Journal of Biological Systems, Vol. 27, No. 4 (2019) 1–24

DOI: 10.1142/S0218339019400011

World Scientific
www.worldscientific.com

IDENTIFYING OPTIMAL VACCINATION STRATEGIES VIA ECONOMIC AND EPIDEMIOLOGICAL MODELING

HENRY ZHAO
Department of Economics, Princeton University
Princeton, New Jersey, USA
hz5@princeton.edu

ZHILAN FENG
Department of Mathematics, Purdue University
West Lafayette, Indiana, USA
zfeng@math.purdue.edu

Received 28 February 2019
Accepted 10 May 2019
Published

Vaccination is among the most effective means of preventing and controlling infectious disease outbreaks. Mathematical models can be used to identify the optimal allocation of vaccine among various groups when host populations are heterogeneous. Population heterogeneity may affect individual decision-making and government policy. We show that mixing among sub-populations can profoundly influence the optimal vaccination allocation. Centralized and decentralized programs are examined, accounting for individual behavior and economic constraints. We also compare approaches to modeling transitions between epidemiological classes by epidemiological and economic modelers, and identify key differences.

Keywords: Epidemiological Models; Optimal Vaccine Allocation; Preferential Mixing; Assortative Interaction; Stable Game Equilibrium.

1. Introduction

When mathematical models are used to evaluate vaccination strategies, results have been sensitive to population heterogeneity.[1–4] A quantity that is commonly used to design disease control policies is the basic reproduction number, denoted by $\mathcal{R}_0$, which has the property that outbreaks are possible if and only if $\mathcal{R}_0 > 1$. If programs for disease control are implemented, the corresponding quantity is the effective reproduction number, denoted by $\mathcal{R}_e$. Typically $\mathcal{R}_e < \mathcal{R}_0$, as we consider only effective disease control strategies.

1.1. *Epidemiology*

The meta-population model presented in this section was studied by Feng *et al.*,[1] who demonstrated how population heterogeneity may affect vaccination strategies.

This is an susceptible-infected-recovered (SIR) type of compartmental model for the transmission of pathogens causing infectious diseases when immunity by virtue of either infection or immunization is permanent (but the same framework can be applied to other models including SEIR, SIS, etc.). The meta-population (a population of populations) consists of n subgroups or sub-populations connected by a mixing matrix or network, which describes person-to-person contacts within and between subgroups.

Let $S_i(t)$, $I_i(t)$, $R_i(t)$ denote the numbers of susceptible, infectious, recovered/immune individuals in group i at time t. Then $N_i = S_i + I_i + R_i$ is the total population size of group i, and p_i represents the proportion of group i that is immunized, $i = 1, 2, \ldots, n$. The model reads:

$$\begin{aligned} \frac{dS_i}{dt} &= \mu N_i(1 - p_i) - (\lambda_i(t) + \mu)S_i, \\ \frac{dI_i}{dt} &= \lambda_i(t)S_i - (\gamma + \mu)I_i, \\ \frac{dR_i}{dt} &= \mu N_i p_i + \gamma I_i - \mu R_i, \quad i = 1, 2, \ldots, n, \end{aligned} \tag{1.1}$$

with initial conditions $S_i(0) > 0$, $I_i(0) \geq 0$ and > 0 for some i, $R_i(0) \geq 0$. The parameter γ is the *per capita* recovery rate, μ is the *per capita* birth and death rates, which are assumed to be equal (so that the total population in group i remains constant). The function $\lambda_i(t)$ represents the force of infection for susceptible members of group i given by

$$\lambda_i = \sigma a_i \sum_{j=1}^{n} c_{ij} \frac{I_j}{N_j}, \tag{1.2}$$

where a_i denotes the average number of contacts an individual in sub-population i has per unit of time (which represents the activity level of group i), and σ is the probability of infection per contact with an infectious individual. The parameter c_{ij} is the proportion of the contacts that members of the ith sub-population have with members of the jth, defined by

$$c_{ij} = \epsilon_i \delta_{ij} + (1 - \epsilon_i) f_j, \tag{1.3}$$

where ϵ_i is the fraction of their contacts that members of the ith sub-population reserve for others in their own sub-population (called preference), δ_{ij} is the Kronecker data,

$$f_j = \frac{(1 - \epsilon_j) a_j N_j}{\sum_k (1 - \epsilon_k) a_k N_k}, \tag{1.4}$$

is the contribution of members of the jth sub-population to the pool of unreserved contacts, and a_j is the *per capita* contact rate (called activity). The function c_{ij} allows mixing to range from proportionate (all $\epsilon_i = 0$) to preferential (some $0 < \epsilon_i \leq 1$).[1,2,4–7]

A fundamental difference between models with and without heterogeneity in terms of optimal immunization strategy is illustrated in Feng *et al.*[1] For example, consider the case of $n = 2$ groups. In the absence of heterogeneity (i.e., parameters for the two groups are identical), $p_i = p$ $(i = 1, 2)$, and the basic and effective reproduction numbers for both groups and the meta-population are

$$\mathcal{R}_{0i} = \mathcal{R}_0 = \frac{\sigma a}{\gamma + \mu}, \quad \mathcal{R}_{ei} = \mathcal{R}_e = \mathcal{R}_0(1 - p), \quad i = 1, 2. \tag{1.5}$$

The threshold population immunity is

$$p_{ic} = p_c = 1 - \frac{1}{\mathcal{R}_0}, \tag{1.6}$$

such that $\mathcal{R}_e < 1$ if and only if $p > p_c$, which eventually leads to disease elimination. That is, for disease elimination, the optimal immunization proportions in the two sub-populations are $p_i > p_{ic}$ with p_{ic} being given in (1.6).

When heterogeneity is considered, however, p_{1c} and p_{2c} can differ. In fact, the basic reproduction number for group i is $\mathcal{R}_{0i} = \sigma a_i/(\gamma + \mu)$, and the effective reproduction number for group i is $\mathcal{R}_{ei} = \mathcal{R}_{0i}(1 - p_i)$, $i = 1, 2$. But, for the meta-population, the effective reproduction number is given by

$$\mathcal{R}_e = \frac{1}{2}[A + D + \sqrt{(A - D)^2 + 4BC}], \tag{1.7}$$

where

$$\begin{aligned} A &= \mathcal{R}_{01}c_{11}(1 - p_1), \quad B = \mathcal{R}_{01}c_{12}(1 - p_1), \\ C &= \mathcal{R}_{02}c_{21}(1 - p_2), \quad D = \mathcal{R}_{02}c_{22}(1 - p_2). \end{aligned} \tag{1.8}$$

Any combination of vaccine coverage (p_1, p_2) for which $\mathcal{R}_e(p_1, p_2) < 1$ can lead to elimination of disease. Feng *et al.*[1] showed that there are infinitely many critical combinations (p_{1c}, p_{2c}) that have the property that $\mathcal{R}_e(p_1, p_2) < 1$ for all $p_1 > p_{1c}$ and $p_2 > p_{2c}$, including the homogeneous allocation

$$p_{1c}^* = p_{2c}^* = 1 - \frac{1}{\mathcal{R}_0}. \tag{1.9}$$

The minimum number of vaccine doses required to achieve disease elimination depends on the mixing matrix (c_{ij}). Most interestingly, the homogeneous allocation given in (1.9) may not be the optimal strategy.

The thick (red) curve in Fig. 1 corresponds to $\mathcal{R}_e(p_{1c}, p_{2c}) = 1$. This is for the case of $N_1 = N_2$, so the total vaccine doses $p_1N_1 + p_2N_2 = (p_1 + p_2)N_1$ is minimized if and only if $p_1 + p_2$ is minimized. All points (p_1, p_2) on the dotted line (tangent to the $\mathcal{R}_e = 1$ curve at the star) correspond to same doses, including the homogeneous allocation (p_{1c}^*, p_{2c}^*) (intersection of the dotted and dashed lines). Note that all points on a line perpendicular to the dashed line correspond to equal-dose strategies, and that a line closer to the origin corresponds to a smaller number of doses. Clearly, there is a point $(\tilde{p}_1, \tilde{p}_2)$, marked by a solid dot, at which $\mathcal{R}_e = (\tilde{p}_1, \tilde{p}_2) = 0.85 < 1$ so it provides a more effective strategy than (p_{1c}^*, p_{2c}^*) (equal number of doses but

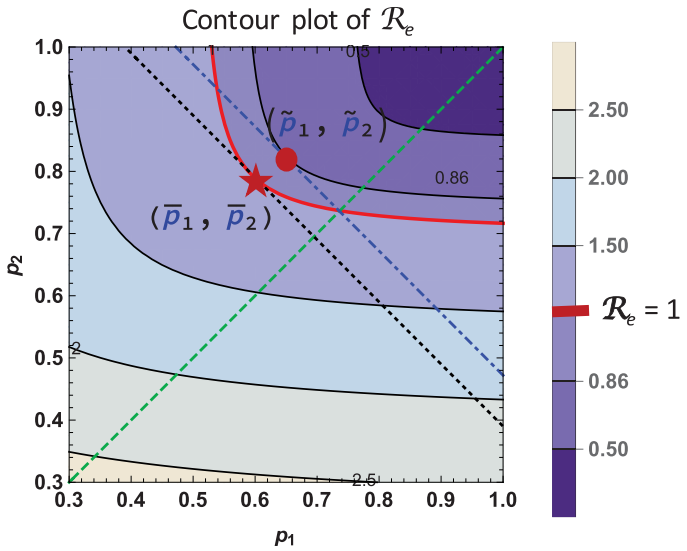


Fig. 1. Contour curves of $\mathcal{R}_e(p_1, p_2)$. At the point marked by a solid dot, $\mathcal{R}_e = 0.85 < 1$. At the point marked by a star, $\mathcal{R}_e = 1$ but the required vaccine doses is smaller than the homogeneous coverage (p^*_{1c}, p^*_{2c}) (the intersection of the dashed and the dot-dashed lines).

a smaller $\mathcal{R}_e$). At the point $(\bar{p}_1, \bar{p}_2)$ marked by the star, $\mathcal{R}_e = 1$, but the sum $\bar{p}_1 + \bar{p}_2 < p^*_{1c} + p^*_{2c}$. Thus, $(\bar{p}_1, \bar{p}_2)$ is also a more effective allocation than (p^*_{1c}, p^*_{2c}) (equal $\mathcal{R}_e$ but a smaller $p_1 + p_2$). Other parameter values used for this figure are: $\beta = 0.05, \gamma = 1/7, \mu = 1/(365 \times 70), a_1 = 8, a_2 = 12, N_1 = N_2 = 500, \epsilon_1 = \epsilon_2 = 0.6$. The time unit is in days.

