



**HAL**  
open science

# The Probability of Epidemic Burnout in the Stochastic SIR Model with Demography

Todd L. Parsons, Benjamin M. Bolker, Jonathan Dushoff, David J D Earn

► **To cite this version:**

Todd L. Parsons, Benjamin M. Bolker, Jonathan Dushoff, David J D Earn. The Probability of Epidemic Burnout in the Stochastic SIR Model with Demography. 2023. hal-04190409

**HAL Id: hal-04190409**

**<https://hal.science/hal-04190409>**

Preprint submitted on 29 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Probability of Epidemic Burnout in the Stochastic SIR Model with Demography

Todd L. Parsons<sup>a</sup>, Benjamin M. Bolker<sup>b,c</sup>, Jonathan Dushoff<sup>b,d</sup>, and David  
J. D. Earn<sup>c,d,\*</sup>

<sup>a</sup>Laboratoire de Probabilités, Statistique et Modélisation (LPSM), Sorbonne  
Université, CNRS UMR 8001, Paris, France

<sup>b</sup>Department of Biology, McMaster University

<sup>c</sup>Department of Mathematics & Statistics, McMaster University

<sup>d</sup>M.G. DeGroote Institute for Infectious Disease Research, McMaster  
University

\*Corresponding author: [earn@math.mcmaster.ca](mailto:earn@math.mcmaster.ca)

August 29, 2023

## Abstract

We present a new approach to computing the probability of epidemic “burnout”, *i.e.*, the probability that a newly emergent pathogen will go extinct after a major epidemic. Our analysis is based on the standard stochastic formulation of the Susceptible-Infected-Removed (SIR) epidemic model including host demography (births and deaths), and corresponds to the standard SIR ordinary differential equations (ODEs) in the infinite population limit. Exploiting a boundary layer approximation to the ODEs and a birth-death process approximation to the stochastic dynamics within the boundary layer, we derive convenient, fully analytical approximations for the burnout probability. We demonstrate—by comparing with computationally demanding individual-based stochastic simulations and with semi-analytical approximations derived previously—that our fully analytical approximations are highly accurate for biologically plausible parameters. We show that the probability of burnout always decreases with increased mean infectious period. However, for typical biological parameters, there is a relevant local minimum in the probability of persistence as a function of the basic reproduction number  $\mathcal{R}_0$ . For the shortest infectious periods, persistence is least likely if  $\mathcal{R}_0 \approx 2.57$ ; for longer infectious periods, the minimum point decreases to  $\mathcal{R}_0 \approx 2$ . For typical acute immunizing infections in human populations of realistic size, our analysis of the SIR model shows that burnout is almost certain in a well-mixed population, implying that susceptible recruitment through births is insufficient on its own to explain disease persistence.

# 1 Introduction

It is well known that solutions of the standard ordinary differential equations (ODEs) describing a Susceptible-Infected-Removed (SIR) epidemic with host births and deaths (aka “vital dynamics” or “demography”) eventually converge on a globally asymptotically stable equilibrium [9]. Approach to the endemic equilibrium (EE) typically occurs via damped oscillations, motivating the use of the SIR model with demography as a basis for models of observed recurrent epidemics observed of childhood infections such as measles [1, 6, 18, 33, 48]. For many biologically reasonable parameter values and population sizes, however, the troughs of these oscillations pass through infected-host densities corresponding to a small fraction of an individual—the so-called “atto-fox problem” [54]—calling into question the appropriateness of the deterministic SIR model.

Here, we estimate the probability that a pathogen disappears at the end of a major epidemic in a stochastic individual-based SIR model, in a population of finite size. In the large population limit, the densities of each type ( $S$ ,  $I$ ,  $R$ ) are asymptotically deterministic and governed by the standard SIR ODEs [21]. We will refer to pathogen extinction soon after introduction as *fizzle*, whereas if the pathogen escapes fizzle, we will refer to extinction at the end of a major epidemic as *epidemic burnout*<sup>1</sup>, following the terminology of [15]. We will say that the pathogen *persists* if it has a subsequent epidemic wave, although it is worth mentioning that we always expect eventual extinction in a stochastic model with a finite population [37]. Figure 1 shows sample paths of the proportion of infected individuals for the stochastic SIR model (together with the trajectory obtained from the ODE), illustrating fizzle, burnout, and persistence.

The problem of epidemic burnout has been of ongoing interest [1, 10, 14], *e.g.*,

“The question ‘will the agent go extinct after the first outbreak?’ cannot be answered within the context of a deterministic description. So we would like to be able to switch back to a stochastic description at the end of the epidemic outbreak. While it is well known how to calculate the probability of extinction from a branching process in a constant environment [...], it seems difficult to do so when environmental quality (from the point of view of the agent, *i.e.*, the presence of susceptibles!) is improving linearly at a certain rate.” [14, p. 42]

and has been previously approached via perturbation methods [51, 61] and by hybrid analytical-numerical approaches [4]:

1. van Herwaarden [61] starts from a large population diffusion approximation to the Markov chain formulation of the SIR model (see §2.1 below). Under the assumption that the individual mortality rate is low, a highly accurate approximation to the solution of the infinite-population limit SIR ODEs is obtained, which is in turn used to estimate the point of entry to a boundary layer where the number of infected individuals is very small. In the boundary layer, the backward equation for the diffusion

---

<sup>1</sup>While “fade-out” (or “fadeout”) is commonly used to describe this extinction, *e.g.*, [1, §2.3], we find it conceptually useful to follow [15] in distinguishing between extinction after a first major epidemic versus that occurring after multiple epidemics, and reserve the term “fadeout” for the latter.

72 approximation<sup>2</sup> is tractable, and is used to obtain an analytical approximation to the  
73 burnout probability ([61, Eq. (5.13)], which requires the numerical evaluation of an  
74 integral). It is, to quote [14, p. 42], “an ingenious piece of work”, although challenging  
75 to interpret for those lacking expertise in the approach.

- 76 2. By contrast, Meerson and Sasorov [51] retain the discrete population model. They  
77 estimate the probability of extinction as the probability of reaching the state with only  
78 one infective individual (weighted by the expected number of returns to this state<sup>3</sup>)  
79 times the probability that a single infective recovers before transmitting to any other  
80 individuals. They approximate this probability by the product of the expected total  
81 time (summed over multiple returns) in the state with a single infected individual and  
82 the disease recovery rate (which is the rate of going extinct given that there is only one  
83 infected individual). The time in the single-infective state is characterized by linear  
84 equations obtained by integrating the forward equations (see, *e.g.*, [39, §14.2]) for all  
85 transient states over all time, for which an approximate solution is found via a WKB  
86 ansatz (see, *e.g.*, [7, Chapter 10]) in the large population limit (*i.e.*, a diffusion approx-  
87 imation is introduced implicitly). Under these assumptions, the burnout probability is  
88 shown to decay exponentially in the population size, with a constant of proportionality  
89 that is approximated analytically in the parameter regime where the initial exponen-  
90 tial growth rate of infected individuals greatly exceeds the *per capita* turnover rate  
91 (equivalent to  $\beta - \gamma \gg \mu$  in our formulation below). While providing coarser estimates  
92 than van Herwaarden [61], this approach yields a deterministic approximation to the  
93 most probable trajectory to pathogen extinction via a Hamiltonian formalism (see,  
94 *e.g.*, [28] and [13, Exercise 5.7.36]). Like van Herwaarden [61], the approximation of  
95 Meerson and Sasorov [51] involves an integral that cannot be evaluated analytically  
96 and presents a non-trivial numerical problem due to singularities in the integrand.
- 97 3. More recently, after identifying discrepancies between the analytical results of van  
98 Herwaarden [61] and Meerson and Sasorov [51] and the results of simulations, especially  
99 at smaller values of the expected population size  $n$ , Ballard et al. [4] introduced a  
100 computational approach that scales as  $\mathcal{O}(n^2)$ . As in the previous approaches, Ballard  
101 et al. use the solution of the SIR ODEs—now evaluated numerically and summed with  
102 a higher-order Gaussian correction [21, Theorem 11.2.3]—to identify the point of entry  
103 into a boundary layer, where a simplified form of the Markov chain is then simulated  
104 to estimate the probability of burnout.

105 The approximations of van Herwaarden [61] and Meerson and Sasorov [51] are summa-  
106 rized in §2.3 of Ballard et al. [4]. We compare the performance of these approximations with  
107 that of a new analytical approximation that we have derived in the spirit of the quote from  
108 Diekmann and Heesterbeek [14] above. Like van Herwaarden [61] and Ballard et al. [4], we  
109 use the SIR ODEs to approximate the stochastic SIR trajectories outside a boundary layer.  
110 Then, inside the boundary layer, we use a time-inhomogeneous birth-and-death process that

---

<sup>2</sup>See *e.g.*, [39] or [24] for a discussion of the forward and backward diffusion equations; [29] is an excellent introduction to boundary-layer methods for Markov chains.

<sup>3</sup>In practice, there is negligible probability of returning to the state with one infective after an excursion to a state with many infectives.

111 approximates the true stochastic dynamics more accurately than the diffusion approximation  
 112 of van Herwaarden [61] (in Appendix D, we obtain the expression from van Herwaarden [61]  
 113 as an approximation to ours). Our approach is simpler and more intuitive than the diffusion  
 114 approximation, and—in contrast to all previous work—we obtain fully analytical expressions  
 115 that are numerically stable and can be computed without recourse to numerical evaluation  
 116 of integrals. Our approach yields expressions for the probability of persistence after any  
 117 number of epidemic waves, and is also more amenable to generalizations and rigorous proof  
 118 than diffusion approximations; indeed, while we do not discuss the matter in detail here, the  
 119 boundary-layer diffusions of van Herwaarden [61] correspond to large population approxima-  
 120 tions for the branching processes we consider here (similar to limits in Feller [22], Lamperti  
 121 [46]). We will present rigorous proofs for the heuristics derived in this paper in a companion  
 122 manuscript.

## 123 2 Methods

### 124 2.1 Model

125 We consider the spread of an infectious disease in a discrete population in which births  
 126 balance deaths on average, so there is a well-defined **expected population size**  $n$ . We  
 127 consider a sequence of models indexed by  $n$ , and for the  $n$ th model denote by  $S_n(t)$ ,  $I_n(t)$   
 128 and  $R_n(t)$  the numbers of individuals at time  $t$  who are susceptible, infected, and removed,  
 129 respectively. The total population size is

$$130 \quad N_n(t) = S_n(t) + I_n(t) + R_n(t). \quad (1)$$

131 Thus, at every time  $t$  we have

$$132 \quad \mathbb{E}[N_n(t)] = n, \quad (2)$$

133 where the expectation is taken over realizations of the stochastic process.

134 Births of new susceptible individuals and deaths occur at *per capita* rate  $\mu$ , independent  
 135 of disease status. Infected individuals recover at rate  $\gamma$ , and new infections occur according  
 136 to the law of mass action in a well-mixed population, *i.e.*, at rate

$$137 \quad \frac{\beta S_n(t) I_n(t)}{n}. \quad (3)$$

138 Since the demographic and epidemiological rates depend only on the state of the system  
 139 at the current time, our sequence is an ensemble of Markov chain models (indexed by the  
 140 expected total population size  $n$ ).

141 Following a common convention in probability theory, we use upper case for functions,  
 142 and lower case for indices and the values of functions at a given time. We index the functions  
 143 by expected population size because we need to consider the limit of the sequence of functions  
 144 as  $n \rightarrow \infty$ , whereas we use subscripts on function values to specify time, *e.g.*,  $s_0 = S_n(0)$ .

145 The model structure is indicated in a compartmental flow chart in Figure 2, and the  
 146 nature and rates of each type of event are summarized in Table 1.

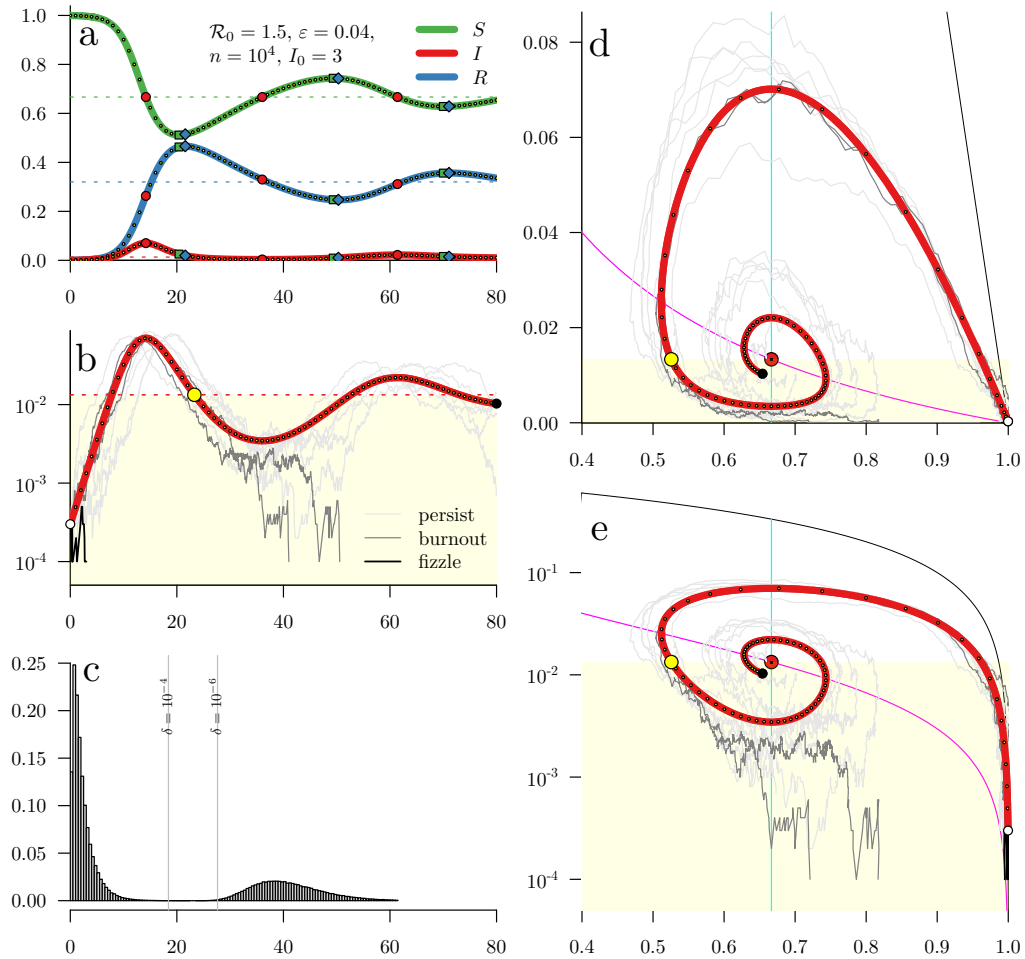


Figure 1: Sample paths of the stochastic SIR model (Figure 2) and the ODE (5) showing fizzle, burnout, and persistence. **(a)** The frequencies of susceptible, infected, and removed individuals in the ODE (symbols indicate the critical points of the curve of the corresponding colour). Dashed lines indicate the endemic equilibrium (8) of the deterministic model (5). **(b)** The proportion of infected individuals as a function of time. The boundary layer—inside which we approximate the stochastic dynamics with a birth-death process—is shaded in yellow, and the first point at which the deterministic trajectory enters the boundary layer is indicated with a heavy yellow dot. **(c)** Probability density of the time to extinction, estimated from  $10^6$  realizations of the stochastic process. The vertical lines show the time  $\tau_\delta$  [Equation (G.11)] for which the probability of fizzle after  $\tau_\delta$  is less than  $\delta$  (the lines correspond to  $\delta = 10^{-4}$  and  $10^{-6}$ ). **(d,e)** Trajectories in the susceptible-infected phase plane with the nullclines; the vertical scale is linear in (d) and logarithmic in (e). The thin black curve is the boundary of the deterministically accessible region, defined by  $S + I = 1$ .

