



Asymptotic Analysis of Optimal Vaccination Policies

Matthew J. Penn¹  · Christl A. Donnelly^{1,2}

Received: 14 June 2022 / Accepted: 24 December 2022

© The Author(s) 2023

Abstract

Targeted vaccination policies can have a significant impact on the number of infections and deaths in an epidemic. However, optimising such policies is complicated, and the resultant solution may be difficult to explain to policy-makers and to the public. The key novelty of this paper is a derivation of the leading-order optimal vaccination policy under multi-group susceptible–infected–recovered dynamics in two different cases. Firstly, it considers the case of a small vulnerable subgroup in a population and shows that (in the asymptotic limit) it is optimal to vaccinate this group first, regardless of the properties of the other groups. Then, it considers the case of a small vaccine supply and transforms the optimal vaccination problem into a simple knapsack problem by linearising the final size equations. Both of these cases are then explored further through numerical examples, which show that these solutions are also directly useful for realistic parameter values. Moreover, the findings of this paper give some general principles for optimal vaccination policies which will help policy-makers and the public to understand the reasoning behind optimal vaccination programs in more generic cases.

Keywords Vaccination · Epidemiology · Epidemics · SIR modelling

1 Introduction

The trajectory of an epidemic can be dramatically changed by the implementation of a vaccination program, as has been shown in the case of COVID-19 (Bloom et al. 2021). These vaccination programs are most effective when they target specific groups in a

✉ Matthew J. Penn
matthew.penn@st-annes.ox.ac.uk

Christl A. Donnelly
christl.donnelly@stats.ox.ac.uk

¹ Department of Statistics, University of Oxford, St Giles', Oxford OX1 3LB, UK

² Department of Infectious Disease Epidemiology, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

population (Fitzpatrick and Galvani 2021), although the optimal targeting strategy is dependent on the properties of the disease and vaccine (Moore et al. 2021). Thus, it is important to have robust methods to determine the optimal strategy whenever a new epidemic emerges.

In recent years, the epidemiological literature has grown rapidly, and a wide range of models have been developed and analysed. These include branching-process models (Pakkanen et al. 2021); network-based models (Bedson et al. 2021); and machine-learning-based models (Muhammad et al. 2021), among many others (Brauer et al. 2019).

However, despite these innovations, compartmental models, where the population is split into a number of subgroups and disease transmission is modelled by a system of differential equations (Abou-Ismaïl 2020), remain a popular choice for epidemiologists and have been widely used for modelling the COVID-19 pandemic (Kong et al. 2022). As discussed in (Kong et al. 2022), a number of different compartment structures have been used, while many authors have also sought to model the effect of government interventions and quarantining procedures (Vardavas et al. 2021; de Camino-Beck 2020; Adhikari et al. 2020).

One such compartmental model that is widely used (Ram and Schaposnik 2021; Acemoglu et al. 2021; Kuniya 2019) is the multi-group SIR (susceptible–infected–recovered) model. This is an extension of the classical SIR model (Kermack and McKendrick 1927) and has been used to model a range of diseases such as measles (Sattenspiel and Dietz 1995), influenza (Brauer 2008) and COVID-19 (Ellison 2020). It provides a general framework with which to assess the effectiveness of different vaccination policies, while also remaining mathematically tractable, allowing theorems about its behaviour to be rigorously proved (Penn and Donnelly 2022). It splits a population up into a number of interconnected subgroups (such as age groups (Longini Jr et al. 1978) and captures the different transmission dynamics between each group. This construction highlights the dual benefit that vaccination can have—vaccines that are infection-reducing directly protect the individuals that are vaccinated while transmission-reducing vaccines can also indirectly protect unvaccinated individuals (Eichner et al. 2017).

This dual benefit can significantly complicate the optimal vaccination problem when there is a negative correlation between the infectiousness of a group and the vulnerability of its members to the disease. Examples of this occur when the population is divided by age for diseases such as COVID-19 (Miura et al. 2021) and seasonal influenza (Molinari et al. 2007). In such cases, the optimal strategy may not be obvious and could be highly dependent on uncertain parameters (Saadi et al. 2021), while the seemingly intuitive solution may be significantly sub-optimal (Delmas et al. 2021). Moreover, the complicated methods used to find the optimal solution, involving solving the adjoint equations derived via Pontryagin's maximum principle (Boutayeb et al. 2021; Lee et al. 2012), mean that the optimal solution may be difficult to understand or qualitatively justify to policy-makers.

When attempting to understand a complicated problem such as finding the optimal vaccination policy, it is often helpful to look at cases with extreme parameter values via asymptotic analysis, which helps the problem to be analytically solvable (at least to leading order). This can help from general principles for optimal vaccination policies.

These principles can then be used both to form heuristics for finding the true optimal policy in a more general setting and also to explain the resultant optimal solution, as it is often comprised of a mixture of policies resulting from these principles.

There have been a number of recent papers that have used asymptotic analysis to derive general principles. Gavish and Katriel (2022) discusses a model with high reproduction numbers and shows that in this case, it is often optimal to vaccinate the less infectious groups in a population. Moreover, Rao and Brandeau (2021), building on the work of Zaric and Brandeau (2001), linearises the model equations and derives a simple knapsack problem, although the solution to this problem is only optimal when considering the short-term evolution of the epidemic. Other special cases are investigated in Duijzer et al. (2018) (which looks at a population with disconnected subgroups) and Duijzer et al. (2016) (which examines the critical vaccination fraction for a population with separable mixing).

Two cases will be considered in this paper, which both provide novel contributions to the literature. Firstly, the case of a population with a small vulnerable subgroup will be analysed, and it will be shown that, in the asymptotic limit (as the size of this population group tends to zero and its vulnerability tends to infinity), any vaccination policy is eventually outperformed by one where this group is vaccinated first. Of course, the concept that vaccinating vulnerable groups is important has been raised in many previous papers, such as Moore et al. (2021) and Dushoff et al. (2007), but the mathematically rigorous asymptotics presented here provide new evidence for the importance of this principle.

The second case to be discussed is that of a small total vaccination supply. The key novel result that will be shown is that (to leading order) the optimal vaccination problem reduces to a linear knapsack problem, which can be easily solved. This knapsack problem differs from the one in Rao and Brandeau (2021) because, by linearising the final size equations rather than the model ODEs (ordinary differential equations), the optimal solutions and predictions of their behaviour are valid for the full evolution of the epidemic, rather than just in the short term. Again, the case of a small vaccine supply has been examined in many papers such as Shim (2021, 2011) and Medlock and Meyers (2009), but these papers have simply analysed the optimisation problem in the standard way, without deriving the explicit leading-order solution as is done in this paper.

In order to prove these results, it is necessary to build on previous literature. A number of results from Penn and Donnelly (2022) (found in Appendix D) are used in the course of the proof alongside some well-established results, such as the final size of an epidemic in SIR-type models (Anderson and May 1992). However, the theorems presented in the main text are completely novel, with their proofs requiring a significant extension of the current literature. In particular, the various propositions in the proofs (found in Appendices A–C) are, to the best of the authors' knowledge, new to the literature. Some of these results, such as, for example, the proof that epidemic final size is continuously dependent on initial conditions and the vaccination policy found in Proposition 5 may also be helpful to those seeking to prove similar results.

The main analytic results will be further investigated through examples, and, in particular, the small supply case will be used to show that it is not always optimal to vaccinate the most infectious group, even when all groups are equally vulnerable. The

UK population's age structure will be used to relate these results to a realistic example, and optimal small-supply vaccination policies will be approximated for diseases with different age-dependent case fatality ratios.

The paper is structured as follows: Firstly, the multi-group SIR model will be introduced. Then, analytic results will be presented in the case of a small vulnerable subgroup, which will be explored through numerical examples. Finally, analytic results related to a small vaccination supply will be presented and again, examples will be used to illustrate the findings.

2 Modelling

2.1 Disease Transmission and Vaccination Model

The model used in this paper is identical to the model presented in Penn and Donnelly (2022), and this section is simply a summary of the modelling section in Penn and Donnelly (2022). The population is divided into n subgroups, and each subgroup i is further divided into six compartments:

S_i := Number of people that are in group i , are susceptible, and are unvaccinated (1)

I_i := Number of people that are in group i , are currently infected, and were infected while unvaccinated (2)

R_i := Number of people that are in group i , are recovered, and were infected while unvaccinated (3)

S_i^V := Number of people that are in group i , are susceptible and are vaccinated (4)

I_i^V := Number of people that are in group i , are infected and were infected after being vaccinated (5)

R_i^V := Number of people that are in group i , are recovered and were infected after being vaccinated. (6)

Using SIR principles, the model becomes

$$\frac{dS_i}{dt} = - \sum_{j=1}^n (\beta_{ij}^1 I_j + \beta_{ij}^2 I_j^V) S_i - \frac{U_i(t) S_i}{N_i - W_i(t)} \quad (7)$$

$$\frac{dI_i}{dt} = \sum_{j=1}^n (\beta_{ij}^1 I_j + \beta_{ij}^2 I_j^V) S_i - \mu_i^1 I_i \quad (8)$$

$$\frac{dR_i}{dt} = \mu_i^1 I_i \quad (9)$$

$$\frac{dS_i^V}{dt} = - \sum_{j=1}^n (\beta_{ij}^3 I_j + \beta_{ij}^4 I_j^V) S_i^V + \frac{U_i(t) S_i}{N_i - W_i(t)} \quad (10)$$

$$\frac{dI_i^V}{dt} = \sum_{j=1}^n (\beta_{ij}^3 I_j + \beta_{ij}^4 I_j^V) S_i^V - \mu_i^2 I_i^V \quad (11)$$

$$\frac{dR_i^V}{dt} = \mu_i^2 I_i^V \quad (12)$$

where

$$W_i(t) := \int_0^t U_i(s) ds, \quad (13)$$

and

$$N_i = S_i(t) + I_i(t) + R_i(t) + S_i^V(t) + I_i^V(t) + R_i^V(t) \quad (14)$$

is the size of group i . Moreover, the β_{ij}^α terms represent transmission from group j to group i and the μ_i^α terms give the infectious period of the relevant individuals in group i .

Here, $U_i(t)dt$ gives the number of individuals in group i that are vaccinated in the small time interval $[t, t + dt]$, and hence, $W_i(t)$ is the number of individuals that have been vaccinated in group i in $[0, t]$. It is assumed that these vaccinations are assigned randomly to the unvaccinated members of group i , so that each vaccine is given to a susceptible member of group i with probability

$$\frac{\text{number of susceptible members}}{\text{number of unvaccinated members}} = \frac{S_i}{N_i - W_i(t)} \quad (15)$$

Thus, the total rate of susceptibles being vaccinated is $\frac{U_i(t)S_i}{N_i - W_i(t)}$.

Note that there is a slight difference between this model and the one commonly found in the literature (in Hansen and Day 2011; Zaman et al. 2008; Kar and Batabyal 2011 among many others) which set the vaccination term equal to $S_i U_i(t)$ instead of $\frac{U_i(t)S_i}{N_i - W_i(t)}$. As discussed in Penn and Donnelly (2022), this corresponds to vaccines that are randomly distributed to the whole population, which can be seen by rewriting the vaccination term as:

$$S_i U_i(t) dt = \frac{S_i}{N_i} \times N_i U_i(t) dt \quad (16)$$

The first term on the right-hand side is then the probability of a randomly chosen member of group i being susceptible, while the second term is the total number of vaccines assigned in a small time interval $[t, t + dt]$, noting that here the dimension of $U_i(t)$ is 1/time (compared to the model used in this paper where the dimension of $U_i(t)$ is population/time), and hence, it is necessary to scale by $N_i dt$ to convert $U_i(t)$ into a number of vaccines.

This is in contrast to the model in this paper which corresponds to vaccines that are randomly distributed only to the unvaccinated population. Penn and Donnelly (2022) provides justification for the use of this “unvaccinated-only model”, which is therefore the one that will be used in this paper. However, they are structurally very similar, and so it would be possible to apply the results in this paper to the more commonly found model.

To deal with the (removable) singularity that can occur when $W_i = N_i$, it is assumed that

$$W_i(t) \leq N_i \quad \forall t \geq 0 \quad \text{and} \quad W_i(t) = N_i \Rightarrow \frac{U_i(t)S_i}{N_i - W_i(t)} = 0 \quad (17)$$

To capture the benefits of vaccination, there are additional constraints put on the β_{ij}^α and μ_j^α terms which are

$$\beta_{ij}^1 \geq \beta_{ij}^2, \beta_{ij}^3 \geq \beta_{ij}^4 \quad \text{and} \quad \mu_i^1 \leq \mu_i^2. \quad (18)$$

Finally, it will be assumed throughout the remainder of this paper that the population sizes are normalised so that

$$\sum_{i=1}^n N_i = 1 \quad (19)$$

Further details are given in Penn and Donnelly (2022).

