–Watson process with offspring probability generating function $q^{-1}g(qs)$. The number of contacts made by the initial infective in the single population epidemic is Poisson with (random) mean $\lambda_L T_I$, so the offspring probability generating function of the Galton–Watson process embedded in the approximating branching process is $g(s) = \phi(\lambda_L(1-s))$, where $\phi(\theta) = E[\exp(-\theta T_I)]$. Thus q is the unique solution in $(0, 1)$ of the equation $\phi(\lambda_L(1-s)) = s$ and the offspring mean for the embedded Galton–Watson process conditioned upon extinction, \tilde{m} say, is given by $\tilde{m} = -\lambda_L \phi^{(1)}(\lambda_L(1-q))$. Further, conditional upon extinction, the mean final size of the approximating branching process is $\tilde{m}/(1-\tilde{m})$. Combining all this with the above von Bahr and Martin-Löf limit theorem and recalling that $R_* = \lambda_G(1 + \mu_{n-1,1})$ yields

$$(3.23) \quad R_* \sim \lambda_G \left\{ 1 + \frac{q\tilde{m}}{1-\tilde{m}} + (1-q)\rho(n-1) \right\} \quad \text{as } n \rightarrow \infty.$$

Thus, in contrast to the situation when $\lambda_L < 1$, for $\lambda_L > 1$ global epidemics can always occur if n is sufficiently large, whatever the value of λ_G (provided it is not zero).

In the critical case $\lambda_L = 1$, the mean final size of the approximating branching process is infinite, so again global epidemics can always occur provided that n is sufficiently large.

Another way of viewing the above is to assume that λ_L is fixed and examine the behavior, as $n \rightarrow \infty$, of the critical value, λ_G^{crit} say, of λ_G for global epidemics to be possible. It follows from the preceding arguments that if $\lambda_L < 1$, then $\lambda_G^{\text{crit}} = O(1)$ as $n \rightarrow \infty$, whilst if $\lambda_L > 1$, then $\lambda_G^{\text{crit}} = O(n^{-1})$ as $n \rightarrow \infty$. This corresponds to the amplification effect discussed for the multigroup Reed–Frost epidemic in Section 2.5.

We can use our model to study the efficacy of various vaccination strategies. For example, as in Section 2.5, consider two vaccination policies: a local one in which a fixed proportion, θ say, of groups is completely vaccinated and a global

one in which a proportion θ of susceptibles in every group is vaccinated. For convenience we suppose that θn is an integer. Under both policies the rate at which a given infective makes outside infections is $(1 - \theta)\lambda_G$. However, in the local policy such an infection is with a group having n susceptibles, so $R_* = (1 - \theta)\lambda_G(1 + \mu_{n-1, 1})$, but in the global policy it is with a group having $(1 - \theta)n$ susceptibles, so $R_* = (1 - \theta)\lambda_G(1 + \mu_{(1-\theta)n-1, 1})$. Clearly $\mu_{n-1, 1} > \mu_{(1-\theta)n-1, 1}$ so the global policy will be more effective in preventing the spread of a global epidemic.

3.4. Final outcome of a multigroup epidemic. In this subsection we consider the final outcome of the multigroup epidemic as m , the number of groups, becomes large. In Section 3.3.1 we examined the final outcome of a nonglobal epidemic; here we shall be concerned with what happens in the event of a global epidemic. Our argument will be heuristic, a formal proof being delayed until Section 4.2.

Let z be the expected proportion of initial susceptibles that are infected by a global epidemic. Thus z can be interpreted as the probability that a given initial susceptible, who is not in one of the initially infectious groups, is ultimately infected by the epidemic.

Fix attention on a single group that initially contained no infectives. We can decompose the ultimate spread of infection within that group by first determining which of the initial susceptibles are infected from outside the group, and then letting these individuals initiate a single population epidemic among the remaining susceptibles in the group.

This decomposition justifies the approach taken by a number of applied authors, who treated the probability π of an external (primary) contact as independent of the internal (secondary) process that they analyzed (see Sections 1.3, 3.2 and 5.1).

Let \tilde{T} be the total person time units of infection present in the population at large over the whole course of the epidemic. Then for large m , $\tilde{T} \sim NzE[T_I]$. At any time a given susceptible in the group under consideration is being infected from outside the group with intensity $N^{-1}\lambda_G$ per outside infective. Thus, as $m \rightarrow \infty$, each given susceptible in the group independently avoids infection from outside with probability $\pi = \exp(-\lambda_G z E[T_I])$. It follows that the ultimate spread of infection within the group has the same distribution as that of the extended model of Addy, Longini and Haber (1991) described in Section 3.2. Hence, the mean final size of the epidemic within the group is given by setting $a = 0$ in (3.10). However, the mean final size also equals zn , since z is the expected proportion of susceptibles that are ultimately infected. Thus we can deduce the equation

$$(3.24) \quad nz = n - \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n-i} \pi^i \alpha_i,$$

which, since $\pi = \exp(-\lambda_G z E[T_I])$, is an implicit equation for z . Clearly $z = 0$ is always a solution of (3.24). We now show that there is a (unique) second solution in $(0, 1)$ if and only if $R_* > 1$.

It is convenient to rearrange (3.24) into

$$(3.25) \quad n(1-z) = \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n-i} \exp(-\lambda_G z E[T_I] i) \alpha_i.$$

We proved in Section 3.2 that $\alpha_i > 0$, $i = 1, 2, \dots$, so the right-hand side of (3.25) is a convex function of z . Thus (3.25) has at most two solutions since its left-hand side is linear in z . Further, by examining the values at $z = 0$ of the derivatives with respect to z of the two sides of (3.25), we see that there is a second solution if and only if

$$(3.26) \quad \lambda_G E[T_I] \sum_{i=1}^n \frac{(n-1)!}{(n-i)!} q_i^{n-i} i \alpha_i > 1.$$

Now

$$(3.27) \quad \begin{aligned} \sum_{i=1}^n \frac{(n-1)!}{(n-i)!} q_i^{n-i} i \alpha_i &= \sum_{i=1}^n \frac{(n-1)!}{(n-i)!} q_i^{n-i} \alpha_i (n - (n-i)) \\ &= \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n-i} \alpha_i - \sum_{i=1}^{n-1} \frac{(n-1)!}{(n-i-1)!} q_i^{n-i} \alpha_i. \end{aligned}$$

From (3.6), the second sum on the right-hand side of (3.27) is $n - 1 - \mu_{n-1,1}$. The first sum can be evaluated by recalling that $\alpha_i = G_{i-1}(1|V)$, where the sequence V is given by $v_i = q_{i+1} = \phi(\lambda_L(i+1))$, $i = 0, 1, \dots$. We obtain

$$(3.28) \quad \begin{aligned} \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n-i} \alpha_i &= \sum_{i=1}^n \frac{n!}{(n-i)!} v_{i-1}^{n-i} G_{i-1}(1|V) \\ &= n \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!} v_i^{n-1-i} G_i(1|V) \\ &= n, \end{aligned}$$

where in the last step we have used the recursive definition (3.1) of the Gontcharoff polynomials $G_i(x|V)$, $i = 0, 1, \dots$. Putting all this together, we obtain from (3.27) that

$$(3.29) \quad \sum_{i=1}^n \frac{(n-1)!}{(n-i)!} q_i^{n-i} i \alpha_i = 1 + \mu_{n-1,1}.$$

Hence from (3.26) and the expression for R_* given in (3.14), (3.24) has a solution in $(0, 1)$ if and only if $R_* > 1$. When $R_* > 1$ the solution of (3.24) in $(0, 1)$ gives the expected proportion of initial susceptibles ultimately infected by a global epidemic.

As noted earlier, in the event of a global epidemic the final size in a group that did not have initial infectives is distributed as the final size of the extended model of Addy, Longini and Haber (1991) with $\pi = \exp(-\lambda_G z E[T_I])$. This distribution may be calculated by using (3.11) or (3.12). Figure 2 illustrates, for various values of λ_L and λ_G , the final size distribution in a group

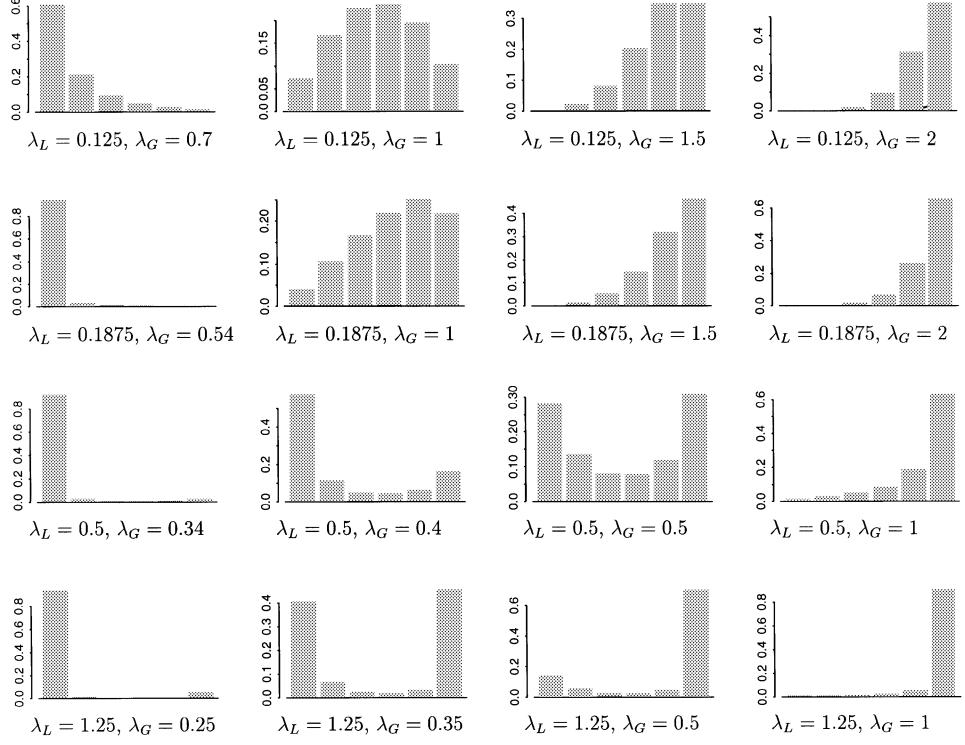


FIG. 2. *The final size distribution in a group when the infectious period T_I follows an exponential distribution with mean 1 and the group size $n = 5$. Note that the critical values of λ_G are (for each row, starting at the top) $\lambda_G^{\text{crit}} = 0.6336, 0.5321, 0.3296, 0.2429$.*