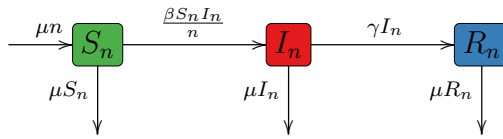


Figure 2: Compartmental model for an SIR epidemic with demography. Labels on the arrows correspond to individual jump rates between states. For simplicity, the model is defined so that births/immigrations on average balance deaths, so that the expected total population size ( $\mathbb{E}[N_n(t)] = n$ ) is fixed.

Table 1: Event types in the stochastic SIR model.

Event type	Rate	Transitions
birth/immigration	$\mu N_n$	$S_n \rightarrow S_n + 1$
transmission	$\beta S_n I_n / N_n$	$S_n \rightarrow S_n - 1, I_n \rightarrow I_n + 1$
recovery	$\gamma I_n$	$I_n \rightarrow I_n - 1, R_n \rightarrow R_n + 1$
susceptible death	$\mu S_n$	$S_n \rightarrow S_n - 1$
infected death	$\mu I_n$	$I_n \rightarrow I_n - 1$
removed death	$\mu R_n$	$R_n \rightarrow R_n - 1$

## 2.2 Deterministic Approximation

In the limit of large population size, the stochastic SIR model (Figure 2, Table 1) is well-approximated by deterministic ODEs. More precisely, writing

$$X_n = \frac{S_n}{n}, \quad Y_n = \frac{I_n}{n}, \quad Z_n = \frac{R_n}{n}, \quad (4)$$

in the limit  $n \rightarrow \infty$ , the frequencies  $(X_n(t), Y_n(t), Z_n(t))$  converge (almost surely on finite time intervals [21]) to the solution  $(X(t), Y(t), Z(t))$  of the ODEs,

$$\frac{dX}{dt} = \mu(1 - X) - \beta XY, \quad (5a)$$

$$\frac{dY}{dt} = (\beta X - \gamma - \mu)Y, \quad (5b)$$

$$\frac{dZ}{dt} = \gamma Y - \mu Z. \quad (5c)$$

Formally, to make this connection, one must be careful to have a sensible relationship between the initial conditions for the stochastic processes and the initial conditions for the ODEs. For example, given an initial state  $(X(0), Y(0), Z(0))$  for the ODEs, if one takes

$$(X_n(0), Y_n(0), Z_n(0)) = \frac{1}{n} ([n X(0)], [n Y(0)], [n Z(0)]) \quad (6)$$

then the theorem applies. More generally, one must choose initial conditions  $(X_n(0), Y_n(0), Z_n(0))$  for the stochastic processes such that the limits  $\lim_{n \rightarrow \infty} X_n(0)$ , *etc.* exist, and one must take

these limits as initial conditions for the ODEs (see Theorem 11.2.1 in Ethier and Kurtz [21, p. 456]; Example B on p. 453 of Ethier and Kurtz [21] illustrates how the SIR model without demography relates to the hypotheses of the theorem, and Chapter 5 in [2] provides a pedagogical introduction to Kurtz’s results in the context of epidemic models).

The trajectories of the deterministic SIR model (5) always converge to a globally asymptotically stable (GAS) equilibrium point, which can be shown via a combination of the Poincaré Bendixson Theorem and Dulac’s criterion [34] or via a Lyapunov function [43]. The nature of the asymptotic state is determined by the **basic reproduction number** (the expected total number of new infections caused by a single infective individual introduced into a naïve population),

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \mu}. \quad (7)$$

If  $\mathcal{R}_0 \leq 1$  then the GAS fixed point is the **disease free equilibrium**,  $(x, y) = (1, 0)$ , whereas if  $\mathcal{R}_0 > 1$  then all solutions converge—either via damped oscillations or monotonically—to an **endemic equilibrium**,

$$(x_\star, y_\star) = \left( \frac{1}{\mathcal{R}_0}, \varepsilon \left( 1 - \frac{1}{\mathcal{R}_0} \right) \right), \quad (8)$$

where

$$\varepsilon = \frac{\mu}{\gamma + \mu} \quad (9)$$

gives the mean infectious period as a fraction of the mean host lifetime. Our analysis requires that  $\varepsilon$  is small but not too small ( $\frac{1}{n} \ll \varepsilon \ll 1$ ), which is true for a wide variety of common acute immunizing infections (see Table 2). The upper bound ( $\varepsilon \ll 1$ ) is essential so we can justify perturbation expansions in  $\varepsilon$ . The lower bound ( $\frac{1}{n} \ll \varepsilon$ ) is equivalent to  $n\varepsilon \gg 1$ , which ensures that the number of infectives at equilibrium ( $ny_\star$ ) is substantially greater than 1 (from Equation (8),  $ny_\star \sim n\varepsilon$ ). The ODEs continue to provide a good approximation to the epidemic dynamics until the **prevalence**  $y$  (the proportion of hosts that are infected) becomes small; we take “small” to mean that  $y$  is less than the equilibrium prevalence  $y_\star$  (8). Thus, we take the **boundary layer**—within which the dynamics must be treated stochastically—to be the region of the phase plane where  $y < y_\star$  (in Appendix D, we also give approximations independent of the specific choice of boundary layer).

The need to analyze the dynamics differently within the boundary layer is especially clear if we consider the introduction of a single infected individual into a fully susceptible population. If  $\mathcal{R}_0 > 1$  then in the ODE system (5)  $Y(t)$  will deterministically increase, whereas in the stochastic model  $Y_n(t)$  will fizzle with probability  $1/\mathcal{R}_0$  [5]; *i.e.*, the ODE (5) fails to capture the dynamics of the stochastic model (Figure 2) when there are few infectives. We therefore use a birth-and-death process to approximate the dynamics of the number of infected hosts when that number is small (in contrast, susceptibles can be assumed to remain sufficiently abundant that we can always use the deterministic approximation  $X(t)$ ).

## 2.3 Birth-and-Death Process Heuristic

New infections occur at rate

$$\frac{\beta S_n(t)}{n} I_n(t) = \beta X_n(t) I_n(t) \approx \beta X(t) I_n(t), \quad (10)$$



201 while the number infected decreases by one due to recovery or death at rate

$$202 \quad (\gamma + \mu)I_n(t). \quad (11)$$

203 When there are few infected hosts ( $I_n(t) < ny_*$ ), we approximate  $I_n(t)$  by a birth and death  
204 process with time-inhomogeneous rates  $b(t)$  and  $d(t)$ , where

$$205 \quad b(t) = \beta X(t), \quad (12)$$

$$206 \quad d(t) = \gamma + \mu. \quad (13)$$

207 Note that when  $X(t)$  equals  $x_*$  (the classical herd immunity threshold),  $b(t) = d(t)$ , and  
208 the birth and death process transitions from subcritical to supercritical. Unlike in models  
209 without demography, the birth of new susceptible individuals ensures that a population will  
210 eventually cross the herd immunity threshold. Therefore, even if the number of infected  
211 hosts initially declines it can eventually grow exponentially, if the infection survives until  
212  $X(t) > x_*$ .

213 We can estimate the survival probability for this branching process, and thus the persis-  
214 tence probability, using the following result.

215 **Theorem 1** (Kendall (1948) [40]). *Let  $\mathcal{K}(t)$  be a birth and death process with time-inhomogeneous*  
216 *per-capita birth rate  $b(t)$  and death rate  $d(t)$ . The probability of eventual extinction starting*  
217 *from one individual at time 0 is*

$$218 \quad q = \left( 1 + \frac{1}{\int_0^\infty e^{-\int_0^t [b(s) - d(s)] ds} d(t) dt} \right)^{-1}. \quad (14)$$

219 *The extinction probability starting from  $k$  individuals is  $q^k$ .*

220 Consequently, the probability of indefinite persistence (a branching process will either go  
221 extinct or grow indefinitely), starting from  $k$  individuals at time 0, is

$$222 \quad \mathbb{P}\{\mathcal{K}(\infty) > 0\} = 1 - q^k. \quad (15)$$

223 To complete our persistence probability estimate, we need an expression for the propor-  
224 tion susceptible at time  $t$  ( $X(t)$  in Equation (12)). As suggested visually by the example  
225 shown in Figure 1, inside the boundary layer ( $y < y_*$ ), both the deterministic and the  
226 stochastic trajectories spend most of their time at prevalences much lower than  $y_*$  (note the  
227 log scale in the subfigures (b) and (e)). Consequently, we can approximate  $X(t)$  by solving  
228 Equation (5) with  $Y(0) = 0$ . Thus, we set

$$229 \quad \frac{dX}{dt} \approx \mu(1 - X), \quad (16)$$

230 and solve this approximate equation as if it were exact to obtain

$$231 \quad X(x_0, t) \approx 1 - (1 - x_0)e^{-\mu t}. \quad (17)$$

232 Here,  $x_0$  is the fraction susceptible at the initial time  $t = 0$ , and we write  $X(x_0, t)$  to  
 233 emphasize the dependence on the initial state. We also write  $q(x_0)$  for the value of  $q$  in  
 234 Equation (14) obtained by taking  $b(t) = \beta X(x_0, t)$ .

235 We first apply this branching process approximation to a population at the disease-free  
 236 equilibrium (DFE). Thus, we set  $x_0 = 1$  in Equation (17), which yields  $X(1, t) \equiv 1$ ; hence we  
 237 have a time-homogeneous branching process in this case, and the integral in Equation (14) is  
 238 easily evaluated and yields  $q(1) = \frac{1}{\mathcal{R}_0} = x_*$ . Considering a small number of initially infective  
 239 individuals,  $I_n(0) = k$ , we recover the classical expression for the establishment probability  
 240 [5], that is, the probability that the pathogen does not fizzle:

$$241 \quad p_k = 1 - x_*^k. \quad (18)$$

242 We now use Kendall's  $q$  (14) to compute the burnout probability. Assuming that the  
 243 pathogen does not fizzle, the number of infected hosts will rapidly exceed  $ny_*$  individuals,<sup>4</sup>  
 244 at which point the densities of both susceptible and infective hosts are well-approximated by  
 245 the ODEs (5). To proceed, we need a formula for the fraction of hosts that are susceptible  
 246 when the trajectory enters the boundary layer at the end of an epidemic; we denote this  
 247 fraction  $x_{\text{in}}$  to emphasize that it refers to the susceptible proportion upon entry *into* the  
 248 boundary layer. In [60], assuming  $\varepsilon$  is small,<sup>5</sup> we derive an approximate expression for  
 249 the fraction susceptible,  $X(y, x_i)$ , as a function of the fraction infected ( $y$ ) and the initial  
 250 fraction susceptible ( $x_i$ ). Using that approximation, we have

$$251 \quad x_{\text{in}} = X(y_*, x_i) \approx -x_* W_0(-\mathcal{R}_0 x_i e^{-\mathcal{R}_0(x_i - y_*)}) + \varepsilon e^{\mathcal{R}_0 y_*} (E_1(\mathcal{R}_0 y_*) - E_1(\mathcal{R}_0 \bar{y}_0)). \quad (19)$$

253 Here,  $W_0$  denotes the principal branch of the Lambert  $W$ -function<sup>6</sup> [12],  $E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt$   
 254 is the exponential integral function [56, 8.2.1] and  $\bar{y}_0$  is the **peak prevalence** in the limit  
 255  $\varepsilon \rightarrow 0$ , *i.e.*, it is the maximum fraction infected in the SIR model *without* vital dynamics,

$$256 \quad \bar{y}_0 = x_i - x_* (1 + \ln(x_i/x_*)). \quad (21)$$

257 (See *e.g.*, [47] for a derivation of  $\bar{y}_0$ .) Taking  $x_i = 1$  corresponds to the invasion of a novel  
 258 pathogen into an epidemiologically naïve population (*i.e.*, at the DFE); in Appendix B we  
 259 give an iterative scheme for  $x_{i,j}$ , an “effective initial fraction susceptible” that, substituted

---

<sup>4</sup>More precisely, for any  $y < \bar{y}_0$  [see Equation (21)], conditional on not fizzling, the probability that  $I_n(t)$  hits 0 before hitting  $yn$  is exponentially small in  $n$  with exponential rate depending on  $y$  (for a rigorous demonstration see [59, Supplementary Information §8.2]; [58] gives explicit higher order terms for the SIS model).

<sup>5</sup>In [60], we use  $\epsilon = \varepsilon/\mathcal{R}_0$  rather than  $\varepsilon$  as the small parameter, because using  $\epsilon$  leads to simpler expressions (see *e.g.*, (24) below for an example). Here, however, we analyze the dependence of our expressions on the epidemiologically relevant parameters  $\mathcal{R}_0$  and  $\varepsilon$  and have re-written expressions from [60] accordingly.