1.2. *Economics*

Similar studies on the role of mixing in vaccination strategies have been conducted with a focus on Economics, in which non-random mixing is referred to as assortative interactions. For example, Galeotti and Rogers[3] studied vaccination strategies when interactions are positive or negative (i.e., assortative or dis-assortative). In this study, two scenarios are discussed, one centralized and other decentralized. The authors aimed to answer the following questions:

(Q1) What information does the government need to efficiently decide the optimal vaccination policy to eliminate an infection?
(Q2) Should vaccination be allocated randomly or concentrated in one group?
(Q3) How should the planner's action depend on the structure of interactions?

The model considered by Galeotti and Rogers[3] is a system of ODEs for a network consisting of two groups indexed by $g \in \{A, B\}$. Denote the proportion of the non-immunized and infected agents in group g by $\rho_g(t)$. Assume that a proportion π_g

of group g has immunity. The ODE system for $\rho_g(t)$ has the following form:

$$\frac{d\rho_g}{dt} = (1-\rho_g)\nu k\theta_g - \rho_g\delta, \quad g = A, B, \tag{1.10}$$

where ν denotes the per-contact rate of infection, k denotes the number of neighbors that each individual has, δ is the *per capita* rate at which individuals in the infected state I transition to the susceptible state S, and θ_g denotes the probability that a randomly contacted individual in group g is infected, defined by

$$\begin{aligned} \theta_A &= \beta(1-\pi_A)\rho_A + (1-\beta)(1-\pi_B)\rho_B, \\ \theta_B &= \beta(1-\pi_B)\rho_B + (1-\beta)(1-\pi_A)\rho_A. \end{aligned} \tag{1.11}$$

The parameter β denotes the proportion of interactions that are with individuals in the same group while the remaining interactions are with individuals in the other group. It measures the extent of positively ($\beta > 1/2$) or negatively ($\beta < 1/2$) assortative interactions. This parameter plays a similar role as ϵ_i in the mixing matrix (c_{ij}) for model (1.1).

It is shown in Galeotti and Rogers[3] that elimination of disease depends on the quantity

$$\lambda = \frac{\nu k}{\delta},$$

which they define as the intensity of infection. It plays the same role as the basic reproduction number $\mathcal{R}_0$ in model (1.1). They also show that immunization strategies can differ dramatically in the case of positively versus negatively assortative interactions. Note that λ are the same for the two groups.

We point out that the main results presented in Galeotti and Rogers[3] are for the special case when heterogeneity between the two groups in the transmission rate (ν), connectivity or activity (k), assortativity (β), and group size is ignored. In this paper, we study a more general model by incorporating heterogeneity in these factors.

Our paper is organized as follows. In Sec. 2, we present the generalized model of (1.10) and derive the effective reproduction number $\mathcal{R}_v$. We show that $\mathcal{R}_v = 1$ is a threshold for the elimination of disease by considering model equilibria and their stability. Section 2.1 focuses on identification of the optimal vaccine allocation. We discuss the results and compare those from the two frameworks (using assortativity or preferential mixing for interactions between groups).

2. A Generalization of the Model (1.10) and (1.11)

To incorporate heterogeneity, whose importance for disease transmission and control Feng *et al.*[1] demonstrated in model (1.10) and (1.11), we allow group-specific transmission rate (ν_g), connectivity/activity (k_g), assortativity (β_g), and population size (N_g), $g = A, B$. Because of the added heterogeneity, model analysis and

the identification of optimal immunization strategies is more challenging. But it is important to compare results from simpler (1.10) and (1.11) and more general models, and to determine if the more general model provides any new insights in answering the questions **Q1**–**Q3** mentioned in Sec. 1. As the epidemiologic model includes heterogeneity, this generalization will also permit us to compare these parallel literatures.

In models (1.10) and (1.11), it is assumed that the two groups have identical transmission rates ν, activities k, and assortativities β. If we allow these parameters to differ between groups, the model can be written as

$$\frac{d\rho_g}{dt} = (1-\rho_g)\nu_g\theta_g k_g - \rho_g\delta, \quad g = A, B, \tag{2.1}$$

where

$$\begin{aligned} \theta_A &= \beta_A(1-\pi_A)\rho_A + (1-\beta_A)(1-\pi_B)\rho_B, \\ \theta_B &= \beta_B(1-\pi_B)\rho_B + (1-\beta_B)(1-\pi_A)\rho_A. \end{aligned} \tag{2.2}$$

Define

$$\lambda_g = \frac{\nu_g k_g}{\delta}, \quad g = A, B, \tag{2.3}$$

which represents the intensity of infection for group g and can differ between groups if k_g or ν_g are different. All parameters and variables are listed in Table 1.

The pathogen can be eliminated if the disease-free equilibrium, denoted by $E_0 = (0, 0)$, is stable. Let $J(E_0)$ denote the Jacobian matrix of (2.1) at E_0. Then,

$$J(E_0) = B - D, \tag{2.4}$$

where

$$\begin{aligned} B &= \begin{bmatrix} \nu_A k_A \beta_A (1-\pi_A) & \nu_A k_A (1-\beta_A)(1-\pi_B) \\ \nu_B k_B (1-\beta_B)(1-\pi_A) & \nu_B k_B \beta_B (1-\pi_B) \end{bmatrix} \quad \text{and} \\ D &= \begin{bmatrix} \delta & 0 \\ 0 & \delta \end{bmatrix}. \end{aligned} \tag{2.5}$$

Table 1. Definition of variables and parameters used in Eqs. (1.10) and (1.11), where $g = A, B$.

Symbol	Description
$\rho_g(t)$	Proportion of the unimmunized people in group g that are infected
π_g	Proportion of group g immunized (assumed to be independent of t)
k_g	Number of neighbors per person in group g has (or contact number)
ν_g	Infection rate per contact for individuals in group g
θ_g	Probability that a person in group g contacts an infected person
λ_g	$= \nu_g k_g/\delta$. Intensity of infection for group g
δ	*per capita* recovery rate
N_g	Population size of group g
$I_g(t)$	$= (1-\pi_g)N_g\rho_g(t)$. Number of infected at time t in group g

Defining $H = BD^{-1}$,

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \doteq \begin{bmatrix} \lambda_A\beta_A(1-\pi_A) & \lambda_A(1-\beta_A)(1-\pi_B) \\ \lambda_B(1-\beta_B)(1-\pi_A) & \lambda_B\beta_B(1-\pi_B) \end{bmatrix}. \tag{2.6}$$

Then H is a non-negative matrix (all elements are non-negative). From the Perron–Frobenius theorem, all eigenvalues of $J(E_0)$ have negative real parts if and only if the dominant eigenvalue of H is less than 1. Let $\mathcal{R}_e = r(H)$ denote the larger eigenvalue of H. Then

$$\mathcal{R}_e = \frac{1}{2}\big(h_{11} + h_{22} + \sqrt{(h_{11}-h_{22})^2 + 4h_{12}h_{21}}\big), \tag{2.7}$$

where h_{ij} is defined in (2.6) as functions of model parameters including vaccine coverage π_g.

Note that $\mathcal{R}_e < 1$ if and only if the disease-free steady-state E_0 is locally asymptotically stable; and thus, the pathogen will die out. Using the approach of Reluga *et al.*,[8] we can show that $\mathcal{R}_e < 1$ implies global stability of E_0.

In addition, because all entries of H decrease with π_A and π_B, we know from Perron–Frobenius Theorem that $r(H)$ (i.e., $\mathcal{R}_e$) will be reduced when π_A or π_B is increased. Therefore, we can determine the optimal immunization strategy by identifying the combination of (π_A^*, π_B^*) for which $\mathcal{R}_e(\pi_A^*, \pi_B^*) = 1$ with the smallest number of vaccine doses.

2.1. *Disease dynamics and optimal immunization strategy*

Let N_g denote the population size of group g. Then the total number of immunized people, a function of π_g, is

$$M(\pi_A, \pi_B) = \pi_A N_A + \pi_B N_B. \tag{2.8}$$

Thus, an optimal vaccine allocation (π_A^*, π_B^*) is a solution to the following optimization problem:

$$\textbf{Minimize } M(\pi_A, \pi_B) = \pi_A N_A + \pi_B N_B, \quad \textbf{subject to } \mathcal{R}_e(\pi_A, \pi_B) = 1, \tag{2.9}$$

with $0 \le \pi_g \le 1$.

Let $\nabla F = (\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2})$ denote the gradient vector of a scalar function $F(x_1, x_2)$. Note that $\nabla M = (N_A, N_B)$. An optimal solution (π_A^*, π_B^*) can be obtained by solving jointly the equations

$$\begin{aligned} (N_A, N_B) + \omega\nabla\mathcal{R}_e(\pi_A, \pi_B) &= 0, \\ \mathcal{R}_e(\pi_A, \pi_B) &= 1, \end{aligned} \tag{2.10}$$

where ω is a constant (Lagrange multiplier).