when the infectious period T_I follows an exponential distribution with mean 1 and the group size $n = 5$. Figure 2 also gives, for each choice of λ_L , the critical value λ_G^{crit} of λ_G for global epidemics to be possible. Notice the difference in the shape of the distribution according to whether the local reproductive ratio $R_L = (n - 1)\lambda_L$ is less than or greater than 1, that is, according to whether the within-group epidemic is below or above its threshold. When $R_L < 1$ the distribution is unimodal for all values of $\lambda_G > \lambda_G^{\text{crit}}$, with the mode increasing from 0 for values of λ_G just greater than λ_G^{crit} to $n (= 5)$ for sufficiently large values of λ_G . When $R_L > 1$ the distribution is initially bimodal as λ_G is increased from λ_G^{crit} , but becomes unimodal with the mode either at or close to n for sufficiently large values of λ_G . (In our examples the mode is always at $n = 5$ but this is unlikely to be the case in general.) The shape of the group final size distribution can be explained in terms of the threshold behavior of the single-group epidemic. When $R_L < 1$, only minor epidemics will occur within a group, but as λ_G increases so does the number of group members infected from the population at large, and hence also the size of the epidemic within the group. When $R_L > 1$, the within-group epidemic is above thresh-

old, so major epidemics can occur as soon as $\lambda_G > \lambda_G^{\text{crit}}$. Thus the distribution is bimodal, being a mixture of two components, one corresponding to a minor epidemic and the other to a major epidemic. Again, as λ_G increases, so does the number of outside infections, and eventually the minor epidemic component will disappear.

3.5. Unequal group sizes. We now consider the situation in which the group sizes are not all equal. For $n = 1, 2, \dots$, let m_n be the number of groups of size n . Let $m = \sum_{n=1}^{\infty} m_n$ be the total number of groups and $N = \sum_{n=1}^{\infty} nm_n$ be the total number of individuals. As before, the infectious periods of different infectives are independently and identically distributed according to a random variable T_I , and throughout its infectious period a given infective makes contact with each other susceptible at rate λ_G/N , and additionally with each susceptible in its own group at rate λ_L . We examine the asymptotic situation in which the number of groups m tends to infinity in such a way that $m_n/m \rightarrow h_n$, $n = 1, 2, \dots$, and $\sum_{n=1}^{\infty} h_n = 1$. Thus, for $n = 1, 2, \dots$, h_n is the asymptotic proportion of groups of size n . Let $\mu_h = \sum_{n=1}^{\infty} nh_n$ be the asymptotic mean group size and assume that $\mu_h < \infty$.

The initial stages of the multigroup epidemic can be approximated by a multitype branching process, in which the units are single-group epidemic processes, the offspring of a given unit are those groups that are directly infected by infectives in that unit and type corresponds to group size. Again the approximation can be made precise in the limit as $m \rightarrow \infty$ by using the coupling argument of Ball (1983b) and Ball and Donnelly (1995). Label the types $1, 2, \dots$, according to group size and let $\Lambda = [\lambda_{ij}]$ be the offspring mean matrix of the embedded multitype Galton–Watson process. Thus λ_{ij} is the expected number of type j groups infected by infectives from a type i group single population epidemic.

Let $T^{(i)}$ and $T_A^{(i)}$ be, respectively, the final size and severity of a single population epidemic in which initially there are 1 infective and $i - 1$ susceptibles. As in Section 3.3.1, the total number of outside infections emanating from a type i group follows a Poisson distribution with random mean $\lambda_G T_A^{(i)}$. The probability that a given outside infection is with an individual in a group of size j is $jm_j/N = jh_j/\mu_h$. Hence

$$\begin{aligned} \lambda_{ij} &= \lambda_G \mathbb{E}[T_A^{(i)}] jh_j/\mu_h \\ (3.30) \quad &= \lambda_G (1 + \mathbb{E}[T^{(i)}]) \mathbb{E}[T_I] jh_j/\mu_h \\ &= \lambda_G (1 + \mu_{i-1,1}) \mathbb{E}[T_I] jh_j/\mu_h, \end{aligned}$$

using the Wald identity for epidemics.

The multiplicative structure of the matrix Λ given by (3.30) implies that its maximal eigenvalue is

$$(3.31) \quad R_* = \lambda_G \mathbb{E}[T_I] \mu_h^{-1} \sum_{n=1}^{\infty} (1 + \mu_{n-1,1}) nh_n.$$

By standard branching process theory a global epidemic (corresponding to nonextinction of the approximating multitype branching process) has nonzero probability of occurring if and only if $R_* > 1$. Formulae implicitly giving the probability of a global epidemic and properties of a nonglobal epidemic can be derived as in Section 3.3.1.

Note that again R_* is of the form $R_* = \mu R_G$, where $R_G = \lambda_G E[T_I]$ is the basic reproductive ratio for the model in which all the groups are of size 1, and $\mu = \mu_h^{-1} \sum_{n=1}^{\infty} (1 + \mu_{n-1,1}) nh_n$ is the size-biased mean clump size. The formula for μ uses the fact that if π_i (as in Sections 1 and 2) is the probability that an individual chosen at random from the population is in a group of size i , then

$$(3.32) \quad \pi_i = \mu_h^{-1} i h_i, \quad i = 1, 2, \dots$$

[thus $\mu = \sum_{n=1}^{\infty} (1 + \mu_{n-1,1}) \pi_n$]. Indeed, using the size-biased sampling, the initial stages of the epidemic can be approximated by a single-type branching process (in which the units are single-group epidemic processes) whose offspring distribution is Poisson with random mean, which is a mixture of $T_A^{(1)}, T_A^{(2)}, \dots$ with respective mixing probabilities π_1, π_2, \dots . Note that this second, single-type approximation avoids any difficulties caused by the possibility of there being infinitely many types in the multitype approximation.

We now turn to the final outcome of a global epidemic. Let z be the probability that a randomly chosen initial susceptible is ultimately infected by the epidemic and, for $n = 1, 2, \dots$, let z_n be the same probability for a randomly chosen initial susceptible in a group of size n . The size-biased sampling implies that

$$(3.33) \quad z = \mu_h^{-1} \sum_{n=1}^{\infty} n z_n h_n.$$

Fix attention on a group of size n that did not contain any initial infectives. Arguing as in Section 3.4, the probability that a given susceptible in that group avoids infection from outside is $\pi = \exp(-\lambda_G z E[T_I])$, and the expected final size of the epidemic within that group is $\tilde{\mu}_{n,0}$. Thus, using (3.10),

$$(3.34) \quad n z_n = n - \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n-i} \pi^i \alpha_i, \quad n = 1, 2, \dots$$

Summing (3.34) over n and using (3.33) yields

$$(3.35) \quad z = 1 - \sum_{n=1}^{\infty} \mu_h^{-1} h_n \sum_{i=1}^n \frac{n!}{(n-i)!} q_i^{n-i} \pi^i \alpha_i,$$

which, since $\pi = \exp(-\lambda_G z E[T_I])$, is an implicit equation for z . Clearly, $z = 0$ is always a solution of (3.35) and similar arguments to those used in Section 3.4 show that there is a (unique) second solution in $(0, 1)$ if and only if $R_* > 1$. When $R_* > 1$, the root of (3.35) in $(0, 1)$ gives the expected proportion of initial susceptibles that are ultimately infected by a global epidemic. As in Section 3.5, the total spread of infection within a group not having initial infectives has the same distribution as in the extended model of Addy, Longini and Haber (1991), with $\pi = \exp(-\lambda_G z E[T_I])$.

3.6. Threshold parameter for the proliferation of infectious individuals. As noted in Section 1.3, Becker and Dietz (1995) consider a model for highly infectious diseases in which it is assumed that once one individual in a group (household) is infected then so is everyone else in that group. This assumption corresponds to setting $\lambda_L = \infty$ in our model. Becker and Dietz (1995) derive two threshold parameters for their model: one, which they call R_0 , for the proliferation of infectious individuals and another, which they call R_{H0} , for the proliferation of infected households. Clearly, R_{H0} is the same as our R_* and it is easily checked that setting $\lambda_L = \infty$ in (3.31) yields Becker and Dietz's formula for R_{H0} . We now determine a threshold parameter for the proliferation of infectious individuals for our model, which, to avoid confusion with our earlier notation, we denote by R_I .