<sup>6</sup>If  $\mathcal{E}(z) = ze^z$ , Lambert's  $W$ -function  $W(z)$  [12] solves the “left-sided” inverse relation  $\mathcal{E}(W(z)) = z$ . This equation has countably many solutions, each corresponding to branches  $W_i$  of the  $W$ -function; we will need the two real branches,  $W_0$ , which maps  $[-\frac{1}{e}, \infty)$  to  $[-1, \infty)$ , and  $W_{-1}$ , which maps  $[-\frac{1}{e}, 0)$  to  $(-\infty, -1]$ . For these two branches,  $W_i$  is a *partial* “right-sided” inverse function for  $\mathcal{E}(z)$ :

$$\begin{aligned} W_{-1}(\mathcal{E}(z)) &= z & \text{if } z \leq -1 \\ W_0(\mathcal{E}(z)) &= z & \text{if } z \geq -1. \end{aligned} \quad (20)$$

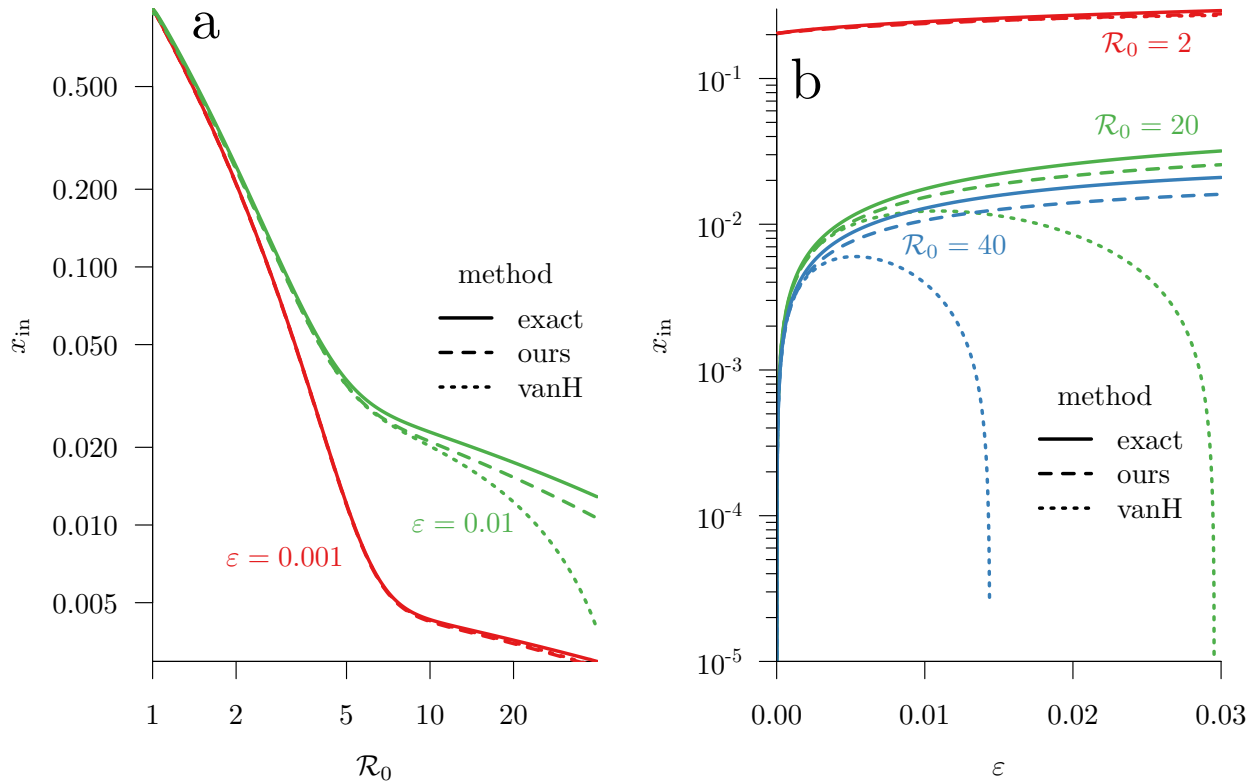


Figure 3: Susceptible proportion ( $x_{\text{in}}$ ) upon entry into the boundary layer ( $y < y_*$ ). **(a)**  $x_{\text{in}}$  as a function of  $\mathcal{R}_0$  (7). **(b)**  $x_{\text{in}}$  as a function of  $\varepsilon$  (9). The exact value of  $x_{\text{in}}$  [obtained by numerically solving the SIR ODEs (5)] is shown with solid curves, our approximation [Equation (19)] is shown with dashed curves, and the approximation of van Herwaarden [61] is shown with dotted curves. Based on Equation (C.3), the minimum  $\mathcal{R}_0$  for which our approximation of  $x_{\text{in}}$  [Equation (19)] is valid is  $\approx e^{2\varepsilon}$  (i.e., 1.02027 for  $\varepsilon = 0.01$  and 1.0020027 for  $\varepsilon = 0.001$ ).

260 for  $x_{\text{in}}$  in (19) and (21), gives the fraction susceptible at the end of the  $j^{\text{th}}$  epidemic wave  
 261 after invasion at the DFE. We compare our approximation of  $x_{\text{in}}$  for  $x_i = 1$  to the value  
 262 obtained by numerically integrating the SIR ODEs (5) in Figure 3 and discuss its domain of  
 263 applicability in Appendix C.

264 If we now take  $t = 0$  to be the end of a major epidemic, *i.e.*, the time when the infected  
 265 host density falls below  $y_*$  and  $x_0 = x_{\text{in}}$ , then the density of infected hosts is small, and the  
 266 density of susceptible hosts is well-approximated by  $X(x_{\text{in}}, t)$  (in a companion manuscript,  
 267 we give rigorous statements and justifications for these results; here we will content ourselves  
 268 with showing that our analytical results closely match the results of individual-based simula-  
 269 tions). We can thus estimate the **conditional burnout probability**—*i.e.*, the probability  
 270 of burnout conditional on not fizzling—by

271 
$$q(x_{\text{in}})^{ny_*}. \quad (22)$$

272 and the **conditional persistence probability** by

$$273 \quad 1 - q(x_{\text{in}})^{ny_*}. \quad (23)$$

274 In Appendix A (below) we compute an exact expression for  $q(x_{\text{in}})$ ,

$$275 \quad q(x_{\text{in}}) = \left( 1 + \frac{\varepsilon}{z^{-a} e^z g(a, z)} \right)^{-1} \quad (24a)$$

$$276 \quad \text{where} \quad z = \frac{\mathcal{R}_0}{\varepsilon} (1 - x_{\text{in}}), \quad (24b)$$

$$277 \quad \text{and} \quad a = \frac{\mathcal{R}_0}{\varepsilon} (1 - x_*) , \quad (24c)$$

278 where  $g$  denotes the lower incomplete gamma function<sup>7</sup> [56, 8.2.1]; we use the nonstandard  
 279 notation  $g$  to avoid confusion with our recovery rate parameter  $\gamma$ . In Appendix A, we derive  
 280 an approximation that is extremely accurate for small values of  $\varepsilon$ :

$$281 \quad q(x_{\text{in}}) \approx \left( 1 + \frac{1}{\sqrt{\frac{2\pi}{\varepsilon(\mathcal{R}_0-1)}} \left( \frac{1-x_*}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon} (1-x_*)} e^{\frac{\mathcal{R}_0}{\varepsilon} (x_*-x_{\text{in}})}} \right)^{-1}. \quad (25)$$

282 We emphasize that this expression is elementary and numerically stable.

283 Thus, the **burnout probability**—*i.e.*, the probability of not fizzling (18) but disap-  
 284 pearing after an epidemic—is

$$285 \quad p_k q(x_{\text{in}})^{ny_*}, \quad (26)$$

286 where  $n$  is the expected total population size,  $y_*$  is the equilibrium prevalence (8),  $q$  is the  
 287 probability of eventual extinction (under post-epidemic conditions) starting from one infected  
 288 individual (14), and  $k$  is the initial number of infected individuals. Our exact expression  
 289 for  $q(x_{\text{in}})$  is given in Equation (24). Similarly, the **persistence probability**—*i.e.*, the  
 290 probability of not fizzling ( $p_k$ ) and then *not* burning out after a first epidemic (23)—is

$$291 \quad \mathcal{P}_1(\mathcal{R}_0, \varepsilon, n, k) = p_k (1 - q(x_{\text{in}})^{ny_*}). \quad (27)$$

292 More generally, the probability of persisting beyond the  $m^{\text{th}}$  epidemic wave is

$$293 \quad \mathcal{P}_m(\mathcal{R}_0, \varepsilon, n, k) = p_k \prod_{j=1}^m (1 - q(x_{\text{in},j})^{ny_*}). \quad (28)$$

294 where

$$295 \quad x_{\text{in},j} = X(y_*, x_{i,j}) \quad (29)$$

296 (see Equation (19) and Appendix B). For biologically reasonable values of  $\varepsilon$ ,  $\mathcal{R}_0$ , and  $n$ , we  
 297 find that the difference between  $\mathcal{P}_1(\mathcal{R}_0, \varepsilon, n, k)$  and  $\mathcal{P}_m(\mathcal{R}_0, \varepsilon, n, k)$  is negligible (not shown),

---

<sup>7</sup> $g(a, z) = \int_0^z t^{a-1} e^{-t} dt$  is proportional to the cumulative distribution function for the gamma distribu-  
 tion. We use this fact to compute  $g(a, z)$  accurately in our burnout R package, mentioned in Footnote 8.

298 because  $q(x_{\text{in},j}) \ll 1$  for  $j \geq 2$ . Intuitively, because the troughs between epidemics get  
 299 shallower and shallower, an invading disease that survives burnout is almost certain to persist  
 300 through many more cycles.

301 Thus, in the §3 (below) we focus on burnout after the initial epidemic when a novel disease  
 302 invades a fully susceptible population. There, we use our accurate, numerically stable,  
 303 and computationally efficient approximation for  $q(x_{\text{in}})$  (25), obtained via Equations (19)  
 304 and (A.8) to compute the probability of burnout.

### 305 **3 Results**

306 Figure 4 shows that our analytical approximation for the persistence probability (27) agrees  
 307 very well with the same probability estimated from large numbers of simulations. The  
 308 probability is shown as a function of the basic reproduction number ( $\mathcal{R}_0$ ) with fixed mean  
 309 infectious period ( $\varepsilon = 0.01$ ). The panels differ only in the underlying expected population  
 310 size (ranging from  $n = 10^4$  to  $10^7$ ). For each value of  $\mathcal{R}_0$ , the simulation-based persistence  
 311 probability was estimated from  $10^7$  individual-based stochastic realizations of the model  
 312 (Figure 2, Table 1). Note that  $\varepsilon = 0.01$  corresponds to an infectious period that is 1% of  
 313 the average host lifetime, far longer than is realistic for most acute immunizing infections;  
 314 however, our approximation only improves for smaller  $\varepsilon$ .

315 Our simple approximation for Kendall’s  $q$  [Equation (25)] allows us to easily and quickly  
 316 explore the conditional and unconditional probability of pathogen extinction across the entire  
 317 range of biologically plausible values of  $\mathcal{R}_0$  and  $\varepsilon$ . Figure 5 shows a contour plot of the  
 318 persistence probability (this graph would have required years of computer time to produce  
 319 from simulations). As was observed previously [4, 20], Figure 5 indicates that the burnout  
 320 probability is non-monotone in  $\mathcal{R}_0$  for  $\varepsilon \lesssim 0.016$ . In this range of  $\varepsilon$ , the probability  
 321 of persistence is lowest for basic reproduction numbers in the range  $2 \lesssim \mathcal{R}_0 \lesssim 2.57$ , and  
 322 increases rapidly with increasing  $\mathcal{R}_0$ . We compute a linear approximation to the value of  
 323  $\mathcal{R}_0$  at which the persistence probability is minimized in Appendix E [the upper limit of  
 324 2.57 for the range of  $\mathcal{R}_0$  is the limit as  $\varepsilon \rightarrow 0$  in Equation (E.5)]; Figure 5 shows that this  
 325 linear approximation performs very well over the range where the persistence probability is  
 326 non-monotonic. Less intuitively, the persistence probability increases for small  $\mathcal{R}_0$  (below  
 327 the red curve in Figure 5) as  $\mathcal{R}_0$  decreases to one. We note, however, that except for very  
 328 large expected population size  $n$ , the secondary peak in the persistence probability—which  
 329 occurs for  $1 < \mathcal{R}_0 \lesssim 2$ —remains small (*cf.* Figure 4), except for pathogens with extremely  
 330 long infectious periods. Figure 5 also suggests that the probability of persistence always  
 331 increases with increasing  $\varepsilon$ , which we confirm analytically in Appendix F.

### 332 **4 Discussion**

333 The problem of infectious disease persistence following a major epidemic [1; 14, p.42; 42,  
 334 p.451; 15; 20; 10] is important for identifying characteristics of pathogens that can success-  
 335 fully invade, and is related to the notion of a “critical community size” required for a disease  
 336 to persist in the long term [5].