2.1.1. *The case of $N_A = N_B$*

System (2.10) can be used to obtain the optimal solution, π_A^* and π_B^*, numerically. Figure 2 illustrates optimal vaccine allocations for several values of β_A and β_B. The values of (λ_A, λ_B) are: $(2, 2.5)$ in (a), (e) and (f), $(3.5, 2)$ in (b) and (c), $(2.5, 3)$ in (d). Population sizes are $N_A = N_B = 500$. Six sets of β_A and β_B values (as indicated in the figures) are used to capture various types of optimal immunization strategies.

Graphically, a solution to the equations in (2.10) is an intersection of the constraint curve $\mathcal{R}_e(\pi_A, \pi_B) = 1$ (represented by the thick curve) with a contour line given by $\pi_A N_A + \pi_B N_B = c$, which is equivalent to a contour line of $\pi_B = -\frac{N_B}{N_A}\pi_A + \frac{c}{N_B}$ (in the unit square), with the smallest c (the blue dashed line). The dot marks the point representing the optimal vaccine allocation (i.e., proportions π_A^* and π_B^* of groups A and B immunized to attain $\mathcal{R}_e = 1$ using the least amount of vaccine).

We observe in Fig. 2 that, depending on the relative strength of transmission in the two groups (λ_A and λ_B) and values for assortativity (β_A and β_B), the optimal immunization strategy (π_A^*, π_B^*), marked by the dot, can appear either on the boundary of the unit square $(\pi_A, \pi_B) \in [0, 1]^2$ or as an interior point with $0 < \pi_g^* < 1$. For example, in (a) and (b), we have $\pi_A^* = 0$ and $\pi_B^* = 0$, respectively. In (c) and (d), we have $\pi_A^* = 1$ and $\pi_B^* = 1$, respectively. These four cases cover various combinations of β_A and β_B including both positively and negatively assortative interactions. In (e) and (f), both β_A and β_B are greater than $1/2$, and $0 < \pi_g^* < 1$ for $g = A, B$. Other parameter values are: $k_1 = 5, k_2 = 3, d = 1/5, N_A = N_B = 500$. From $\lambda_g = \nu_g k_g / d$, the values of ν_g, $g = A, B$, can be determined.

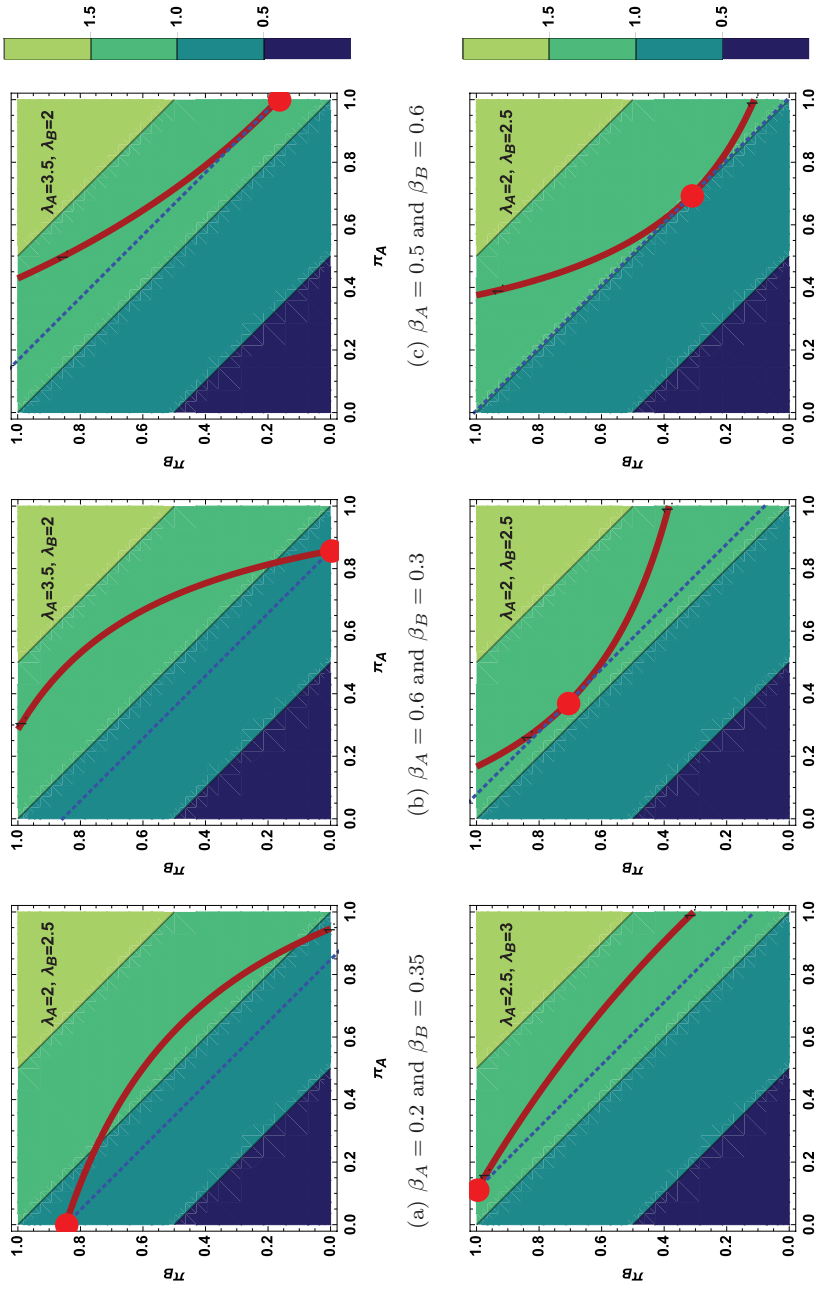
Remark. Unlike the homogeneous case, $\beta_A = \beta_B$, studied by Galeotti and Rogers,[3] there is no longer a single threshold $\beta = 1/2$ to separate optimal immunization strategies for positively or negatively assortative interactions. Also, the optimal allocation for the case $\beta > 1/2$ need not be $\pi_A^* = \pi_B^*$. Thus, heterogeneity substantially complicates the answer to question **Q2**.

2.1.2. *The case of $N_A \neq N_B$*

In Fig. 2, the population sizes for the two groups are equal ($N_A = N_B = 500$). Figure 3 illustrates the role of heterogeneous sub-population sizes ($N_A \neq N_B$) using the same values for other parameters as in Fig. 2(e). Because $\lambda_A < \lambda_B$ and $\beta_A < \beta_B$, the vaccine coverage is lower in group A than group B. However, when $N_B > N_A$, the optimal solution has a higher coverage in group A.

2.1.3. *Numerical exploration of model solutions*

Optimal immunization strategies based on solving the optimization problem (2.9), represented graphically in Fig. 2, identify combinations of immunization proportions



(a) $\beta_A = 0.2$ and $\beta_B = 0.35$ (b) $\beta_A = 0.6$ and $\beta_B = 0.3$ (c) $\beta_A = 0.5$ and $\beta_B = 0.6$

(d) $\beta_A = 0.45$ and $\beta_B = 0.48$ (e) $\beta_A = 0.6$ and $\beta_B = 0.65$ (f) $\beta_A = 0.8$ and $\beta_B = 0.45$

Fig. 2. The optimal solutions π_A^* and π_B^* of the problem (2.9) for system (2.1). Six scenarios corresponding to values of λ_g and β_g, $g = A, B$, are demonstrated in (a)–(f). The optimal solution is marked by the dot, the intersection of the contour curve of $\mathcal{R}_e = 1$ (thick curve) and contour line $\pi_B = -\frac{N_A}{N_B}\pi_A + \frac{c}{N_B}$ with the smallest c (dashed line). Evidently, the optimal strategy can appear either on one of the 4 boundary lines of the unit square (a)–(d) or as an interior point (e) and (f).

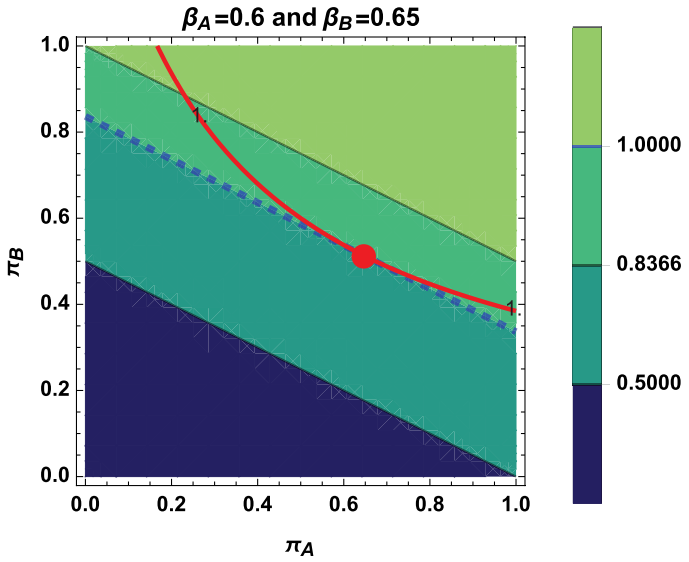


Fig. 3. Similar to Fig. 2(e), but for the case when N_A and N_B differ: $N_A = 500 < N_B = 1000$. However, this figure shows that $\pi_A^* = 0.65 > \pi_B^* = 0.51$. Other parameter values are the same as in Fig. 2(e).