Consider the branching process approximation of Section 3.3.1 and call an initial infective in a group a *primary* case and all subsequent infectives in that group *secondary* cases. Thus both primary and secondary cases can give rise to further primary cases, but only primary cases can give rise to secondary cases. We now consider a two-type branching process, in which types 1 and 2 denote primary and secondary cases, respectively. Then, using the size-biased sampling, it is easily seen that the mean matrix for the above two-type branching process is given, in the notation of Section 3.5, by

$$(3.36) \quad M = \begin{bmatrix} R_G & \mu - 1 \\ R_G & 0 \end{bmatrix}.$$

The threshold parameter R_I is given by the maximal eigenvalue of M and a simple calculation shows that

$$(3.37) \quad R_I = \frac{R_G}{2} \left(1 + \sqrt{1 + 4(\mu - 1)/R_G} \right).$$

Again, it is easily checked that setting $\lambda = \infty$ in (3.37) yields the corresponding expression for R_0 given in Becker and Dietz (1995). Note that we only had to find the maximal eigenvalue of a 2×2 matrix, whereas Becker and Dietz's approach required the maximal eigenvalue of an $N \times N$ matrix (where N denoted their largest group size), though, of course, there is a lot of structure in their $N \times N$ matrix, which is owing to our size-biased sampling. The same comment also applies to our derivation of R_* . Note also that $R_I = 1$ if and only if $R_* = 1$, so the critical values of $R_* = 1$ shown in Figure 1 also apply to $R_I = 1$.

For homogeneously mixing epidemic models it is well known that the critical fraction, v^* say, of susceptibles that have to be vaccinated to make a supercritical epidemic critical is

$$(3.38) \quad v^* = 1 - 1/R_0.$$

Becker and Dietz (1995) show that (3.38) still holds for their model if individuals are vaccinated independently with probability v^* and, in our notation, R_0 is replaced by R_I . However, this is *not* the case for our model. Equation

(3.38) holds for Becker and Dietz's model since vaccinating individuals independently with probability v reduces both R_G and $\mu - 1$ *pro rata* to $(1 - v)R_G$ and $(1 - v)(\mu - 1)$, respectively. The second of these reductions does not hold for our model since the mean size of a single population epidemic is not linear in the initial number of susceptibles.

In a later paper, Becker and Hall (1996) consider the household threshold parameter R_{H0} and associated vaccination strategies for the spread of an epidemic among a population of households made up of individuals of p different types, labelled $1, \dots, p$. The methodology of Section 3.5 can be extended to encompass this situation, using appropriate size-biased sampling. To determine the household threshold parameter R_{H0} , we consider a p -type branching process, where p is the number of types of individuals present in the population, whilst to determine the individual threshold parameter R_I we consider a $2p$ -type branching process, where again we distinguish between primary and secondary cases. In order to derive explicit expressions for the threshold parameters we need results for multitype epidemics analogous to those given for single population epidemics in Section 3.2. These can be found, for example, in Ball (1986), Picard and Lefèvre (1990) or Ball and Clancy (1993). The details are rather involved and will be published separately. Becker and Hall (1996) sidestep these complications by labelling households according to the type of epidemic that occurs in them (so that the type space, \mathcal{T} say, for households soon gets rather large) and then defining their model in terms of the mean number of type τ households generated by a type i individual, for $\tau \in \mathcal{T}$ and $i = 1, \dots, p$. They then get more explicit results by again making the very special assumption that once one individual in a household is infected, then the whole household becomes infected. Thus the extension of our methodology to the multitype setting is concerned with models that are described at a more basic level (i.e., in terms of individual infectious periods and individual-to-individual infection rates) than those considered by Becker and Hall (1996).

4. Embedding representations of the final size of the epidemic.

4.1. Embedding and the asymptotic distribution of final size. Showing that the epidemic process and its final size in a population can be constructed by sampling an appropriate embedding process at suitably defined stopping times may be of interest in itself since it yields an alternative way of constructing some aspects of the epidemic process, but it also turns out to be an efficient tool for studying the distribution of the final size and related quantities, in particular in asymptotic situations. In the present case, the construction and methods of Scalia-Tomba (1985, 1990) can be used, with minor modifications, to show that the final size of the epidemic, in a large population, is either small, with probabilities related to the approximating branching process, or large, with an approximately normal distribution around the mean expected from the deterministic approximation.

The general idea of the construction is to create a process describing the number of individuals in the population who would become infected, with

infection being considered as coming from outside the population, and then creating the epidemic within the population by letting the infectious individuals in the population define the amount of infection to which the remaining susceptibles will be exposed. The final size of the epidemic will then typically be characterized by a balance equation stating that the epidemic stops when the total “infection pressure” generated by those infected in the population (including initial infectives) becomes equal to the infection pressure needed to infect the same individuals. One then proceeds to show that the embedding process is asymptotically Gaussian and that the balance equation translates into a first-crossing problem for the embedding process. Finally, some further calculations are needed to clarify the “either small or large” character of the epidemic process.

4.2. The case of distributed infectious period and households. We will first carry out the construction for the situation considered in Section 3.1, in which the population is composed of a priori defined groups or households with given sizes.

4.2.1. The basic household process $(R(t), A(t))$. Let $(R(t), A(t))$ describe what has happened to a household of size n , with no initial infectives, after having been subjected to global infection pressure $t \geq 0$, that is, when it has been exposed to t time units of global infection. Here R is the number of household members having had the disease [in the terminology of Section 3.2, R is the final size when there are 0 initial infectives and $\pi = \exp(-\lambda_G t/N)$] and A is the cumulative sum of the infectious periods of these individuals. We assume that, as far as the process $(R(t), A(t))$ is concerned, local (within-household) infections are instantaneous. Thus R and A are constant except for a finite number of (simultaneous) jumps, corresponding to infection from outside of a susceptible individual; they are nondecreasing with $0 \leq R \leq n$ and $0 \leq A \leq \Sigma$, where Σ is the sum of n independent copies of T_I . The R and A components are strongly correlated (as T and T_A in Section 3.2) and may even be equal if T_I takes a constant value t_I (Reed–Frost case).

It may be useful to have a more concrete construction of (R, A) . Label the individuals in the household $1, 2, \dots, n$. For $k = 1, 2, \dots, n$, let individual k be endowed with random variables $(Q_G^{(k)}, Q_L^{(k)}, T_I^{(k)})$, where Q_G and Q_L are the thresholds for global and local infections, respectively. Thus Q_G is the total time units of global infection that has to be present before a given individual is globally infected, so Q_G follows a negative exponential distribution with rate λ_G/N . Similarly, Q_L is the total time units of local infection that has to be present before a given individual is locally infected, so Q_L follows a negative exponential distribution with rate λ_L . All these random variables, whether for the same or different individuals, are assumed to be independent.

To construct the associated realization of (R, A) , first, the n Q_G -values are marked on the t -axis. The first jump of (R, A) occurs at the smallest of these, that is, at the least amount of global infection necessary to infect an individual in our previously completely susceptible household. This infected individual will initiate an epidemic among the remaining $n - 1$ susceptibles that is deter-

mined by the values of $\{Q_T^{(k)}, T_I^{(k)}\}$ [see, e.g., the construction of Sellke (1983) as described in Ball (1986)]. Let T and T_A be, respectively, the total size and severity of this epidemic, where T includes the initial infective. Then the size of the first jump of (R, A) is (T, T_A) . Next, the $T - 1$ marks corresponding to the individuals who are no longer susceptible should be deleted from the t -axis. The next jump of the process will then occur at the smallest remaining mark, at which point one of the individuals not infected by the epidemic corresponding to the first mark will be globally infected. This individual will initiate an epidemic among the other remaining susceptibles. The total size and severity of this second epidemic is the size of the second jump of (R, A) and so on. This view of (R, A) yields, for instance, easy estimates of the increments of the process, since these depend on finding at least one of the original marks in the time interval considered.

The results in Section 3.2 [in particular (3.9) and (3.10)] are directly interpretable in terms of (R, A) . We have $E[R(t)] = \tilde{\mu}_{n,0}$, with $\pi = \exp(-\lambda_G t/N)$ [we will use the notation $\tilde{\mu}_{n,0}(\pi)$ in the sequel, to make the dependence on π explicit], and $E[A(t)] = E[T_I]E[R(t)]$, since Wald's identity for epidemics applies for any fixed number of initial infectives. Let us further, for $t, s \geq 0$, denote $\text{cov}(R(t), R(s))$ by $c_n^R(t, s)$, $\text{cov}(A(t), A(s))$ by $c_n^A(t, s)$ and $\text{cov}(R(t), A(s))$ by $c_n^B(t, s)$. For $t = s$, the covariances can be derived from (3.9). For $t \leq s$, say, one may, at least in theory, use the properties of the exponential distributions to derive the covariances: given $(t, R(t), A(t))$, $R(s) - R(t)$ and $A(s) - A(t)$ will have the same distribution as $(R(s-t), A(s-t))$ in a household starting with $n - R(t)$ susceptibles. However, it will be seen in the sequel that explicit determination of the covariance functions is not essential for the derivation of the main results; they will only be needed explicitly for the case $t = s$, for a particular choice of t .

4.2.2. Embedding the epidemic process. Assume that each household in the population has a process $(R_i(t), A_i(t))$, $i = 1, \dots, m$, of the type described above. Let $R_\bullet(t) = \sum R_i(t)$ and $A_\bullet(t) = \sum A_i(t)$. Assume also that an initial amount T_0 of infectious time is applied to the initially totally susceptible population. We can now define a sequence of stochastic times in which to consider $(R_\bullet(t), A_\bullet(t))$ (these correspond roughly to a description of the epidemic by cumulated generations, with anticipated local or within household infections):

$$\begin{aligned} T_0 &\rightarrow R_\bullet(T_0), A_\bullet(T_0), \\ T_1 &= T_0 + A_\bullet(T_0) \rightarrow R_\bullet(T_1), A_\bullet(T_1), \\ &\vdots \\ T_{k+1} &= T_0 + A_\bullet(T_k) \rightarrow R_\bullet(T_{k+1}), A_\bullet(T_{k+1}), \\ &\vdots \end{aligned}$$

Thus T_1 is the total amount of infection that has been generated in the population after the local household epidemics initiated by the initial T_0 units of infectious time have occurred. These T_1 units of infection may create further

global infections which may in turn give rise to further local infections, after which there will have been a total of T_2 units of infectious time generated in the population. The process continues until the additional infectious time created by a set of local infections is not enough to give rise to further global infections. Consequently, the above sequence stops at $T_\infty := \min\{t \geq 0: t = T_0 + A_\bullet(t)\}$ (see Figure 3). Then $R_\bullet(T_\infty)$ represents the final size of the epidemic in the population and $T_\infty = A_\bullet(T_\infty) + T_0$ represents its severity.