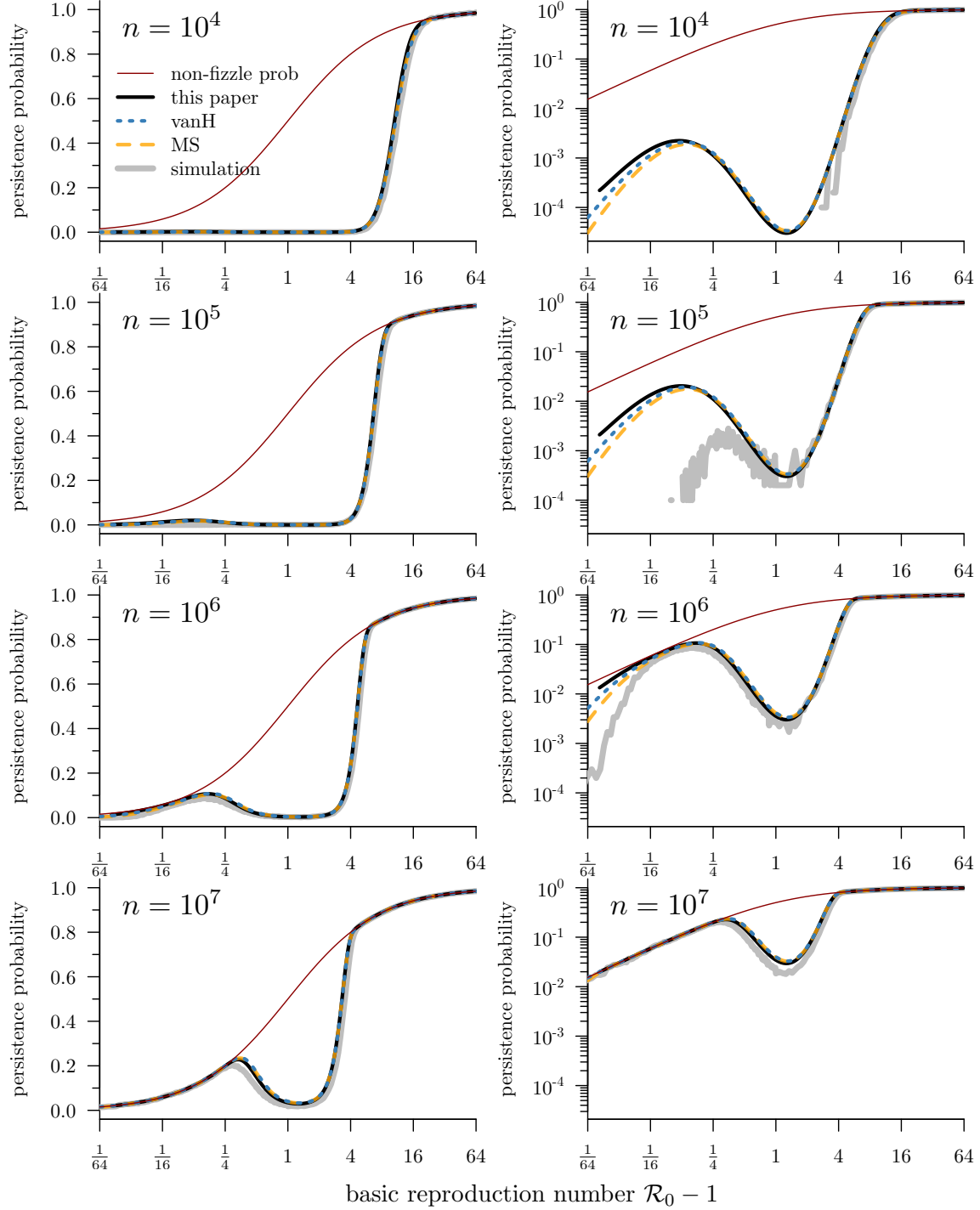


Figure 4: Persistence probability as a function of the basic reproduction number  $\mathcal{R}_0$ , for population sizes ranging from  $n = 10^4$  to  $10^7$ . The vertical scale is linear in the left column and logarithmic in the right column; the horizontal scale is logarithmic (in  $\mathcal{R}_0 - 1$ ) in all panels. (The horizontal axis range is from  $\mathcal{R}_0 - 1 = \frac{1}{64} = 0.015625$  to 64, but our approximation is valid only for  $\mathcal{R}_0 - 1 \gtrsim 0.02027$ ; see Equation (C.3).) The initial state is  $(S_n(0), I_n(0), R_n(0)) = (n-1, 1, 0)$ . The mean infectious period as a fraction of mean lifetime is  $\varepsilon = 0.01$ , which is unrealistically long for most infections (Table 2), but the agreement between the analytical approximation (27) and numerical simulations (Appendix G) is *better* for smaller  $\varepsilon$ . In addition to our analytical approximation (27), we show the semi-analytical approximations of Meerson and Sasorov (MS [51]) and van Herwaarden (vanH [61]). The thin red curve shows the probability of not fizzling,  $1 - x_*$ .

### contours of persistence probability

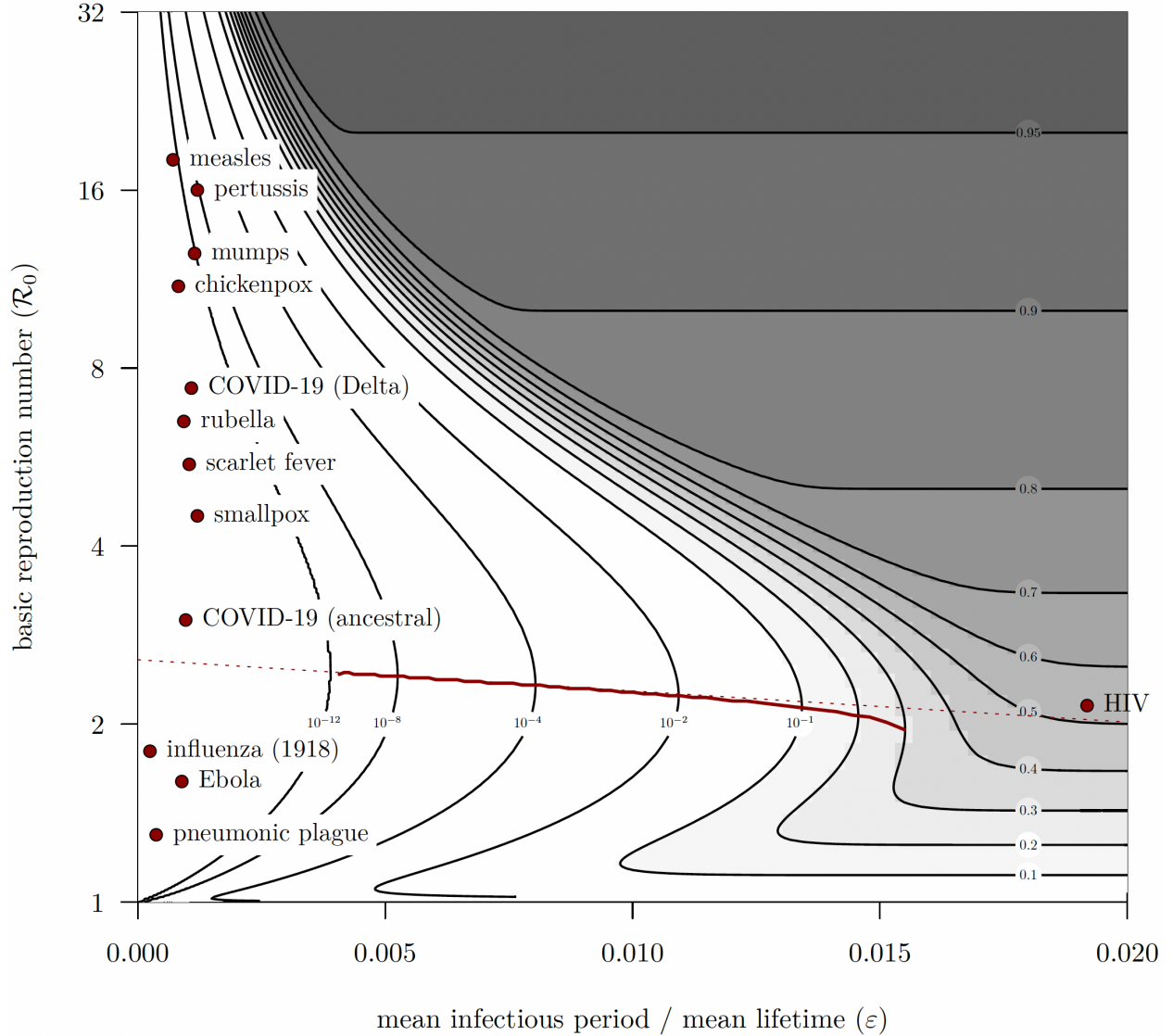


Figure 5: Probability of persistence after a large epidemic [ $\mathcal{P}_1$ , Equation (27)] as a function of basic reproduction number ( $\mathcal{R}_0$ ) and mean infectious period as a proportion of mean lifetime ( $\varepsilon$ ), for population size  $n = 10^6$ . The initial state is assumed to be a single individual introduced into a fully susceptible population ( $I_n(0) = k = 1$ ,  $S_n(0) = n - k$ ). Positions for the red dots for infectious diseases of humans are from Table 2 (to avoid text overlap, measles is shifted up by 1 to 18, pertussis down by 1 to 16, and COVID-19 (Delta) up by 0.6 to 7.4). The solid red curve shows the local minimum of persistence probability, and the dotted red line shows the analytical approximation (E.5) to the local minimum.

Table 2: Representative parameters for acute immunizing infections (and HIV for comparison).

Disease	$\mathcal{R}_0$	$T_{\text{lat}}$ [days]	$T_{\text{inf}}$ [days]	$\varepsilon \times 10^3$	Source
measles	17	8	5	0.71	[1]
pertussis	17	8	14	1.2	[1]
mumps	12	15	6	1.1	[1]
chickenpox	11	10	5	0.82	[1]
COVID-19 (Delta)	6.8	5.8	14	1.1	[47]
rubella	6.5	10	7	0.93	[1]
scarlet fever	5.5	1.5	18	1	[1]
smallpox	4.5	15	7	1.2	[44]
COVID-19 (ancestral)	3	3.7	14	0.97	[47]
HIV	2.2	87	270	19	[35]
influenza (1918)	1.8	2	2.5	0.25	[1, 52]
Ebola	1.6	9.3	7	0.89	[62]
pneumonic plague	1.3	4.3	2.5	0.37	[23]

The basic reproduction number ( $\mathcal{R}_0$ ), mean latent period ( $T_{\text{lat}}$ ), and mean infectious period ( $T_{\text{inf}}$ ) are taken from the cited sources. The dimensionless parameter  $\varepsilon$  is defined in Equation (9) in terms of the recovery rate ( $\gamma$ ) and birth-death rate ( $\mu$ ) in the SIR model. We associate  $1/\gamma$  with the mean generation interval of the SEIR model, i.e.,  $1/\gamma = T_{\text{lat}} + T_{\text{inf}}$  [11, 45], set  $\mu = 0.02/\text{year}$  to mimic human birth and death rates, and compute  $\varepsilon = \mu/(\gamma + \mu)$ . Where original sources present a range, we have listed the midpoint. Many of the estimates come from Anderson and May [1] ( $\mathcal{R}_0$  is taken from Table 4.1 [1, p. 70]; the mean latent and infectious periods come from Table 3.1 [1, p. 31]). All the diseases listed in this table are shown in Figure 5.



337 Given sufficient computing resources, it is possible to estimate the persistence probabil-  
 338 ity for a given model from large numbers of stochastic, individual-based simulations. The  
 339 grey curves in Figure 4 show this probability estimated from simulations of the SIR model.  
 340 Figure 4 also shows the probability estimated using previous analytical methods [51, 61]  
 341 (blue and orange curves) and our new approximation (black curves). All three analytical  
 342 approaches yield similar results<sup>8</sup>, and differences in the estimated probabilities can be seen  
 343 only on a logarithmic scale in the limit as  $\mathcal{R}_0 \rightarrow 1^+$  (*e.g.*, for  $\mathcal{R}_0 \lesssim 1.05$  in Figure 4), where  
 344 all of these approximations<sup>9</sup> are technically invalid: in a stochastic, finite population model,  
 345 as  $\mathcal{R}_0 \rightarrow 1^+$  there is no phase during which the deterministic model is a good approximation,  
 346 and the distinction between fizzle, burnout and fadeout breaks down [55]. Analysis of the  
 347 limit  $\mathcal{R}_0 \rightarrow 1^+$  could improve understanding of the process of eradication as the magnitude  
 348 of control measures is increased—and for this reason we will discuss approximations that  
 349 are more appropriate for  $\mathcal{R}_0 \sim 1$  in future work—but for the burnout problem on which we  
 350 focus here, the limit  $\mathcal{R}_0 \rightarrow 1^+$  is of limited interest.

351 While our approximation agrees closely with previous work [51, 61] for ranges of  $\mathcal{R}_0$  that  
 352 are biologically relevant, there are several important theoretical and practical advantages of  
 353 our approach; our analysis:

- 354 • is more precise and easier to justify theoretically since it is based directly on the  
 355 underlying stochastic process rather than a diffusion approximation;
- 356 • is simpler and easier to understand, and thus easier to apply to models that are more  
 357 complex than the SIR model considered here;
- 358 • yields fully analytical approximations that are numerically stable, unlike the previous  
 359 analytical approaches [51, 61], which depend on non-trivial numerical integrations with  
 360 singular integrands;
- 361 • predicts the persistence probability after an arbitrary number of epidemic waves.

362 We expand on these points below.

363 We have obtained useful analytical estimates [Equations (24), (25), (27) and (28)] of  
 364 the SIR epidemic burnout and persistence probabilities in a well-mixed population, via a  
 365 hybrid use of ODEs when prevalence is high and time-dependent branching processes when  
 366 prevalence is low. As noted after Equation (28), the probability of burning out in each  
 367 subsequent epidemic trough after persisting through the first is negligibly small for the SIR  
 368 model. In future work, we will apply our method to more detailed models that account  
 369 for disease-induced mortality, decay of immunity, vaccination, super-spreading, and other  
 370 factors.

371 By coupling methods for birth-and-death processes [3] with functional central limit theo-  
 372 rems for ODE approximations [21], we are able to support our approximations with rigorous  
 373 error estimates (which we will present in a technical companion paper). Rigorous proofs

---

<sup>8</sup>We have implemented all three approximations in an open-source R package, which we used to create our figures. The package is available at <https://github.com/davidearn/burnout>.

<sup>9</sup>Differences between our approximation and those reported in [51, 61] as  $\mathcal{R}_0 \rightarrow 1^+$  are at least in part because they use  $\mu$  rather than  $\varepsilon$  as the small parameter, and consequently predict persistence for  $\beta/\gamma > 1$  rather than  $\beta/(\gamma + \mu) > 1$ .

374 are much more difficult for asymptotic methods applied to the partial differential equation  
 375 (PDE) corresponding to the diffusion approximation [51, 61] and its boundary-layer ap-  
 376 proximations, the latter being degenerate elliptic PDEs on a manifold with corners and/or  
 377 boundaries.

378 Our time-dependent branching process approach (§2.3) also yields analytical results that  
 379 are more amenable to computation than previous approximations [51, 61]. Our application  
 380 of Laplace’s method to approximate the integral in Kendall’s  $q$  [Equation (14)] is particu-  
 381 larly useful. Equation (25) for the conditional burnout probability provides a fully analyt-  
 382 ical formula—not requiring the numerical evaluation of integrals as in previous approaches  
 383 [51, 61]—that can be evaluated without numerical instabilities and agrees very well with  
 384 numerical simulations across a wide range of biologically plausible values of  $\mathcal{R}_0$  and  $\varepsilon$ . The  
 385 convenience and speed of our simple analytical expression for the persistence probability  
 386 (27) also allows us to obtain results for larger population sizes than are tractable via hybrid  
 387 numerical methods [4], and facilitates efficient exploration of more of the parameter space  
 388 (though with less accuracy at smaller population sizes).

