Stochasticity and Heterogeneity in Host-Vector Models: Supplementary Information

By Alun L. Lloyd[1,2], Ji Zhang[1], A. Morgan Root[2]

[1]*Biomathematics Graduate Program and* [2]*Department of Mathematics*
North Carolina State University, Raleigh, NC 27695
email: alun_lloyd@ncsu.edu
tel. (919) 515-1910, fax (919) 515-1909

1. Generating Function and Invasion Probability with a Fixed Duration of Infection

Assuming that the duration of infection is fixed leads to a Poisson distribution for the secondary infections. Such a Poisson distribution has generating function given by $G(s) = \exp(\mu(s-1))$ (Grimmett and Stirzaker, 1992), where μ is the mean number of secondary infections.

In the setting of a directly transmitted infection, equation (3.1) of the main text can be used to find the probability of extinction for the branching process, starting with one infectious individual, which (as it is throughout this paper) is written as s. This gives the transcendental equation

$$s = \exp\{R_0(s-1)\}. \tag{A.1}$$

As usual, s is the smallest non-negative solution of this equation. The probability of a major outbreak is then given by $1-s$.

In the branching process description of the well-mixed host-vector model, the number of secondary infections amongst vectors (respectively hosts) due to an infective host (respectively vector) is Poisson distributed, with mean R_0^{HV} (respectively R_0^{VH}) (Diekmann and Heesterbeek, 2000). Thus an equation of the form

$$s = \exp\left(R_0^{HV}\left[\exp\left\{R_0^{VH}(s-1)\right\} - 1\right]\right) \tag{A.2}$$

or

$$s = \exp\left(R_0^{VH}\left[\exp\left\{R_0^{HV}(s-1)\right\} - 1\right]\right) \tag{A.3}$$

must be solved to find the extinction (and hence invasion) probability if there is initially one infective indivdual present.

2. Multi-type Model: Calculation of Generating Functions

Individuals give rise to secondary infections of type i, for $i = 1, 2, \ldots, n$, according to independent Poisson processes of rates β_i. We write the individuals' numbers of

secondary infections of type i as the random variables X_i. Individuals' infectious periods, denoted by the random variable T, follow some distribution whose average is $\tau = 1/\gamma$. The average number of secondary infections of type i is then given by $R_0^i = \beta_i \tau = \beta_i/\gamma$.

(*a*) *Fixed Duration of Infection*

If infection lasts exactly τ time units, then the X_i are independent and Poisson distributed, with means $\beta_i \tau$. Since the X_i are independent, their joint probability mass function is given by the product

$$P(X_1 = x_1, X_2 = x_2, \dots X_n = x_n) = \prod_{i=1}^{n} \frac{e^{-\beta_i \tau} (\beta_i \tau)^{x_i}}{x_i!}. \tag{A.1}$$

The probability generating function factors into a product of generating functions, each of which is that of the appropriate Poisson distribution

$$G_{X_1, X_2, \dots, X_n}(s_1, s_2, \dots, s_n) = \prod_{i=1}^{n} e^{\beta_i \tau (s_i - 1)} = \prod_{i=1}^{n} e^{R_0^i (s_i - 1)}. \tag{A.2}$$

If the duration of infection T follows some distribution with density function $f(t)$, the joint probability mass function can be found by conditioning on T:

$$\begin{aligned} P(X_1 = x_1, \dots X_n = x_n) &= \int P(X_1 = x_1, X_2 = x_2, \dots X_n = x_n | \, T = t) f(t) \, dt \\ &= \int \left(\prod_{i=1}^{n} \frac{e^{-\beta_i t} (\beta_i t)^{x_i}}{x_i!} \right) f(t) \, dt. \end{aligned} \tag{A.3}$$

(*b*) *Constant Recovery Rates*

In the case where T is exponentially distributed with mean τ, we have that $f(t) = \gamma \exp(-\gamma t)$ and so

$$P(X_1 = x_1, \dots X_n = x_n) = \gamma \left(\prod \frac{\beta_i^{x_i}}{x_i!} \right) \int t^{\sum x_i} e^{-(\gamma + \sum \beta_i) t} \, dt. \tag{A.4}$$

Here, all sums and products are taken over i running from 1 to n. In what follows, whenever the index and range of a sum or product is omitted, it is to be taken that the index is i and runs from 1 to n.