4.2.3. Asymptotic distribution of the embedding process and of the final size of the epidemic. For $n = 1, 2, \dots$, let m_n be the number of households of size n , m the total number of households, $N = \sum nm_n$ the total number of individuals in the population, $\theta_n = m_n/m$ the proportions of households of size n and $\bar{m}_1 = \sum n\theta_n < \infty$. Then $E[R_\bullet(t)] = \sum m_n \hat{\mu}_{n,0}(\exp(-\lambda_G t/N))$. In order to handle the bivariate character of (R_\bullet, A_\bullet) , let us use the Cramér–Wold device and define, for $(\alpha, \beta) \in \mathbb{R}^2$,

$$(4.1) \quad Z_m^{(\alpha, \beta)}(t) = \frac{1}{\sqrt{m}} \left(\alpha(R_\bullet(Nt) - E[R_\bullet(Nt)]) + \beta(A_\bullet(Nt) - E[A_\bullet(Nt)]) \right).$$

Proceeding as in Scalia-Tomba (1990), under the further condition that $\sum n^2\theta_n < \infty$, it can be shown that, as $m \rightarrow \infty$, $Z_m^{(\alpha, \beta)}$ converges in distribution, on $D[0, \infty)$ with the Skohorod topology, to a Gaussian process with mean 0 and covariance function

$$(4.2) \quad \gamma^{(\alpha, \beta)}(t, s) = \sum \theta_n (\alpha^2 c_n^R(t, s) + \alpha\beta(c_n^B(t, s) + c_n^B(s, t)) + \beta^2 c_n^A(t, s)),$$

where π , at times t and s , now equals $\exp(-\lambda_G t)$ and $\exp(-\lambda_G s)$, respectively.

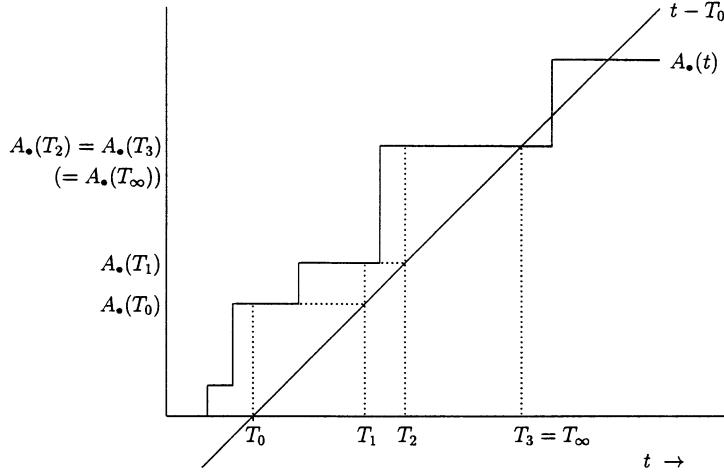


FIG. 3. Determination of the severity of the epidemic from the embedded process $A_\bullet(t)$.

Now let $r(t) = \sum \theta_n \tilde{\mu}_{n,0}(\exp(-\lambda_G t))$, $a(t) = E[T_1]r(t)$, $\tilde{r}(t) = r(t)/\tilde{m}_1$ and $\tilde{a}(t) = a(t)/\tilde{m}_1$. Then, letting $(\alpha, \beta) = (0, 1)$, we have that, as $m \rightarrow \infty$,

$$(4.3) \quad \tilde{A}_m(t) = \frac{1}{\sqrt{m}}(A_\bullet(Nt) - ma(t))$$

converges weakly to a Gaussian process with mean 0 and covariance function $\gamma^{(0,1)}$. Assume now that $T_0/N \rightarrow \mu_0 > 0$ as $m \rightarrow \infty$. Then

$$(4.4) \quad \frac{T_\infty}{N} = \min \left\{ t: t = \frac{T_0}{N} + \tilde{a}(t) + \frac{\sqrt{m}}{N} \tilde{A}_m(t) \right\},$$

and, since $(\sqrt{m}/N)\tilde{A}_m(t)$ converges uniformly to 0 on any compact subset of $[0, \infty)$, we have that $T_\infty/N \rightarrow \tau(\mu_0) := \min\{t: t = \mu_0 + \tilde{a}(t)\}$. We may then conclude that $Z_m^{(\alpha, \beta)}(T_\infty/N)$ converges in distribution to $Z_m^{(\alpha, \beta)}(\tau(\mu_0))$ for all (α, β) , which means that the vector

$$(4.5) \quad \frac{N}{\sqrt{m}} \left(\frac{R_\bullet(T_\infty)}{N} - \tilde{r}\left(\frac{T_\infty}{N}\right), \frac{A_\bullet(T_\infty)}{N} - \tilde{a}\left(\frac{T_\infty}{N}\right) \right)$$

converges to a bivariate normal distribution with mean 0 and covariance matrix $M(\mu_0)$ with elements $M_{11} = \sum \theta_n c_n^R(\tau(\mu_0), \tau(\mu_0); \pi = \exp(-\lambda_G \tau(\mu_0)))$, and $M_{12} = M_{21}$ and M_{22} of similar form, with c_n^B and c_n^A replacing c_n^R . By using the identities satisfied by T_∞ and by $\tau(\mu_0)$ [see Scalia-Tomba (1990)], this result can be recast into the convergence in distribution of the vector

$$(4.6) \quad \sqrt{m} \left(\frac{R_\bullet(T_\infty)}{N} - \tilde{r}(\tau(\mu_0)), \frac{A_\bullet(T_\infty)}{N} - \tilde{a}(\tau(\mu_0)) \right)$$

to a bivariate normal distribution with mean 0 and covariance matrix $\tilde{m}_1^{-2}(1 - \tilde{a}'(\tau(\mu_0)))^{-2} A M(\mu_0) A^T$, where $A_{11} = 1 - \tilde{a}'(\tau(\mu_0))$, $A_{12} = \tilde{r}'(\tau(\mu_0))$, $A_{21} = 0$ and $A_{22} = 1$.

This is the basic result on asymptotic normality of the final size of the epidemic, around the value predicted by ‘‘deterministic’’ considerations, when the epidemic is started by a large amount of initial infection [some algebra will show that the definitions of $\tau(\mu_0)$ and, consequently, of $\tilde{r}(\tau(\mu_0))$ and $\tilde{a}(\tau(\mu_0))$ agree with (3.35), when $\mu_0 = 0$].

The most interesting case to study is, however, when T_0 remains fixed as $m \rightarrow \infty$, corresponding to few initial infectives. One must then combine the branching process approximations of Sections 3.3.1 and 3.5 with the asymptotic normality results shown above. Once again, the strategy in Scalia-Tomba (1985, 1990), of studying the final size distribution in different ranges of values, may be followed. Let us, for simplicity, denote the final size of the epidemic in a population with m households by T_m and assume that the epidemic is above threshold (otherwise, the results in Sections 3.3.1 and 3.5 account for the whole asymptotic distribution). The branching process approximations of Sections 3.3.1 and 3.5 then show that $\Pr\{T_m = k\} \rightarrow p(k)$, for $k = 0, 1, \dots$, where $p(\cdot)$ is the distribution of the total size in the approximating branching process. This distribution has total mass $\tau < 1$, say, corresponding to the event

of extinction of the approximating branching process. It therefore remains to show that the remaining probability mass $1 - \tau$ is concentrated around the deterministic solution for a large epidemic [see (3.35); the solution corresponds to $\rho = \tilde{r}(\tau(0))$, with the convention that the non-zero solution should be taken when $\mu_0 = 0$]. One then starts by studying $\Pr\{k < T_m < a_m\}$, where $\{a_m\}$ satisfies $a_m \rightarrow \infty$ but $a_m/m \rightarrow 0$ as $m \rightarrow \infty$, with the aim of showing that

$$(4.7) \quad \lim_{k \rightarrow \infty} \lim_{m \rightarrow \infty} \Pr\{k < T_m < a_m\} = 0.$$

The coupling construction of an approximating branching process by Ball and Donnelly (1995) (see Section 3.3.1), combined with the introduction of a lower bounding branching process [cf. Whittle (1955), Ball and Clancy (1992) and Andersson (1993)], can be used for this purpose. The approximating branching process $B_U(t)$, say, is always larger than the infectives process $I_m(t)$, since every contact is considered as a new individual in B_U , but some contacts do not yield new infectives in I_m , since contacts may occur with already infected or removed individuals. This mechanism amounts to a thinning of the branching process, with thinning probabilities depending on the total progeny up to the time point considered for the contact. To make things simpler, one can therefore apply thinning to each contact in the branching process, with fixed probability $\varepsilon > 0$, which will overestimate the “true” thinning probabilities as long as the total number of individuals ever having been infected is less than εN , thus constructing a second branching process $B_L(t)$, for which we will have $B_L(t) \leq I_m(t) \leq B_U(t)$, at least as long as the total epidemic is less than εN . If we denote the distribution functions of final size (total progeny) by F_L , F_m , and F_U , respectively, we will then have $F_U(i) \leq F_m(i) \leq F_L(i)$, for all $i \leq \varepsilon N$. Thus, $0 \leq F_m(a_m) - F_m(k) \leq F_L(a_m) - F_U(k) \leq \tau(\varepsilon) - F_U(k)$, where $\tau(\varepsilon)$ is the extinction probability in the ε -thinned process. However, since $\tau(\varepsilon) \rightarrow \tau$ as $\varepsilon \rightarrow 0$ and $F_U(k) \rightarrow \tau$ as $k \rightarrow \infty$, one obtains the desired result.