389 As is suggested visually by Figure 5, and proved in Appendix F, the persistence proba-  
 390 bility increases with infectious period ( $\varepsilon$ ) across all values of  $\mathcal{R}_0$ . For any given infectious  
 391 period, one viable life history strategy for persistence is a high  $\mathcal{R}_0$  (dark grey shading in  
 392 Figure 5). In addition to this high  $\mathcal{R}_0$  strategy, for a limited range of longer infectious pe-  
 393 riods ( $0.01 \lesssim \varepsilon \lesssim 0.016$ ) there is a second life-history strategy that promotes persistence:  
 394  $\mathcal{R}_0$  close to but greater than one (from Figure 5 we can infer that  $\varepsilon \gtrsim 0.01$  is necessary for  
 395 a pathogen with  $\mathcal{R}_0 < 2$  to have even a 10% chance of avoiding fizzle or burnout, whereas  
 396 the secondary peak in persistence probability disappears for  $\varepsilon \gtrsim 0.016$ ). In Appendix E,  
 397 we use our analytical results to compute a linear approximation to the value of  $\mathcal{R}_0 > 1$  at  
 398 which the burnout probability is maximized [Equation (E.5)]. This shows excellent agree-  
 399 ment with the numerical results over the range of  $\varepsilon$  for which the secondary peak exists  
 400 and the burnout probability is numerically distinguishable from 1 (in Figure 5 the dotted  
 401 red curve is the approximation and the solid red curve is the numerically computed exact  
 402 value). Intriguingly, with the exception of the ancestral strain of SARS-CoV-2—which has  
 403 been replaced by variants with much higher  $\mathcal{R}_0$ —the endemic infectious diseases of humans  
 404 listed in Table 2 roughly divide into high and low  $\mathcal{R}_0$  strategies.

405 These strategies can be interpreted in terms of the **herd immunity threshold**,  $x_*$ , *i.e.*,  
 406 the minimum proportion susceptible at which the epidemic can grow from a small number  
 407 of infections. When  $\mathcal{R}_0$  is large, the susceptibility threshold  $x_*$  is low, allowing the fraction  
 408 susceptible to rapidly reach the threshold. When  $\mathcal{R}_0$  is low, there is a larger reservoir of  
 409 susceptible hosts at the end of the first major epidemic, which reduces the wait until the herd  
 410 immunity threshold is crossed. In either case, a longer infectious period (larger  $\varepsilon$ ) allows the  
 411 pathogen to “wait out” the period of herd-immunity. This non-monotonicity of the burnout  
 412 probability as a function of  $\mathcal{R}_0$  was previously observed [4, 20], and the maximum burnout  
 413 probability was conjectured to occur for  $\mathcal{R}_0 \approx 3$  [20] or  $\mathcal{R}_0 = 2$  [4]. We have shown that,  
 414 in fact, the value of  $\mathcal{R}_0$  at which the probability is maximized is a decreasing function of  
 415  $\varepsilon$  (solid red curve in Figure 5). The probability-maximizing  $\mathcal{R}_0$  varies from  $\mathcal{R}_0 \simeq 2.57$  for  
 416  $\varepsilon \rightarrow 0$  [Equation (E.5)] to  $\mathcal{R}_0 \simeq 2$  for  $\varepsilon \simeq 0.016$ ; for larger  $\varepsilon$ , the persistence probability  
 417 increases monotonically with  $\mathcal{R}_0$ .

418 These results also have evolutionary implications: reduced virulence may be associated

419 with longer infectious periods (*e.g.*, if fewer hosts die while infectious), thereby reducing  
420 the probability of burnout. This suggests a mechanism—distinct from the population ge-  
421 netics/weak selection arguments presented in [59]—that could explain how natural selection  
422 may favour strains with longer infectious periods, rather than the maximal  $\mathcal{R}_0$  strain, in  
423 finite populations. We are currently working on a multi-strain model that properly captures  
424 this possibility.

425 While the qualitative inferences we have made from analysis of the stochastic SIR model  
426 are suggestive of general processes, and—as we observe above—could have interesting im-  
427 plications, further research is needed to determine if they really do generalize broadly. Most  
428 acute immunizing infections afflicting human populations have short infectious periods and  
429 moderate  $\mathcal{R}_0$  values, and with these constraints our analysis of the stochastic SIR model  
430 indicates that extinction of the pathogen at the end of the first major epidemic is almost  
431 certain in a well-mixed population.

432 Figure 5 makes clear that the stochastic SIR model is insufficient on its own to explain  
433 population persistence; it is essential to consider additional mechanisms, *e.g.*, waning im-  
434 munity or antigenic evolution resulting in effective loss of host immunity [19], effects in  
435 a meta-population [16, 30], long-lived carrier infections (see [31] for a recent survey), or  
436 zoonotic reservoirs [19, 32].

437 Multi-type or non-Markovian birth-and-death processes [36, 53], combined with more  
438 complicated compartmental models or renewal equation models with more general generation  
439 intervals [11] may allow our approach to be extended to models incorporating, *e.g.*, latent  
440 periods and asymptomatic and carrier infections, or greater or lesser variability in infectious  
441 periods. A more difficult problem is to consider pathogen persistence in a meta-community  
442 of linked sites [16, 50], or other structured populations, rather than a well-mixed population.  
443 Smaller local community sizes tend to make local extinction more likely, whereas asynchrony  
444 in epidemic dynamics could allow pathogens to reinvade following a local extinction [17]. Are  
445 these processes adequate to plausibly explain the persistence of pathogens? Is the existence  
446 of low/high  $\mathcal{R}_0$  strategies generic, or an artifact of the SIR compartmental model? Are  
447 longer infectious periods always favourable for pathogen persistence? We will pursue these  
448 questions in future work.

## 449 Acknowledgements

450 This project was was partially supported by the CNRS IEA grant “Structured Popula-  
451 tions, Epidemics, and Control Strategies (SPECS)”. DJDE, JD and BMB were supported  
452 by NSERC.

## References

- [1] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford, 1991.
- [2] H. Andersson and T. Britton. *Stochastic Epidemic Models and their Statistical Analysis*, volume 151 of *Lecture notes in statistics*. Springer, New York, 2000.
- [3] F. Ball. Coupling methods in epidemic theory. In D. Mollison, editor, *Epidemic Models: Their Structure and Relation to Data*, pages 34–52. Cambridge University Press, 1995.
- [4] P. G. Ballard, N. G. Bean, and J. V. Ross. The probability of epidemic fade-out is non-monotonic in transmission rate for the Markovian SIR model with demography. *J. Theor. Biol.*, 393:170–178, 2016.
- [5] M. S. Bartlett. Deterministic and stochastic models for recurrent epidemics. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 4, pages 81–108, 1956.
- [6] M. S. Bartlett. Measles periodicity and community size. *Journal of the Royal Statistical Society Series A*, 120:48–70, 1957.
- [7] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. McGraw-Hill, New York, 1978.
- [8] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken, New Jersey, anniversary edition, 2012.
- [9] F. Brauer, C. Castillo-Chavez, and Z. Feng. *Mathematical models in epidemiology*, volume 32. Springer, 2019.
- [10] T. Britton, T. House, A. L. Lloyd, D. Mollison, S. Riley, and P. Trapman. Five challenges for stochastic epidemic models involving global transmission. *Epidemics*, 10:54–57, 2015.
- [11] D. Champredon, Dushoff J., and D. J. D. Earn. Equivalence of the Erlang SEIR epidemic model and the renewal equation. *SIAM Journal on Applied Mathematics*, 78(6):3258–3278, 2018. doi: 10.1137/18M1186411. URL <https://epubs.siam.org/doi/10.1137/18M1186411>.
- [12] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert  $W$  function. *Adv. Comput. Math.*, 5(1):329–359, Dec 1996. ISSN 1572-9044. doi: 10.1007/BF02124750. URL <https://doi.org/10.1007/BF02124750>.
- [13] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Applications of Mathematics*. Springer, New York, 2nd edition, 1998. ISBN 0387984062. URL <http://www.loc.gov/catdir/enhancements/fy0815/97045236-d.html>.
- [14] O. Diekmann and J. A. P. Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. John Wiley & Sons, LTD, New York, 2000.

- 489 [15] J. Dushoff. Incorporating stochasticity in simple models of disease spread. In A. B.  
490 Gumel and S. Lenhart, editors, *Modeling paradigms and analysis of disease transmission*  
491 *models*, volume 75. American Mathematical Soc., 2010.
- 492 [16] D. J. D. Earn, P. Rohani, and B. T. Grenfell. Persistence, chaos and synchrony in ecology  
493 and epidemiology. *Proc. R. Soc. Lond. B*, 265(1390):7–10, 1998. doi: 10.1098/rspb.1998.  
494 0256. URL <https://royalsocietypublishing.org/doi/10.1098/rspb.1998.0256>.
- 495 [17] D. J. D. Earn, S. A. Levin, and P. Rohani. Coherence and conservation. *Science*, 290  
496 (5495):1360–1364, 2000. doi: 10.1126/science.290.5495.1360. URL <http://science.sciencemag.org/content/290/5495/1360>.  
497
- 498 [18] D. J. D. Earn, P. Rohani, B. M. Bolker, and B. T. Grenfell. A simple model for  
499 complex dynamical transitions in epidemics. *Science*, 287(5453):667–670, 2000. doi:  
500 10.1126/science.287.5453.667. URL [http://science.sciencemag.org/content/287/](http://science.sciencemag.org/content/287/5453/667)  
501 [5453/667](http://science.sciencemag.org/content/287/5453/667).
- 502 [19] D. J. D. Earn, J. Dushoff, and S. A. Levin. Ecology and evolution of the flu. *Trends in*  
503 *Ecology and Evolution*, 17(7):334–340, 2002. doi: 10.1016/S0169-5347(02)02502-8.
- 504 [20] D. J. D. Earn, D. Champredon, M. deJonge, S. Drohan, K. Hempel, C. Molina, I. Papst,  
505 D. Rosati, B. Bolker, and J. Dushoff. The Puzzling Persistence of Invading Pathogens.  
506 *Poster presentation at the 2014 Ecology and Evolution of Infectious Diseases Conference*  
507 *“Multi-Scale Mechanisms of Disease Emergence and Control”*, Colorado State Univer-  
508 *sity*, 2014.
- 509 [21] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*.  
510 John Wiley and Sons, New York, 1986.
- 511 [22] W. Feller. Diffusion processes in genetics. In J. Neyman, editor, *Proc. Second Berkeley*  
512 *Symp. on Math. Statist. and Prob., July 31-August 12, 1950*, pages 227–246. Univ. of  
513 Calif. Press, Berkeley, Calif., 1951.
- 514 [23] R. Gani and S. Leach. Epidemiologic determinants for modeling pneumonic plague  
515 outbreaks. *Emerging infectious diseases*, 10(4):608–614, 04 2004. doi: 10.3201/eid1004.  
516 030509. URL <https://pubmed.ncbi.nlm.nih.gov/15200849>.
- 517 [24] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural*  
518 *Sciences*. Springer, Berlin Heidelberg, 2004.
- 519 [25] D. T. Gillespie. A general method for numerically simulating the stochastic time evo-  
520 lution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–434,  
521 1976.
- 522 [26] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal Of*  
523 *Physical Chemistry*, 81(25):2340–2361, 1977.
- 524 [27] D. T. Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical*  
525 *Chemistry*, 58:35–55, 2007.

- 526 [28] R. Graham and T. Tél. Existence of a potential for dissipative dynamical systems.  
527 *Phys. Rev. Lett.*, 52:9–12, Jan 1984. doi: 10.1103/PhysRevLett.52.9. URL <https://link.aps.org/doi/10.1103/PhysRevLett.52.9>.  
528
- 529 [29] J. Grasman and O. A. van Herwaarden. *Asymptotic Methods for the Fokker-Planck*  
530 *Equation and the Exit Problem in Applications*. Springer, Berlin Heidelberg, 1999.
- 531 [30] B. T. Grenfell and J. Harwood. (Meta)population dynamics of infectious diseases.  
532 *Trends Ecol. Evol.*, 12(10):395–399, 1997.
- 533 [31] K. Hampson and D. Haydon. Persistent pathogens and wildlife reservoirs. *Science*, 374  
534 (6563):35–36, 2021.
- 535 [32] D. T. Haydon, S. Cleaveland, L. H. Taylor, and M. K. Laurenson. Identifying reservoirs  
536 of infection: a conceptual and practical challenge. *Emerg. Infect. Dis.*, 8(12):1468–1473,  
537 2002.
- 538 [33] K. Hempel and D. J. D. Earn. A century of transitions in New York City’s measles  
539 dynamics. *Journal of the Royal Society of London, Interface*, 12(106):20150024, 2015.  
540 doi: 10.1098/rsif.2015.0024.
- 541 [34] H. W. Hethcote, H. W. Stech, and P. van den Driessche. Periodicity and stability  
542 in epidemic models: a survey. In *Differential equations and applications in ecology,*  
543 *epidemics, and population problems*, pages 65–82. Elsevier, 1981.
- 544 [35] T. D. Hollingsworth, R. M. Anderson, and C. Fraser. HIV-1 Transmission, by Stage of  
545 Infection. *The Journal of Infectious Diseases*, 198(5):687–693, 09 2008. ISSN 0022-1899.  
546 doi: 10.1086/590501. URL <https://doi.org/10.1086/590501>.
- 547 [36] P. Jagers. *Branching Processes with Biological Applications*. Wiley, London, 1975.
- 548 [37] P. Jagers. Stabilities and instabilities in population dynamics. *J. Appl. Prob.*, 29:  
549 770–780, 1992.
- 550 [38] P. Johnson. *adaptivetau: Tau-Leaping Stochastic Simulation*, 2019. URL [https://](https://CRAN.R-project.org/package=adaptivetau)  
551 [CRAN.R-project.org/package=adaptivetau](https://CRAN.R-project.org/package=adaptivetau). R package v. 2.2-3.
- 552 [39] S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes*. Academic Press,  
553 San Diego, 1981.
- 554 [40] D. G. Kendall. On the generalized “Birth-and-Death” process. *Ann. Math. Stat.*, 19  
555 (1):1–15, 1948.
- 556 [41] J. Kevorkian and J. D. Cole. *Multiple Scale and Singular Perturbation Methods*.  
557 Springer-Verlag New York Inc., New York, 1996.
- 558 [42] Aaron A King, Sourya Shrestha, Eric T Harvill, and Ottar N Bjørnstad. Evolution of  
559 acute infections and the invasion-persistence trade-off. *The American Naturalist*, 173  
560 (4):446–455, 2009.