The integral is easily carried out, giving

$$P(X_1 = x_1, \dots X_n = x_n) = \gamma \left(\prod \frac{\beta_i^{x_i}}{x_i!} \right) \frac{(\sum x_i)!}{(\gamma + \sum \beta_i)^{1 + \sum x_i}}. \tag{A.5}$$

The generating function is then

$$\begin{aligned} & G_{X_1,X_2,\ldots,X_n}(s_1,s_2,\ldots,s_n) \\ = & \sum_{x_1,x_2,\ldots,x_n} \left(\prod s_i^{x_i}\right) \gamma \left(\prod \frac{\beta_i^{x_i}}{x_i!}\right) \frac{(\sum x_i)!}{(\gamma+\sum \beta_i)^{1+\sum x_i}} \\ = & \sum_{x_1,x_2,\ldots,x_n} \gamma \left(\prod \frac{(\beta_i s_i)^{x_i}}{x_i!}\right) \frac{(\sum x_i)!}{(\gamma+\sum \beta_i)^{1+\sum x_i}} \\ = & \sum_{x_1,x_2,\ldots,x_n} \gamma \left(\prod (\beta_i s_i)^{x_i}\right) \frac{1}{(\gamma+\sum \beta_i)^{1+\sum x_i}} \binom{\sum x_i}{x_1\ x_2\ \ldots\ x_n} \end{aligned} \quad \text{(A.6)}$$

Here the indices x_j each run over the non-negative integers and the last term is a multinomial coefficient. We rewrite this sum by collecting terms for which $\sum x_i$ takes values 0,1,2,..., and then write $j = \sum x_i$, giving

$$\begin{aligned} & G_{X_1,\ldots,X_n}(s_1,s_2,\ldots,s_n) \\ = & \sum_{j=0}^{\infty} \sum_{x_1,\ldots,x_n} \gamma \left(\prod (\beta_i s_i)^{x_i}\right) \frac{1}{(\gamma+\sum \beta_i)^{1+j}} \binom{j}{x_1\ x_2\ \ldots\ x_n}. \end{aligned} \quad \text{(A.7)}$$

Here, the sum over the $x_1, x_2, \ldots, x_n$ is taken over all non-negative x_i such that $\sum_{i=1}^n x_i = j$. The multinomial theorem then gives

$$\begin{aligned} G_{X_1,\ldots,X_n}(s_1,s_2,\ldots,s_n) &= \sum_{j=0}^{\infty} \frac{\gamma(\sum \beta_i s_i)^j}{(\gamma+\sum \beta_i)^{1+j}} \\ &= \frac{\gamma}{\gamma+\sum \beta_i} \frac{1}{1-\frac{\sum \beta_i s_i}{\gamma+\sum \beta_i}} \\ &= \frac{\gamma}{\gamma+\sum \beta_i - \sum \beta_i s_i} \\ &= \frac{1}{1+\sum_{i=1}^n \beta_i \tau (1-s_i)}. \end{aligned} \quad \text{(A.8)}$$

Finally, since $R_0^i = \beta_i \tau$,

$$G_{X_1,X_2,\ldots,X_n}(s_1,s_2,\ldots,s_n) = \frac{1}{1+\sum_{i=1}^n R_0^i (1-s_i)}. \quad \text{(A.9)}$$

This is the generating function of a multivariate geometric distribution (Griffiths, 1972).

(*c*) *Fixed Durations of Infection*

As discussed above, when durations of infection are fixed, the generating functions $G_{H_i}(s_{H_1},\ldots,s_{H_m},s_{V_1},\ldots,s_{V_n})$ and $G_{V_j}(s_{H_1},\ldots,s_{H_m},s_{V_1},\ldots,s_{V_n})$ can be written as products of simpler generating functions, and equations (4.6) and (4.7) from the main text reduce to

$$G_{H_i V_1}(s_{V_1}) \cdots G_{H_i V_n}(s_{V_n}) = s_{H_i},\ i = 1,\ldots,m \quad \text{(A.10)}$$

$$G_{V_j H_1}(s_{H_1}) \cdots G_{V_j H_m}(s_{H_m}) = s_{V_j},\ j = 1,\ldots,n. \quad \text{(A.11)}$$

Here, $G_{H_iV_j}(s)$ is the generating function describing the number of secondary infections of vector type j that arise from a host of type i, and similarly for $G_{V_jH_i}(s)$. These generating functions take the form $\exp\{R_0(s-1)\}$, with the appropriate $R_0^{H_iV_j}$ or $R_0^{V_jH_i}$.