The remaining range, as long as $\{a_m\}$ is taken so that $a_m/\sqrt{m} \rightarrow \infty$, can be studied using the Gaussian process approximation [see Scalia-Tomba (1985)] to show that the crossing condition, equivalent to achieving the final size, can only be fulfilled in a $O(\sqrt{m})$ neighborhood of the “deterministic” value. Having thus accounted for the whole asymptotic probability mass, one now proceeds by showing that $\Pr\{(T_m - \rho N)/\sqrt{N} \in K\} = \Pr\{(T_m - \rho N)/\sqrt{N} \in K \mid T_m > a_m\}\Pr\{T_m > a_m\} \approx \Pr\{(T_m - \rho N)/\sqrt{N} \in K \mid T_m > a_m\}(1 - \tau)$, for large N and $K \subset \mathbb{R}$ bounded. The final step consists in showing that the epidemic process, conditioned on $T_m > a_m$, that is, on having a large epidemic, again follows the Gaussian approximation derived above, now with $\mu_0 = 0$. However, the conditioning event involves members from at most a_m households and times, used as arguments in the $Z_m^{(\alpha, \beta)}$ process, of order $O(a_m/m)$. The effect on the (conditional) limit law of $Z_m^{(\alpha, \beta)}$ will be vanishingly small and the limit law will be unchanged, at least as long as the removal of any set of a_m households from the total set of m households does not affect asymptotic proportions or means. This last requirement is equivalent to the uniform in-

tegrability of the sequence of household size proportion distributions, indexed by m .

4.3. The case of fixed infection probabilities. In the case studied in Sections 2.1 and 2.2, in which the infection probabilities are fixed and independent, it is possible to construct an embedding process directly based on the clumps formed by the local infection process. Let $\{C_k\}$ denote the total number of local components of size $k = 1, \dots, N$ that have been formed by local infection, in a population of size N . Closely following Scalia-Tomba (1985, 1990), we now construct an epidemic between components, with susceptibility and infectivity proportional to size. To each component of size k , we attach a threshold variable with geometric distribution with “success probability” = $1 - (1 - p_G)^k$, representing the number of individual infection attempts necessary to infect the component. We denote these variables by $\{Q_{kj}\}$, $1 \leq j \leq C_k$, $k = 1, \dots, N$. We now define processes $X_{kj}(t) = 1_{\{Q_{kj} \leq t\}}$ and $X_k(t) = \sum_j X_{kj}(t)$, which represent the numbers of k -components that have been infected after t infection attempts on the population. Finally, we define $X(t) = \sum_k kX_k(t)$, the total number of individuals infected after t infection attempts. We now construct the generations of the epidemic process by considering $X(t)$ at suitably defined random times. Assuming that the epidemic is started by m_0 initial infectives global to the population, we set $T_1 = m_0$, $T_2 = m_0 + X(T_1)$ and, in general, $T_{k+1} = m_0 + X(T_k)$. These times form an increasing sequence which stops at $T_\infty = \min\{t: t = m_0 + X(t)\}$. The final size of the epidemic in the population is then $X(T_\infty)$.

We would now want to consider the asymptotic situation $N \rightarrow \infty$, $p_G \approx \lambda_G/N$, local infection probabilities fixed and m_0 either fixed or increasing with N . Except for the randomness of $\{C_k\}$, the problem is similar to the situations studied in Scalia-Tomba (1985, 1990). It can therefore be expected that similar results will be valid, modified only by the additional randomness generated by $\{C_k\}$. However, in models like the great circle (Section 2.1) or the epidemic among giants (Section 2.3), there will potentially be an infinite number of types (sizes) of local components. Work is in progress on how best to resolve the technical problems arising in these situations; the results will be published separately.

5. Applications.

5.1. Estimating R_* from household total size data. A number of authors [e.g., Longini and Koopman (1982), Becker (1989) and Addy, Longini and Haber (1991)] have previously considered household infection data, using a secondary attack rate that corresponds to our λ_L . They do not explicitly model the spread of the infection through the wider population, simply assuming that each individual is exposed to the same probability of external infection (via a primary attack rate).

The present paper provides a framework for modelling the spread through the whole population that is consistent with these previous household models

(see Section 3.4). We can thus estimate parameters of the internal process similarly to those authors and additionally relate the probability of external infection to our global infection rate parameter λ_G , and thus estimate the overall group-to-group reproductive ratio R_* .

In this subsection we describe a method for estimating the threshold parameter R_* when the available data are the total number of individuals in each group that are ultimately infected by the epidemic. We shall assume that the number of groups m is large, that a global epidemic has occurred and that the distribution of the infectious period is known. Our method is based on Addy, Longini and Haber (1991); see Becker (1989) for alternative approaches.

Consider first the extended model of Addy, Longini and Haber (1991) and suppose that the total size of such epidemics in a number of independent groups is known. Let λ_L denote the individual-to-individual local infection rate and let λ_G denote the individual-to-population global infection rate as defined in Section 3.1. Addy, Longini and Haber (1991) give an algorithm for obtaining maximum likelihood estimates of the local infection rate λ_L and the probability π that a random individual escapes infection. [The likelihood is straightforward to compute numerically using (3.12).] In our situation the epidemic total sizes in different groups are not mutually independent, but if the number of groups is large, the total sizes will be approximately independent in the event of a global epidemic. Thus estimates for λ_L and π can be derived using the method of Addy, Longini and Haber (1991). An estimate for z can then be obtained using (3.35), allowing λ_G to be estimated from the equation $\pi = \exp(-\lambda_G z E[T_I])$. An estimate for R_* can then be obtained from (3.31).

As a simple example, we consider data on the spread of an influenza epidemic in Tecumseh, Michigan, analyzed in Addy, Longini and Haber (1991). The data do not exactly fit our situation since (a) they are combined data over two separate epidemics, (b) only 10% of households are included, (c) households of more than five individuals are omitted and (d) it is likely that some susceptibles were in fact immune or at least highly resistant to the strain involved in the outbreak (Klaus Dietz and Jim Koopman, personal communication), which will lead to underestimation of R_* [see Dietz (1993)]. The present analysis should therefore be viewed as simply illustrative of our methodology. The data are shown in Table 1.

Addy, Longini and Haber (1991) considered two possible distributions for the infectious period, namely, $T_I \equiv 4.1$ days and T_I follows a gamma distribution with probability density function $f(t) = c^2 t \exp(-ct)$, $t > 0$, where $c = 2/4.1 \approx 0.49$. For the model with a constant infectious period, Addy, Longini and Haber (1991) obtained the estimates $\hat{\lambda}_L = 0.0423$ and $\hat{\pi} = 0.8677$, from which $\hat{p}_L = 0.1592$, $\hat{\lambda}_G = 0.1950$ and $\hat{z} = 0.1775$. From these we can calculate $\hat{R}_G = 0.7995$, $\hat{\mu} = 1.4145$ and hence $\hat{R}_* = \mu R_G = 1.1309$. For the model with a gamma distributed infectious period, Addy, Longini and Haber (1991) obtained the estimates $\hat{\lambda}_L = 0.0446$ and $\hat{\pi} = 0.8674$, from which $\hat{p}_L = 0.1605$, $\hat{\lambda}_G = 0.1955$ and $\hat{z} = 0.1775$; from which $\hat{R}_G = 0.8006$, $\hat{\mu} = 1.4102$ and $\hat{R}_* = 1.1303$. Note that the two models give very similar estimates of R_* .

TABLE 1
Observed distribution of influenza A(H3N2) infections in 1977–1978 and 1980–1981 combined epidemics in Tecumseh, Michigan

No. infected	No. of susceptibles* per household				
	1	2	3	4	5
0	110	149	72	60	13
1	23	27	23	20	9
2		13	6	16	5
3			7	8	2
4				2	1
5					1
Total	133	189	108	106	31

*The criterion for classifying individuals as susceptible is a pre-season hemagglutination inhibition test detecting no antibody in a dilution of 1 in 128 or less. Households with more than five susceptibles are deleted from all analyses. [From Addy, Longini and Haber (1991).]

and other parameters. Also, the observed proportion of initial susceptibles ultimately infected by the disease is $250/1414 = 0.1768$, which is in close agreement with the estimate $\hat{z} = 0.1775$ fitted from both models.

The moral appears to be that the data are inadequate to discriminate between rival models for within-group contacts, but the parameters that determine the overall spread of the epidemic— μ and R_G , and hence $R_* = \mu R_G$ —are not sensitive to this inadequacy. Similar conclusions can be drawn from the comparison by Islam, O'Shaughnessy and Smith (1995) of their fit of the somewhat extreme inverse Gertsbakh model (see Section 5.2.2) with the results of Haber, Longini and Cotsonis (1988) for some similar sets of household epidemic data.

The method of Addy, Longini and Haber (1991) also yields approximate confidence sets for (λ_L, π) . Thus an approximate confidence interval for R_* could be obtained, since R_* is a function of λ_L and π . The above method will always yield an estimate of R_* that is larger than 1. This is because (3.35) is essentially deterministic, and in a deterministic model an initial trace of infection can only lead to a nonzero proportion of the population ultimately being infected if the model is above threshold. Thus the above method of estimating R_* should only be used if there is a good reason to believe that a global epidemic has occurred.

5.2. Vaccination: the equalizing strategy. The fundamental aim of a vaccination program must be to reduce the basic reproductive ratio R_* to below unity. In Section 2.5 we examined the implications of this for a simple example of large groups of equal sizes. Now, having seen (Section 3) that the relation

$R_* = \mu R_G$ holds for a wider set of models, we return to examine the question of optimal vaccination strategies in more generality. We shall compare different strategies that vaccinate a fixed proportion v of the population.