- 561 [43] A. Korobeinikov and G. C. Wake. Lyapunov functions and global stability for SIR,  
562 SIRS, and SIS epidemiological models. *Appl. Math. Lett.*, 15(8):955 – 960, 2002.
- 563 [44] O. Krylova. *Predicting epidemiological transitions in infectious disease dynamics: Small-*  
564 *pox in historic London (1664-1930)*. PhD, McMaster University, Canada, 2011.
- 565 [45] O. Krylova and D. J. D. Earn. Effects of the infectious period distribution on predicted  
566 transitions in childhood disease dynamics. *Journal of the Royal Society of London,*  
567 *Interface*, 10:20130098, 2013. doi: 10.1098/rsif.2013.0098.
- 568 [46] J. Lamperti. The limit of a sequence of branching processes. *Z. Wahrsch. Verw. Gebi-*  
569 *ete*, 7(4):271–288, 1967. doi: 10.1007/BF01844446. URL [https://doi.org/10.1007/](https://doi.org/10.1007/BF01844446)  
570 [BF01844446](https://doi.org/10.1007/BF01844446).
- 571 [47] Zachary Levine and David J. D. Earn. Face masking and COVID-19: Potential effects  
572 of variation on transmission dynamics. *J. R. Soc. Interface*, 19:20210781, 2022. URL  
573 <https://dx.doi.org/10.1098/rsif.2021.0781>.
- 574 [48] W. London and J. A. Yorke. Recurrent outbreaks of measles, chickenpox and mumps. I.  
575 Seasonal variation in contact rates. *American Journal of Epidemiology*, 98(6):453–468,  
576 1973.
- 577 [49] J. Ma and D. J. D. Earn. Generality of the final size formula for an epidemic of a newly  
578 invading infectious disease. *Bulletin of Mathematical Biology*, 68(3):679–702, 2006. doi:  
579 10.1007/s11538-005-9047-7.
- 580 [50] C. McCluskey and D. J. D. Earn. Attractivity of coherent manifolds in metapopulation  
581 models. *J. Math. Biol.*, 62(4):509–541, 2011. doi: 10.1007/s00285-010-0342-z.
- 582 [51] B. Meerson and P. V. Sasorov. WKB theory of epidemic fade-out in stochastic popula-  
583 tions. *Phys. Rev. E*, 80(4):041130, 2009.
- 584 [52] C. E. Mills, J. M. Robins, and M. Lipsitch. Transmissibility of 1918 pandemic influenza.  
585 *Nature*, 432(7019):904–906, 2004.
- 586 [53] C. J. Mode. *Multitype Branching Processes: Theory and Applications*. Elsevier, New  
587 York, 1971.
- 588 [54] D. Mollison. Dependence of epidemic and population velocities on basic parameters.  
589 *Math. Biosci.*, 107:255–287, 1991.
- 590 [55] I. Nåsell. The threshold concept in stochastic epidemic and endemic models. In D. Mol-  
591 lison, editor, *Epidemic Models: Their Structure and Relation to Data*, pages 71–83.  
592 Cambridge University Press, 1995.
- 593 [56] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark. *NIST Handbook of*  
594 *Mathematical Functions*. Cambridge University Press, 2010.
- 595 [57] R. E. O’Malley, Jr. *Singular Perturbation Methods for Ordinary Differential Equations*,  
596 volume 89. Springer, 1991.

- 597 [58] T. L. Parsons. Invasion probabilities, hitting times, and some fluctuation theory for the  
598 stochastic logistic process. *J. Math. Biol.*, 77(4):1193–1231, 2018. ISSN 1432-1416. doi:  
599 10.1007/s00285-018-1250-x. URL <https://doi.org/10.1007/s00285-018-1250-x>.
- 600 [59] T. L. Parsons, A. Lambert, T. Day, and S. Gandon. Pathogen evolution in finite popu-  
601 lations: slow and steady spreads the best. *J. Royal Soc. Interface*, 15(147), 2018. ISSN  
602 1742-5689. doi: 10.1098/rsif.2018.0135. URL [http://rsif.royalsocietypublishing.](http://rsif.royalsocietypublishing.org/content/15/147/20180135)  
603 [org/content/15/147/20180135](http://rsif.royalsocietypublishing.org/content/15/147/20180135).
- 604 [60] Todd L. Parsons and David J. D. Earn. Analytical approximations for the phase  
605 plane trajectories of the SIR model with vital dynamics. preprint, [https://cnrs.](https://cnrs.hal.science/hal-04178969)  
606 [hal.science/hal-04178969](https://cnrs.hal.science/hal-04178969), 2023.
- 607 [61] O. A. van Herwaarden. Stochastic epidemics: the probability of extinction of an infec-  
608 tious disease at the end of a major outbreak. *Journal of Mathematical Biology*, 35(7):  
609 793–813, 1997.
- 610 [62] Z.S. Y. Wong, C. M. Bui, A. A. Chughtai, and C. R. Macintyre. A systematic review  
611 of early modelling studies of Ebola virus disease in West Africa. *Epidemiol. Infect.*, 145  
612 (6):1069–1094, 2017.



## A Computing the Epidemic burnout Probability

To apply Kendall's  $q$  (14) to the problem of epidemic burnout, we need to compute the integral

$$\int_0^\infty e^{-\int_0^t [\beta X(x_{\text{in}}, s) - (\gamma + \mu)] ds} (\gamma + \mu) dt = \int_0^\infty \exp\left(-\int_0^\tau \left[\mathcal{R}_0 X\left(x_{\text{in}}, \frac{\sigma}{\gamma + \mu}\right) - 1\right] d\sigma\right) d\tau, \quad (\text{A.1})$$

where, on the right hand side, we use the mean duration of infection  $(1/(\gamma + \mu))$  as the time unit and write  $\sigma = (\gamma + \mu)s$ ,  $\tau = (\gamma + \mu)t$ . Recalling (17), we can write

$$X(\sigma) \equiv X\left(x_{\text{in}}, \frac{\sigma}{\gamma + \mu}\right) = 1 - (1 - x_{\text{in}})e^{-\varepsilon\sigma}, \quad (\text{A.2})$$

and hence

$$X'(\sigma) = \varepsilon(1 - x_{\text{in}})e^{-\varepsilon\sigma} = \varepsilon(1 - X(\sigma)). \quad (\text{A.3})$$

Now, to evaluate the inner integral in (A.1), we make a change of variables, using  $x$  as the variable of integration:

$$\begin{aligned} \int_0^\tau \left[\mathcal{R}_0 X\left(x_{\text{in}}, \frac{\sigma}{\gamma + \mu}\right) - 1\right] d\sigma &= \int_{X(0)}^{X(\tau)} [\mathcal{R}_0 x - 1] \frac{1}{\frac{dx}{d\sigma}} dx = \int_{x_{\text{in}}}^{X(\tau)} \frac{\mathcal{R}_0 x - 1}{\varepsilon(1 - x)} dx \\ &= -\frac{\mathcal{R}_0}{\varepsilon}(X(\tau) - x_{\text{in}}) - \frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star) \ln \frac{1 - X(\tau)}{1 - x_{\text{in}}}. \end{aligned} \quad (\text{A.4})$$

Changing variables in a similar way, we have

$$\int_0^T \exp\left(-\int_0^\tau \left[\mathcal{R}_0 X\left(x_{\text{in}}, \frac{\sigma}{\gamma + \mu}\right) - 1\right] d\sigma\right) d\tau \quad (\text{A.5a})$$

$$= \int_0^T \exp\left(\frac{\mathcal{R}_0}{\varepsilon}(X(\tau) - x_{\text{in}}) + \frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star) \ln \frac{1 - X(\tau)}{1 - x_{\text{in}}}\right) d\tau \quad (\text{A.5b})$$

$$= \int_{x_{\text{in}}}^{X(T)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x - x_{\text{in}})} \left(\frac{1 - x}{1 - x_{\text{in}}}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} \frac{dx}{\varepsilon(1 - x)}. \quad (\text{A.5c})$$

We are interested in the probability of ultimate extinction, which corresponds to taking the limit as  $T \rightarrow \infty$ , or, equivalently,  $X(T) \rightarrow 1$ , giving us

$$\int_{x_{\text{in}}}^1 e^{\frac{\mathcal{R}_0}{\varepsilon}(x - x_{\text{in}})} \left(\frac{1 - x}{1 - x_{\text{in}}}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} \frac{dx}{\varepsilon(1 - x)} \quad (\text{A.6a})$$

$$= \frac{1}{\varepsilon} e^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_{\text{in}})} \left(\frac{\mathcal{R}_0}{\varepsilon}(1 - x_{\text{in}})\right)^{-\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} \int_0^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_{\text{in}})} e^{-x} x^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star) - 1} dx \quad (\text{A.6b})$$

$$= \frac{1}{\varepsilon} e^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_{\text{in}})} \left( \frac{\mathcal{R}_0}{\varepsilon} (1-x_{\text{in}}) \right)^{-\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} g \left( \frac{\mathcal{R}_0}{\varepsilon} (1-x_*) , \frac{\mathcal{R}_0}{\varepsilon} (1-x_{\text{in}}) \right), \quad (\text{A.6c})$$

where we recall  $g$  denotes the lower incomplete gamma function. Equation (24) follows immediately.

**Asymptotics for Small  $\varepsilon$**  We may also write the integral (A.6a) as

$$\frac{1}{\varepsilon} \int_{x_{\text{in}}}^1 \frac{1}{1-x} e^{\frac{\mathcal{R}_0}{\varepsilon} \left( x - x_{\text{in}} + (1-x_*) \ln \left( \frac{1-x}{1-x_{\text{in}}} \right) \right)} dx = \frac{1}{\varepsilon} \int_{x_{\text{in}}}^1 h(x) e^{\frac{\phi(x)}{\varepsilon}} dx, \quad (\text{A.7})$$

for  $h(x) = \frac{1}{1-x}$  and  $\phi(x) = \mathcal{R}_0 \left( x - x_{\text{in}} + (1-x_*) \ln \left( \frac{1-x}{1-x_{\text{in}}} \right) \right)$ . Assuming  $\varepsilon$  is small, we can apply Laplace's method [7, §6.4]: provided  $x_{\text{in}} \leq x_*$ ,  $\phi(x)$  has its maximum at  $x = x_*$ , so the above is asymptotically equal to

$$\frac{1}{\varepsilon} \sqrt{\frac{2\pi\varepsilon}{|\phi''(x_*)|}} h(x_*) e^{\frac{\phi(x_*)}{\varepsilon}} = \sqrt{\frac{2\pi}{\varepsilon(\mathcal{R}_0-1)}} e^{\frac{\mathcal{R}_0}{\varepsilon}(x_*-x_{\text{in}})} \left( \frac{1-x_*}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)}, \quad (\text{A.8})$$

yielding (25).

**Remark 1.** Note that, since  $x_{\text{in}} < x_*$ ,

$$\begin{aligned} 0 < -\mathcal{R}_0 \int_{x_{\text{f}}}^{x_*} \ln(1-t) dt &= \mathcal{R}_0 \left( (x_* - x_{\text{in}}) + (1-x_*) \ln(1-x_*) - (1-x_{\text{in}}) \ln(1-x_{\text{in}}) \right) \\ &< \mathcal{R}_0 \left( (x_* - x_{\text{in}}) + (1-x_*) \ln(1-x_*) - (1-x_*) \ln(1-x_{\text{in}}) \right) = \phi(x_*), \end{aligned}$$

so the Laplace approximation, and thus the original integral (A.6a), are both exponentially large in  $\varepsilon^{-1}$ .

## B Subsequent epidemic waves

In [60], we derive an iterative scheme to compute “effective initial conditions” for every epidemic wave following initial disease invasion. Writing  $x_{i,j}$  for the fraction susceptible at the start of the  $j^{\text{th}}$  epidemic wave, we find our trajectory approximations agree very closely with the “exact” value obtained by solving the SIR ODEs (5) numerically, starting from the DFE.

Setting  $x_{i,1} = 1$ , we iteratively obtain  $x_{i,j+1}$  from  $x_{i,j}$  by computing

$$x_{\text{f},j} = -x_* W_0 \left( \mathcal{E}(-x_{i,j}/x_*) \right), \quad (\text{B.1a})$$

$$\bar{y}_{0,j} = x_{i,j} - x_* \left( 1 + \ln(x_{i,j}/x_*) \right). \quad (\text{B.1b})$$

$$x_{i,j+1} = 1 + (1-x_*) W_0 \left( \mathcal{E} \left( -\frac{1-x_{\text{f},j}}{1-x_*} \right) \right). \quad (\text{B.1c})$$

The intermediate quantities in this recurrence relation,  $x_{\text{f},j}$  and  $\bar{y}_{0,j}$ , are the final fraction susceptible and maximal fraction infected, respectively, for the SIR model *without* vital dynamics ( $\varepsilon = 0$ ) with initial condition  $(x_{i,j}, 0)$ .

## C The Domain of Applicability of the Approximation (19) to $x_{\text{in}}$

The refined trajectory approximation that yields Equation (19) is derived in [60] under the assumption that  $\mathcal{R}_0$  is large. Despite this, we find that the approximation to  $x_{\text{in}}$  obtained from it (19) performs very well for all but values of  $\mathcal{R}_0$  very close to 1 or very large values of  $\varepsilon > 0$  (see Figure 3.). In particular,  $W_0(x)$  is undefined for  $x < -e^{-1}$ , so we must have

$$-\mathcal{R}_0 e^{-\mathcal{R}_0(1-y_\star)} > -e^{-1}, \quad (\text{C.1})$$

or, expanding and rearranging,

$$\varepsilon < 1 - \frac{\ln \mathcal{R}_0}{\mathcal{R}_0 - 1} = 1 + \frac{x_\star \ln x_\star}{1 - x_\star}. \quad (\text{C.2})$$

Alternately, we can find an approximate lower bound for  $\mathcal{R}_0$ ,

$$\mathcal{R}_0 > e^{2\varepsilon} \quad (\text{C.3})$$

by observing that  $1 - \frac{\ln x}{x-1} \leq \frac{1}{2} \ln x$ . To derive this latter inequality, note that both sides approach a limit of 0 as  $x \rightarrow 1$ , whereas

$$\frac{d}{dx} \left( 1 - \frac{\ln x}{x-1} - \frac{1}{2} \ln x \right) = \frac{1}{(x-1)^2} \left( \ln x - \frac{x^2-1}{2x} \right). \quad (\text{C.4})$$

Again,  $\ln x - \frac{x^2-1}{2x}$  vanishes at  $x = 1$ , whereas

$$\frac{d}{dx} \left( \ln x - \frac{x^2-1}{2x} \right) = -\frac{(x-1)^2}{2x} \leq 0, \quad (\text{C.5})$$

so  $\ln x - \frac{x^2-1}{2x} \leq 0$  for  $x \geq 1$ , and thus  $\frac{d}{dx} \left( 1 - \frac{\ln x}{x-1} - \frac{1}{2} \ln x \right) \leq 0$  also, proving the desired inequality.