2.2 Optimisation Problem

The optimal vaccination problem considered in this paper aims to find the vaccination policy, U , which minimises a weighted sum of the total number of infections in each group. Thus, the problem is:

$$\min \left\{ \sum_{i=1}^n p_i \left(R_i(\infty) + \kappa_i R_i^V(\infty) \right) : \sum_{i=1}^n U_i(t) \leq A(t), \quad \sum_{i=1}^n W_i(t) \leq B(t), \right. \\ \left. U_i(t) \geq 0, \quad W_i(t) \leq N_i \quad \forall t \geq 0 \right\}. \quad (20)$$

Here, $A(t)$ represents the maximal vaccination rate, $B(t)$ represents the maximal vaccine supply and $R_i(\infty)$ and $R_i^V(\infty)$ are the limiting values of $R_i(t)$ and $R_i^V(t)$ as $t \rightarrow \infty$. The weights p_i and $p_i \kappa_i$ could be interpreted in a number of ways, depending on the quantity of interest. For example, $p_i = \kappa_i = 1$ if one wanted to minimise infections, or p_i and $p_i \kappa_i$ could be the case fatality ratio of unvaccinated and vaccinated members of group i , respectively, if one wanted to minimise deaths. However, it is important to note that $\kappa_i \leq 1$ for each i as vaccinated members of the population should be no more vulnerable to the disease than their unvaccinated counterparts.

It is helpful to define $H(U)$ to be the objective function—that is

$$H(U) = \sum_{i=1}^n p_i \left(R_i(\infty) + \kappa_i R_i^V(\infty) \right), \quad (21)$$

where R_i and R_i^V are found from solving the model equations with vaccination policy given by U .

It will be assumed throughout this paper that all “feasible” U are sufficiently smooth for all the quoted theorems to hold. In general, this does not significantly restrict U —for example, the results in Penn and Donnelly (2022) simply require that each $U_i(t)$ is bounded and Lebesgue integrable, while Theorems 1 and 2 require only that U has finite support. Moreover, it is assumed that $B(t)$ is non-decreasing (as total supply should not decrease over time) and piecewise differentiable.

3 Results

3.1 A Small, Vulnerable Subgroup

Consider the case where one of the groups in the population (which, without loss of generality, will be assumed to be group 1) is very small and vulnerable. That is, the population N_1 satisfies

$$N_1(\epsilon) = \epsilon \ll 1 \quad (22)$$

while the weights satisfy

$$p_1(\epsilon) = p_1 \quad \text{and} \quad p_i(\epsilon) = p_i^* \epsilon \quad \forall i \neq 1 \quad (23)$$

for some constants p_1 and p_i^* . It will be assumed that all κ_i are constant. In this setting, group 1 contains a very small proportion of the population, but each member of group 1 is much more vulnerable than the rest of the population.

Thus, this case is practically valid when there is a small subsection of the population that carries the majority of the vulnerability to a disease. As will be discussed further in Section 3.2.3, this has applicability to diseases such as COVID-19, where the majority of the deaths occur significantly older people, while it could also apply to diseases where there are rare conditions that cause a minority of people to be much more vulnerable.

It is mathematically convenient to rescale the parameters p_i so that only p_1 depends on ϵ . This can be done by multiplying all the p_i terms by $\frac{1}{p_1 \epsilon}$ so that

$$\tilde{p}_1(\epsilon) = \frac{1}{\epsilon} \quad \text{and} \quad \tilde{p}_i(\epsilon) = \frac{p_i^*}{p_1} := \tilde{p}_i \quad \forall i \neq 1. \quad (24)$$

This leads to an equivalent optimisation problem in the sense that the optimal vaccination policy will be the same. This occurs because the only change to the objective function is a scalar multiplication of $\frac{1}{p_1\epsilon}$ to each of the terms. Note that while this multiplicative factor tends to infinity as ϵ tends to 0, this system is only analysed for nonzero values of ϵ , and hence, this rescaling is valid.

3.1.1 Analytic Results

The first result presented in this section shows that, in the limit of a group with small size and large vulnerability (with the total cost of the whole group being infected, $N_1 \tilde{p}_1$, remaining constant) any fixed vaccination policy where the vulnerable group is not vaccinated first will eventually (that is, for sufficiently small ϵ) be outperformed by a similar policy where the vulnerable group is vaccinated first.

Group 1 will be given a population size $N_1 = \epsilon$ and an infection cost $\tilde{p}_1 = \frac{1}{\epsilon}$ (recall that the \tilde{p}_i represent the rescaled values of p_i , and so it is acceptable that $\tilde{p}_1 > 1$ for small ϵ). It will be assumed that the initial conditions in the group are proportional to ϵ , so that there exists some $\sigma \in (0, 1]$ such that the initial susceptible population is $\sigma\epsilon$ and the initial infected population is $(1 - \sigma)\epsilon$.

Before stating the full theorem, it is helpful to explain the various constraints and variables that will be introduced. Define, for each value of $\epsilon \geq 0$, $U(t; \epsilon)$ to be the “fixed” vaccination policy where group 1 is not vaccinated first. Of course, the vaccination policy cannot be completely fixed, as the size, ϵ , of group 1 is decreasing, and so it will simply be assumed that the vaccines given out to each group satisfies

$$|W_i(t; \epsilon) - W_i(t; 0)| < \epsilon \quad \forall t \geq 0 \quad \text{and} \quad \forall i \in \{1, \dots, n\} \quad (25)$$

Note that all groups are allowed to have small changes in the number of vaccinations they receive—this allows, for example, for vaccinations that would have been given to group 1 being reassigned as group 1’s population shrinks.

Moreover, to reduce the lengths of the proofs, it will be assumed that U has uniformly bounded finite support—that is, there is some constant t_U such that for each $i \in \{1, \dots, n\}$,

$$t > t_U \Rightarrow U_i(t; \epsilon) = 0 \quad \forall t, \epsilon \geq 0 \quad (26)$$

In order for group 1 to not be vaccinated first in the limit as $\epsilon \rightarrow 0$, there must be some time τ at which some fixed proportion w of the other groups have been vaccinated, while at least some fixed proportion $(1 - \alpha)$ of group 1 has not been vaccinated. That is,

$$W_1(\tau; \epsilon) < \alpha\epsilon \quad \text{and} \quad \sum_{i=1}^n W_i(\tau; \epsilon) > w. \quad (27)$$

One can also define a vaccination policy $\tilde{U}(t; \epsilon)$ where group 1 is vaccinated first. This will be done by re-directing all vaccinations from the $U(t; \epsilon)$ policy to group 1

until it is fully vaccinated, and keeping the same vaccination policy after group 1 is fully vaccinated (ignoring any vaccines that $U(t; \epsilon)$ assigns to group 1 after this time).

To ensure convergence of the model at $\epsilon = 0$, given $\Pi(\epsilon)$ defined by

$$\Pi(\epsilon) := \left\{ i : \exists t \geq 0 \text{ s.t. } I_i(t; \epsilon) > 0 \right\}, \quad (28)$$

it will be assumed that $\Pi(\epsilon) = \{1, \dots, n\}$ for all $\epsilon > 0$ (as any groups which never suffer any infections can be ignored) and that $\Pi(0) = \{2, \dots, n\}$. While this second condition may not be strictly necessary for the theorem to hold, it is unrestrictive and ensures convergence—if this were not the case, then it would be possible that infection in some set of groups were seeded only by group 1. Thus, when $\epsilon = 0$, these groups would suffer no infections, while for any $\epsilon > 0$, they would have an epidemic of size independent (at leading order) of ϵ .

The final condition on the model is that the people in group 1 can be infected by other groups and that vaccinated members of group 1 gain protection from this infection. That is, there is some $i \in \{1, \dots, n\}$ such that

$$\beta_{1i}^1 > \beta_{1i}^3 \geq 0. \quad (29)$$

This is an important condition, as if people group 1 could only be infected by other members of group 1 the total number of infections in group 1 would decay as $\epsilon \rightarrow 0$, meaning that it would no longer necessarily be optimal to vaccinate group 1 first (as most people in group 1 would not catch the disease anyway for small ϵ).

With these considerations, Theorem 1 can now be stated.

Theorem 1 *Suppose that for all $\epsilon > 0$,*

$$N_1(\epsilon) = \epsilon, \quad S_1(0; \epsilon) = \epsilon\sigma, \quad I_1(0; \epsilon) = (1 - \sigma)\epsilon \quad \text{and} \quad \tilde{p}_1(\epsilon) = \frac{1}{\epsilon} \quad (30)$$

for some $\sigma \in (0, 1)$ and that all other parameter values and initial conditions are independent of ϵ .

Consider any vaccination policy with uniformly bounded finite support given by $U(t; \epsilon)$ and suppose that there exists fixed $\alpha, \tau, w > 0$ such that

$$W_1(\tau; \epsilon) < \alpha\epsilon \quad \text{and} \quad \sum_{i=1}^n W_i(\tau; \epsilon) > w \quad \forall \epsilon > 0. \quad (31)$$

Define a new policy, $\tilde{U}(t; \epsilon)$, given by

$$\tilde{U}_1(t; \epsilon) = \begin{cases} \sum_{i=1}^n U_i(t) & \text{if } \sum_{i=1}^n W_i(t; \epsilon) \leq \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (32)$$

and, for $i \neq 1$,

$$\tilde{U}_i(t; \epsilon) = \begin{cases} 0 & \text{if } \sum_{i=1}^n W_i(t; \epsilon) \leq \epsilon \\ U_i(t; \epsilon) & \text{otherwise} \end{cases}. \quad (33)$$

Suppose that for each $i \in \{1, \dots, n\}$ and $t \geq 0$,

$$|W_i(t; 0) - W_i(t; \epsilon)| < \epsilon. \quad (34)$$

Define

$$\Pi(\epsilon) := \{i : \exists t \geq 0 \text{ s.t. } I_i(t; \epsilon) > 0\} \quad (35)$$

and suppose that $\Pi(\epsilon) = \{1, \dots, n\}$ for any $\epsilon > 0$ and that $\Pi(0) = \{2, \dots, n\}$. Finally, suppose that there exists an $i \in \{2, \dots, n\}$ such that

$$\beta_{1i}^1 > \beta_{1i}^3 \geq 0. \quad (36)$$

Then, the policy \tilde{U} is feasible and for sufficiently small ϵ ,

$$H(U(t; \epsilon)) > H(\tilde{U}(t; \epsilon)). \quad (37)$$

For the second theorem, it is helpful to note that, using the results in Penn and Donnelly (2022), if one defines

$$\chi(t) := \begin{cases} A(t) & \text{if } \int_0^t A(s)ds < B(t) \\ \min(A(t), B'(t)) & \text{if } \int_0^t A(s)ds \geq B(t) \end{cases}, \quad (38)$$

then (assuming that there is an optimal solution, and under mild smoothness conditions on U , A and B) there must be an optimal solution satisfying

$$\sum_{i=1}^n W_i(t) = \max \left(\int_0^t \chi(s)ds, 1 \right). \quad (39)$$

The following theorem then proves that the limiting optimal vaccination policy vaccinates the vulnerable group as quickly as possible. To reduce the length of the proof, it will be assumed that $\sigma = 1$, so that (in the small ϵ limit) all members of group 1 can be vaccinated before being infected.

Theorem 2 *With the definitions of Theorem 1, suppose additionally that*

$$\sum_{j=2}^n (\beta_{1j}^1 - \beta_{1j}^3) I_j(0; \epsilon) > 0. \quad (40)$$

That is, the initial difference between the infective force on vaccinated and unvaccinated members of the population is positive. Suppose further that

$$\sigma = 1. \quad (41)$$

Suppose an optimal vaccination policy for each ϵ is given by $\bar{U}(t; \epsilon)$ and suppose that $\bar{U}(t; \epsilon)$ has uniformly bounded finite support. Then, there exists an η depending only on α , τ , w and the model parameters such that, for any U satisfying the condition (31) as defined in Theorem 1

$$\epsilon \in (0, \eta) \Rightarrow H(U) > H(\bar{U}). \quad (42)$$

Moreover, there is a sequence of optimal vaccination policies, $\bar{U}(t; \epsilon)$, which satisfies

$$\lim_{\epsilon \rightarrow 0} \left(\frac{\bar{W}_1(t; \epsilon)}{\epsilon} \right) = 1 \quad \forall t \quad \text{s.t.} \quad \int_0^t \chi(s) ds > 0. \quad (43)$$

Note that the existence of an optimal vaccination policy has been assumed in the statement of this theorem. The authors believe that an optimal policy should exist, as Proposition 5 in the appendices can be used to show that $H(U)$ is continuous. However, more care would need to be taken with the smoothness assumptions on U to create a rigorous proof of this.

Theorems 1 and 2 are proved in the appendices.

3.1.2 Examples

To illustrate these analytic results, consider a simple two-group example. Suppose that group 1 is small, vulnerable, and non-infectious, while group 2 is large, invulnerable and infectious. These groups could be interpreted as “old” and “young”, respectively, although there is no specific physical situation being modelled here.