As noted in Section 2.5, any vaccination strategy will reduce R_G simply *pro rata*, so the difference between strategies will lie in how they affect the (size-biased) mean component size μ . In practice, of course, we do not know the component sizes at the time of vaccination, and even for small group sizes the evaluation of the distribution and mean of the component size in general requires quite complicated iterative calculation (see Section 3.2).

When we consider more general within-group distributions, one straightforward strategy that suggests itself is to leave the numbers of susceptibles in each group as nearly equal as possible. We call this the *equalizing strategy*, and conjecture that it is optimal for the groups model for any infectious period distribution.

5.2.1. *The equalizing strategy for groups or households.* It is easy to see that the equalizing strategy will be optimal if and only if, for all n , two groups of n susceptibles contribute less to μ than a pair of groups of sizes $n - 1$ and $n + 1$. Because the probability of a global infection hitting a group with n susceptibles is proportional to n , this condition is equivalent to the sequence $(n\mu_n)$ being convex; here $\mu_n = \mu_{1, n-1}$ is the mean size of an outbreak in a group of n susceptibles which is started by just one of them becoming infected. The condition that $(n\mu_n)$ be convex is in turn equivalent to the requirement that the second difference $D_n \equiv n\mu_n - 2(n - 1)\mu_{n-1} + (n - 2)\mu_{n-2}$ be greater than or equal to 0 for all n .

It is easily shown that the equalizing strategy is optimal for the simple “all-or-none case” where within-group outbreaks are either of size 1, with probability q_1 , or of size n (this can be thought of as arising when the infectious period is either of length 0 or ∞ , with respective probabilities q_1 and $1 - q_1$). For this case, $\mu_n = q_1 + n(1 - q_1)$, whence $D_n = 2 - 2q_1 \geq 0$.

Note that the model of Becker and Dietz (1995) is the even more special “all” case ($q_1 = 0$, and therefore $D_n = 2$), where within-group infectivity is so high that no one escapes. [The details of their calculations are a little complicated because they work in terms of the non-size-biased distribution $\{h_k\}$ rather than $\{\pi_k\}$ (see Section 2.3), but it is not difficult to check that their conclusions are consistent with the optimal strategy being to minimize μ , and that this is achieved by using the equalizing strategy.]

The all-or-none case is one in which the within-group infections by an individual are maximally correlated. The other extreme, for models with a general infectious period, is the independent links or Reed–Frost case considered in Section 2. For this, we have calculated D_n for $n = 2, \dots, 15$ using Maple (see Figure 4), and the conjecture appears to hold for all these values, with a pattern suggesting that it is likely to hold for all n .

Indeed, on the basis of this, and a similar plot for the epidemic with exponentially distributed infectious period (for $n = 2, \dots, 8$), we conjecture that in fact $D_n \geq 2 - 2q_1$ for all n , so that the simple all-or-none case is the lower bound.

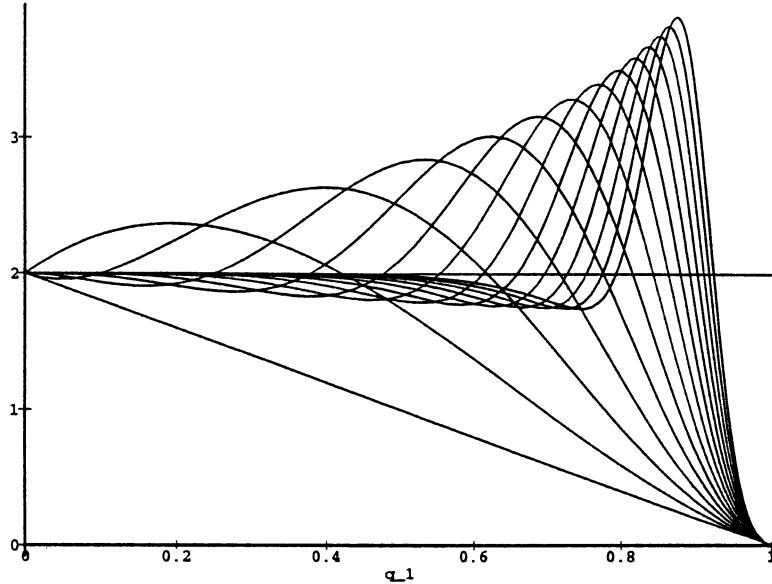


FIG. 4. The second difference D_n , conjectured to be greater than or equal to $2 - 2q_1$, plotted against q_1 for $n = 2, \dots, 15$ for the independent links (Reed-Frost) case.

Our calculations for the independent links and exponential infectious period models used the techniques of Section 3.2, particularly (3.6). For small values of n , we can work out explicit expressions for D_n for a general infectious period in terms of the sequence (q_k) , where q_k is as before [see following (3.6)] the probability that an infective will infect none of a set of k susceptibles in the same group. Let $d_k \equiv q_k - q_{k+1}$ be greater than or equal to 0 because (q_k) is a monotone nonincreasing sequence. We find $D_2 = 2d_0$, $D_3 = 2d_0 + 6d_0d_1$ and $D_4 = 2d_0 + 12d_0(q_1d_1 + d_0d_2 + 2d_1d_2)$, in each case greater than or equal to our conjectured minimum of $2d_0$.

In the remainder of this subsection, we prove that the equalizing strategy is optimal for a variant of the groups model in which individual contacts are *negatively correlated* (Section 5.2.2), for the case of large groups (Section 5.2.3), and, in Section 5.2.4, for the great circle model introduced in Section 2.3.

Note that for the different class of models where global infections choose groups with equal probabilities, instead of individuals with equal probabilities, it is easy to show that the equalizing strategy is *not* optimal. For this case, the appropriate second difference is $D_n = \mu_n + \mu_{n-1} - 2\mu_{n-2}$, and it is quite easy to show that $D_2 = 1 - 2q_1$, which is less than 0 for $q_1 > 1/2$, and that, in general, $D_n = -1/n < 0$ at $q_1 = 1$.

5.2.2. The equalizing strategy for the Gertsbakh model and its inverse. Gertsbakh (1977) introduced a model in which each individual makes exactly one (potentially infectious) contact. This does not seem realistic in the context of epidemics, but is of some theoretical interest in that it provides an example

with negative correlation between the contacts made by an individual, which is not possible for our basic general infectious period model as defined in Section 3.1. Gertsbakh (1977) also considered the inverse of this model, in which there is exactly one contact to each individual, and this inverse model has recently been fitted to data by Islam, O'Shaughnessy and Smith (1995). Both models can be generalized, replacing “exactly one contact” by “one contact with probability $(n - 1)p$, otherwise none” [we choose this parameterization so that p is the probability of contacting any one specific individual; $p \leq 1/(n - 1)$].

For the inverse model, with one initial infective in a group of size n , the probability that the total size of the group epidemic will be k is

$$(5.1) \quad p'_k = \binom{n-1}{k-1} p(kp)^{k-2}(1-kp)^{n-k} \quad (1 \leq k \leq n)$$

[Islam, O'Shaughnessy and Smith (1995)]. It is not easy to calculate the mean outbreak size μ_n from (5.1). However, it follows from a simple and general result (see Appendix) that μ_n is the same as for the forward model, for which calculations turn out to be easier. We can write down the corresponding probability that the total size of the group epidemic will be k for the forward model,

$$(5.2) \quad p_k = \frac{(n-1)!}{(n-k)!} (p^{k-1} - (n-k)p^k) \quad (1 \leq k \leq n).$$

From (5.2) it is straightforward to show that (for both the Gertsbakh model and its inverse)

$$(5.3) \quad \mu_n = \sum_{k=0}^{n-1} \frac{(n-1)!}{(n-k-1)!} p^k.$$

From this, matching powers of p in $D_n \equiv n\mu_n + (n-2)\mu_{n-2} - 2(n-1)\mu_{n-1}$, we have that

$$(5.4) \quad D_n = \sum_{k=1}^{n-1} \frac{(n-2)!}{(n-k-1)!} k(k+1)p^k.$$

This is a sum of nonnegative terms and is therefore greater than or equal to its first term, $2p = 2 - 2q_1$, so our conjecture holds for both the Gertsbakh model and its inverse.

5.2.3. The equalizing strategy for large groups. We next consider the vaccination problem in the limiting case where groups are large (see Sections 2.4 and 2.5). Let $r_i = (n_i - 1)p_L \approx n_i p_L$ denote the local reproductive ratio in group i , whose size is n_i , and let $R_* = R_G\mu$ be the overall reproductive ratio as before. Then we can distinguish three cases:

1. *Safe*: $R_* < 1$ and $r_i < 1$ for all groups; only minor local outbreaks can occur.
2. *Restricted*: $R_* < 1$, but $r_i > 1$ for some groups; the overall (between-group) epidemic is below threshold, but an outbreak can affect significant proportions [$O(n_i)$] within a small number of groups.

3. *Generalized: $R_* > 1$:* a global outbreak affecting a significant proportion of the whole population [$O(N)$] can occur.

This classification follows that of Watson (1972) except that we have changed the name for case (1): Watson called this localized, which given our use of “local” in this paper would more appropriately be applied to case (2). We could further subdivide (3) according to whether $r_i > 1$ or $r_i < 1$ for most groups: this determines whether the outbreaks in each group are driven by local or global contacts.

In case (2), it is arguable whether we need to vaccinate. If we wish to, then the requirement is to reduce each group below its local threshold, which requires that we vaccinate a proportion $v_i \geq 1 - 1/r_i$ in group i . Note that in case (2), $R_* = R_G\mu < 1$ although some groups are above their local threshold and therefore $\mu \gg 1$; thus we must have $R_G = o(1)$ in this case.