## D Boundary Layer Independent Estimates

Thus far, we have computed the burnout probability via a specific, but arbitrary choice of boundary layer  $y_\star$ , and explicit solutions for  $x_{\text{in}}$ , the fraction susceptible when first entering the boundary layer under the ODE approximation (5). Here, we consider an alternative approach, using results from [60] to *implicitly* characterize  $x_{\text{in}}$ . In conjunction with (A.6a), this allows us – at the cost of a small loss of precision – to give expressions for the extinction and persistence probabilities that is independent of the precise choice of threshold, provided the threshold is  $\mathcal{O}(\varepsilon)$ . In addition to be of interest in and of themselves, we use them in Appendix E to compute the value of  $\mathcal{R}_0$  maximizing the burnout probability and also to show how one derives the result of [61] as an approximation to Equation (24).

In [60], we use the method of matched asymptotic expressions [41, 57] to derive analytical approximations to the phase-plane trajectories of the SIR model with vital dynamics, *i.e.*,

696 expressions  $Y(x)$  and  $X(y)$  expressing the density of infected hosts as a function of the  
 697 density of susceptible hosts and vice versa. In the boundary later we obtain lowest and first  
 698 order approximations to  $Y(x)$ : the lowest order approximation is

$$699 \quad Y(x) \approx \bar{y}_0 \left( \frac{1-x_f}{1-x} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x_f-x)}, \quad (\text{D.1})$$

700 whereas the refined estimate is

$$701 \quad Y(x) \approx \left( \frac{1}{x_f} - 1 \right) (x_* - x_f) \left( \frac{1-x_f}{1-x} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{-\frac{\mathcal{R}_0}{\varepsilon}(x-x_f) + \left(\frac{1}{x_f}-1\right)^{-1} \mathcal{Y}_{x_f}^1(1)}, \quad (\text{D.2})$$

702 where

$$703 \quad x_f = -x_* W_0(-\mathcal{R}_0 e^{-\mathcal{R}_0}), \quad (\text{D.3})$$

704 is the final size of the SIR epidemic without vital dynamics [49] and

$$705 \quad \mathcal{Y}_{x_f}^1(1) = \int_{x_f}^1 \left[ \left( \frac{x_*}{t} - 1 \right) \left( \frac{1}{u} - 1 \right) \frac{1}{1-u+x_* \ln u} - \left( \frac{1}{x_f} - 1 \right) \frac{1}{t-x_f} \right] dt \quad (\text{D.4a})$$

$$706 \quad \approx \left( \frac{1}{x_f} - \frac{x_*}{x_* - x_f} \right) \ln x_f - \frac{x_*}{x_* - x_f} \left( \frac{1}{x_f} - 1 \right). \quad (\text{D.4b})$$

707 A very closely related expression (using  $\mu$  rather than  $\varepsilon$  as the small parameter) is derived  
 708 in [61].

709 Recalling that,  $Y(x_{\text{in}}) = y_*$ , evaluating either of (D.1) or (D.2) at  $x = x_{\text{in}}$  gives us a  
 710 relation between  $x_{\text{in}}$ ,  $x_f$ , and  $y_*$ . From the former (D.1), we have

$$711 \quad \left( \frac{1}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{-\frac{\mathcal{R}_0}{\varepsilon}x_{\text{in}}} = \frac{y_*}{\bar{y}_0} \left( \frac{1}{1-x_f} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{-\frac{\mathcal{R}_0}{\varepsilon}x_f}, \quad (\text{D.5})$$

712 whereas the latter (D.2) gives us

$$\left( \frac{1}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{-\frac{\mathcal{R}_0}{\varepsilon}x_{\text{in}}} = \frac{y_*}{(1-x_f) \left( \frac{x_*}{x_f} - 1 \right)} \left( \frac{1}{1-x_f} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{-\frac{\mathcal{R}_0}{\varepsilon}x_f + \left(\frac{1}{x_f}-1\right)^{-1} \mathcal{Y}_{x_f}^1(1)}. \quad (\text{D.6})$$

713 Substituting expression (D.5) into the integral (A.6a) and proceeding as in Appendix A  
 714 gives  
 715

$$716 \quad \int_{x_{\text{in}}}^1 \left( \frac{1-x}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x-x_{\text{in}})} \frac{dx}{\varepsilon(1-x)} \quad (\text{D.7a})$$

$$717 \quad = \frac{y_*}{\bar{y}_0} \int_{x_{\text{in}}}^1 \left( \frac{1-x}{1-x_f} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x-x_f)} \frac{dx}{\varepsilon(1-x)} \quad (\text{D.7b})$$

$$718 \quad = \frac{1}{\varepsilon} \frac{y_*}{\bar{y}_0} \left( \frac{\mathcal{R}_0}{\varepsilon} (1-x_f) \right)^{-\frac{\mathcal{R}_0}{\varepsilon}(1-x_*)} e^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_f)} g \left( \frac{\mathcal{R}_0}{\varepsilon} (1-x_*), \frac{\mathcal{R}_0}{\varepsilon} (1-x_{\text{in}}) \right). \quad (\text{D.7c})$$

719 Now, for  $\frac{a}{z}$  fixed, as  $a \rightarrow \infty$ ,  $g(a, z) \sim \Gamma(a) - z^a e^{-z}$  (see [56, 8.11.6]) and thus

$$720 \quad z^{-a} e^z (g(a, z) - g(a, z')) \sim \left(\frac{z'}{z}\right)^a e^{z-z'} - 1. \quad (\text{D.8})$$

721 Applying this with  $a = \frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)$ ,  $z = \frac{\mathcal{R}_0}{\varepsilon}(1 - x_f)$ , and  $z' = \frac{1 - x_{\text{in}}}{\varepsilon}$ , we see that the error in  
722 replacing  $x_{\text{in}}$  by  $x_f$  in the incomplete gamma function above is equal to

$$723 \quad \frac{1}{\varepsilon} \frac{y_\star}{\bar{y}_0} \left(\frac{1 - x_{\text{in}}}{1 - x_f}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} e^{-\frac{\mathcal{R}_0}{\varepsilon}(x_f - x_{\text{in}})} = \frac{1}{\varepsilon} \frac{y_\star}{\bar{y}_0} e^{\frac{\mathcal{R}_0}{\varepsilon} \left[ (1 - x_\star) \ln\left(\frac{1 - x_{\text{in}}}{1 - x_f}\right) - (x_f - x_{\text{in}}) \right]} \quad (\text{D.9a})$$

$$724 \quad = \frac{1}{\varepsilon} \frac{y_\star}{\bar{y}_0} e^{\frac{\mathcal{R}_0}{\varepsilon} \left[ (1 - x_\star) \ln\left(1 + \frac{x_f - x_{\text{in}}}{1 - x_f}\right) - (x_f - x_{\text{in}}) \right]} \quad (\text{D.9b})$$

$$725 \quad = \frac{1}{\varepsilon} \frac{y_\star}{\bar{y}_0} e^{\frac{\mathcal{R}_0}{\varepsilon} \left[ (x_f - x_{\text{in}}) \frac{x_f - x_\star}{1 - x_f} + \mathcal{O}(\varepsilon^2) \right]}. \quad (\text{D.9c})$$

726 Both  $x_f - x_{\text{in}}$  and  $y_\star$  are  $\mathcal{O}(\varepsilon)$  whereas  $\frac{x_f - x_\star}{1 - x_f}$  is  $\mathcal{O}(1)$ , so this error is  $\mathcal{O}(1)$ . Thus in absolute  
727 terms the error is not small. However, as we observed in Appendix A above, the integral  
728 (D.7a) is exponentially large in  $\varepsilon^{-1}$ , so the error is negligible *relative* to this leading term  
729 (indeed, replacing the incomplete gamma function by  $\Gamma\left(\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)\right)$  produces a similarly  
730 negligible error). We can also replace  $x_{\text{in}}$  by  $x_f$  the Laplace approximation with negligible  
731 error:

$$\int_{x_{\text{in}}}^1 \left(\frac{1 - x}{1 - x_{\text{in}}}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x - x_{\text{in}})} \frac{dx}{\varepsilon(1 - x)} \approx \frac{y_\star}{\bar{y}_0} \sqrt{\frac{2\pi}{\varepsilon(\mathcal{R}_0 - 1)}} \left(\frac{1 - x_\star}{1 - x_f}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x_\star - x_f)}, \quad (\text{D.10})$$

732 Similarly, repeating the same argument using the higher order expression (D.6) gives

$$734 \quad \int_{x_{\text{in}}}^1 \left(\frac{1 - x}{1 - x_{\text{in}}}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x - x_{\text{in}})} \frac{dx}{\varepsilon(1 - x)} \quad (\text{D.11a})$$

$$735 \quad \approx \frac{1}{\varepsilon} \frac{y_\star}{(1 - x_f) \left(\frac{1}{\mathcal{R}_0 x_f} - 1\right)} \left(\frac{\mathcal{R}_0}{\varepsilon}(1 - x_f)\right)^{-\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_f) + \left(\frac{1}{x_f} - 1\right)^{-1} \mathcal{Y}_{x_f}^1(1)} \quad (\text{D.11b})$$

$$\begin{aligned} & \times g\left(\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star), \frac{\mathcal{R}_0}{\varepsilon}(1 - x_f)\right) \\ 736 \quad & \approx \frac{y_\star}{(1 - x_f) \left(\frac{1}{\mathcal{R}_0 x_f} - 1\right)} \sqrt{\frac{2\pi}{\varepsilon(\mathcal{R}_0 - 1)}} \left(\frac{1 - x_\star}{1 - x_f}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x_\star - x_f) + \left(\frac{1}{x_f} - 1\right)^{-1} \mathcal{Y}_{x_f}^1(1)}. \end{aligned} \quad (\text{D.11c})$$

737 Now, we recall that the burnout probability is

$$738 \quad q(x_{\text{in}})^{ny_\star} = \left(1 + \frac{1}{\int_{x_{\text{in}}}^1 \left(\frac{1 - x}{1 - x_{\text{in}}}\right)^{\frac{\mathcal{R}_0}{\varepsilon}(1 - x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x - x_{\text{in}})} \frac{dx}{\varepsilon(1 - x)}}\right)^{-ny_\star} \quad (\text{D.12a})$$

$$= e^{-ny_\star \ln \left( 1 + \frac{1}{\int_{x_{\text{in}}}^1 \left( \frac{1-x}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x-x_{\text{in}})} \frac{dx}{\varepsilon(1-x)}} \right)} \quad (\text{D.12b})$$

$$\approx e^{-\frac{ny_\star}{\int_{x_{\text{in}}}^1 \left( \frac{1-x}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x-x_{\text{in}})} \frac{dx}{\varepsilon(1-x)}}} \quad (\text{D.12c})$$

In Remark 1, we showed that  $\int_{x_{\text{in}}}^1 \left( \frac{1-x}{1-x_{\text{in}}} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x-x_{\text{in}})} \frac{dx}{\varepsilon(1-x)}$  is exponentially large in  $\varepsilon$ , and thus that the error in making the last approximation above is exponentially small.

Substituting any of the expressions (D.7c), (D.10), (D.11b), or (D.11c) for the integral, we see that the terms  $y_\star$  cancel, giving us an approximate expression for the burnout probability that does not depend on the specific choice of threshold, only upon its order of magnitude,  $\varepsilon$ :

$$q(x_{\text{in}})^{ny_\star} \approx e^{-\frac{n\varepsilon\bar{y}_0}{\left( \frac{\mathcal{R}_0}{\varepsilon}(1-x_f) \right)^{-\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_f)} g\left( \frac{\mathcal{R}_0}{\varepsilon}(1-x_\star), \frac{\mathcal{R}_0}{\varepsilon}(1-x_f) \right)} \quad (\text{D.13a})$$

$$\approx e^{-n\bar{y}_0 \sqrt{\frac{\varepsilon(\mathcal{R}_0-1)}{2\pi}} \left( \frac{1-x_f}{1-x_\star} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x_f-x_\star)}} \quad (\text{D.13b})$$

or

$$q(x_{\text{in}})^{ny_\star} \approx e^{-\frac{n\varepsilon(1-x_f) \left( \frac{1}{\mathcal{R}_0 x_f} - 1 \right)}{\left( \frac{\mathcal{R}_0}{\varepsilon}(1-x_f) \right)^{-\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_f) + \left( \frac{1}{x_f} - 1 \right)^{-1} \mathcal{Y}_{x_f}^1(1)} g\left( \frac{\mathcal{R}_0}{\varepsilon}(1-x_\star), \frac{\mathcal{R}_0}{\varepsilon}(1-x_f) \right)} \quad (\text{D.14a})$$

$$\approx e^{-n(1-x_f) \left( \frac{1}{\mathcal{R}_0 x_f} - 1 \right) \sqrt{\frac{\varepsilon(\mathcal{R}_0-1)}{2\pi}} \left( \frac{1-x_f}{1-x_\star} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x_f-x_\star) - \left( \frac{1}{x_f} - 1 \right)^{-1} \mathcal{Y}_{x_f}^1(1)}} \quad (\text{D.14b})$$

respectively.

**Remark 2.** If in Equation (D.14a) we approximate  $g\left(\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star), \frac{\mathcal{R}_0}{\varepsilon}(1-x_f)\right)$  by  $\Gamma\left(\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)\right)$  (i.e., if we approximate the integral up to  $\frac{\mathcal{R}_0}{\varepsilon}(1-x_f)$  by the integral over the whole real line, introducing an error of  $\mathcal{O}(\varepsilon)$ ), we obtain an expression for the burnout probability equivalent to that from [61] (up to minor differences resulting from using different small parameters,  $\mu$  and  $\varepsilon$ ).