For the two host, one vector model, equations (A.10 and A.11) become

$$G_{H_1V}(s_V) = s_{H_1} \tag{A.12}$$

$$G_{H_2V}(s_V) = s_{H_2} \tag{A.13}$$

$$G_{VH_1}(s_{H_1})G_{VH_2}(s_{H_2}) = s_V, \tag{A.14}$$

and so the invasion probabilities can be found in terms of the solution of

$$G_{VH_1}(G_{H_1V}(s_V))G_{VH_2}(G_{H_2V}(s_V)) = s_V. \tag{A.15}$$

With the exponential form of the generating functions, this gives

$$\exp\left(R_0^{VH_1}\left[\exp\left\{R_0^{H_1V}(s_V-1)\right\}-1\right]+R_0^{VH_2}\left[\exp\left\{R_0^{H_2V}(s_V-1)\right\}-1\right]\right)=s_V. \tag{A.16}$$

Invasion probabilities obtained from the numerical solution of (A.16) are shown in figure (A.1), together with probabilities obtained from simulation of a corresponding stochastic model. The figures show the probability of a major outbreak following the introduction of a single infective vector, for different values of the parameter γ_1, which specifies the fraction of bites that are made on hosts of type 1. Very good agreement is seen between the two sets of results in each case.

We notice that, if we are in a situation for which invasion is possible, the invasion probabilities in figure A.1 are greater than those in the corresponding figure in the main text, in agreement with the well-known result that invasion probabilities are higher in a model with fixed durations of infection than in the corresponding constant recovery model.

3. Variability About the Endemic Equilibrium

Figure (A.2) presents results that correspond to figure (7) of the main text, but for population sizes that are ten times as large as was used there. This leads to a reduction in the variability seen about the average, and we also see that the moment equations are now able to provide an estimate of variability over the entire range of γ_1 values.

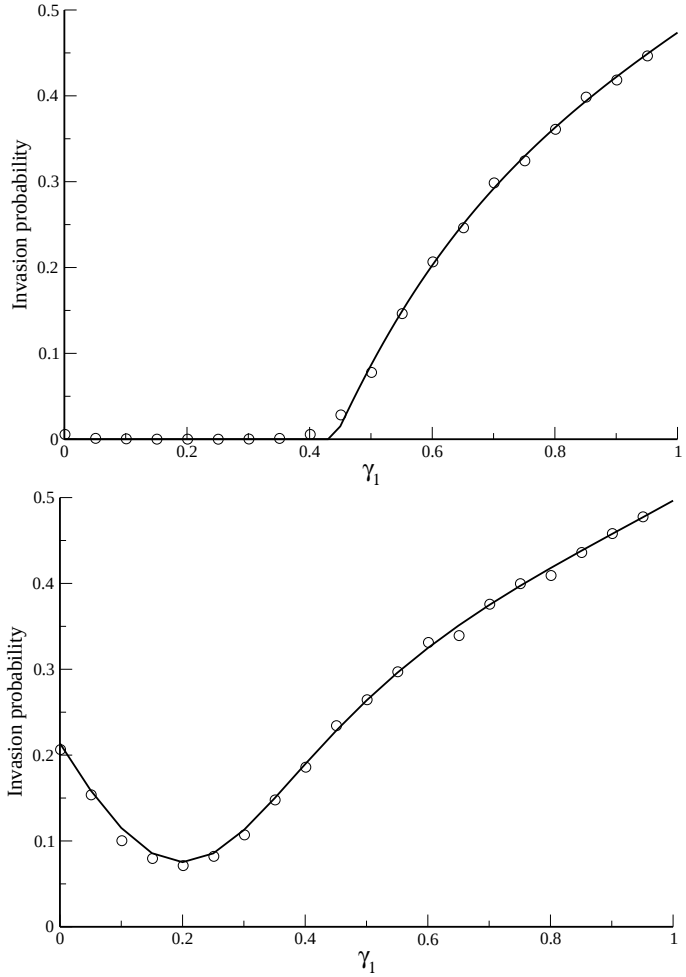

Figure A.1. Major outbreak probability, following the introduction of infection with a single infective vector, in a two host, one vector model, as a function of the vector's preference for the first type of host. A fraction γ_1 of bites are made on hosts of type 1, $\gamma_2 = 1 - \gamma_1$ are made on hosts of type 2. The solid curve denotes the probability as found from the solution of equation (A.16), symbols are probabilities estimated from numerical simulation, using 10000 realizations for each parameter value. For this figure, a constant infectious period assumption was made, with both vector and host being infectious for exactly 7 time units. Parameter values are as follows: $p = 0.2$, $q = 0.15$, $k = 0.5$, $H_1 = 200$, $H_2 = 800$. For the upper panel, $V = 2000$ and for the lower panel, $V = 3000$.