Suppose the transmission matrices are given by

$$\beta^1 = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \beta^2 = \chi \beta^1 \quad \beta^3 = \rho \beta^1 \quad \text{and} \quad \beta^4 = \chi \rho \beta^1 \quad (44)$$

for some parameters χ and ρ which will be varied. This corresponds to the case of vaccination having (independently) an effectiveness χ at stopping people being infected and ρ at stopping infected people transmitting the disease. Moreover, suppose that

$$\mu_i^\alpha = 1 \quad \forall i, \alpha \quad (45)$$

and

$$N_1 = \epsilon, \quad \tilde{p}_1 = \frac{1}{\epsilon}, \quad \kappa_1 = 1 \quad N_2 = 1, \quad \tilde{p}_2 = p^* \quad \text{and} \quad \kappa_2 = 1, \quad (46)$$

for some parameter p^* that will be varied. Finally, suppose that the initial conditions are

$$S_1(0; \epsilon) = \epsilon, \quad I_1(0; \epsilon) = 0, \quad S_2(0; \epsilon) = 1 - I^* \quad \text{and} \quad I_2(0; \epsilon) = I^*, \quad (47)$$

for some parameter I^* that will be varied and that the vaccination constraints are given by:

$$A(t) = 1 \quad \text{and} \quad B(t) = \max(t, 1). \quad (48)$$

Consider therefore a vaccination policy where group 2, the infectious group, is vaccinated first (and hence, as $B(\infty) = N_2$, it is the only group that is vaccinated). That is,

$$U_1(t; \epsilon) = 0 \quad \text{and} \quad U_2(t; \epsilon) = \begin{cases} 1 & \text{if } t \leq 1 \\ 0 & \text{otherwise} \end{cases}. \quad (49)$$

Hence, with \tilde{U} defined as in Theorem 1, one has

$$\tilde{U}_1(t; \epsilon) = \begin{cases} 1 & \text{if } t \leq \min(1, \epsilon) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \tilde{U}_2(t; \epsilon) = \begin{cases} 1 & \text{if } t \in (\epsilon, 1] \\ 0 & \text{otherwise} \end{cases}. \quad (50)$$

Figure 1 shows a comparison of the objective values $H(U(t; \epsilon))$ and $H(\tilde{U}(t; \epsilon))$ for different values of ϵ . As expected, when $\epsilon = 1$, vaccinating the more infectious group first is optimal (as they have the same vulnerability in this case), while for ϵ smaller than around 0.1, it becomes more effective to vaccinate the vulnerable group first, illustrating the results of Theorem 1.

It is useful to consider the approximate smallness of ϵ required in Theorem 1. That is, how small ϵ needs to be in order for $\tilde{U}(t; \epsilon)$ to be the better vaccination policy. To explore this, define, for each value of I^* and p^* ,

$$\epsilon^*(I^*, p^*) := \inf \left(\left\{ \epsilon : H(\tilde{U}(t; \epsilon)) > H(U(t; \epsilon)) \right\} \cup \{1\} \right). \quad (51)$$

That is, $\epsilon^*(I^*, p^*)$ is the smallest value of ϵ such that vaccinating group 2 first is better than the \tilde{U} policy, with a cut-off value at 1 (as it is possible that for some parameter values, the \tilde{U} policy is always better).

Figure 2 shows the behaviour of $\epsilon^*(I^*, p^*)$. As expected, ϵ^* is decreasing in I^* —this is because when there are fewer initial infectives, there is more time to vaccinate the infectious group before the epidemic has a chance to grow, reducing the peak of the epidemic. Moreover, ϵ^* is decreasing in p^* , as higher values of p^* mean that the number of infections in group 2 is valued higher.

Moreover, Fig. 2 suggests that, for each fixed p^* , ϵ^* is uniformly bounded below for all I^* . Indeed, this is expected as when I^* is very small, there are negligible infections within the interval $t \in [0, 1]$ and so the vaccination policies U and \tilde{U} are in effect being carried out in a completely uninfected population. As the R_0 (that is, the initial

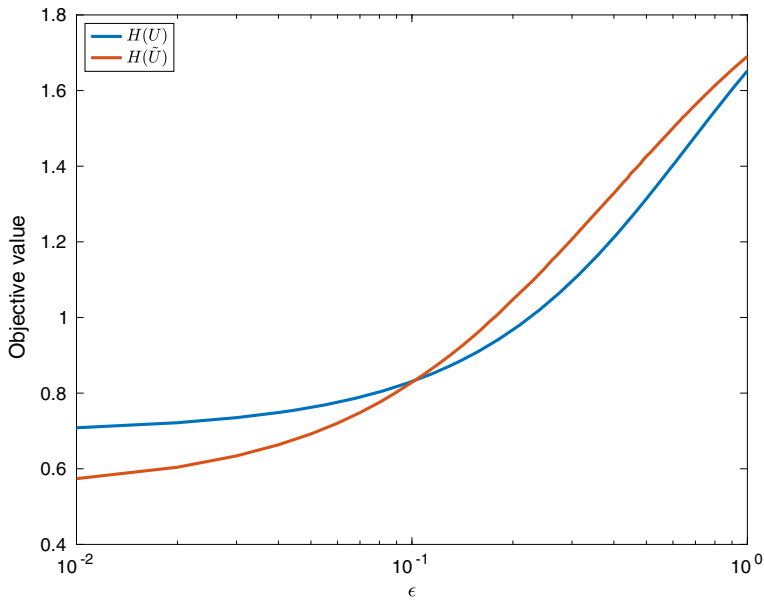


Fig. 1 (Color Figure Online) A comparison of the two vaccination policies, $U(t; \epsilon)$ (where the infectious group is vaccinated first) and $\tilde{U}(t; \epsilon)$ (where the vulnerable group is vaccinated first) for different values of ϵ . Note that here, $I^* = 0.01$, $\chi = \rho = 0.5$ and $p^* = 1$

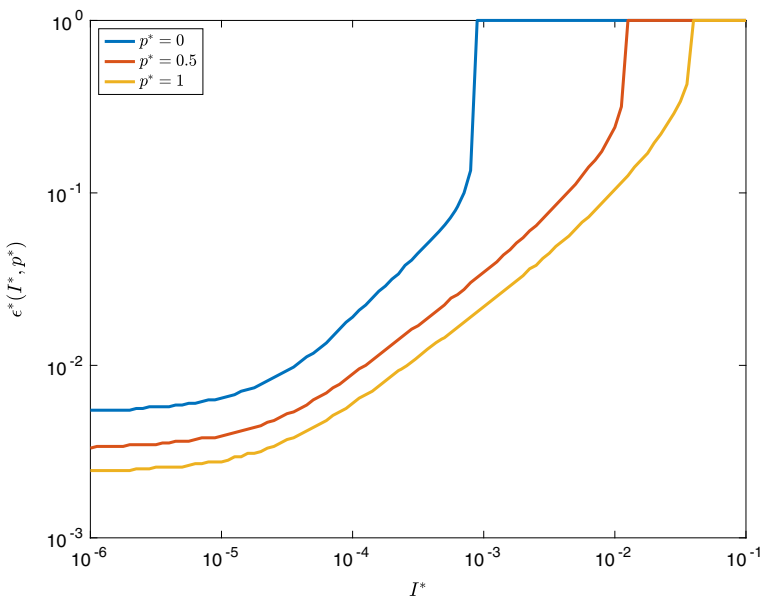


Fig. 2 (Color Figure Online) A plot of $\epsilon^*(I^*, p^*)$, the highest value of ϵ for which U is a better vaccination policy than \tilde{U} . Note that ϵ^* is capped at 1, so that a value of 1 indicates that there were no values found of ϵ^* such that U was the better policy. Note that here, $\chi = \rho = 0.5$

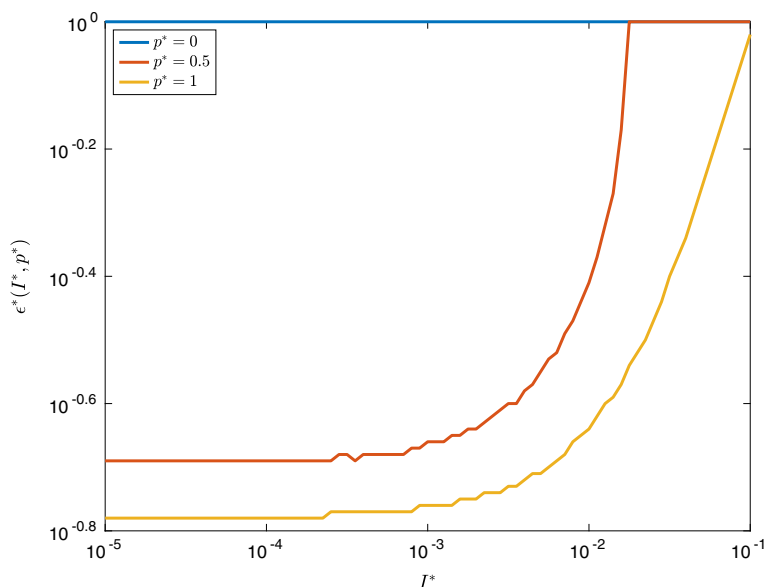


Fig. 3 (Color Figure Online) A plot of $\epsilon^*(I^*, p^*)$, the highest value of ϵ for which U is a better vaccination policy than \tilde{U} , in the case of complete vaccination effectiveness (so $\chi = \rho = 0$). Note that because the values of the objective function are $O(I^*)$, there is some numerical instability which has caused some non-smoothness of the plot

growth rate of the disease) number of a fully vaccinated population (in this case) is greater than 1, $I(t; \epsilon)$ will reach an $O(1)$ value regardless of the vaccination policy. Thus, while decreasing I^* will increase the time to reach this $O(1)$ value, it will not significantly change the final infections in the epidemic, and hence, ϵ^* should converge to a fixed value for small I^* .

When the fully vaccinated population has an R_0 lower than 1, the difference between U and \tilde{U} is more distinct. Indeed, provided I^* is small enough for vaccination to be completed before many infections have occurred, one would expect $O(I^*)$ infections in group 2 in either of the two vaccination policies (for sufficiently small ϵ), as in both policies, the size of the infected compartment will be decreasing after the vaccination has been completed. However, in the U case, one would expect $O(I^*\epsilon)$ infections in total in group 1 (as there is an $O(I^*)$ infection force on a group of size $O(\epsilon)$ for $O(1)$ time), while in the \tilde{U} case, one would expect $O(I^*\epsilon^2)$ infections in total in group 1, as the population of this group is only of size $O(\epsilon)$ for $O(\epsilon)$ time. This behaviour is illustrated in Fig. 3, which shows that ϵ^* converges to significantly higher values than in Fig. 2—indeed, in the case that $p^* = 0$, it appears that U is never optimal for any $\epsilon \leq 1$.

3.2 A Small Vaccination Supply

In this section, the case of a small, immediately available vaccine supply will be considered. In this case, it will be possible to analytically derive the optimal vaccination policy (in the limit of small supply).

This case may be particularly relevant if there was an outbreak of a disease where a vaccine already existed (so that some vaccinations are available immediately), but where supplies were limited, and scaling production would take a significant amount of time. An example of this can be found in the recent monkeypox outbreak (Mahase 2022) where the UK initially purchased 20 000 smallpox vaccines. This small figure—not even enough to vaccinate 0.1% of the UK population (UN 2019)—would certainly fall within the small vaccination supply case.

Moreover, one can use the results in this section regardless of the time at which vaccinations become available (that is, they are not only relevant at the start of an epidemic). This would be of practical use whenever vaccine production is slow, or when the disease is sufficiently mild (or vaccine production is sufficiently expensive) that a large-scale vaccination program is not deemed economically feasible.

3.2.1 Analytic Results

To state the analytic result from this section, it is helpful to define

$$\beta'_{ij} = \begin{cases} \beta_{ij}^1 & \text{if } i, j \leq n \\ \beta_{i(n-j)}^2 & \text{if } i \leq n < j \leq 2n \\ \beta_{(n-i)j}^3 & \text{if } j \leq n < i \leq 2n \\ \beta_{(n-i)(n-j)}^4 & \text{if } n < i, j \leq 2n \end{cases} \quad (52)$$

This large transmission matrix captures the dynamics of all $2n$ susceptible and infectious groups in the model (both vaccinated and unvaccinated). Indeed, after vaccination has been completed, there is no movement from S_i to S_i^V so β' allows for the model to be considered as a $2n$ -group SIR model without vaccination. Thus, in particular, one can derive a simple final size relation for the total number of infections in the epidemic. Similarly, define

$$\mu'_i = \begin{cases} \mu_i^1 & \text{if } i \leq n \\ \mu_{(i-n)}^2 & \text{if } n < i \leq 2n \end{cases} \quad (53)$$

and

$$p'_i = \begin{cases} p_i & \text{if } i \leq n \\ \kappa_{(i-n)} p_{(i-n)} & \text{if } n < i \leq 2n \end{cases} \quad (54)$$

In this case of small supply, it is possible to effectively differentiate the final size of the epidemic with respect to the vaccination policy and use the resultant linear

approximation to form a simple knapsack problem for the optimal vaccination policy. This will involve writing the objective in the form:

$$H(\mathbf{U}(t; \epsilon)) = H(\mathbf{0}) + \mathbf{y}^T \mathbf{W}(\tau(\epsilon); \epsilon) + o(\epsilon) \quad (55)$$

where \mathbf{W} is the final vaccination amounts in each group. To define the gradient, \mathbf{y} , it is necessary to use the inverse of a matrix \mathbf{Q} given by

$$Q_{ij} = \frac{1}{1 - e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)}} \left[\delta_{ij} + \frac{S_i(0; 0) e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)}}{\mu'_j} \beta'_{ij} \right], \quad (56)$$

where as before, the variables $f_i(t; \eta)$ indicate the value of the relevant model variable at time t , given that the parameter ϵ is equal to η , and δ_{ij} is the Kronecker delta. Then, \mathbf{y} is defined by:

$$\mathbf{x} = \mathbf{Q}^{-T} \mathbf{p}' \quad \text{and} \quad y_i = \frac{S_i(0; 0)}{N_i} (x_{i+n} - x_i) \quad \forall i \in \{1, \dots, n\}. \quad (57)$$

These definitions allow for the theorem to be stated.