(π_A^*, π_B^*) that reduce $\mathcal{R}_v$ to 1 with minimal vaccine doses. Additional insights can be obtained by examining the numerical solutions of system (2.1). But note first that the effect of immunization coverage π_g on disease prevalence cannot be assessed using the quantity ρ_g, which is the fraction infected among unimmunized people. Consequently, even if ρ_g is large, the number of infected (I_g) could be very small when if the proportion unimmunized was small. The prevalence can be represented by $I_g/N_g = (1 - \pi_g)\rho_g$, $g = A, B$.

Figure 4 illustrates numerical solutions of system (2.1) to compare the effect of various immunization programs on disease prevalence. Other parameter values are chosen to correspond with those in Fig. 2(e). In this case, the optimal strategy to eliminate the pathogen corresponds to the threshold immunization proportions $\pi_A^* = 0.37$ and $\pi_B^* = 0.71$. That is, for all π_g with $\mathcal{R}_v(\pi_A, \pi_B) = 1$, $\pi_A^* + \pi_B^*$ is the smallest, and that $\mathcal{R}_e(\pi_A, \pi_B) < \mathcal{R}_e(\pi_A^*, \pi_B^*)$ for all $\pi_g > \pi_g^*$, $g = A, B$.

For the top panel A, we fix $\pi_B = \pi_B^* = 0.71$ and vary π_A. We change the plots in the fractions infected in group A (I_A/N_A, dot-dashed), group B (I_B/N_B, dotted), and their average (solid) over time. The plots in this panel show that, for $\pi_A < \pi_A^*$, prevalence stabilizes at a positive level, which decreases with π_A (see (A1–A3)), and the pathogen is eliminated when π_A exceeds $\pi_A^* = 0.37$ (see A4). The plots (B1–B3) show solutions for the same values of $\mathcal{R}_e$ as those in (A1–A3), respectively, but with different combinations of π_g. We observe that, while the average prevalence is similar between panels A and B, the endemic levels in the two groups differ considerably.

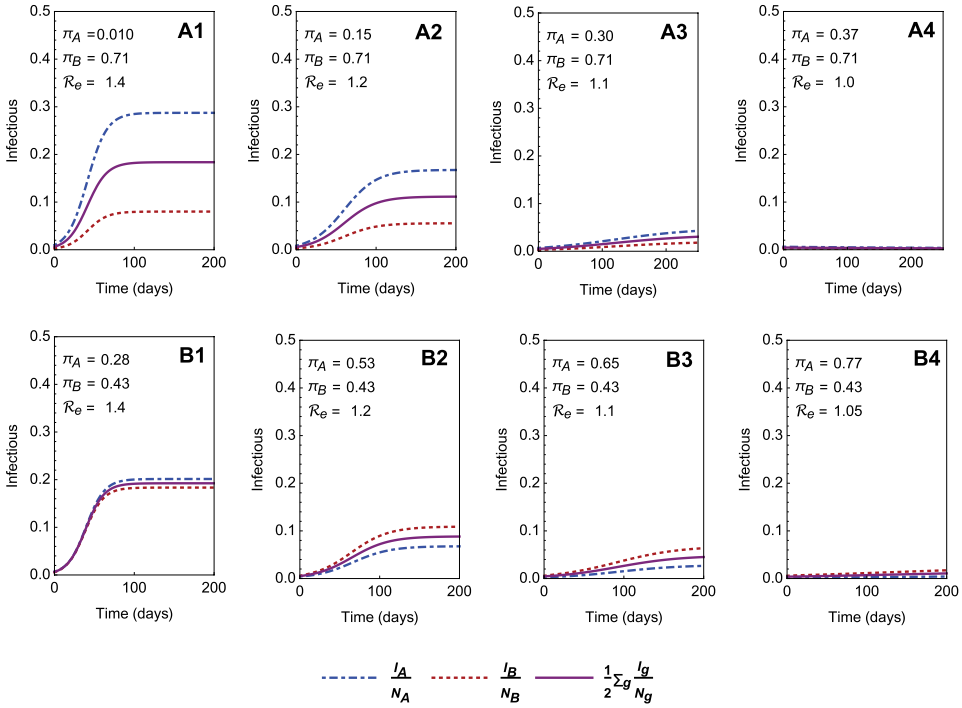


Fig. 4. Numerical solutions of the system (2.1) with parameter values corresponding to Fig. 2(e). The vertical axis represents the proportion of infections $I_g/N_g = (1-\pi_g)\rho_g$, $g = A, B$.

2.2. *Comparison of models with and without heterogeneity*

We include in this section immunization strategies deduced from the simpler model (1.10) and (1.11) using the same analysis; that is, solution of problem (2.9).

First, because groups A and B have identical parameter values except π_g, we can simplify the equation $\mathcal{R}_e(\pi_A, \pi_B) = 1$, which from (2.7) yields:

$$\sqrt{(h_{11}-h_{22})^2 + 4h_{12}h_{21}} = 2 - (h_{11}+h_{22}). \tag{2.11}$$

By squaring both sides of the above equation and simplifying further, we obtain the following equation (note that this may introduce an additional solution):

$$1 - (h_{11}+h_{22}) = -h_{11}h_{22} + h_{12}h_{21}, \tag{2.12}$$

which, in the original notation, is

$$G \doteq [\lambda_A\beta_A(1-\pi_A) + \lambda_B\beta_B(1-\pi_B)] + (1-\pi_A)(1-\pi_B)\lambda_A\lambda_B(1-\beta_A-\beta_B) = 1. \tag{2.13}$$

Using $\lambda_A = \lambda_B = \lambda$ and $\beta_A = \beta_B = \beta$, condition (2.13) reduces to

$$1 - \lambda\beta(2-\pi_A-\pi_B) = (1-\pi_A)(1-\pi_B)\lambda^2(1-2\beta), \tag{2.14}$$

which is Eq. (9) in Galeotti and Rogers.[3]

Then, replacing $\mathcal{R}(\pi_A, \pi_B) = 1$ by $G(\pi_A, \pi_B) = 1$ in (2.10), where G is defined in (2.13), we can determine the optimal solution by solving the system:

$$\begin{aligned} (N, N) + \omega \nabla G(\pi_A, \pi_B) = 0, \\ G(\pi_A, \pi_B) = 1, \end{aligned} \tag{2.15}$$

with $\lambda_A = \lambda_B = \lambda$ and $\beta_A = \beta_B = \beta$. Note that, in this case,

$$\nabla G = (-\lambda\beta - (1 - \pi_B)\lambda^2[1 - 2\beta], -\lambda\beta - (1 - \pi_A)\lambda^2[1 - 2\beta]). \tag{2.16}$$

Using the first equation in (2.15) and eliminating ω, we arrive at the following system, equivalent to (2.15):

$$\begin{aligned} \lambda^2(1 - 2\beta)(\pi_A - \pi_B) = 0, \\ G(\pi_A, \pi_B) = 1. \end{aligned} \tag{2.17}$$

Because $\lambda > 0$, the first equation in (2.17) implies that for all $\beta \neq 1/2$,

$$\pi_A^* = \pi_B^* \doteq x. \tag{2.18}$$

Thus, all optimal solutions satisfying $0 < \pi^* < 1$ can be determined by solving the following equation for x:

$$\lambda\beta(1 - x) + \lambda\beta(1 - x) + (1 - x)(1 - x)\lambda^2(1 - 2\beta) = 1, \tag{2.19}$$

which can be written as

$$a_2 x^2 + a_1 x + a_0 = 0, \tag{2.20}$$

where

$$\begin{aligned} &a_0 = \lambda^2(1 - 2\beta), \quad a_1 = -2\lambda\beta - 2\lambda^2(1 - 2\beta), \\ &a_2 = 2\lambda\beta + \lambda^2(1 - 2\beta) - 1. \end{aligned} \tag{2.21}$$

It can be verified that $a_1^2 - 4a_2a_0 = 4\lambda^2(1 - \beta)^2 \geq 0$. Thus, the quadratic equation (2.20) has two real solutions given by

$$x_\pm = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2a_0}}{2a_2} = \frac{\beta + \lambda(1 - 2\beta) \pm \sqrt{(1 - \beta)^2}}{\lambda(1 - 2\beta)}. \tag{2.22}$$

After simplification, we obtain

$$x_- = 1 - \frac{1}{\lambda}, \quad x_+ = 1 - \frac{1}{\lambda(1 - 2\beta)}. \tag{2.23}$$

Recall that $x_\pm = \pi_A^* = \pi_B^*$ given in (2.23) are possible solutions only for $\beta \neq 1/2$. In addition, we need to ensure that $x_\pm \in [0, 1]$.

It is easy to show that, when $\pi_A = \pi_B = 0$, $\mathcal{R} = \lambda$. This implies that, if $\lambda < 1$, E_0 is stable, so there is no need to vaccinate. Thus, we assume that $\lambda > 1$. Then,

it is clear from (2.23) that

$$0 < x_- < 1 \quad \text{for all } \beta \neq \frac{1}{2}. \tag{2.24}$$

It can also be verified that

$$0 < x_+ < 1 \quad \text{if and only if } \frac{1}{2}\left(1 + \frac{1}{\lambda}\right) < \beta < 1. \tag{2.25}$$

Therefore, for $\frac{1}{2}(1 + \frac{1}{\lambda}) < \beta < 1$, both x_- and x_+ are possible solutions. We then need to determine which is smaller. Note that

$$x_+ - x_- = \frac{2(\beta - 1)}{\lambda(2\beta - 1)} < 0 \quad \text{for } \frac{1}{2}\left(1 + \frac{1}{\lambda}\right) < \beta < 1.$$

That is, $x_+ < x_-$. It follows that x_+ is the optimal solution.