We can also have $R_G = o(1)$ in case (3), provided that $r_i > 1$ for some groups. This is the interesting case already referred to in Sections 2.4 and 3.3.2, where the large amplification from local infections in such groups puts the population as a whole above threshold although the global infection probability p_G is extremely small. In this case, it is easy to see that the equalizing strategy will, as we increase the proportion vaccinated v , take us to case (2) before we reach the fully satisfactory situation of case (1).

To keep matters simple, we shall make the stronger requirement that our vaccination strategy must take us to case (1), where not even a restricted outbreak can occur. We can then prove that the equalizing strategy is optimal, as follows.

We assume that we have m groups of sizes n_i , $i = 1, \dots, m$, giving a total population $N = \sum_i n_i$. We consider the asymptotic regime where the sizes tend to infinity with the relative sizes fixed, that is, $n_i = f_i N$, where $\{f_i\}$ is a fixed distribution, and we scale p_G and p_L so that the global reproductive ratio $R_G = N p_G$ and mean within-group reproductive ratio $\bar{R}_L = N p_L/m$ (note that for once this is not a size-biased mean) remain fixed. We could also let the number of groups $m \rightarrow \infty$, provided we keep the distribution of group sizes fixed.

As usual, we consider a strategy that vaccinates a proportion v of the population, asking what distribution of the vaccine among groups will achieve $R_* < 1$ for minimal v . We shall use primes to denote modified values of parameters after vaccination, including s'_i to denote the number of susceptibles in group i after vaccination.

Since we require $r'_i < 1$ for all i , we can use the branching process approximation for within-group contacts to give us $\mu'_i = 1/(1 - r'_i)$. Hence μ' , the size-biased mean of the μ'_i 's, is $\sum_i \mu'_i (s'_i / \sum_j s'_j)$. Now $r'_i = s'_i p_L$ and $R'_G = p_G \sum_j s'_j$, whence $R'_* = R'_G \mu' = c \sum_i \mu'_i r'_i = c \sum_i r'_i / (1 - r'_i)$, where $c = p_G/p_L$. Also, $\sum_i r'_i = p_L \sum_i s'_i = (1 - v)N p_L$, that is, $\sum r'_i \propto 1 - v$.

Thus our vaccination problem is equivalent to maximizing $\sum r'_i$ subject to keeping $\sum r'_i / (1 - r'_i)$ fixed and subject to the constraints $r'_i \leq r_i$ (i.e., $s'_i \leq n_i$; violating this constraint would require a negative number of vaccinations in

group i). Since $f(x) = x/(1-x)$ is convex over the range of interest $(0,1)$, the solution to this is to take all the r'_i 's equal, as far as possible, and since $s'_i \propto r'_i$, this is the equalizing strategy. By “as far as possible” we mean that if the constraints $s'_i \leq n_i$ prevent us from making the s'_i 's equal, we leave $s'_i = n_i$ in all groups up to a certain size n_v and take $s'_i = n_v$ wherever $n_i > n_v$.

To find the value of v required for the straightforward case where we can make the s'_i 's equal, we note that their value will then be $(1-v)N/m$ and hence the r'_i 's will all equal $(1-v)\bar{R}_L$. Also $R'_G = (1-v)R_G$ and so $R'_* = R'_o\mu' = (1-v)R_G/(1-(1-v)\bar{R}_L)$, whence the condition $R'_* < 1$ yields

$$(5.5) \quad v > 1 - 1/R_0,$$

where $R_0 \equiv R_G + \bar{R}_L$, the mean reproductive ratio for an individual from a randomly chosen household.

In conclusion, we note that our analysis only required consideration of the branching process regime at the start of a potential outbreak. One consequence of this [see, e.g., Mollison (1995), Section 2.3] is that our results will hold for the deterministic version of our large group model; this gives an indirect way of recovering the deterministic result of May and Anderson (1984) [see also Hethcote and Van Ark (1986)].

5.2.4. The equalizing strategy for the great circle model. We conclude by proving that the equalizing strategy is optimal for the great circle model: that is, the optimal policy is to spread the vaccinations around the circle as evenly as possible.

We consider then a population of N individuals spaced equally around a circle. If we vaccinate a fixed number m of individuals, so that $v = m/N$, then the (non-size-biased) mean length of the intervals of susceptibles between these will be $\tau \equiv (N-m)/m = (1/v) - 1$. Now consider choosing a susceptible at random and then looking at the numbers of susceptibles T_+ and T_- , respectively, to its right and left between it and the next vaccinated individual. Let $T = 1 + T_+ + T_-$. Then $\mathbf{E}[T]$ is the size-biased mean for a group (interval) of susceptibles.

Consider first the case where local contacts always infect, so that $\mathbf{E}[T]$ will also be the mean clump size for the epidemic. That $\mathbf{E}[T]$ is minimal when T is as near constant as possible, that is, it is $= \tau$ when τ is an integer, and has a distribution concentrated on the two integers either side of τ otherwise, is a well-known result for renewal processes: it can be thought of as saying that waiting times for buses will be minimal if they are scheduled at equal intervals.

What we prove is a generalization of this result that takes account of our actual local infection process. It turns out that we can allow a more general local infection process than the basic great circle model in which infections by different individuals are independent.

Thus, secondly, consider a model in which the local outbreak caused by an individual i , in the absence of vaccination, has an arbitrary distribution on intervals containing that individual. Suppose that it consists of C_+ individuals

to the right of i and C_- individuals to the left, so that its total size is $C = 1 + C_+ + C_-$. Let $p_r = P(C_+ \geq r)$.

We are now ready to put the vaccination process and the infection process together; we need of course to assume that these are independent. When we include the information on vaccinated intervals, the local outbreak caused by i becomes $D = 1 + R + L$, where $R = \min(T_+, C_+)$ and $L = \min(T_-, C_-)$. Our target is to find the distribution of T that minimizes $\mu = \mathbf{E}[D]$.

Now

$$\mathbf{E}[R] = \sum_{r=1}^{\infty} P(R \geq r) = \sum_{r=1}^{\infty} P(C_+ \geq r)P(T_+ \geq r) = \sum_{r=1}^{\infty} p_r P(T_+ \geq r).$$

Next comes the crucial step: $P(T_+ < r) = \sum_{j=1}^r P(T_+ = j-1)$. But, $T_+ = j-1$ only if the individual at j (i.e., j steps to the right of individual i) is vaccinated, and this has probability $1/\tau = m/(N-m)$. Hence $P(T_+ = j-1) \leq 1/\tau$ and therefore $P(T_+ \geq r) \geq 1 - \sum_{j=1}^r 1/\tau = 1 - r/\tau$. Of course we also have $P(T_+ \geq r) \geq 0$, so that $\mathbf{E}[R] = \sum_{r=1}^{\infty} p_r P(T_+ \geq r)$ will be minimized by taking $P(T_+ \geq r) = \max(1 - r/\tau, 0)$. A mirror argument for $\mathbf{E}[L]$ leads to the corresponding condition $P(T_- \geq r) = \max(1 - r/\tau, 0)$.

To see that these minima are uniquely attained when the distribution of T is concentrated on $[\tau]$ (the integer part of τ) and $[\tau] + 1$, note that equality in the argument of the last paragraph [turning “only if” into “if and only if” and hence giving $P(T_+ = j-1) = 1/\tau$] requires that it is impossible to have two vaccinated individuals within the range $j = 0$ to $[\tau] - 1$; hence $T \geq [\tau]$. And $T \leq [\tau] + 1$ because otherwise it would be possible for T_+ (and T_-) to be $= [\tau] + 1$, which would contradict $P(T_+ \geq r) = \max(1 - r/\tau, 0)$.

We have thus shown that the equalizing vaccination strategy is optimal for the generalized great circle model, in which the local outbreak caused by an individual takes an arbitrary distribution on the intervals containing the individual.

APPENDIX

Mean size of inverse epidemics. We prove here a general lemma, used in Section 5.2.1, which tells us that the mean size of an epidemic started by a randomly chosen individual is the same as for the model where we invert the contact structure.

LEMMA 1. *Consider any directed graph Γ on a finite number of nodes N . Let $\mu(\Gamma)$ be the mean size of the set of points that can be reached from a randomly chosen initial node of Γ by following links of the graph. Let Γ' denote the inverse of Γ , in which the direction of every link is reversed. Then $\mu(\Gamma') = \mu(\Gamma)$.*

REMARK. Γ can either be a specific given graph or, as in our epidemic models here, a random graph, with any desired distribution as long as it is (a.s.) finite.

PROOF OF LEMMA 1. Suppose first that Γ is a specific given graph. For any pair of nodes i, j of Γ , let $I_{ij} = 1$ if j can be reached from i , = 0 otherwise. Then the size of the set of points that can be reached from i is $\sum_j I_{ij}$ and hence the mean size if i is chosen randomly, that is, from the uniform distribution over the set of nodes, is $\mu(\Gamma) = (1/N) \sum_i \sum_j I_{ij}$.

Similarly, for the inverse graph, $\mu(\Gamma') = (1/N) \sum_j \sum_i I'_{ji}$, but i can be reached from j in Γ' if and only if j can be reached from i in Γ . Therefore $I'_{ji} = I_{ij}$ for all i and j , whence $\mu(\Gamma') = \mu(\Gamma)$.

Finally, if Γ is a random graph, we simply take expectations of $\mu(\Gamma)$ and $\mu(\Gamma')$ with respect to its distribution. \square

Acknowledgments. We are grateful to Klaus Dietz and Niels Becker for stimulating discussions, for comments on an earlier draft of this paper and for showing us preprints of their related work (see Sections 1.3 and 3.6), and to the Isaac Newton Institute for Mathematical Sciences in Cambridge, during whose Epidemic Models program in 1993 this work began.