4. Ross Model with Host Immunity

The Ross model assumes that hosts are immediately susceptible upon recovery, i.e. that there is no host immunity. This would correspond to an SIS model in the setting of a directly transmitted infection. The simplest way to account for host immunity is to add a "recovered and immune" class of hosts to the model; we denote the number of such individuals by R. Infective hosts move into this compartment upon

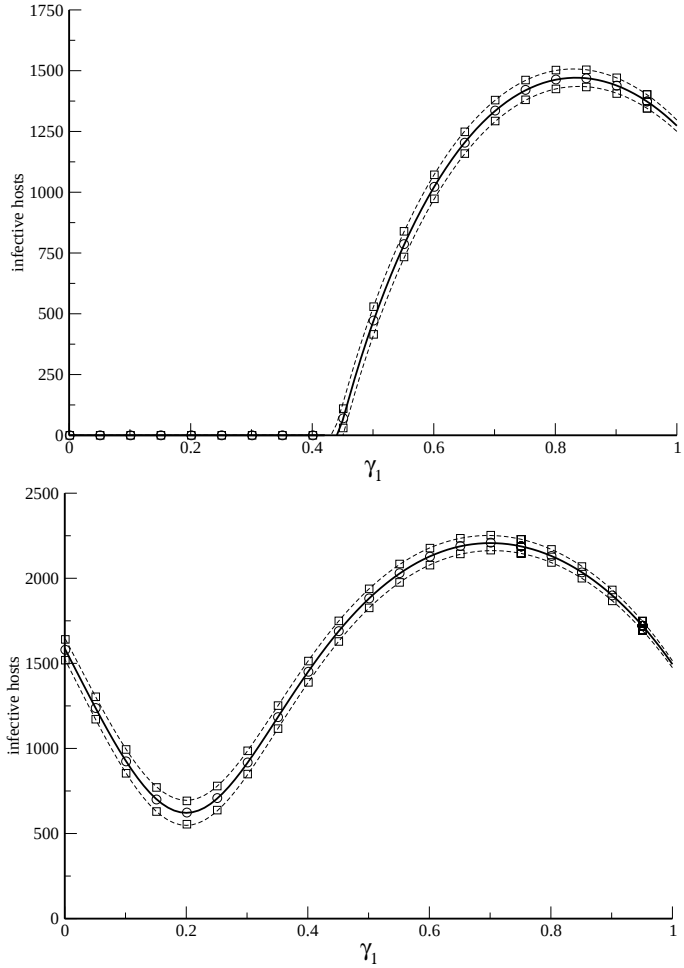

Figure A.2. Average behavior and variation about the average in the two host, one vector model, in terms of the vector's preference for the first host. As in the previous figure, a fraction γ_1 of bites are made on hosts of type 1, $\gamma_2 = 1 - \gamma_1$ are made on hosts of type 2. Solid curves denote the average total number of infective hosts, $E(Y_1 + Y_2)$. The variability of realizations about this average is indicated by the dashed curves, which depict mean $\pm$ standard deviation of the total number of infective hosts. The standard deviation is calculated as the square root of $\mathrm{Var}(Y_1) + \mathrm{Var}(Y_2) + 2\mathrm{cov}(Y_1, Y_2)$. All curves are calculated using moment equations closed by means of the multivariate normal approximation. Symbols denote the corresponding quantities calculated from 1000 realizations of the stochastic model. Parameter values are as in figure (7), except that the host and vector population sizes are ten times as large. The complete set of parameter values is as follows: $p = 0.2$, $q = 0.15$, $k = 0.5$, $\xi = \delta = 1/7$, $H_1 = 2000$, $H_2 = 8000$. For the upper panel, $V = 20000$ and for the lower panel, $V = 30000$.

recovery and then return to the susceptible class as they lose immunity. Thus we have an SIRS model.