## E The $\mathcal{R}_0$ Maximizing the Probability of Burnout

Using the simplified expression for burnout probability (D.13b), we can obtain an approximation to the value of  $\mathcal{R}_0$  that maximizes the probability of burnout linear in  $\varepsilon$  which is highly accurate across the range of values of  $\varepsilon$  for which the burnout probability is non-monotone. The expression (D.13b) is minimized when

$$\bar{y}_0 \sqrt{\frac{\varepsilon(\mathcal{R}_0-1)}{2\pi}} \left( \frac{1-x_f}{1-x_\star} \right)^{\frac{\mathcal{R}_0}{\varepsilon}(1-x_\star)} e^{\frac{\mathcal{R}_0}{\varepsilon}(x_f-x_\star)} \quad (\text{E.1})$$

is maximized, or equivalently, when its partial derivative with respect to  $\mathcal{R}_0$  is equal to zero. Computing the partial derivative and collecting terms of like order in  $\varepsilon$ , we seek  $\mathcal{R}_0$  such

765 that

$$766 \quad \frac{1}{\varepsilon} \left[ \ln \left( \frac{\mathcal{R}_0 + W_0(-\mathcal{R}_0 e^{-\mathcal{R}_0})}{\mathcal{R}_0 - 1} \right) - \frac{1}{\mathcal{R}_0} \right] + \frac{\ln \mathcal{R}_0}{\mathcal{R}_0(\mathcal{R}_0 - 1 - \ln \mathcal{R}_0)} + \frac{\sqrt{\mathcal{R}_0 - 1}}{2} = 0. \quad (\text{E.2})$$

767 An analytical closed form solution does not appear to exist, but one can use a formal asymptotic series expansion  $\mathcal{R}_0 = \sum_{j=0}^{\infty} r_j \varepsilon^j$  to obtain a polynomial approximation in  $\varepsilon$  to arbitrarily large degree (here, we content ourselves with a linear approximation). Substituting  
769 this series into Equation (E.2) and collecting terms of order  $\varepsilon^{-1}$  and order one, we obtain

$$771 \quad \ln \left( \frac{r_0 + W_0(-r_0 e^{-r_0})}{r_0 - 1} \right) - \frac{1}{r_0} = 0 \quad (\text{E.3})$$

$$772 \quad -\frac{1 + (r_0^2 - r_0 + 1)W_0(-r_0 e^{-r_0})}{r_0^2(r_0 - 1)(1 + W_0(-r_0 e^{-r_0}))} r_1 + \frac{\sqrt{r_0 - 1}}{2} + \frac{\ln r_0}{r_0(r_0 - 1 - \ln r_0)} = 0. \quad (\text{E.4})$$

773 We may solve Equation (E.3) by Newton iteration to find the unique root  $r_0 = 2.572629848$ ,  
774 which we use to solve Equation (E.4) to find  $r_1 = -27.71866282$ , giving us the linear approximation  
775

$$776 \quad \arg \max_{\mathcal{R}_0 > 1} q(x_{\text{in}})^{ny_*} \approx 2.572629848 - 27.71866282\varepsilon. \quad (\text{E.5})$$

777 We compare this linear approximation to the numerically determined minimum in Figure 5.

## 778 **F The Burnout Probability is a Decreasing Function** 779 **of $\varepsilon$**

780 In what follows, we show that  $\frac{\partial q(x_{\text{in}})}{\partial \varepsilon} \leq 0$ , from which we conclude that  $q(x_{\text{in}})$  is decreasing  
781 as  $\varepsilon$  increases, for all values of  $\mathcal{R}_0$ . If we set

$$a = \frac{\mathcal{R}_0}{\varepsilon}(1 - x_*) \quad \text{and} \quad z = \frac{\mathcal{R}_0}{\varepsilon}(1 - x_{\text{in}}),$$

782 then we can write Equation (24) as

$$783 \quad q(x_{\text{in}}) = \left( 1 + \frac{\varepsilon}{e^z z^{-a} g(a, z)} \right)^{-1}, \quad (\text{F.1})$$

784 whence

$$785 \quad \frac{\partial q(x_{\text{in}})}{\partial \varepsilon} = -q(x_{\text{in}})^2 \left( \frac{1}{e^z z^{-a} g(a, z)} - \frac{\varepsilon \frac{\partial}{\partial \varepsilon} [e^z z^{-a} g(a, z)]}{e^{2z} z^{-2a} g(a, z)^2} \right). \quad (\text{F.2})$$

786 The first term in the large brackets on the right hand side is always positive, so the result  
787 follows if one can show that  $\varepsilon \frac{\partial}{\partial \varepsilon} [e^z z^{-a} g(a, z)] \leq 0$ . Applying the chain rule gives

$$788 \quad \varepsilon \frac{\partial}{\partial \varepsilon} [e^z z^{-a} g(a, z)] = \varepsilon \frac{\partial z}{\partial \varepsilon} \frac{\partial}{\partial z} [e^z z^{-a} g(a, z)] + \varepsilon \frac{\partial a}{\partial \varepsilon} \frac{\partial}{\partial a} [e^z z^{-a} g(a, z)] \quad (\text{F.3a})$$

$$789 \quad = -\left( z + \mathcal{R}_0 \frac{\partial x_{\text{in}}}{\partial \varepsilon} \right) \frac{\partial}{\partial z} [e^z z^{-a} g(a, z)] - a \frac{\partial}{\partial a} [e^z z^{-a} g(a, z)]. \quad (\text{F.3b})$$

790 Recalling that  $g(a, z) = \int_0^z t^{a-1} e^{-t} dt$ , the latter is equal to

$$791 \quad -\left(z + \mathcal{R}_0 \frac{\partial x_{\text{in}}}{\partial \varepsilon}\right) \left(e^z z^{-a} g(a, z) \left(1 - \frac{a}{z}\right) + \frac{1}{z}\right) - ae^z z^{-a} \left(\int_0^z t^{a-1} e^{-t} \ln t dt - g(a, z) \ln z\right). \quad (\text{F.4})$$

792 Integrating by parts in the rightmost term, this becomes

$$793 \quad -\left(z + \mathcal{R}_0 \frac{\partial x_{\text{in}}}{\partial \varepsilon}\right) \left(e^z z^{-a} g(a, z) \left(1 - \frac{a}{z}\right) + \frac{1}{z}\right) - ae^z z^{-a} \int_0^z \frac{g(a, t)}{t} dt$$

$$794 \quad = -\mathcal{R}_0 \frac{\partial x_{\text{in}}}{\partial \varepsilon} \left(e^z z^{-a} g(a, z) \left(1 - \frac{a}{z}\right) + \frac{1}{z}\right) - e^z z^{-a+1} g(a, z) - 1$$

$$795 \quad - e^z z^{-a} \int_0^z \left(\frac{a}{t} g(a, t) - \frac{a}{z} g(a, z)\right) dt \quad (\text{F.5})$$

797 Now,  $g(a, z) \geq 0$ , whereas  $\frac{a}{z} = \frac{1-x_\star}{1-x_{\text{in}}} \leq 1$ , since  $x_{\text{in}} < x_\star$ , so  $1 - \frac{a}{z} \geq 0$  and, since  $g(a, z)$  is an  
798 increasing function of  $z$ , we have

$$799 \quad \int_0^z \left(\frac{a}{t} g(a, t) - \frac{a}{z} g(a, z)\right) dt \geq \int_0^z \left(\frac{a}{t} - \frac{a}{z}\right) g(a, t) dt \geq 0. \quad (\text{F.6})$$

800 Thus, provided  $\frac{\partial x_{\text{in}}}{\partial \varepsilon} \geq 0$ ,  $\varepsilon \frac{\partial}{\partial \varepsilon} [e^z z^{-a} g(a, z)] \leq 0$ , as required.

801 Finally, from Equation (19), we see that

$$802 \quad \frac{\partial x_{\text{in}}}{\partial \varepsilon} = \lim_{\varepsilon \rightarrow 0} e^{\mathcal{R}_0 y_\star} (E_1(\mathcal{R}_0 y_\star) - E_1(\mathcal{R}_0 \bar{y}_0)) \geq 0, \quad (\text{F.7})$$

803 since  $y_\star \leq \bar{y}_0$  and  $E_1(x)$  is a decreasing function of  $x$ .

## 804 G Simulations

805 **Stochastic simulation algorithm** Exact realizations of the stochastic SIR model (Fig-  
806 ure 2 and Table 1) can be obtained using the standard Gillespie algorithm [25, 26]. If we  
807 denote the various event rates  $a_i$  (*e.g.*,  $a_1 = \mu n$ , *etc.*) then the total event rate is  $a = \sum_i a_i$ .  
808 The time to the next event is drawn from an exponential distribution with mean  $1/a$ , and  
809 the event is taken to be of type  $i$  with probability  $a_i/a$ . This algorithm scales with expected  
810 population size  $n$  and is prohibitively slow when running large numbers of simulations with  
811  $n \gtrsim 10^5$ . We therefore used the adaptive  $\tau$ -leaping approximation [27], as implemented in  
812 the `adaptivetau` R package [38]. The key idea in this approach is to identify, at any point of  
813 the simulation, a time  $\tau$  over which the various event rates can be considered approximately  
814 constant, and then determine the number of events of each type that can be expected over  
815 this time interval. We then “leap forward” by time  $\tau$  rather than treating events individually.

816 **Estimating the Required Number of Simulations** To determine the number of sim-  
817 ulations required to estimate the epidemic burnout probability to a given accuracy, we use  
818 the central limit theorem. Suppose we run  $m$  independent simulations. Let

$$819 \quad \mathbb{1}_i = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ simulation ends in burnout, and} \\ 0 & \text{otherwise.} \end{cases} \quad (\text{G.1})$$



820 Then, the law of large numbers [8, §6] tells us that

$$821 \quad \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \mathbb{1}_i = \mathbb{E}[\mathbb{1}_1] = q, \quad (\text{G.2})$$

822 where  $q$  is Kendall's  $q$  (14). Consequently,  $q_m = \frac{1}{m} \sum_{i=1}^m \mathbb{1}_i$  is an unbiased estimator [8,  
823 p.483] of  $q$ . Let  $\mathcal{E} = q - q_m$  be the error in our estimates. Then, the central limit theorem  
824 [8, §27] tells us that  $\sqrt{m}\mathcal{E} = \frac{1}{\sqrt{m}} \sum_{i=1}^m (\mathbb{1}_i - q)$  converges to a normal distribution with the  
825 same variance as  $\mathbb{1}_1 - q$ , *i.e.*,  $\sqrt{m}\mathcal{E}$  converges in distribution to a normal random variable  
826 with variance

$$827 \quad \sigma^2 = \mathbb{E}[(\mathbb{1}_1 - q)^2] = \mathbb{E}[\mathbb{1}_1^2 - 2q\mathbb{1}_1 + q^2] \quad (\text{G.3a})$$

$$828 \quad = \left(1^2 \cdot q + 0^2 \cdot (1 - q)\right) - 2q\left(1 \cdot q + 0 \cdot (1 - q)\right) + q^2 \quad (\text{G.3b})$$

$$829 \quad = q(1 - q) \leq \frac{1}{4}, \quad (\text{G.3c})$$

830 where the inequality (G.3c) follows because  $0 \leq q \leq 1$ . In particular, for large  $m$ , the  
831 expected squared error is  $\mathbb{E}[\mathcal{E}^2] \lesssim \frac{1}{4m}$ , and thus, to have  $\mathbb{E}[\mathcal{E}^2] \leq \delta$ , we perform at least  
832  $m = \lceil \frac{1}{4\delta} \rceil$  runs.

833 **Fizzle vs. Epidemic burnout** To efficiently distinguish fizzles from epidemic burnout,  
834 we use Equation (15) to estimate a time  $\tau_\delta$  (measured in units of the mean infectious period  
835  $1/(\gamma + \mu)$ ) such that the probability that, starting from  $k$  infected individuals and  $x_i = 1$ ,  
836 a sample path in which infective individuals are still present at time  $\tau_\delta$  eventually fizzles is  
837 less than  $\delta$ . Let  $T_k$  be the (random) time of fizzle starting from  $k$  individuals. Then,

$$838 \quad \mathbb{P}\{T_k > t\} = \mathbb{P}\{Z(t) > 0 \mid Z(0) = k\} \quad (\text{G.4})$$

$$839 \quad = 1 - \left(1 + \frac{1}{\int_0^t e^{-\int_0^s [\beta - (\gamma + \mu)] du} (\gamma + \mu) ds}\right)^{-k} \quad (\text{G.5})$$

$$840 \quad = 1 - \left(1 + \frac{1}{\frac{1}{\mathcal{R}_0 - 1} (1 - e^{-(\beta - \gamma - \mu)t})}\right)^{-k}. \quad (\text{G.6})$$

841 Now, because fizzle is not a certainty,

$$842 \quad \lim_{t \rightarrow \infty} \mathbb{P}\{T_k > t\} = 1 - x_\star^k > 0. \quad (\text{G.7})$$

843 To determine  $\tau_\delta$ , we condition on eventual fizzle to estimate its time of occurrence:

$$844 \quad \mathbb{P}\{T_k > t \mid T_k < \infty\} = \frac{\mathbb{P}\{T_k > t\} - \mathbb{P}\{T_k = \infty\}}{\mathbb{P}\{T_k < \infty\}} \quad (\text{G.8})$$

$$845 \quad = \frac{\left(\frac{1}{\mathcal{R}_0}\right)^k - \left(1 + \frac{1}{\frac{1}{\mathcal{R}_0 - 1} (1 - e^{-(\beta - \gamma - \mu)t})}\right)^{-k}}{\left(\frac{1}{\mathcal{R}_0}\right)^k} \quad (\text{G.9})$$

846

$$= 1 - \left( 1 + \frac{1}{\frac{\mathcal{R}_0}{\mathcal{R}_0 - 1} (1 - e^{-(\beta - \gamma - \mu)t})} \right)^{-k}. \quad (\text{G.10})$$

847 Solving for  $\mathbb{P}\{T_k > \tau_\delta \mid T_k < \infty\} = \delta$  yields

848

$$\tau_\delta = \frac{1}{\mathcal{R}_0 - 1} \ln \left( \frac{(1 - \delta)^{-\frac{1}{k}} - \frac{1}{\mathcal{R}_0}}{(1 - \delta)^{-\frac{1}{k}} - 1} \right). \quad (\text{G.11})$$

849 Choosing a suitably small  $\delta$ , we assume that any sample path in which infective individuals  
850 are still present at  $\tau_\delta$  will not fizzle.