REFERENCES

- ADDY, C. L., LONGINI, I. M. and HABER, M. (1991). A generalized stochastic model for the analysis of infectious disease final size data. *Biometrics* **47** 961–974.
- ANDERSSON, M. (1993). The final size of a multitype chain-binomial epidemic process. M.Sc. dissertation, Technical Report 1993:31, Dept. Mathematics, Chalmers Univ. Technology, Göteborg, Sweden.
- BAILEY, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and Its Applications*. Griffin, London.
- BALL, F. G. (1983a). A threshold theorem for the Reed–Frost chain-binomial epidemic. *J. Appl. Probab.* **20** 153–157.
- BALL, F. G. (1983b). The threshold behavior of epidemic models. *J. Appl. Probab.* **20** 227–241.
- BALL, F. G. (1986). A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. *Adv. in Appl. Probab.* **18** 289–310.
- BALL, F. G. (1996). Threshold behaviour in stochastic epidemics among households. *Applied Probability 1. Lecture Notes in Statist.* **114** 253–266. Springer, Berlin.
- BALL, F. G. and CLANCY, D. (1992). The final outcome of a generalised stochastic multitype epidemic model. Technical Report 92-4, Nottingham Statistics Group.
- BALL, F. G. and CLANCY, D. (1993). The final size and severity of a generalised stochastic multitype epidemic model. *Adv. in Appl. Probab.* **25** 721–736.
- BALL, F. G. and DONNELLY, P. J. (1995). Strong approximations for epidemic models. *Stochastic Process Appl.* **55** 1–21.
- BALL, F. G. and NÅSELL, I. (1994). The shape of the size distribution of an epidemic in a finite population. *Math. Biosci.* **123** 167–181.
- BARBOUR, A. D. and MOLLISON, D. (1989). Epidemics and random graphs. In *Stochastic Processes in Epidemic Theory. Lecture Notes in Biomath.* **86** 86–89. Springer, Berlin.
- BARTLETT, M. S. (1957). Measles periodicity and community size. *J. Roy. Statist. Soc. Ser. A* **120** 48–70.
- BARTOSZYŃSKI, R. (1972). On a certain model of an epidemic. *Applicationes Mathematicae* **13** 139–151.
- BECKER, N. G. (1989). *Analysis of Infectious Disease Data*. Chapman and Hall, London.
- BECKER, N. G. and DIETZ, K. (1995). The effect of the household distribution on transmission and control of highly infectious diseases. *Math. Biosci.* **127** 207–219.
- BECKER, N. G. and HALL, R. (1996). Immunization levels for preventing epidemics in a community of households made up of individuals of different types. *Math. Biosci.* **132** 205–216.

- BOLKER, B. M., ALTMANN, M., AUBERT, M., BALL, F. G., BARLOW, N. D., BOWERS, R. G., DOBSON, A. P., ELKINGTON, J. S., GARNETT, G. P., GILLIGAN, C. A., HASSELL, M. P., ISHAM, V., JACQUEZ, J. A., KLECKOWSKI, A., LEVIN, S. A., MAY, R. M., METZ, J. A. J., MOLLISON, D., MORRIS, M., REAL, L. A., SATTENSPIEL, L., SWINTON, J., WHITE, P. and WILLIAMS, B. G. (1995). Group report: spatial dynamics of infectious diseases in natural populations. In *Ecology of Infectious Diseases in Natural Populations* (B. T. Grenfell and A. P. Dobson, eds.) 399–420. Cambridge Univ. Press.
- BOLLOBÁS, B. (1985). *Random Graphs*. Academic Press, London.
- DALEY, D. J. (1967). Some aspects of Markov chains in queueing theory and epidemiology. Ph.D. thesis, Cambridge Univ.
- DALEY, D. J. and GANI, J. (1994). A deterministic general epidemic model in a stratified population. In *Probability, Statistics and Optimisation* (F. P. Kelly, ed.) 117–132. Wiley, Chichester.
- DALY, F. (1979). Collapsing supercritical branching processes. *J. Appl. Probab.* **16** 732–739.
- DE KOELIER, A. A., DIEKMANN, O. and REIJNDERS, P. J. H. (1995). A mechanistic model to describe the spread of phocid distemper. Report AM-R9514, CWI, Amsterdam.
- DIEKMANN, O., HEESTERBEEK, J. A. P. and METZ, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28** 365–382.
- DIETZ, K. (1993). The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research* **2** 23–41.
- DOWNTON, F. (1972). A correction to “The area under the infectives trajectory of the general stochastic epidemic.” *J. Appl. Probab.* **9** 873–876.
- GANI, J. and SHANBHAG, D. N. (1974). An extension of Raikov’s theorem derivable from a result in epidemic theory. *Z. Wahrsch. Verw. Gebiete* **29** 33–37.
- GERTSBAKH, I. B. (1977). Epidemic process on a random graph: some preliminary results. *J. Appl. Probab.* **14** 427–438.
- GONTCHAROFF, W. (1937). *Détermination des Fonctions Entières par Interpolation*. Hermann, Paris.
- HABER, M., LONGINI, I. M. and COTSONIS, G. A. (1988). Models for the statistical analysis of infectious disease data. *Biometrics* **44** 163–173.
- HANSKI, I. and GILPIN, M. (1991). Metapopulation dynamics: brief history and conceptual domain. *Biology Journal of the Linnean Society* **42** 3–16.
- HETHCOTE, H. W. (1978). An immunization model for a heterogeneous population. *Theoret. Population Biol.* **14** 338–349.
- HETHCOTE, H. W. and VAN ARK, J. W. (1986). Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation and immunization programs. *Math. Biosci.* **84** 84–118.
- ISLAM, M. N., O’SHAUGHNESSY, C. D. and SMITH, B. (1996). A random graph model for the final size distribution of household infections. *Statistics in Medicine* **15** 837–843.
- JAGERS, P. (1975). *Branching Processes with Biological Applications*. Wiley, New York.
- KAREIVA, P. (1990). Population dynamics in spatially complex environments: theory and data. *Philos. Trans. Roy. Soc. London Ser. B* **330** 175–190.
- KINGMAN, J. F. C. (1993). *Poisson Processes*. Clarendon Press, Oxford.
- KURTZ, T. G. (1970). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* **8** 344–356.
- KURTZ, T. G. (1981). *Approximation of Population Processes*. SIAM, Philadelphia.
- LEFÈVRE, C. and PICARD, PH. (1990). A non-standard family of polynomials and the final size distribution of Reed–Frost epidemic processes. *Adv. in Appl. Probab.* **22** 25–48.
- LONGINI, I. M. and KOOPMAN, J. S. (1982). Household and community transmission parameters from final distributions of infections in households. *Biometrics* **38** 115–126.
- MAY, R. M. and ANDERSON, R. M. (1984). Spatial heterogeneity and the design of immunization programs. *Math. Biosci.* **72** 83–111.
- MOLLISON, D. (1972). The rate of propagation of simple epidemics. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **3** 579–614. Univ. California Press, Berkeley.

- MOLLISON, D. (1995). The structure of epidemic models. In *Epidemic Models: Their Structure and Relation to Data* (D. Mollison, ed.) 17–33. Cambridge Univ. Press.
- MOLLISON, D. and LEVIN, S. A. (1995). Spatial dynamics of parasitism. In *Ecology of Infectious Diseases in Natural Populations* (B. T. Grenfell and A. Dobson, eds.) 384–398. Cambridge Univ. Press.
- NÅSELL, I. (1995). The threshold concept in stochastic epidemic and endemic models. In *Epidemic Models: Their Structure and Relation to Data* (D. Mollison, ed.) 71–83. Cambridge Univ. Press.
- PICARD, PH. and LEFÈVRE, C. (1990). A unified analysis of the final size and severity distribution in collective Reed–Frost epidemic processes. *Adv. in Appl. Probab.* **22** 269–294.
- POST, W. M., DEANGELIS, D. L. and TRAVIS, C. C. (1983). Endemic disease in environments with spatially heterogeneous host populations. *Math. Biosci.* **63** 289–302.
- RUSHTON, S. and MAUTNER, A. J. (1955). The deterministic model of a simple epidemic for more than one community. *Biometrika* **42** 126–132.
- SCALIA-TOMBA, G. (1985). Asymptotic final size distribution for some chain-binomial processes. *Adv. in Appl. Probab.* **17** 477–495.
- SCALIA-TOMBA, G. (1990). On the asymptotic final size distribution of epidemics in heterogeneous populations. In *Stochastic Processes in Epidemic Theory. Lecture Notes in Biomath.* **86** 189–196. Springer, Berlin.
- SELLKE, T. (1983). On the asymptotic distribution of the size of a stochastic epidemic. *J. Appl. Probab.* **20** 390–394.
- VON BAHR, B. and MARTIN-LÖF, A. (1980). Threshold limit theorems for some epidemic processes. *Adv. in Appl. Probab.* **12** 319–349.
- WATSON, R. K. (1972). On an epidemic in a stratified population. *J. Appl. Probab.* **9** 659–666.
- WAUGH, W. A. O'N. (1958). Conditioned Markov processes. *Biometrika* **45** 241–249.
- WHITTLE, P. (1955). The outcome of a stochastic epidemic—a note on Bailey's paper. *Biometrika* **42** 116–122.

FRANK BALL
 DEPARTMENT OF MATHEMATICS
 UNIVERSITY PARK
 NOTTINGHAM NG7 2RD
 ENGLAND
 E-MAIL: fgb@maths.nott.ac.uk

DENIS MOLLISON
 DEPARTMENT OF ACTUARIAL MATHEMATICS
 AND STATISTICS
 RICCARTON
 EDINBURGH EH14 4AS
 SCOTLAND
 E-MAIL: denis@ma.hw.ac.uk

GIANPAOLO SCALIA-TOMBA
 DEPARTMENT OF MATHEMATICAL METHODS
 AND MODELS
 UNIVERSITA LA SAPIENZA
 VIA A SCARPA 10
 00161 ROMA
 ITALY
 E-MAIL: gianpi@dmmm.uniroma1.it