Theorem 3 Suppose that, for all $\epsilon > 0$

$$B(t; \epsilon) = \epsilon \quad \forall t \geq 0. \quad (58)$$

and that all other parameter values and initial conditions are independent of ϵ . Suppose that $A(t)$ is a continuous function with

$$A(0) > 0 \quad (59)$$

and that the matrix M is invertible. For sufficiently small ϵ , define

$$\tau(\epsilon) := \inf \left\{ t : \int_0^t A(s) ds = \epsilon \right\}. \quad (60)$$

Suppose that \mathbf{U} satisfies the condition

$$\sum_{i=1}^n U_i(s) = \min \left(\int_0^t \chi(s) ds, 1 \right), \quad (61)$$

where χ is defined in (B169). Then, for sufficiently small ϵ , the objective function is given by:

$$H(\mathbf{U}(t; \epsilon)) = H(\mathbf{0}) + \mathbf{y}^T \mathbf{W}(\tau(\epsilon); \epsilon) + o(\epsilon). \quad (62)$$

Moreover, if there is a unique element of \mathbf{y} equal to the minimum of \mathbf{y} then the optimal vaccination policy (to leading order in ϵ) is uniquely given by:

$$U_i(t; \epsilon) = \begin{cases} A(t) & \text{if } i = \min\{y_i : i \in \{1, \dots, n\}\} \text{ and } \int_0^t A(s)ds < \epsilon \\ 0 & \text{otherwise} \end{cases}. \quad (63)$$

The second part of the theorem assumes a unique minimal element of \mathbf{y} . This is not guaranteed to happen, and if there were multiple groups with equal values of \mathbf{y} , this would mean that the effectiveness of vaccinating these groups would be equal to $O(\epsilon)$. However, any sets of parameters satisfying this condition would be unstable to small perturbations (as a trivial example, consider perturbing the initial susceptible populations $S_i(0, 0)$ of the groups with a minimal values of y_i). Thus, in any practical scenario, the probability that the best estimates of the parameters give multiple minimal values of y_i is very small.

Theorem 3 is proved in the appendices.

3.2.2 Vaccinating a Homogeneous Population

To illustrate the effectiveness of this approximation, consider first an example of a homogeneous population (so $n = 1$). Consider the case where $\beta^1 = \beta$, $\beta^2 = \beta^3 = 0.5\beta$ and $\beta^4 = 0.25\beta$ for some parameter β that will be varied. Suppose moreover that

$$N_1 = \mu_1^1 = \mu_1^2 = p_1 = \kappa_1 = A(t) = 1, \quad S_1(0) = 1 - 10^{-4} \quad \text{and} \quad I_1(0) = 10^{-4}. \quad (64)$$

Finally, suppose $B(t) = \epsilon$ where ϵ will be varied.

Figure 4 shows a comparison of the predicted and actual change in number of deaths, ρ_1 for two values of ϵ . It illustrates that, even when $\epsilon = 0.1$, a relatively large value, \mathbf{y} gives a good approximation of the true value (found by simulation). Moreover, when $\epsilon = 0.01$, the two lines are almost indistinguishable. This is useful validation for the approximation, as the correction term was simply proved to be $o(\epsilon)$ rather than, for example, $O(\epsilon^2)$, and so it is encouraging that the predictions are so close.

An interesting property of Fig. 4 is that the value of β for which vaccination is most effective appears to be very close to $S(0)\beta = 1$ (as $S(0) \sim 1$). Note that here, as $\mu = 1$, this is equal to the initial reproduction number of the disease. This has the perhaps surprising consequence that if one has a set of disconnected, equally vulnerable subgroups, a small vaccination supply should be assigned to a group with initial reproduction number close to 1, rather than giving it to the group with the highest value of β (that is, the most group with the most infectious individuals). This result is in line with the findings of Gavish and Katriel (2022), which showed that vaccinating less infectious groups can be more effective, and is an important consideration for vaccination policy planning.

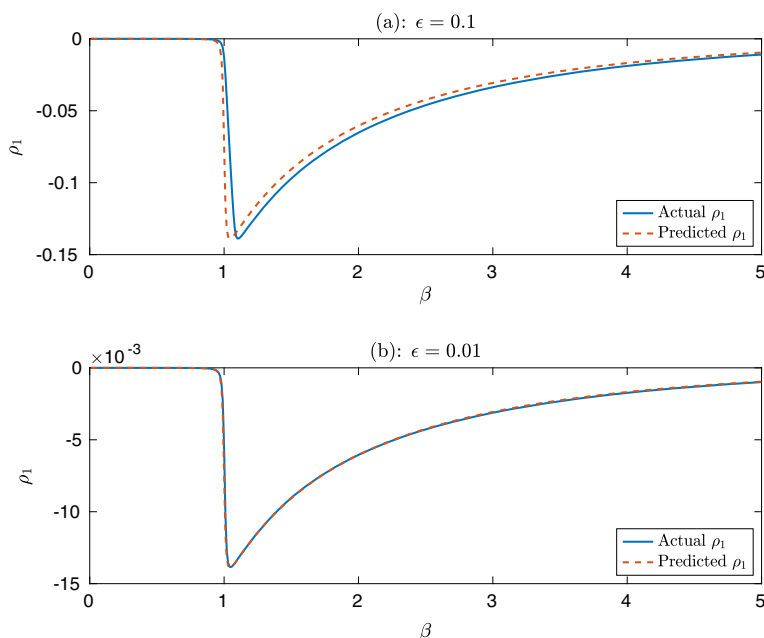


Fig. 4 (Color Figure Online) A comparison of the predicted and actual values of the change in deaths, ρ_1 , in the case of a homogeneous population for two different values of vaccination supply, ϵ and for different values of infectivity, β . Note the different scales on the two y axes

3.2.3 Application to Age-Structured Populations

Consider assigning a small quantity of vaccinations to an age-structured population, using the example of the UK. The disease model has been estimated using the inter-age-group contact matrices Λ from Prem et al. (2017), alongside population estimates N from UN (2019). As in Prem et al. (2017), this gives a transmission matrix of

$$\beta_{ij}^1 = \beta \frac{\Lambda_{ij}}{N_j} \quad (65)$$

for some scalar parameter β . As in the previous section, it will be assumed that

$$\mu_i^\alpha = 1 \quad \forall i, \alpha \quad (66)$$

and

$$\beta^2 = 0.5\beta^1, \quad \beta^3 = 0.5\beta^1 \quad \text{and} \quad \beta^4 = 0.25\beta^1. \quad (67)$$

It will also be assumed that the initial infected population is small, so that, for each i

$$S_i(0; \epsilon) = (1 - 10^{-4})N_i \quad \text{and} \quad I_i(0; \epsilon) = 10^{-4}N_i. \quad (68)$$

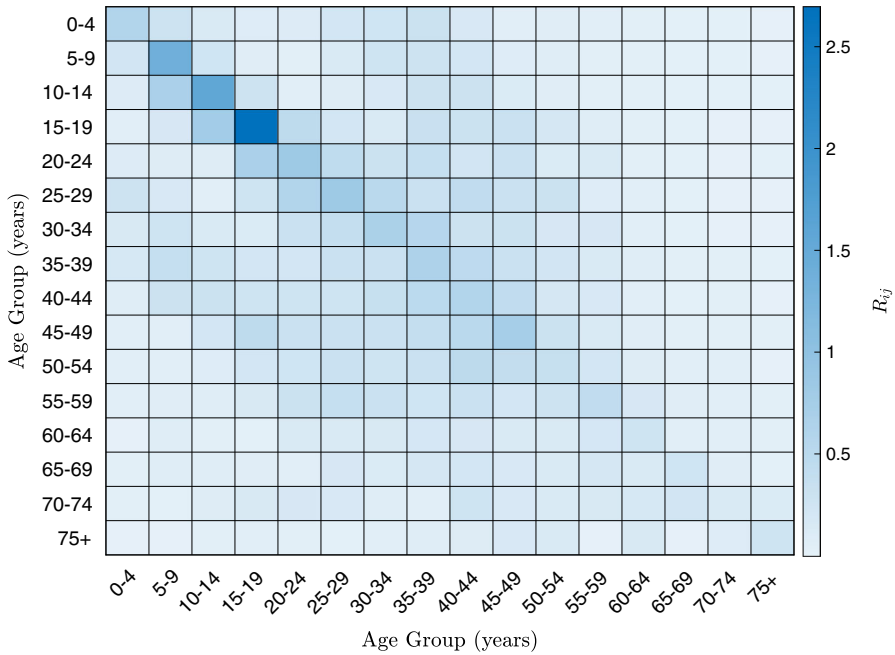


Fig. 5 (Color Figure Online) A heatmap of the next-generation matrix for the age-structured UK population

In the following examples, β will be chosen so that the disease-free next-generation matrix of a completely unvaccinated population, given by

$$R_{ij} = \frac{N_i \beta_{ij}^1}{\mu_j^1} = \beta_{ij}^1 \quad (69)$$

has a spectral radius (that is, largest eigenvalue) equal to 4. This sets the R_0 number in the overall population to be 4. To illustrate the population structure, Fig. 5 shows a heatmap of the matrix R_{ij} . This highlights the strongly assortative nature of the contacts (that is, members of a subgroup are most likely to be contacts with members of their own subgroup), while also showing that contacts are lower for older age groups.

Now, two different age-dependent case-fatality ratios will be considered—uniform case-fatality and approximate COVID-19 case fatality, taken from Dyer (2021). In both cases, it will be assumed that vaccination reduces the case fatality ratio by 90% (following the results of Dyer (2021) for the COVID-19 vaccines) so that $\kappa_i = 0.1$ for all i . However, it is worth emphasising that this model is simply based on real-world data and does not seek to accurately model the COVID-19 pandemic.

Figure 6 shows the effectiveness of vaccinating each age group in the two different cases, as a proportion of the optimal effectiveness. Note that here the proportion of effectiveness of assigning vaccine to group i is given by $\frac{y_i}{\min_j(y_j)}$, as each y_j is non-positive. It highlights that the significantly higher mortality rates for COVID-19 for the

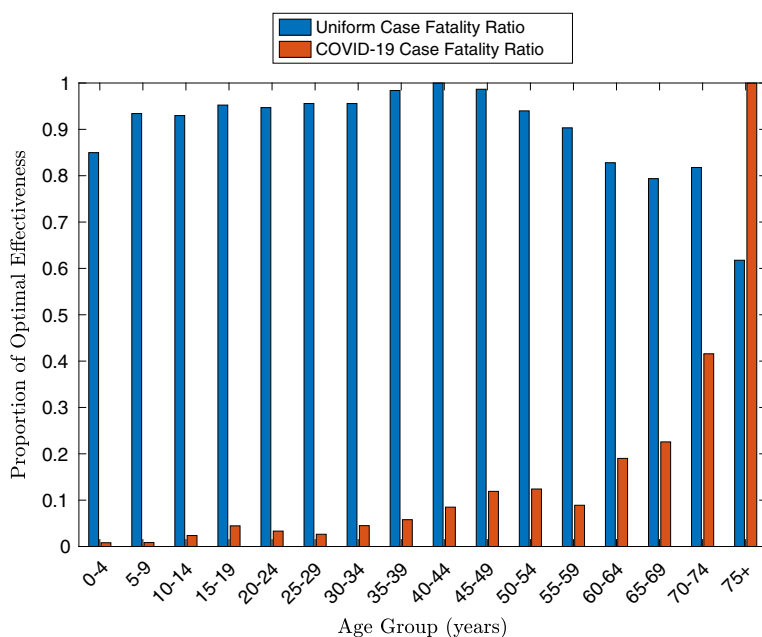


Fig. 6 (Color Figure Online) The effectiveness of assigning a small quantity of vaccines to each age group as a proportion of the optimal effectiveness

older age groups mean that vaccinating them is much more effective than vaccinating the other age groups. This is an example of Theorems 1 and 2, as the oldest age group makes up a relatively small percentage (around 9%) of the population, but, if one scales p such that it has median value 1, the $p_i N_i$ value for the oldest age group is approximately 20, and so is definitely $O(1)$ rather than $O(\epsilon)$.

A perhaps surprising exception to the general correlation between effectiveness and mortality is the relatively low effectiveness of vaccinating the 55–59-year-old age group, which is lower than the 45–49-year-old and 50–54-year-old groups. This illustrates the non-intuitive nature that optimal vaccination policies can take, and the importance of investigating their behaviour fully. The main reason for this low effectiveness is that, while the 55–59-year-old age group is more vulnerable to COVID-19 than the younger groups, according to Prem et al. (2017), they have much less contact with the 75+-year-old age group, and thus, vaccinating this group provides significantly less secondary protection to most vulnerable members of the population. The authors speculate that this could be due to a significant number of the parents of the 55–59-year-old age group having died (particularly in comparison with the younger groups), reducing their links with the 75+-year-old age group. Moreover, those in the 55–59-year-old age group may also not be old enough to have many 75+-year olds in their social circles (in comparison with members of older groups). However, further investigation would be needed to justify this claim.