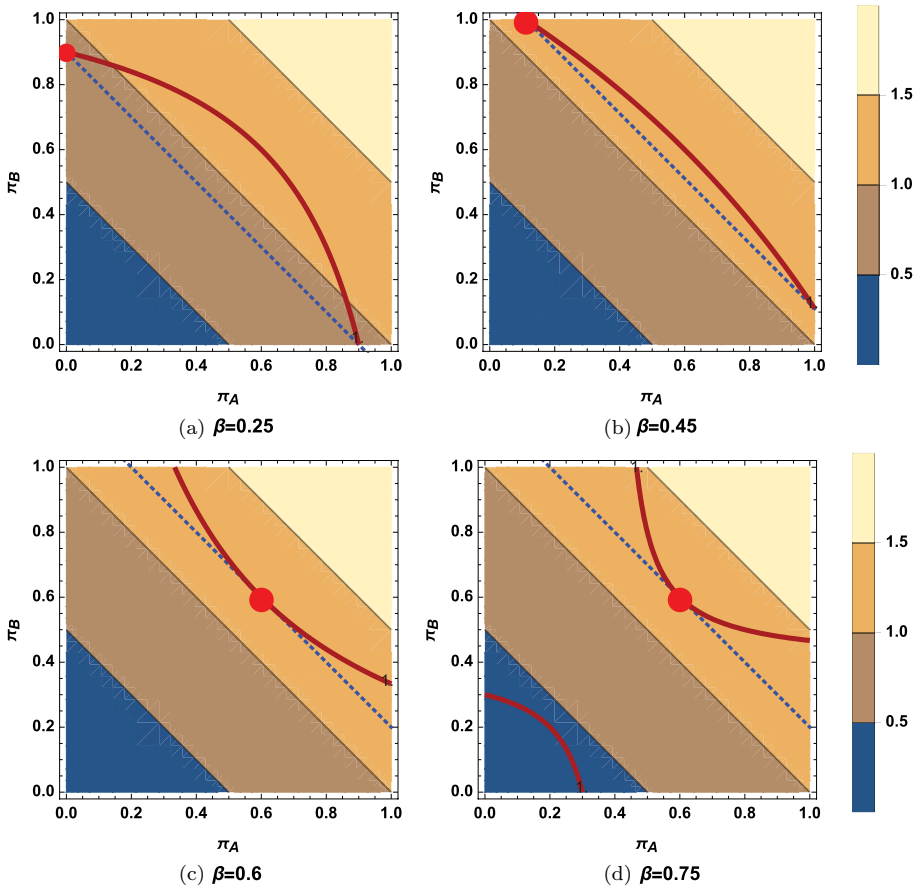


Fig. 5. The optimal solutions (the dots) for β in different intervals are separated by thresholds: (a) and (b) are for $\beta < 1/2$, (c) is for $1/2 < \beta < (1 + 1/\lambda)/2$, and (d) is for $\beta > (1 + 1/\lambda)/2$. Contour lines of the function $\pi_A + \pi_B$ are also plotted. We assumed that $\pi_A \leq \pi_B$. Note the symmetry in all curves and lines due to the lack of heterogeneity.

Figure 5 illustrates the solutions under various conditions. The scenarios presented are based on values of β separated by the following thresholds:

$$\beta_1 = \frac{1}{2}, \quad \beta_2 = \frac{1}{2}\left(1 + \frac{1}{\lambda}\right).$$

The thick curve corresponds to $G(\pi_A, \pi_B) = 1$ and the straight lines are contour lines $\pi_A + \pi_B = c$ for different c values. In (a) and (b), $\beta < \beta_1$, and for the optimal solution, we have either $\pi_A^* = 0$ (see (a)) or $\pi_B^* = 1$ (see (b)). Recall that we assumed that $\pi_A \leq \pi_B$. In (c), $\beta_1 < \beta < \beta_2$, in which case $\pi_A^* = \pi_B^* = x_-$ is the only solution. In (d), both x_- and x_+ are solutions (thus there are two curves corresponding to $G = 1$), but the optimal solution is $\pi_A^* = \pi_B^* = x_+$.

3. Decentralized Games

The previous sections covered the optimal solution for a central government wishing to vaccinate. However, in the absence of such an authority, people will make their own individually optimal choices. This would almost certainly result in societally suboptimal levels of vaccination, because people do not internalize the negative externalities that they impose on others.

Following Galeotti and Rogers,[3] people choose whether or not to be vaccinated before interacting with others. Individuals derive a flow utility of one when healthy and zero when sick. Assuming that people do not have preferences about when they are sick, the utility for each person can be measured by the proportion of time that they are sick. Thus, with cost $c > 0$ of being immunized, individuals who choose to be vaccinated will ensure themselves a utility of $1 - c_g$ because — assuming vaccination is 100% efficacious — they never get sick, but must pay the cost. Alternatively, unvaccinated individuals in group g who spend ρ_g proportion of their time sick will have a utility of $1 - \rho_g$.

3.1. *Characterization of a stable interior game equilibrium*

We show in this section that, depending on parameter values, there can be a unique stable interior equilibrium with $0 < \pi_g < 1$, $g = A, B$, or boundary equilibrium points including $\pi_A \in (0, 1)$, $\pi_B = 1$ and $\pi_A = 1$, $\pi_B \in (0, 1)$. For vaccination to be necessary in group g, assume that

$$\lambda_g(1 - c_g) > 1, \quad g = A, B. \tag{3.1}$$

Let $\Pi^* = (\pi_A^*, \pi_B^*)$ denote an interior equilibrium (i.e., $0 < \pi_g^* < 1$ for $g = A, B$). Similar to Galeotti and Rogers,[3]

$$\pi_g > 0 \implies \rho_g < c_g \quad \text{and} \quad \pi_g < 1 \implies \rho_g > c_g. \tag{3.2}$$

Thus, for π_g^* to be an interior game equilibrium, the endemic equilibrium (ρ_A^*, ρ_B^*) satisfies

$$\rho_g^* = c_g, \quad g = A, B. \tag{3.3}$$

3.1.1. *Existence of an interior game equilibrium* Π^*

For general values of π_g, a disease equilibrium $\rho_g^* \in (0,1)$ of the extended system (2.1) is obtained by setting the right-hand sides of the equations equal to zero, which is equivalent to a positive solution of the following system:

$$\begin{aligned} F(\rho_A,\rho_B) &= (1-\rho_A)\lambda_A(\beta_A(1-\pi_A)\rho_A + (1-\beta_A)(1-\pi_B)\rho_B) - \rho_A = 0,\\ G(\rho_A,\rho_B) &= (1-\rho_B)\lambda_B(\beta_B(1-\pi_B)\rho_B + (1-\beta_B)(1-\pi_A)\rho_A) - \rho_B = 0. \end{aligned} \tag{3.4}$$

To find an interior game equilibrium $\pi_g^* \in (0,1)$, we can solve a similar set of equations as in (3.4) but with $\rho_g^* = c_g$ for $g = A, B$, and with $F(\pi_A,\pi_B)$ and $G(\pi_A,\pi_B)$ being functions of π_g:

$$\begin{aligned} F(\pi_A,\pi_B) &= (1-c_A)\lambda_A(\beta_A(1-\pi_A)c_A + (1-\beta_A)(1-\pi_B)c_B) - c_A = 0,\\ G(\pi_A,\pi_B) &= (1-c_B)\lambda_B(\beta_B(1-\pi_B)c_B + (1-\beta_B)(1-\pi_A)c_A) - c_B = 0. \end{aligned} \tag{3.5}$$

For ease of presentation, introduce the notation

$$\hat{\lambda}_g = \lambda_g(1-c_g), \quad g = A, B, \tag{3.6}$$

and let

$$H_A = \frac{\beta_B}{\hat{\lambda}_A} - \frac{1-\beta_A}{\hat{\lambda}_B}\frac{c_B}{c_A}, \quad H_B = \frac{\beta_A}{\hat{\lambda}_B} - \frac{1-\beta_B}{\hat{\lambda}_A}\frac{c_A}{c_B}. \tag{3.7}$$

Lemma 3.1. *Let* H_g *be as defined in* (3.7). *Assume that* $\hat{\lambda}_g = \lambda_g(1-c_g) > 1$ *and* $\beta_A + \beta_B \neq 1$. *Then a unique interior game equilibrium* (π_A^*, π_B^*) *exists if the following conditions hold*:

(i)

$$\frac{1-\beta_A}{\beta_B} < \frac{c_A\hat{\lambda}_B}{c_B\hat{\lambda}_A} < \frac{\beta_A}{1-\beta_B}, \quad \text{if } \beta_A + \beta_B > 1; \tag{3.8}$$

$$\frac{\beta_A}{1-\beta_B} < \frac{c_A\hat{\lambda}_B}{c_B\hat{\lambda}_A} < \frac{1-\beta_A}{\beta_B}, \quad \text{if } \beta_A + \beta_B < 1; \tag{3.9}$$

(ii)

$$\frac{H_g}{\beta_A + \beta_B - 1} < 1, \quad g = A, B. \tag{3.10}$$

Moreover, the game equilibrium is given by:

$$\pi_A^* = 1 - \frac{H_A}{\beta_A + \beta_B - 1}, \quad \pi_B^* = 1 - \frac{H_B}{\beta_A + \beta_B - 1}. \tag{3.11}$$

Proof. Note that (π_A^*, π_B^*) is a solution of system (3.5), which is linear in $1-\pi_A$ and $1-\pi_B$. The condition $\beta_A + \beta_B \neq 1$ implies that the matrix