We assume that a host loses immunity at constant rate μ, so that $1/\mu$ is the average duration of immunity. As μ tends to infinity, i.e. as the duration of immunity

Event	Transition	Rate at which event occurs
Infection of Host	$S \to S-1,\ Y \to Y+1$	$\alpha IS/H$
Recovery of Host	$Y \to Y-1,\ R \to R+1$	ξY
Host loses immunity	$S \to S+1,\ R \to R-1$	$\mu R = \mu(H-S-Y)$
Infection of Vector	$I \to I+1$	$\beta(V-I)Y/H$
Death of Vector	$I \to I-1$	δI

Table A.1. *Events of the model with host immunity and their rates of occurrence.*

tends to zero, the Ross model is recovered. Notice that we still do not include a description of the demography of the hosts, although this could be added in a straightforward way.

Writing the number of susceptible hosts as S, the stochastic host immunity model is described by the rates given in the table. We remark that our model assumes that the size of the host population remains constant, so that $H = S + Y + R$. Thus R can be rewritten as $H - S - Y$ and so the model need only track the three quantities, S, Y and I.

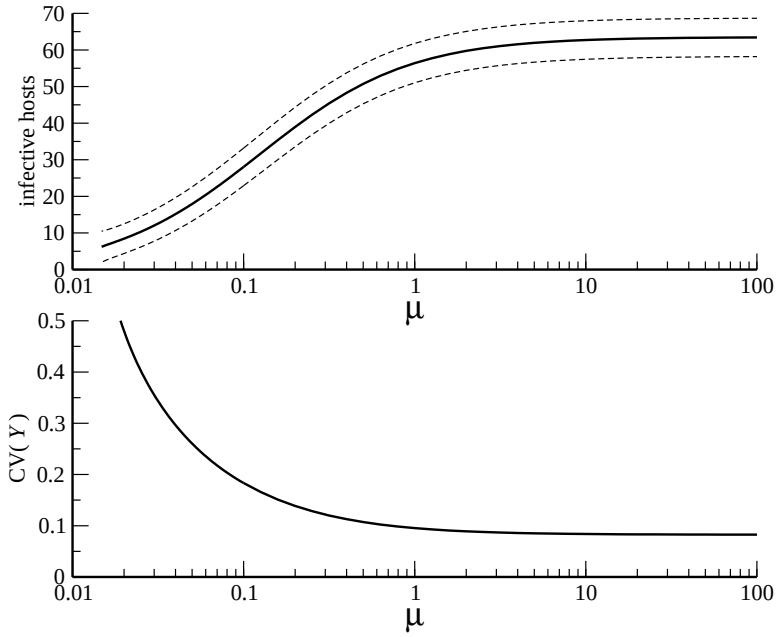

Figure A.3. Moments of the stochastic host immunity model in its quasi-stationary distribution, calculated using the moment equations for a range of values of the rate of loss of host immunity, μ. The upper panel shows the mean number of infective hosts (heavy solid curve) and mean $\pm$ standard deviation (light dashed curves). The lower panel shows the coefficient of variation of the number of infective hosts. Parameter values are as follows: $p = 0.2$, $q = 0.15$, $k = 0.5$, $\xi = \delta = 1/7$, $H = 100$, and $V = 1000$.

Figure (A.3) compares the behavior of the host immunity model for a range of values of the rate, μ, at which immunity is lost. Except for the μ parameter,

this figure employs the same set of parameter values as figure (2) of the main text. Immunity means that individuals take no part in the infection process while in the R compartment, reducing the pool of hosts found in the S and Y classes. As a result, the number of infective hosts, Y, is decreased, with this effect being most pronounced when the rate of loss of immunity, μ, is small (see the upper panel of figure A.3). Consequently, the impact of stochasticity is greater in this model compared to the Ross model, particularly when μ is small (see the lower panel of figure (A.3). Notice that, even though the variance of Y is not seen to change much with μ in the figure, these similar sizes of variation are being seen relative to a much smaller mean as μ becomes small. As mentioned above, large values of μ correspond to the Ross model, and we see that the variability and prevalence of infection observed corresponds to those seen in figure (2).

References

Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical Epidemiology of Infectious Diseases*. Chichester: John Wiley & Son.

Griffiths, D. A. 1972 A bivariate birth-death process which approximates to the spread of a disease involving a vector. *J. Appl. Prob.* **9**, 65–75.

Grimmett, G. R. & Stirzaker D. R. 1992 *Probability and Random Processes*. Oxford University Press.