In the case of uniform mortality, the vaccination policy becomes even less intuitive, as Fig. 6 shows that the optimal age group to vaccinate is the 40–44-year olds. Indeed,

from Fig. 5, it may seem that the 15–19-year-old group would be the best group to vaccinate, as they have the highest overall transmission—that is, the maximum value of

$$\text{Total infectious force of group } j := \sum_{i=1}^{16} R_{ij}. \quad (70)$$

However, if instead, one considers

$$\text{Total external infectious force of group } j := \sum_{i=1, i \neq j}^{16} R_{ij}, \quad (71)$$

then it is the 35–39 and the 40–44 age groups which have the highest values. This can be considered in conjunction with the results of the previous subsection, which showed that vaccinating groups with R_0 numbers close to 1 is optimal for disconnected populations. Indeed, the “secondary effect” of vaccinations (that is, the number of people who are not vaccinated, but are protected from the disease because of vaccines given to others) can be higher for groups with lower internal infectious force, particularly when their external infectious force is higher.

Finally, it is useful to again explore the range of values for ϵ for which \mathbf{y} gives a good approximation of the true number of infections. As the minimum (scaled so that the total population size is 1) value of N_i is 0.0498 in this case, ϵ will be tested at 0.0498. The results of this are shown in Fig. 7, which again illustrates the effectiveness of this approximation. Indeed, the largest error across either case is of order 10^{-4} , which in turn is of order $\epsilon^2 \mathbf{y}$. This suggests that the $o(\epsilon)$ correction term in Theorem 3 is significantly smaller than ϵ , which increases the usefulness of this approximation. However, further investigation is needed to determine whether this correction is of $O(\epsilon^2 \mathbf{y})$ for all parameter values.

4 Discussion

This paper has shown two general principles for optimal vaccination policies by looking at the asymptotic behaviour of the optimal policy in the case of extreme parameters. Firstly, it has shown that small, vulnerable groups should in general be vaccinated first, regardless of the overall timetable of vaccination. This is an important result as it requires very little data on the population—merely the case fatality ratios and populations of the different subgroups—and in particular needs no forecasting of future transmission trends or vaccine supply.

The analytically derived results (in the limiting case) also show that the effect of vaccinating this small group far outweighs the effect of vaccinating any of the other groups. Indeed, if the size of the vulnerable group is $O(\epsilon)$ and the case fatality ratio of the other groups is $O(\epsilon)$, then Theorem 1 shows that vaccinating the vulnerable group will lead to an $O(\epsilon)$ decrease in the number of fatalities, while vaccinating the same number of people from another group will only decrease this by $O(\epsilon^2)$. As discussed in

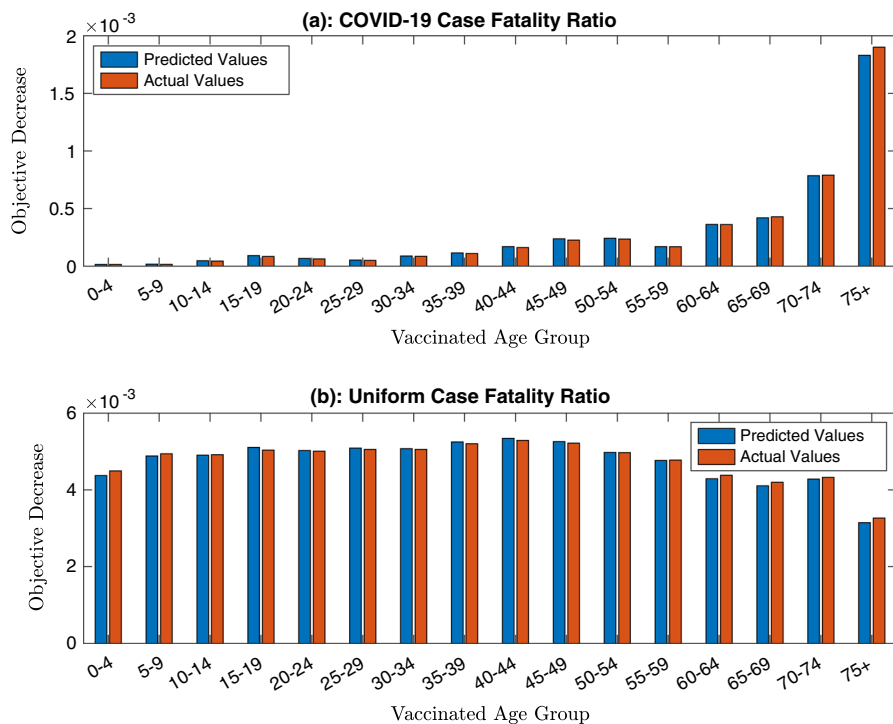


Fig. 7 (Color Figure Online) A comparison of the predicted and simulated change in the objective function when vaccinating each individual group at $\epsilon = 0.0498$. Both the cases of a COVID-19 case fatality ratio (in **a**) and a uniform case fatality ratio (in **b**) are presented

Sect. 3.2.3, this result is of practical importance for diseases such as COVID-19, where the majority of the fatalities would be from certain age groups within the population. In particular, it provides strong evidence for the importance of sharing vaccines on a global scale, as this is the only way to ensure that vaccinations can be given to all people who are most vulnerable to the disease.

However, this result should be used with caution, as it certainly does not imply that a population should always be vaccinated in order of decreasing vulnerability to the disease. The optimal vaccination policy is, in general, a balance between directly protecting the vulnerable by vaccinating them and by indirectly protecting them by vaccinating those groups with the highest infectiousness. This is shown in Fig. 6 by the fact that, when a COVID-19 case fatality ratio is used, the relative effectiveness of vaccinating each age group does not decrease everywhere with age. The results of Theorems 1 and 2 simply provide a principle that in the asymptotic limit, the optimal strategy is to vaccinate small, vulnerable groups first. In the absence of data on vaccination effectiveness (which is crucial in determining whether indirectly protecting the more vulnerable population may be better), this provides a mathematically sound justification for beginning with the most vulnerable members of a population while gathering data to determine the rest of the vaccination policy.

The second principle derived in this paper was a linear approximation to the change in number of fatalities from a disease, which allows for the estimation of the optimal vaccination policy in the case of a small total supply. Again, this principle is flexible, applying for any set of parameters and provides a computationally cheap way of approximating the optimal solution, even for large numbers of groups, as it merely requires the solution of a linear system involving the same number of variables as the number of groups.

A useful feature of this approximation is that it appears to have high accuracy even for reasonably large values of the total supply, such as when 10% of the population can be vaccinated. Figures 4 and 7 show that there is very little deviation between the predicted and actual values of the objective function and so suggest that this is a flexible and widely applicable method of approximation, even when the population contains a large number of subgroups. However, it would be helpful to strengthen the results of Theorem 3 to get a stronger bound on the error for small ϵ to ensure that this similarity holds for all models.

The results of the examples presented in Sect. 4 are also informative for vaccination policy. As shown in Fig. 4, in a completely homogeneous population, vaccination has the most effect when the reproduction number ($\frac{\beta}{\mu}$ in this case) is slightly bigger than 1, with a steep decline in effectiveness for reproduction numbers below 1 and a more gradual decline for large reproduction numbers. This result allows one to consider the “vaccination leverage” of a population—that is, the effectiveness that a small quantity of vaccination can have—and shows that, even in the case of homogeneous case fatality ratios, vaccinating in order of infectiousness may be far from optimal, as it is much more difficult to reduce infections in a highly infectious population.

Indeed, a similar idea was shown to apply when the UK age structure was considered. In the case of uniform case fatality, the optimal group to assign a small amount of vaccinations was the 40–44 age group which, as shown in Fig. 5, is not the most infectious group. This perhaps counter-intuitive result highlights the importance of mathematically justifying the principles one uses to decide on optimal vaccination policies, as “common-sense” arguments may in fact give false conclusions. Communicating such principles to governments and policy-makers will be crucial in future pandemics, particularly ones with more homogeneous case fatality ratios where the optimal policy is not as intuitive as for diseases like COVID-19.

An important limitation of Theorem 3 is that the optimal policies for small vaccination supplies do not necessarily generalise to give the beginning of the optimal vaccination policy in the case of a much larger vaccination supply. Indeed, it is possible to have bifurcations in the optimal vaccination policy as the supply increases—for example, it can become possible to completely avoid an epidemic by vaccinating a large quantity of an infectious group. Thus, while the linear approximation can be a useful starting point when attempting to estimate the optimal strategy, it is important to consider alternatives when a large proportion of the population can be vaccinated.

The results of this paper are only applicable if the trajectory of the disease in question can be well-approximated by multi-group SIR dynamics. In particular, this requires there to be reasonably high levels of the disease in a population [otherwise stochastic dynamics change the epidemic behaviour (Ball and Neal 2002)], and for population

subgroups to be sufficiently large (again to prevent stochasticity dominating). Moreover, the model assumptions would not hold if individuals could be re-infected, or if the effect of vaccination was not eternal (though if the timescale of the epidemic was sufficiently shorter than the timescale of immunity decay, then the model would still provide a good approximation).

A final barrier to using the results in this paper is that estimation error in the model parameters could lead to the optimal solutions being incorrectly calculated. Estimating the β_{ij}^α parameters is particularly complicated, especially in a multi-group setting where it is difficult to establish the chain of transmission between different groups. Because of this, building models based on contact rates between groups [estimated using surveys (Prem et al. 2017)] or proxies such as commuting patterns (Keeling and White 2011) may be the best method, at least to provide priors on the parameters. Theorems 1 and 2 are significantly less susceptible to errors in parameters, as they do not require any of the β_{ij}^α or μ_j^α parameters to be known, although the level of “smallness” of ϵ would vary depending on the disease in question. Theorem 3 is significantly more susceptible to error, as all the model parameters are needed. However, while there may be bifurcations in the optimal strategy, the optimal value of the objective function should depend continuously on the parameters (a fact which could be proved by extending the results of Proposition 5), limiting the effect of small estimation errors.

Despite this, the authors expect that similar results to those presented in Theorem 3 will hold for a very wide class of deterministic models. Essentially, the only necessary characteristic of the model that is required by Theorem 3 is that the objective function, $H(U)$, is a continuously differentiable function of the vaccination policy U in some neighbourhood of $\mathbf{0}$. Indeed, y in Theorem 3 can be replaced by $\nabla H(\mathbf{0})$ in a general setting. Certainly, it should be conceptually simple (though perhaps algebraically complicated) to generalise this result to other compartmental models such as SEIR (Susceptible–Exposed–Infected–Recovered) and even those modelling vector-transmitted diseases.

The authors also expect that Theorem 1 will hold for general models where the effect of vaccination is eternal. The essential points in the proof of Theorem 1 are that vaccinating the small group does not affect the overall vaccination program (to leading order) and that it does have an $O(1)$ effect on the objective function. Both of these should still hold in a wide range of models, although it may be difficult to define the meaning of “very small group” and “very vulnerable group”—particularly in more complicated settings such as individual-based models.

This work could be extended by deriving more principles for extreme parameter values and investigating whether they generalise to realistic model parameters. By combining the existing results in this paper and others such as Gavish and Katriel (2022) with potential new ones, one could create an algorithm that creates good heuristics of optimal vaccination policies that could be used as starting points for accurately approximating the optimal policy for a general parameter set. This could have significant implications for the design of vaccination policies, as it would enable the optimisation problem to be estimated for very complex models, as the time taken to converge to an optimal solution would significantly decrease given good initial heuristics.

5 Conclusion

The results of this paper are summarised below:

- If a sufficiently vulnerable, sufficiently small population exists in a multi-group SIR model, it is optimal to vaccinate this group first.
- For small overall vaccination supplies, the optimal vaccination problem can be well approximated by a simple knapsack problem.
- This linearisation appears to be a good approximation even for relatively large vaccination supplies (such as 10% of the population).
- This linearisation shows that, in the case of uniform case fatality, it is not necessarily optimal to vaccinate the most infectious group.

Funding This work was funded by a DTP Studentship awarded to Matthew Penn by the European Physical Sciences Research Council (EPSRC) through the University of Oxford (<https://www.ukri.org/councils/epsrc/career-and-skills-development/studentships/doctoral-training-partnerships/>). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission. The funders played no specific role in any aspect of this work.

Data Availability Statement The data and code used in this study are available from <https://github.com/mpenn114/AsymptoticVaccination>.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Proof of Theorem 1

Note that, throughout this section, the tilde will be removed from the rescaled p_i terms to reduce notation (and hence all p_i terms used here will be assumed to be rescaled).

Theorem 1 *Suppose that for all $\epsilon > 0$*

$$N_1(\epsilon) = \epsilon, \quad S_1(0; \epsilon) = \epsilon\sigma, \quad I_1(0; \epsilon) = (1 - \sigma)\epsilon \quad \text{and} \quad p_1(\epsilon) = \frac{1}{\epsilon} \quad (\text{A1})$$

for some $\sigma \in (0, 1)$. Suppose that all other parameter values and initial conditions are independent of ϵ .