$\begin{bmatrix} \beta_A c_A & (1-\beta_A)c_B \\ (1-\beta_B)c_A & \beta_B c_B \end{bmatrix}$ is non-singular. Thus, system (3.5) has a unique solution, $1-\pi_A^*$ and $1-\pi_B^*$, from which we obtain:

$$\begin{aligned} \pi_A^* &= 1 - \frac{\det\begin{bmatrix} c_A/\hat{\lambda}_A & (1-\beta_A)c_B \\ c_B/\hat{\lambda}_B & \beta_B c_B \end{bmatrix}}{\det\begin{bmatrix} \beta_A c_A & (1-\beta_A)c_B \\ (1-\beta_B)c_A & \beta_B c_B \end{bmatrix}} = 1 - \frac{H_A}{\beta_A+\beta_B-1}, \\ \pi_B^* &= 1 - \frac{\det\begin{bmatrix} \beta_A c_A & c_A/\hat{\lambda}_A \\ (1-\beta_B)c_A & c_B/\hat{\lambda}_B \end{bmatrix}}{\det\begin{bmatrix} \beta_A c_A & (1-\beta_A)c_B \\ (1-\beta_B)c_A & \beta_B c_B \end{bmatrix}} = 1 - \frac{H_B}{\beta_A+\beta_B-1}. \end{aligned} \tag{3.12}$$

It can be checked that $\pi_g^* < 1$ if and only if $H_g/(\beta_A+\beta_B-1) > 0$ for $g = A, B$, which is true if and only if (3.8) or (3.9) holds. It can also be verified that $\pi_g^* > 0$ if and only if $H_g/(\beta_A+\beta_B-1) < 1$, which is true if and only if (3.10) holds.

This completes the proof. □

Note that conditions (3.8)–(3.10) only define implicitly regions in the (β_A, β_B) plane where the game equilibrium Π^* exists. These conditions can be described by the following four lines:

$$\begin{aligned} &\text{Line 1.} \quad \beta_B = 1 - Q_1\beta_A; \\ &\text{Line 2.} \quad \beta_B = 1 - Q_1Q_2\beta_A; \\ &\text{Line 3.} \quad \beta_B = Q_1(1-\beta_A); \\ &\text{Line 4.} \quad \beta_B = Q_1Q_3(1-\beta_A), \end{aligned} \tag{3.13}$$

where Q_i $(1 \le i \le 3)$ denote the following constants:

$$Q_1 = \frac{\lambda_A(1-c_A)c_B}{\lambda_B(1-c_B)c_A}, \quad Q_2 = \frac{\lambda_B(1-c_B)-1}{\lambda_A(1-c_A)\frac{c_B}{c_A}-1}, \quad Q_3 = \frac{\lambda_B(1-c_B)\frac{c_A}{c_B}-1}{\lambda_A(1-c_A)-1}. \tag{3.14}$$

More specifically, the region where Π^* exists is the intersection of the regions above (below) the four lines if $\beta_A+\beta_B-1 > 0$ ($\beta_A+\beta_B-1 < 0$). To get a better idea what these regions look like, consider the case when the two groups have the same parameter values except c_g and β_g. In this case, it can be verified that, when $\beta_A+\beta_B-1 > 0$, the following relations hold:

(a) if $c_A > c_B$ then $Q_1 < 1$, $Q_2 > 1$, $Q_3 > 1$, $Q_1Q_2 > 1$, $Q_1Q_3 > 1$;
(b) if $c_A < c_B$ then $Q_1 > 1$, $Q_2 < 1$, $Q_3 < 1$, $Q_1Q_2 < 1$, $Q_1Q_3 < 1$.

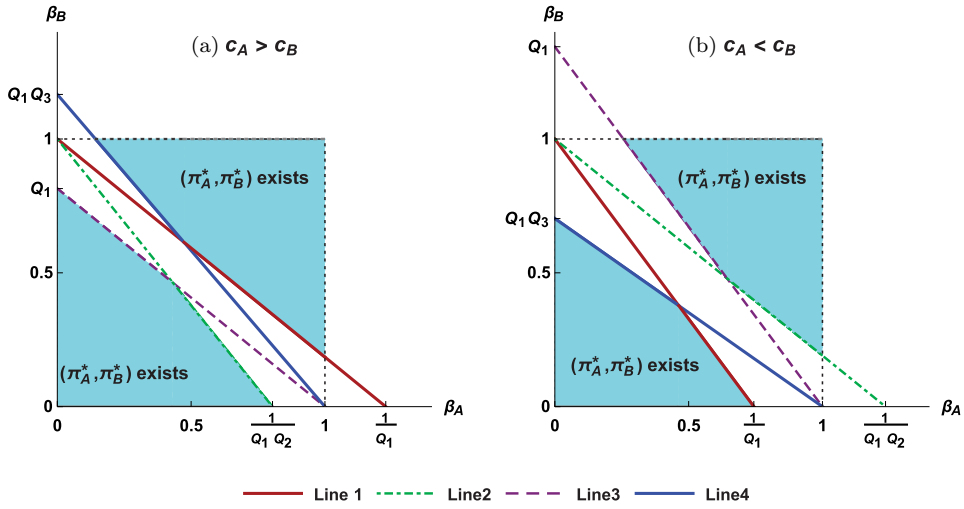


Fig. 6. Depiction of the regions in the (β_A, β_B) plane in which the game equilibrium (π_A^*, π_B^*) exists. The two plots are for the case of (a) $c_A > c_B$ and (b) $c_A < c_B$ with all other parameters having the same values in the two groups. The four lines are defined in (3.13).

When $\beta_A + \beta_B - 1 > 0$, the directions of inequalities involving Q_i are reversed (Fig. 6).

We observed in Fig. 6 that Π^* does not exist for all β_g, and that the size of the region for existence increases as c_A and c_B become similar. The parameter values used are the same as before except that $\nu_A = \nu_B = 0.12$, $k_A = k_B = 4$. In (a) $c_A = 0.12 > c_B = 0.1$, and in (b) $c_A = 0.1 < c_B = 0.13$.

Remark. In the special case, when all parameters are identical in the two groups (except π_g), as in the model of Galeotti and Rogers,[3] $Q_i = 1$ for $i = 1, 2, 3$. Thus, all four lines defined in (3.13) are the same as $\beta_A + \beta_B = 1$ (i.e., $2\beta = 1$ when $\beta_A = \beta_B$). Therefore, Π^* exists for all parameter values except $\beta = 1/2$.

3.1.2. *Stability of* Π^*

For the stability of game equilibrium Π^*, we can examine the derivative of the best response of group B with respect to group A's immunization rate, as done by Galeotti and Rogers,[3] and identify conditions under which

$$\mathcal{P}_B \doteq \left.\frac{d\pi_B^*(\pi_A)}{d\pi_A}\right|_{\Pi^*} > -1. \tag{3.15}$$

Lemma 3.2. *A necessary condition for the interior game equilibrium* $\Pi^* = (\pi_A^*, \pi_B^*)$ *to be stable is the following inequality*:

$$\begin{aligned}(1+\lambda_A\theta_A)(1-\beta_A)(1-\pi_B)c_B((1-\beta_B)c_A - \beta_B c_B)\\ < (1-\beta_B)(1-\pi_A)c_A((1-\beta_A)c_B - \beta_A c_A).\end{aligned} \tag{3.16}$$

Proof. For ease of notation, let $\rho_A(\pi_A)$ and $d\rho_A(\pi_A)/d\pi_A$ denote the following quantities:

$$\begin{aligned}\rho_A(\pi_A) &= \rho_A(\pi_A, \pi_B^*(\pi_A)),\\ \frac{d\rho_A(\pi_A)}{d\pi_A} &= \frac{\partial \rho_A(\pi_A, \pi_B^*(\pi_A))}{\partial \pi_A} + \frac{\partial \rho_A(\pi_A, \pi_B^*(\pi_A))}{\partial \pi_B}\frac{d\pi_B^*(\pi_A)}{d\pi_A}.\end{aligned} \tag{3.17}$$

Differentiating the equation $G(\rho_A, \rho_B) = 0$ in (3.5), evaluating at $\rho_A^* = \rho_A(\pi_A^*)$, and noting that $\rho_B(\pi_A, \pi_B^*(\pi_A)) = c_B$ when π_A is sufficiently close to π_A^*, we have

$$-\beta_B c_B \mathcal{P}_B - (1-\beta_B)\rho_A^* + (1-\beta_B)(1-\pi_A^*)\left.\frac{d\rho_A(\pi_A)}{d\pi_A}\right|_{\Pi^*} = 0. \tag{3.18}$$

Using $\rho_A = \lambda_A\theta_A/(1+\lambda_A\theta_A)$ and the θ_A equation in (2.2), we have

$$\begin{aligned}\left.\frac{d\rho(\pi_A)}{\pi_A}\right|_{\Pi^*} &= \left.\frac{\lambda_A}{(1+\lambda_A\theta_A)^2}\theta_A'(\pi_A)\right|_{\Pi^*}\\ &= \frac{\lambda_A}{(1+\lambda_A\theta_A)^2}\left[-\beta_A c_A + \beta_A(1-\pi_A)\mathcal{P}_B - (1-\beta_A)c_B\left.\frac{d\rho_A(\pi_A)}{d\pi_A}\right|_{\Pi^*}\right].\end{aligned} \tag{3.19}$$

Using (3.18) to eliminate $\left.\frac{d\rho_A(\pi_A)}{d\pi_A}\right|_{\Pi^*}$, we obtain:

$$\mathcal{P}_B = \frac{-(1-\beta_B)(1-\pi_A)\beta_A c_A - (1-\beta_B)c_A\left(\frac{(1+\lambda_A\theta_A)^2}{\lambda_A} - \beta_A(1-\pi_A)\right)}{(1-\beta_A)(1-\beta_B)(1-\pi_A)c_B + \beta_B c_B\left(\frac{(1+\lambda_A\theta_A)^2}{\lambda_A} - \beta_A(1-\pi_A)\right)}. \tag{3.20}$$

To use the condition $\mathcal{P}_B > -1$, we show first that the denominator in (3.20) is positive. It suffices to note that

$$\frac{1+\lambda_A\theta_A}{\lambda_A} = \frac{\theta_A}{\rho_A} = \beta_A(1-\pi_A) + (1-\beta_A)(1-\pi_B)\frac{c_B}{c_A}, \tag{3.21}$$

and thus,

$$\begin{aligned}\frac{(1+\lambda_A\theta_A)^2}{\lambda_A} - \beta_A(1-\pi_A) &= (1+\lambda_A\theta_A)(1-\beta_A)(1-\pi_B)\frac{c_B}{c_A}\\ &\quad + \lambda_A\theta_A\beta_A(1-\pi_A) > 0.\end{aligned} \tag{3.22}$$

From (3.20), using $\mathcal{P}_B > -1$ and the fact that the denominator of (3.20) is positive, we know that a necessary condition for Π^* to be stable is the following inequality:

$$\begin{aligned}&\left[\frac{(1+\lambda_A\theta_A)^2}{\lambda_A} - \beta_A(1-\pi_A)\right]((1-\beta_B)c_A - \beta_B c_B)\\ &\quad < (1-\beta_B)(1-\pi_A)((1-\beta_A)c_B - \beta_A c_A).\end{aligned} \tag{3.23}$$

From (3.22) and (3.23), we arrive at the inequality (3.16).

This completes the proof. □

Remark. The necessary condition (3.16) for the stability of the interior game equilibrium is much simpler in the model considered in Galeotti and Rogers.[3] This is because, when $\lambda_g = \lambda$, $c_g = c$, $\beta_g = \beta$, (3.16) simplifies to:

$$(1 + \lambda\theta)(1 - 2\beta) < (1 - 2\beta), \tag{3.24}$$

which can be true if and only if $1 - 2\beta < 0$ (i.e., $\beta > 1/2$). Thus, a stable interior game equilibrium can be stable only for positive assortative interactions. However, this may not be the case in the extended model (2.1).

Although inequality (3.16) does not provide a simple condition in terms of assortativity, we can examine it numerically (Fig. 7). The contour curve of the function $L(\beta_A, \beta_B)$, which has the property that $L > 0$ is equivalent to the inequality (3.16), where

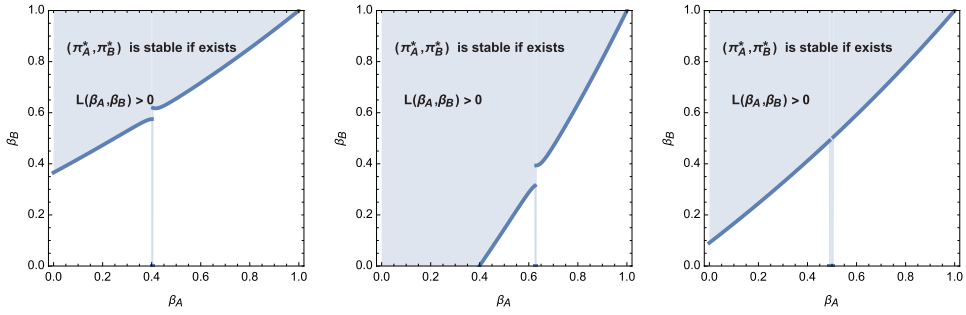
$$\begin{aligned} L(\beta_A, \beta_B) = {} & (1 - \beta_B)(1 - \pi_A)c_A((1 - \beta_A)c_B - \beta_A c_A) \\ & - (1 + \lambda_A\theta_A)(1 - \beta_A)(1 - \pi_B)c_B((1 - \beta_B)c_A - \beta_B c_B). \end{aligned} \tag{3.25}$$

The plots in (a) and (b) use the same parameter values as in Figs. 6(a) and 6(b), respectively, and the plot in (c) is for the case of $c_A = c_B$.

3.1.3. *Effect of immunization on disease equilibrium ρ_g*

When the c_g change, the game equilibrium Π^* may change. Figure 8 shows the dependence of π_A^* and π_B^* on costs c_g. All other parameter values are the same as in Fig. 2(e).

We observe from Figs. 8(a) and 8(b) that, for given c_B, π_A^* decreases with c_A (this is intuitive inasmuch as a higher cost would lead to lower proportion immunized) while π_B^* increases with c_A. Symmetric properties can also be observed (i.e.,



(a) $C_A > C_B$ (b) $C_A < C_B$ (c) $C_A = C_B$

Fig. 7. Plot of the function $L(\beta_A, \beta_B)$ given in (3.25) (the lighter surface) and the constant 1 (the darker plane). The interior game equilibrium Π^* is stable in the region where $L > 0$. The parameter values in (a) are the same as in Fig. 2(e) with $c_A = 0.1$, $c_B = 0.12$. Those in plot (b) are the same except that $c_A = 0.13$, $c_B = 0.1$ and $\nu_1 = \nu_2 = 0.12$. In (c), parameter values are the same as in (a) except that $\nu_A = \nu_B = 0.12$, $k_A = k_B = 5$.

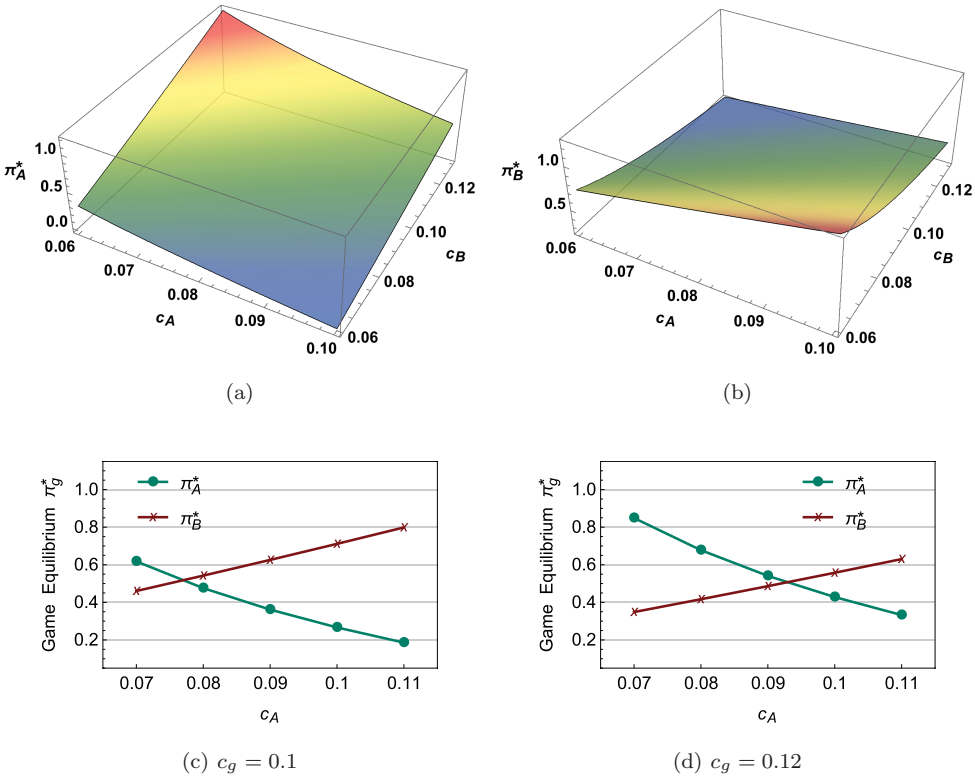


Fig. 8. Plots in (a) and (b) illustrate the stable game equilibrium of immunization π_g^* as a continuous function of the costs c_A and c_B. Plots in (c) and (d) list the values of π_A^* and π_B^* for fixed pairs of (c_A, c_B). Values of other parameters are the same as in Fig. 2(e).

for given c_A, π_A^* increases with c_B while π_B^* decreases with c_B). Plots in (c) and (d) make it easier to compare the relative magnitudes of π_A^* and π_B^* for a fixed pair of (c_A, c_B). For example, plot (c) shows that, for $(c_A, c_B) = (0.1, 0.1)$, $\pi_A^* = 0.27$ while $\pi_B^* = 0.71$. Similarly, plot (d) shows that, for $(c_A, c_B) = (0.08, 0.12)$, $\pi_A^* = 0.68$ while $\pi_B^* = 0.42$.

4. Discussion

We extend the model of Galeotti and Rogers[3] (1.10), who considered two sub-populations that may have different vaccination coverages. Simplifying assumptions in their model include the following: (i) the two sub-populations have the same contact rates (k) and probability of transmission on contact (ν); and (ii) the same level of assortative interactions (β) or preference in mixing with others in the same group. Because of these assumptions, their model is capable of proving only the homogeneous solution for the interior optimal vaccination strategy (π_A^*, π_B^*); i.e., $\pi_A^* = \pi_B^*$. We extend their model by incorporating heterogeneity in several characteristics

known to influence disease transmission and control via immunization, and analyze the resulting model using an approach devised for epidemiological models.

Our model (2.1) is capable of generating heterogeneous interior optimal vaccination strategies, and can produce qualitative and quantitative outcomes very different from the simpler model (1.10). For example, Galeotti and Rogers[3] pointed out that an interior equilibrium solution (π_A^*, π_B^*) of model (1.10) can only happen in symmetric form; i.e., $\pi_A^* = \pi_B^* \in (0, 1)$, and that any asymmetric equilibrium will occur on a boundary; i.e., either $\pi_A = 0$ or $\pi_B = 1$ (assuming that $\pi_A < \pi_B$). In our extended model (2.1), however, there are asymmetric interior equilibria (see Figs. 2(e) and 2(f)).