Consider any vaccination policy given by $U(t; \epsilon)$ and suppose that there exists fixed $\alpha, \tau, w > 0$ such that

$$W_1(\tau; \epsilon) < \alpha\epsilon \quad \text{and} \quad \sum_{i=1}^n W_i(\tau; \epsilon) > w \quad \forall \epsilon > 0. \quad (\text{A2})$$

Define a new policy $\tilde{U}(t; \epsilon)$

$$\tilde{U}_1(t; \epsilon) = \begin{cases} \sum_{i=1}^n U_i(t) & \text{if } \sum_{i=1}^n W_i(t; \epsilon) \leq \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (\text{A3})$$

and, for $i \neq 1$

$$\tilde{U}_i(t; \epsilon) = \begin{cases} 0 & \text{if } \sum_{i=1}^n W_i(t; \epsilon) \leq \epsilon \\ U_i(t; \epsilon) & \text{otherwise} \end{cases}. \quad (\text{A4})$$

Suppose that for each $i \in \{1, \dots, n\}$ and $t \geq 0$,

$$|W_i(t; 0) - W_i(t; \epsilon)| \leq \epsilon. \quad (\text{A5})$$

Define

$$\Pi(\epsilon) := \{i : \exists t \geq 0 \text{ s.t. } I_i(t; \epsilon) > 0\} \quad (\text{A6})$$

and suppose that $\Pi(\epsilon) = \{1, \dots, n\}$ for any $\epsilon > 0$ and that $\Pi(0) = \{2, \dots, n\}$. Finally, suppose that there exists a $i \in \{2, \dots, n\}$ such that

$$\beta_{1i}^1 > \beta_{1i}^3 \geq 0. \quad (\text{A7})$$

Then, the policy \tilde{U} is feasible and for sufficiently small ϵ ,

$$H(U(t; \epsilon)) > H(\tilde{U}(t; \epsilon)). \quad (\text{A8})$$

Proof It is first important to prove that the \tilde{U} is feasible. Firstly,

$$\sum_{i=1}^n \tilde{U}_i(t; \epsilon) \leq \sum_{i=1}^n U_i(t; \epsilon) \quad (\text{A9})$$

which, as U is feasible, means that the supply and rate constraints are satisfied. Moreover, as each $U_i(t; \epsilon) \geq 0$,

$$\tilde{U}_i(t; \epsilon) \geq 0 \quad \forall i \in \{1, \dots, n\}. \quad (\text{A10})$$

Also, for $i \neq 1$,

$$\tilde{U}_i(t; \epsilon) \leq U_i(t; \epsilon) \Rightarrow \tilde{W}_i(t; \epsilon) \leq W_i(t; \epsilon) \leq N_i. \quad (\text{A11})$$

Finally, define

$$t^* := \sup \left\{ t : \sum_{i=1}^n W_i(t; \epsilon) \leq \epsilon \right\} \in \mathbb{R} \cup \{\infty\} \quad (\text{A12})$$

and then

$$U_1(t; \epsilon) \leq \int_0^{t^*} \sum_{i=1}^n U_i(s; \epsilon) ds \leq \epsilon = N_1 \quad (\text{A13})$$

as required. \square

Define $S_i(t; \epsilon)$ to be the number of susceptibles given the parameters $N_1(\epsilon)$, $S_1(0; \epsilon)$ and $I_1(0; \epsilon)$ and the vaccination policy $U(t; \epsilon)$, and define $\tilde{S}_i(t; \epsilon)$ to be the number of susceptibles given the parameters $N_1(\epsilon)$, $S_1(0; \epsilon)$ and $I_1(0; \epsilon)$ and the vaccination policy $\tilde{U}(t; \epsilon)$. Use similar definitions for the other variables in the model.

Proposition 4

Proposition 4 For each $t \geq 0$ and $i \in \{1, \dots, n\}$,

$$|\tilde{W}_i(t; \epsilon) - \tilde{W}_i(0; \epsilon)| \leq 2\epsilon. \quad (\text{A14})$$

Proof Firstly, note that

$$\tilde{W}_1(t; \epsilon) \leq \epsilon \quad (\text{A15})$$

so

$$|\tilde{W}_1(t; \epsilon) - \tilde{W}_1(0; \epsilon)| \leq \epsilon. \quad (\text{A16})$$

Now, suppose that $i \neq 1$. Then, for each ϵ , $t \geq 0$, with t^* defined as in (A12),

$$|W_i(t; \epsilon) - \tilde{W}_i(t; \epsilon)| = \left| \int_0^t U_i(s) ds - \int_{t^*}^{\max(t, t^*)} U_i(s) ds \right|. \quad (\text{A17})$$

If $t < t^*$, then

$$|W_i(t; \epsilon) - \tilde{W}_i(t; \epsilon)| \leq \left| \int_0^t U_i(s) ds \right| \leq \left| \int_0^{t^*} U_i(s) ds \right| \leq \epsilon \quad (\text{A18})$$

while if $t \geq t^*$, then

$$|W_i(t; \epsilon) - \tilde{W}_i(t; \epsilon)| = \left| \int_0^{t^*} U_i(s) ds \right| \leq \epsilon. \quad (\text{A19})$$

Thus, noting

$$W_i(t; 0) = \tilde{W}_i(t; 0) \quad (\text{A20})$$

and using (34),

$$|\tilde{W}_i(t; \epsilon) - \tilde{W}_i(0; \epsilon)| \leq |\tilde{W}_i(t; \epsilon) - W_i(t; \epsilon)| + |W_i(t; \epsilon) - W_i(t; 0)| \leq \epsilon + \epsilon = 2\epsilon \quad (\text{A21})$$

as required. \square

Proposition 5

Next, it is helpful to consider the continuous dependence of the final size of the epidemic on the initial conditions and the vaccination policy. A weaker result is proved in Penn and Donnelly (2022) (and is referenced in this proof as Lemma 14). However, that result only holds for finite times, and extending it to hold for the final sizes requires a significant amount of extra work.

Proposition 5 *Suppose that the U_i have uniformly bounded support for each $\epsilon > 0$. Moreover, for each of the model variables, f_i , suppose that*

$$|f_i(0; \epsilon) - f_i(0; 0)| < K\epsilon \quad (\text{A22})$$

for some constant K and that

$$|W_i(t; \epsilon) - W_i(t; 0)| < K'\epsilon \quad (\text{A23})$$

for some constant K' . Finally, suppose all parameters are independent of ϵ with the exception that $N_1(\epsilon) = \epsilon$. Then, for each $\delta > 0$, there exists some $\Delta > 0$ such that

$$\epsilon \in [0, \Delta] \Rightarrow |f_i(\infty; \epsilon) - f_i(\infty; 0)| < \delta \quad \forall f \in \{I_i(t; \epsilon), I_i^V(t; \epsilon), R_i(t; \epsilon), R_i^V(t; \epsilon)\}. \quad (\text{A24})$$

Note that this holds both in the case of Theorem 1 (where $N_1 \rightarrow 0$, $\Pi(\epsilon) = \{1, \dots, n\}$ for $\epsilon > 0$ and $\Pi(0) = \{2, \dots, n\}$) or, in the case where each N_i is independent of ϵ (by adding a disconnected group of size ϵ).

Proof Choose any $\delta > 0$. Now, it is possible to write the system for \mathbf{I} and \mathbf{I}^V in the form

$$\frac{d\mathbf{J}(t; \epsilon)}{dt} = \mathbf{M}(t; \epsilon)\mathbf{J}(t; \epsilon), \quad (\text{A25})$$

where \mathbf{M} depends on the values of $S(t; \epsilon)$, $S^V(t; \epsilon)$, β_{ij}^α and μ_i^α and

$$\mathbf{J} = \begin{pmatrix} \mathbf{I} \\ \mathbf{I}^V \end{pmatrix}. \quad (\text{A26})$$

Hence, in particular, by using Proposition 4 and Lemma 14 for any fixed $t \geq 0$,

$$\lim_{\epsilon \rightarrow 0} (\mathbf{M}(t; \epsilon)) = \mathbf{M}(t; 0). \quad (\text{A27})$$

Moreover, if the support of each $U_i(t; \epsilon)$ is bounded by t_U (which exists by assumption), then for $t > t_U$, each $S_i(t; \epsilon)$ and $S_i^V(t; \epsilon)$ is non-increasing in t and so $\mathbf{M}(t; \epsilon)$ is non-increasing. As it is bounded below, it therefore must converge to some matrix $\mathbf{M}(\infty; \epsilon)$, and, for $t > t_U$,

$$\frac{d\mathbf{J}(t; \epsilon)}{dt} \geq \mathbf{M}(\infty; \epsilon)\mathbf{J}(t; \epsilon). \quad (\text{A28})$$

Hence, by Lemma 9,

$$\mathbf{J}(t_U + t'; \epsilon) \geq e^{t' \mathbf{M}(\infty; \epsilon)} \mathbf{J}(t_U; \epsilon). \quad (\text{A29})$$

Moreover, by Lemma 11,

$$\lim_{t' \rightarrow \infty} (\mathbf{J}(t_U + t'; \epsilon)) = 0 \quad (\text{A30})$$

and hence (by non-negativity)

$$\lim_{t' \rightarrow \infty} (e^{t' \mathbf{Q}(\infty; \epsilon)} \mathbf{J}(t_U; \epsilon)) = 0. \quad (\text{A31})$$

Now, define

$$\max_{i, \alpha} (\mu_i^\alpha) := \mu_{\max} \quad (\text{A32})$$

and then define

$$\mathcal{M}(\infty; 0) := \mathbf{M}(\infty; 0) + \mu_{\max} \mathcal{I}_{2n}, \quad (\text{A33})$$

where \mathcal{I}_{2n} is the $2n$ by $2n$ identity matrix. Thus, in particular, $\mathcal{M}(\infty; 0)$ is non-negative and so

$$e^{\mathbf{M}(\infty; 0)} = e^{-\mu_{\max}} e^{\mathcal{M}(\infty; 0)} \quad (\text{A34})$$

is non-negative as the exponential of a non-negative matrix is non-negative (as it is a weighted sum of powers of that matrix with positive weights). Thus, by Perron–Frobenius theory, summarised in Berman and Plemmons (1994), there exists a real

non-negative eigenvalue $\lambda(\infty; 0)$ (called the Perron eigenvalue) of $e^{M(\infty; 0)}$ such that any other eigenvalues $\rho(\infty; 0)$ satisfy

$$|\rho(\infty; 0)| \leq |\lambda(\infty; 0)| \quad (\text{A35})$$

so, in particular

$$\Re(\rho(\infty; 0)) \leq \Re(\lambda(\infty; 0)). \quad (\text{A36})$$

□

Claim $0 < |\lambda(\infty; 0)| < 1$

Proof Note that $\lambda(\infty; 0) > 0$, as

$$\text{trace}\left(e^{M(\infty; 0)}\right) \geq \text{trace}\left(e^{-\mu_{\max}} \mathcal{I}_{2n}\right) > 0 \quad (\text{A37})$$

and thus, $\lambda(\infty; 0) \neq 0$.

From Berman and Plemmons (1994), there is a non-negative eigenvector, \mathbf{v} , with eigenvalue $\lambda(\infty; 0)$. Now, \mathbf{v} must be an eigenvector of $M(\infty; 0)$ (as eigenvectors of a matrix and its exponential are the same). Thus, there is some $\lambda^*(\infty; 0)$ such that

$$M(\infty; 0)\mathbf{v} = \lambda^*(\infty; 0)\mathbf{v}. \quad (\text{A38})$$

In particular, writing $\mathbf{v} = (v_1, \dots, v_{2n})^T$

$$\lambda^*(\infty; 0)v_1 = (M(\infty; 0)\mathbf{v})_1 = -\mu_1^1 v_1 \quad (\text{A39})$$

and thus, either $\lambda^*(\infty; 0) = -\mu_1^1 < 0$ or $v_1 = 0$. Suppose first that $\lambda^*(\infty; 0) = -\mu_1^1$. Then, this means that (as the eigenvalues of $e^{M(\infty; 0)}$ are the exponentials of the eigenvalues of $M(\infty; 0)$),

$$|\lambda(\infty; 0)| = |e^{-\mu_1^1}| < 1. \quad (\text{A40})$$

Similarly, $v_{n+1} \neq 0$ implies that

$$|\lambda(\infty; 0)| = |e^{-\mu_1^2}| < 1. \quad (\text{A41})$$

Thus, suppose for the remainder of the proof of this claim that $v_1 = v_{n+1} = 0$. Now, for $i \leq n$, the entries on the i th row of $M(\infty; 0)$ are given by:

$$M(\infty; 0)_{ij} = \begin{cases} S_i(\infty; 0)\beta_{ij}^1 - \delta_{ij}\mu_i^1 & \text{if } j \leq n \\ S_{i-n}(\infty; 0)\beta_{i(j-n)}^3 & \text{if } j > n \end{cases} \quad (\text{A42})$$

and for $i > n$, they are given by

$$M(\infty; 0)_{ij} = \begin{cases} S_i^V(\infty; 0)\beta_{ij}^2 & \text{if } j \leq n \\ S_{i-n}^V(\infty; 0)\beta_{i(j-n)}^4 - \delta_{ij}\mu_i^2 & \text{if } j > n \end{cases}, \quad (\text{A43})$$

where δ_{ij} is the Kronecker delta.