We show that optimal strategies in the new model are no longer determined solely by the nature (positive or negative) of assortative interactions. Consequently, the answers to questions **Q1**–**Q3** are complicated by additional threshold conditions. Both analytic and numerical analyses of the new model are conducted to investigate how vaccination strategies may depend on both disease transmission parameters and costs associated with vaccination and illness. Optimal vaccination strategies under both centralized vaccination programs and decentralized games are discussed.

An interesting question is how are the models formulated by economical modelers (such as the model (1.10) in Galeotti and Rogers[3]) and the model formulated by epidemiological modelers (such as model the simplest SIS or SIR model) related? If they are significantly different, what are the underlying assumptions in addition to those already pointed out? In Appendix A.1, we tried to recover from the equations for proportions (ρ_g) in (1.10) the equations in terms of the numbers of susceptible (S_g), infected (I_g), and immune (R_g). The result seems to suggest that some biologically unreasonable assumptions have been made in model (1.10) (see the remarks at the end of Appendix A.1). One of the key assumptions is that people do not gain immunity from infections while the immunity induced from vaccination does not wane. However, for almost all vaccine preventable infectious diseases, if vaccination induces an immune response, so must infection. If we allow vaccine-induced and naturally-acquired immunity differ (e.g., in the degree or duration of protection), there should be two separate classes, e.g., V (for vaccinated and immune) and R (for recovered with temporary immunity). This is demonstrated in Appendix A.2. Clearly the modified model (A.8) based on a SIR model is more biologically reasonable than model (2.1) based on a SIS model. One of the main differences between the two systems will be in the analysis that involves the endemic equilibrium, such as the effect of vaccination on disease prevalence or the game equilibrium for the decentralized game. More detailed analyses for system (A.8) or a more general model will be presented in future work.

Analysis of System (2.1) suggests that the optimal vaccination strategies (π_A^*, π_B^*) under centralized and decentralized cases may depend on the cost function. If so, this will illustrate the importance of melding Economics and Epidemiology. We have begun studying this and will publish results elsewhere.

Acknowledgments

We thank the reviewers for comments and suggestions that helped us to improve the presentation of our work. ZF's research is partially supported by NSF Grant DMS-1814545.

References

1. Feng Z, Hill AN, Smith PJ, Glasser JW, An elaboration of theory about preventing outbreaks in homogeneous populations to include heterogeneity or preferential mixing, *J Theor Biol* **386**:177–187, 2015.
2. Feng Z, Hill AN, Curns AT, Glasser JW, Evaluating targeted interventions via meta-population models with multi-level mixing, *Math Biosci* **287**:93–104, 2017.
3. Galeotti A, Rogers BW, Strategic immunization and group structure, *Am Econ J Microecon* **5**(2):1–32, 2013.
4. Glasser JW, Feng Z, Omer SB, Smith PJ, Rodewald LE, The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: A modelling study, *Lancet Infect Dis* **16**(5):599–605, 2016.
5. Busenberg S, Castillo-Chavez C, A general solution of the problem of mixing of sub-populations and its application to risk- and age-structured epidemic models for the spread of AIDS, *IMA J Math Appl Med Biol* **8**:1–29, 1991.
6. Jacquez JA, Simon CP, Koopman J, Sattenspiel L, Perry T, Modeling and analyzing HIV transmission: The effect of contact patterns, *Math Biosci* **92**:119–199, 1988.
7. Nold A, Heterogeneity in disease transmission modeling, *Math Biosci* **124**:59–82, 1980.
8. Reluga TC, An SIS epidemiology game with two subpopulations, *J Biol Dyn* **3**(5):515–531, 2009.

Appendix A. Comparison of Modeling Approaches

Epidemiological modelers primarily use the modeling framework represented in model (1.1) whereas economic modelers use the one represented in model (1.10). We illustrate here how models formulated using these approaches may be related.

A.1. *An SIS model with immunized proportions π_g fixed at $t = 0$*

Note that the proportion of unimmunized individuals that are infectious and susceptible at time t can be written as

$$\rho_g(t) = \frac{I_g(t)}{(1-\pi_g)N_g}, \quad 1-\rho_g(t) = \frac{S_g(t)}{(1-\pi_g)N_g}, \quad g = A, B, \tag{A.1}$$

with π_g and N_g being independent of t. Then from the ρ_g and θ_g equations in (1.10) and (1.11), we have

$$\frac{I_A'}{(1-\pi_A)N_A} = \frac{S_A}{(1-\pi_A)N_A}\nu k\left(\beta\frac{I_A}{N_A} + (1-\beta)\frac{I_B}{N_B}\right) - \delta\frac{I_A}{(1-\pi_A)N_A}. \tag{A.2}$$

Using (A.1), we can simplify (A.2) to get

$$\frac{dI_A}{dt} = S_A\nu k\left(\beta\frac{I_A}{N_A} + (1-\beta)\frac{I_B}{N_B}\right) - \delta I_A. \tag{A.3}$$

Using the ρ_B and θ_B equations in (1.10) and (1.11), we can get a similar equation for I_B'. Thus, the model in terms of number of individuals consists of the following equations:

$$\begin{aligned} \frac{dI_A}{dt} &= S_A \nu k \left(\beta \frac{I_A}{N_A} + (1-\beta)\frac{I_B}{N_B}\right) - \delta I_A, \\ \frac{dI_B}{dt} &= S_B \nu k \left(\beta \frac{I_B}{N_B} + (1-\beta)\frac{I_A}{N_A}\right) - \delta I_B, \\ I_g + S_g &= (1-\pi_g)N_g, \quad g = A, B. \end{aligned} \tag{A.4}$$

Because π_g and N_g are constants, system (A.4) implies the initial condition:

$$S_g(0) + I_g(0) = (1-\pi_g)N_g, \quad g = A, B. \tag{A.5}$$

That is, the immunization proportion π_g is set at the initial time.

Remarks. These remarks are based on (A.4) and (A.5). (i) When vaccination is considered, epidemiological models generally include a vaccinated or immune class, typically denoted by V or R. Depending on how vaccination is implemented (e.g., at birth or susceptible individuals at all time), the V equation may not be needed. In any event, there must be at least two differential equations for the three variables S, I, and V plus a constraint (e.g., the total population is constant) for each group. In the case of two groups, at least four differential equations plus two constraints are needed. However, system (A.4) for two groups has only two differential equations and two constraints. The reason is that, in addition to a constant total population size for each group, it is also assumed that π_g and thus $S_g + I_g$ remain constant for each group. Consequently, the proportions immune, V_g, also remain constant. (ii) The model assumes that new infections are only from the S class and that infected people will return to the susceptible class after recovery. If vaccinated people can gain permanent immunity, should people gain any immunity from an infection? (iii) Note that $S_g + I_g = (1-\pi_g)N_g$. Thus, $S_g' = -I_g'$. Therefore, System (A.4) does not include birth or death terms. This suggests that the model is appropriate only for dynamics in the short-term.

A.2. *An SIR type model with temporary immunity from infection*

In model (A.4), all infected people will return to the susceptible class after recovery. If we assume that infections can induce some temporary immunity after recovery, we need to introduce another class, denoted by R_g (different from V_g), and consider a SIR model describing the disease dynamics among unimmunized people. Because $S_g + I_g + R_g = (1-\pi_g)N_g$ is constant, we can ignore the R_g equation by using $R_g = (1-\pi_g)N_g - S_g - I_g$. Let τ denote the *per capita* rate of immunity loss, then

the model reads:

$$\begin{aligned}
\frac{dS_A}{dt} &= -S_A\nu_A k_A\left(\beta_A\frac{I_A}{N_A}+(1-\beta_A)\frac{I_B}{N_B}\right)+\tau R_A,\\
\frac{dI_A}{dt} &= S_A\nu_A k_A\left(\beta_A\frac{I_A}{N_A}+(1-\beta_A)\frac{I_B}{N_B}\right)-\delta I_A,\\
\frac{dS_B}{dt} &= -S_B\nu_B k_B\left(\beta_B\frac{I_B}{N_B}+(1-\beta_B)\frac{I_A}{N_A}\right)+\tau R_B,\\
\frac{dI_B}{dt} &= S_B\nu_B k_B\left(\beta_B\frac{I_B}{N_B}+(1-\beta_B)\frac{I_A}{N_A}\right)-\delta I_B,\\
S_g+I_g+R_g &= (1-\pi_g)N_g,\quad g=A,B.
\end{aligned} \tag{A.6}$$

We now need to introduce a new notation for the proportion of unimmunized people who are susceptible, which we denote by η_g, $g = A, B$. Then

$$\eta_g(t)=\frac{S_g(t)}{(1-\pi_g)N_g},\quad \rho_g(t)=\frac{I_g(t)}{(1-\pi_g)N_g},\quad g=A,B. \tag{A.7}$$

From the equations in (A.6), we obtain the following system for ρ_g and η_g:

$$\begin{aligned}
\frac{d\eta_g}{dt} &= -\eta_g\nu_g k_g\theta_g+\tau(1-\eta_g-\rho_g),\\
\frac{d\rho_g}{dt} &= \eta_g\nu_g k_g\theta_g-\delta\rho_g,\quad g=A,B,
\end{aligned} \tag{A.8}$$

where θ_g denote the same expressions as in (2.2), i.e.,

$$\begin{aligned}
\theta_A &= \beta_A(1-\pi_A)\rho_A+(1-\beta_A)(1-\pi_B)\rho_B,\\
\theta_B &= \beta_B(1-\pi_B)\rho_B+(1-\beta_B)(1-\pi_A)\rho_A.
\end{aligned} \tag{A.9}$$