Now, as $\Pi(0) = \{2, \dots, n\}$, by Lemma 13, it is necessary that

$$J_i(t; 0) > 0 \quad \forall t > 0 \quad \text{and} \quad i \in \{2, \dots, n\}. \quad (\text{A44})$$

Moreover, if $I_i^V(t; 0) = 0$ for some $t > 0$, then, by Lemma 15, as $\Pi(0) = \{2, \dots, n\}$, it is necessary that

$$S_i^V(t; 0)\beta_{ji}^3 = S_i^V(t; 0)\beta_{ji}^4 = 0 \quad \forall j \in \{2, \dots, n\} \quad (\text{A45})$$

and so, if $t \geq t_U$, then this implies

$$S_i^V(\infty; 0)\beta_{ji}^3 = S_i^V(\infty; 0)\beta_{ji}^4 = 0 \quad \forall j \in \{2, \dots, n\}. \quad (\text{A46})$$

Thus, in this case, for $j \notin \{1, n+1\}$

$$M(\infty; 0)_{ij} = -\mu_{(i-n)}^2 \delta_{ij}. \quad (\text{A47})$$

Therefore, suppose $J_i(t_U; 0) = 0$ for some $i \notin \{1, n+1\}$ (and so, necessarily, $i \in \{n+2, \dots, 2n\}$). Then,

$$(M(\infty; 0)v)_i = \sum_{j=1}^{2n} M(\infty; 0)_{ij} v_j \quad (\text{A48})$$

$$= M(\infty; 0)_{i1} v_1 + M(\infty; 0)_{i(n+1)} v_{(n+1)} + M(\infty; 0)_{ii} v_i \quad (\text{A49})$$

$$= -\mu_i^2 v_i \quad (\text{A50})$$

and so

$$|\lambda(\infty; 0)| = |e^{-\mu_i^2}| < 1. \quad (\text{A51})$$

Consequently, this holds if any $J_i(t_U; 0) = 0$. Conversely, suppose that $J_i(t_U; 0) \neq 0$ for all $i \notin \{1, (n+1)\}$. Then, there exists some $\alpha > 0$ and some non-negative vector w such that

$$J(t_U; 0) = \alpha v + w. \quad (\text{A52})$$

Therefore, for any positive integer n ,

$$e^{n\mathbf{M}(\infty;0)} \mathbf{J}(t_U; 0) = e^{n\mathbf{M}(\infty;0)} (\alpha \mathbf{v} + \mathbf{w}) = \lambda(\infty; 0)^n \alpha \mathbf{v} + e^{n\mathbf{M}(\infty;0)} \mathbf{w} \geq \lambda(\infty; 0)^n \alpha \mathbf{v}. \quad (\text{A53})$$

Now, \mathbf{v} is an eigenvector so it has a nonzero component, which means that

$$\left(\lim_{n \rightarrow \infty} (e^{n\mathbf{M}(\infty;0)} \mathbf{J}(t_U; 0)) = \mathbf{0} \right) \Rightarrow \left(\lim_{n \rightarrow \infty} (\lambda(\infty; 0)^n \alpha \mathbf{v}) = \mathbf{0} \right) \Rightarrow \left(|\lambda(\infty; 0)| < 1 \right) \quad (\text{A54})$$

and so $|\lambda(\infty; 0)| < 1$ holds in all cases, which finishes the proof of this claim.

Claim *There exists some constant X independent of δ such that $\int_T^\infty J_i(s; \epsilon) ds \leq X\delta$*

Proof Now, the exponentials of the eigenvalues of $\mathbf{M}(\infty; 0)$ are the eigenvalues of $e^{\mathbf{M}(\infty;0)}$ which means that, if $\eta(\infty; 0)$ is an eigenvalue of $\mathbf{M}(\infty; 0)$ then there exists some $\kappa > 0$ such that

$$|e^{\eta(\infty;0)}| \leq |\lambda(\infty; 0)| < e^{-4\kappa} < 1 \Rightarrow |e^{\Re(\eta(\infty;0))}| < e^{-4\kappa} \Rightarrow \Re(\eta) < -4\kappa \quad (\text{A55})$$

and so all eigenvalues of $\mathbf{M}(\infty; 0)$ have strictly negative real part. Thus, by continuous dependence of eigenvalues on the matrix, as $\mathbf{M}(t; 0)$ converges to $\mathbf{M}(\infty; 0)$ as $t \rightarrow \infty$, there exists some $T > t_U$ such that

$$\Re(\eta(t; 0)) < -2\kappa \quad \forall t > T \quad (\text{A56})$$

where $\eta(t; 0)$ is an eigenvalue of $\mathbf{M}(t; 0)$. Now, fix $\delta > 0$. From Lemma 11, by choosing T to be sufficiently large, one can assume that

$$J_i(T; 0) < \delta \quad \forall i \in \{1, \dots, 2n\}. \quad (\text{A57})$$

Moreover, there exists some Δ (which is dependent on T) such that

$$\Re(\eta(T; \epsilon)) < -\kappa \quad \forall \epsilon \in [0, \Delta]. \quad (\text{A58})$$

Now, similarly, by choosing Δ to be sufficiently small, one can assume that by Lemma 14

$$|J_i(t; \epsilon) - J_i(t; 0)| < \delta \quad \forall t < T \Rightarrow |J_i(T; \epsilon)| < 2\delta \quad \forall i \in \{1, \dots, 2n\} \quad \text{and} \quad \forall \epsilon \in [0, \Delta] \quad (\text{A59})$$

and, for all $t < T$,

$$|R_i(T; \epsilon) - R_i(T; 0)|, |R_i^V(T; \epsilon) - R_i^V(T; 0)| < \delta \quad \forall i \in \{1, \dots, 2n\}, \quad \text{and} \quad \forall \epsilon \in [0, \Delta]. \quad (\text{A60})$$

Now, for any $t > 0$,

$$\mathbf{M}(t + T; \epsilon) \leq \mathbf{M}(T; \epsilon). \quad (\text{A61})$$

Thus, as the solution to the system

$$\frac{dz}{dt} = \mathbf{M}(T; \epsilon)z, \quad z(0) = \mathbf{J}(T; \epsilon) \quad (\text{A62})$$

is

$$z(t) = e^{\mathbf{M}(T; \epsilon)t} \mathbf{J}(T; 0); \quad (\text{A63})$$

one has, by Lemma 9,

$$\mathbf{J}(t + T; \epsilon) \leq e^{\mathbf{M}(T; \epsilon)t} \mathbf{J}(T; 0). \quad (\text{A64})$$

Now, noting that $\mathbf{M}(T; \epsilon)$ is invertible as all its eigenvalues have strictly negative real part, for any $t > 0$

$$\int_T^{t+T} \mathbf{J}(s; \epsilon) ds \leq \int_0^t e^{\mathbf{M}(T; \epsilon)s} \mathbf{J}(T; \epsilon) ds \quad (\text{A65})$$

$$= \mathbf{M}(T; \epsilon)^{-1} (e^{\mathbf{M}(T; \epsilon)t} \mathbf{J}(T; \epsilon) - \mathbf{J}(T; \epsilon)) \quad (\text{A66})$$

and so, taking t to ∞ and noting that all eigenvalues of $e^{\mathbf{M}(T; \epsilon)}$ have real part less than 1 shows that

$$\int_T^\infty \mathbf{J}(s; \epsilon) ds \leq -\mathbf{M}(T; \epsilon)^{-1} \mathbf{J}(T; \epsilon). \quad (\text{A67})$$

Now, each element of $\mathbf{M}(t; \epsilon)$ is uniformly bounded (for any bounded range of ϵ and all $t \geq 0$) as the parameters and variables are uniformly bounded. Thus, by expressing the inverse in terms of determinants of sub-matrices of $\mathbf{M}(t; \epsilon)$ (each of which must be uniformly bounded as $\mathbf{M}(t; \epsilon)$ is uniformly bounded) by Cramer's rule (Blyth and Robertson 2002), one can see that there exists a constant M^* such that for each i and j ,

$$\det(\mathbf{M}(t; \epsilon)) \neq 0 \Rightarrow |\mathbf{M}(t; \epsilon)_{ij}^{-1}| \leq \left| \frac{M^*}{\det(\mathbf{M}(t; \epsilon))} \right|. \quad (\text{A68})$$

Note that

$$|\det(\mathbf{M}(T; \epsilon))| = \left| \prod_{\lambda \text{ eigenvalue of } \mathbf{M}(T; \epsilon)} (\lambda) \right| \geq \kappa^n \quad (\text{A69})$$

because all eigenvalues of $M(T; \epsilon)$ have real part at most $-\kappa$ and hence modulus at least κ . Thus, there exists some constant X (independent of δ) such that for each i and j ,

$$\left| M(T; \epsilon)_{ij}^{-1} \right| \leq \frac{X}{4n}. \quad (\text{A70})$$

Thus, by the conditions on $J(T; \epsilon)$,

$$\int_T^\infty J_i(s; \epsilon) ds \leq X\delta \quad (\text{A71})$$

which completes the proof of this claim. \square

As all the parameters and variables are uniformly bounded for all ϵ , there exists a constant Y (independent of δ) such that

$$\left| \frac{dJ_i}{dt} \right| \leq Y \quad \forall i \in \{1, \dots, 2n\}. \quad (\text{A72})$$

Now, suppose there exists some $J_i(t; \epsilon) > \delta^{\frac{1}{3}}$ for $t > T$ and $\epsilon \in [0, \eta_1]$. Then, by non-negativity of J_i

$$\int_T^\infty J_i(s; \epsilon) ds \geq \int_t^{t+\delta^{\frac{1}{2}}} J_i(s; \epsilon) ds \geq \int_0^{\delta^{\frac{1}{2}}} \delta^{\frac{1}{3}} - Ys ds = \delta^{\frac{5}{6}} - \frac{Y}{2}\delta. \quad (\text{A73})$$

Thus, taking δ sufficiently small such that

$$\delta^{\frac{5}{6}} - \frac{Y}{2}\delta > X\delta \quad (\text{A74})$$

gives a contradiction. This means that, for each $i \in \{1, \dots, 2n\}$

$$J_i(t; \epsilon) \leq \delta^{\frac{1}{3}} \quad \forall t \geq T \quad \text{and} \quad \forall \epsilon \in [0, \Delta] \quad (\text{A75})$$

and hence, combining this with (A59) (and assuming $\delta < 1$ so $\delta < \delta^{\frac{1}{3}}$) shows that

$$|J_i(t; \epsilon) - J_i(t; 0)| \leq \delta^{\frac{1}{3}} \quad \forall t \quad \text{and} \quad \forall i \in \{1, \dots, 2n\}. \quad (\text{A76})$$

Moreover, by (A71), for any $t > 0$

$$|R_i(T+t; \epsilon) - R_i(T+t; 0)| \leq |R_i(T; \epsilon) - R_i(T; 0)| + |R_i(T; 0) - R_i(T+t; 0)| \quad (\text{A77})$$

$$\leq \delta + |R_i(T+t; \epsilon) - R_i(T; \epsilon)| + |R_i(T+t; 0) - R_i(T; 0)| \quad (\text{A78})$$

$$\leq \delta + 2X\mu_i^1(\epsilon)\delta + 2X\mu_i^1(0)\delta \quad (\text{A79})$$

$$\leq X^*\delta \quad (\text{A80})$$

for some constant X^* , alongside an identical result for R_i^V . Combining this with (A60) (and redefining $\delta \rightarrow \delta^3$), the result of the proposition is proved.

Theorem 1

Note that Proposition 5 also holds for the vaccination policies $\tilde{U}(t; \epsilon)$, using Proposition 4. Thus, one can define a function $\delta(\epsilon)$ such that for all sufficiently small ϵ

$$|f_i(t; \epsilon) - f_i(t; 0)|, |\tilde{f}_i(t; \epsilon) - f_i(t; 0)| \leq \delta(\epsilon) \quad \forall f \in \{I, I^V, R, R^V\} \quad (\text{A81})$$

and

$$\delta(\epsilon) = o(1) \quad \text{as } \epsilon \rightarrow 0. \quad (\text{A82})$$

Then, using, for example

$$|R_i(\infty; \epsilon) - \tilde{R}_i(\infty; \epsilon)| \leq |R_i(\infty; \epsilon) - \tilde{R}_i(\infty; 0)| + |\tilde{R}_i(\infty; \epsilon) - \tilde{R}_i(\infty; 0)| \quad (\text{A83})$$

(as $R(\infty; 0) = \tilde{R}(\infty; 0)$) shows that

$$|R_i(\infty; \epsilon) - \tilde{R}_i(\infty; \epsilon)|, |R_i^V(\infty; \epsilon) - \tilde{R}_i^V(\infty; \epsilon)| < 2\delta(\epsilon) \quad \forall \epsilon \in [0, \eta] \quad (\text{A84})$$

which means

$$\left| \sum_{j=2}^n p_j \left(R_j(\infty; \epsilon) + \kappa_j R_j^V(\infty; \epsilon) \right) - \sum_{j=2}^n p_j \left(\tilde{R}_j(\infty; \epsilon) + \kappa_j \tilde{R}_j^V(\infty; \epsilon) \right) \right| = O(\delta). \quad (\text{A85})$$

Thus, the aim of the remainder of the proof is to show that the leading-order changes to $R_1(\infty; \epsilon)$ are of exactly $O(\epsilon)$, and so $p_1 R_1(\infty; \epsilon)$ changes by an $O(1)$ amount, meaning these changes to the objective function will eventually dominate the other changes given in (A85). This can be done by taking advantage of the fact that the quantities $f_1(t; \epsilon)$ are small, and so a linearised version of the equations for group 1 can be used.

Before beginning this process, it is helpful to note the following. From (A56) in the proof of Proposition 5, there exists some $T^* > t_U$ independent of δ and ϵ such that

$$\lambda(T^*; 0) < e^{-2\kappa} < 1 \quad (\text{A86})$$

where $\lambda(T^*; 0)$ is the (necessarily real and non-negative) Perron eigenvalue of $e^{M(T^*; 0)}$ (and is the exponential of the $\eta(\infty; 0)$ referenced in (A56)). Moreover, by the continuity of eigenvalues on the entries of the matrix, there exists some $\Delta > 0$ such that the analogously defined $\lambda(T^*; \epsilon)$ also satisfies

$$\lambda(T^*; \epsilon) < e^{-\kappa} < 1 \quad \forall \epsilon \in [0, \Delta]. \quad (\text{A87})$$

Now, note that, for $t \geq T^* > t_U$, the matrix $M(t; \epsilon)$ and hence the matrix $e^{M(t; \epsilon)}$ is non-increasing. Thus, as $e^{M(t; \epsilon)}$ is non-negative (as proved in Proposition 5), it is necessary from Perron–Frobenius theory (Berman and Plemmons 1994) that its Perron eigenvalue, $\lambda(t; \epsilon)$ satisfies

$$\lambda(t; \epsilon) \leq \lambda(T^*; \epsilon) < e^{-\kappa} < 1. \quad (\text{A88})$$

Then, following the method used to derive (A67), one has, for any $t \geq T^*$

$$\int_t^\infty I_1(t; \epsilon) dt \leq (M(t; \epsilon)^{-1} J(t; \epsilon))_1 \quad \forall \epsilon \in [0, \Delta]. \quad (\text{A89})$$

This is exactly the same equation as (A67), except that here, T^* is independent of δ (as no conditions on $J(T; 0)$ are assumed). Now, note that

$$M(t; 0)_{1j} = -\mu_1^1 \delta_{1j} \quad \text{and} \quad M(t; 0)_{(n+1)j} = -\mu_1^2 \delta_{(n+1),j} \quad (\text{A90})$$

where here δ_{ij} is the Kronecker delta. This means that, for any vector \mathbf{y} , the equation

$$M(t; 0)\mathbf{x} = \mathbf{y} \quad (\text{A91})$$

must satisfy

$$x_1 = \frac{-y_1}{\mu_1^1} \quad x_{n+1} = -\frac{y_{n+1}}{\mu_1^2} \quad \text{and} \quad \mathbf{x} = M^{-1}\mathbf{y}. \quad (\text{A92})$$

Thus, in particular

$$M_{1j}^{-1}(t; 0) = \frac{-1}{\mu_1^1} \delta_{1j} \quad \text{and} \quad M_{(n+1)j}^{-1}(t; 0) = \frac{-1}{\mu_1^2} \delta_{(n+1)j}, \quad (\text{A93})$$

where here δ_{ij} denotes the Kronecker delta. Now, note that, as the inverse of a matrix is a rational function of its entries,

$$M^{-1}(t; 0) = M^{-1}(t; \epsilon) + O(\epsilon) \quad (\text{A94})$$

and hence

$$M_{1j}^{-1}(t; 0) = \frac{-1}{\mu_1^1} \delta_{1j} + O(\epsilon). \quad (\text{A95})$$

Moreover, defining

$$\mu_{\min} := \min\{\mu_i^1, \mu_i^2\}, \quad (\text{A96})$$

there must exist a $T(\epsilon) \in \left(T^*, T^* + \frac{2n}{\delta^{\frac{1}{3}}\mu_{\min}}\right)$ such that for each i ,

$$I_i(T(\epsilon); \epsilon) < \delta^{\frac{1}{3}} N_i(\epsilon). \quad (\text{A97})$$

Otherwise,

$$\sum_{i=1}^n \frac{d}{dt} \left(\frac{R_i(t; \epsilon)}{\mu_i^1 N_i(\epsilon)} + \frac{R_i^V(t; \epsilon)}{\mu_i^2 N_i(\epsilon)} \right) \geq \sum_{i=1}^n \left(\frac{\mu_i^1 I_i(t; \epsilon)}{\mu_i^1 N_i(\epsilon)} + 0 \right) \geq \delta^{\frac{1}{3}} \quad \forall t \in \left(T^*, T^* + \frac{2n}{\delta^{\frac{1}{3}}\mu_{\min}}\right) \quad (\text{A98})$$

and integrating this between T^* and $T^* + \frac{2n}{\delta^{\frac{1}{3}}\mu_{\min}}$ gives

$$\sum_{i=1}^n \left[\frac{R_i\left(T^* + \frac{2n}{\delta^{\frac{1}{3}}\mu_{\min}}; \epsilon\right)}{\mu_i^1 N_i} + \frac{R_i^V\left(T^* + \frac{2n}{\delta^{\frac{1}{3}}\mu_{\min}}; \epsilon\right)}{\mu_i^2 N_i(\epsilon)} \right] \geq \frac{2n\delta^{\frac{1}{3}}}{\delta^{\frac{1}{3}}\mu_{\min}} > \frac{n}{\mu_{\min}}. \quad (\text{A99})$$

Thus, as $\frac{\mu_{\min}}{\mu_i^\alpha} \leq 1$ for each i and α ,

$$\sum_{i=1}^n \left[\frac{R_i\left(T^* + \frac{2n}{\delta^{\frac{1}{3}}(\mu_{\min}+1)}; 0\right) + R_i^V\left(T^* + \frac{2n}{\delta^{\frac{1}{3}}(\mu_{\min}+1)}; 0\right)}{N_i(\epsilon)} \right] > n \quad (\text{A100})$$

which means, for some i

$$\frac{R_i\left(T^* + \frac{2n}{\delta^{\frac{1}{3}}(\mu_{\min}+1)}; 0\right) + R_i^V\left(T^* + \frac{2n}{\delta^{\frac{1}{3}}(\mu_{\min}+1)}; 0\right)}{N_i(\epsilon)} > 1, \quad (\text{A101})$$

which is a contradiction as the total population size in group i cannot exceed $N_i(\epsilon)$ by definition of $N_i(\epsilon)$. Thus, for each $\epsilon \in [0, \Delta]$,

$$\int_{T(\epsilon)}^{\infty} I_1(t; \epsilon) dt \leq (\mathbf{M}(T; \epsilon)^{-1} \mathbf{J}(T(\epsilon); \epsilon))_1 \quad (\text{A102})$$

$$= \begin{pmatrix} O(\epsilon\delta^{\frac{1}{3}}) \\ O(\delta^{\frac{1}{3}}) \\ \cdot \\ \cdot \\ O(\delta^{\frac{1}{3}}) \end{pmatrix} \quad (\text{A103})$$

$$= O(\epsilon\delta^{\frac{1}{3}}) \quad (\text{A104})$$

while similarly

$$\int_{T(\epsilon)}^{\infty} I_1^V(t; \epsilon) dt = O(\epsilon\delta^{\frac{1}{3}}). \quad (\text{A105})$$

Moreover,

$$\int_0^{T(\epsilon)} \delta \epsilon dt = O(\epsilon\delta^{\frac{2}{3}}). \quad (\text{A106})$$

These results allow for the linearisation to be carried out. To reduce notation, define

$$T := T(\epsilon). \quad (\text{A107})$$

Now, to begin the linearisation, define

$$X(t) = \sum_{j=1}^n \left[\beta_{1j}^1 I_j(t; 0) + \beta_{1j}^2 I_j^V(t; 0) \right], \quad (\text{A108})$$

which is the leading-order infective force on group 1. By Proposition 5,

$$X(t) = \sum_{j=1}^n \left[\beta_{1j}^1 I_j(t; \epsilon) + \beta_{1j}^2 I_j^V(t; \epsilon) \right] + O(\delta). \quad (\text{A109})$$

Then, as $S_1(t; \epsilon) \leq \epsilon$,

$$\frac{dI_1}{dt}(t; \epsilon) + \mu_1^1 I_1(t) = S_1(t; \epsilon) X(t) + O(\delta\epsilon). \quad (\text{A110})$$

Now, note that

$$R_1(\infty; \epsilon) = \mu_1^1 \int_0^{\infty} I_1(t; \epsilon) dt \quad (\text{A111})$$

$$= \mu_1^1 \int_0^T I_1(t; \epsilon) dt + \mu_1^1 \int_T^{\infty} I_1(t; \epsilon) dt \quad (\text{A112})$$

$$= \int_0^T \left(S_1(t; \epsilon) X(t) - \frac{dI_1}{dt}(t; \epsilon) + O(\epsilon\delta) \right) dt + O(\delta^{\frac{1}{3}}\epsilon) \quad (\text{A113})$$

$$= I_1(0; \epsilon) - I_1(T) + \int_0^T S_1(t; \epsilon) X(t) dt + O(\delta^{\frac{1}{3}}\epsilon) \quad (\text{A114})$$

$$= I_1(0; \epsilon) + \int_0^T S_1(t; \epsilon) X(t) dt + O(\delta^{\frac{1}{3}}\epsilon). \quad (\text{A115})$$

Now, the equations for I^V are of the same form, but with S^V in place of S and a different leading-order infection function $Y(t)$ given by

$$Y(t) = \sum_{j=1}^n \left[\beta_{1j}^3 I_j(t; 0) + \beta_{ij}^4 I_j^V(t; 0) \right]. \quad (\text{A116})$$

Thus, an analogous derivation (noting that $I^V(0; \epsilon) = 0$) shows that

$$R_1^V(\infty; 0) = \int_0^T Y(t) S_1^V(t; \epsilon) dt + O(\epsilon\delta^{\frac{1}{3}}) \quad (\text{A117})$$

while analogous results hold for \tilde{R}_1 and \tilde{R}_1^V (with \tilde{S}_1 and \tilde{S}_1^V in place of S_1 and S_1^V). Now, note that

$$S_1(t; \epsilon) = S_1(t; \epsilon) \left(\frac{N_1(\epsilon) - W_1(t; \epsilon)}{N_1(\epsilon)} \right) \exp \left[- \sum_{j=1}^n \left(\frac{\beta_{1j}^1 R_j(t; \epsilon)}{\mu_j^1} + \frac{\beta_{1j}^2 R_j^V(t; \epsilon)}{\mu_j^2} \right) \right] \quad (\text{A118})$$

$$= \sigma(N_1(\epsilon) - W_1(t; \epsilon)) \exp \left[- \sum_{j=1}^n \left(\frac{\beta_{1j}^1 R_j(t; \epsilon)}{\mu_j^1} + \frac{\beta_{1j}^2 R_j^V(t; \epsilon)}{\mu_j^2} \right) \right]. \quad (\text{A119})$$

Define

$$P(t) := \exp \left[- \sum_{j=1}^n \left(\frac{\beta_{1j}^1 R_j(t; 0)}{\mu_j^1} + \frac{\beta_{1j}^2 R_j^V(t; 0)}{\mu_j^2} \right) \right] \quad (\text{A120})$$

and then, note that by Proposition 5

$$P(t) = \exp \left[- \sum_{j=1}^n \left(\frac{\beta_{1j}^1 R_j(t; \epsilon)}{\mu_j^1} + \frac{\beta_{1j}^2 R_j^V(t; \epsilon)}{\mu_j^2} \right) \right] + O(\delta) \quad (\text{A121})$$

which means (as $(N_1(\epsilon) - W_1(t; \epsilon)) \leq \epsilon$ and $\sigma < 1$)

$$S_1(t; \epsilon) = \sigma(N_1 - W_1(t; \epsilon)) P(t) + O(\delta\epsilon) \quad (\text{A122})$$

with an identical result for \tilde{S} . It is helpful to note for later that, as $W_1(t; \epsilon) \leq \tilde{W}(t; \epsilon)$, this means that

$$S_1(t; \epsilon) \geq \tilde{S}_1(t; \epsilon) + O(\delta\epsilon). \quad (\text{A123})$$

Now, this means

$$\int_0^T X(t) S_1(t; \epsilon) dt = \int_0^T X(t) \sigma(N_1 - W_1(t; \epsilon)) P(t) dt + O(\epsilon \delta^{\frac{2}{3}}) \quad (\text{A124})$$

and so

$$R_1(\infty; \epsilon) = I_1(0; \epsilon) + \int_0^T X(t) \sigma(N_1 - W_1(t; \epsilon)) P(t) dt + O(\epsilon \delta^{\frac{1}{3}}). \quad (\text{A125})$$

Now, note that

$$\int_0^T X(t) \sigma(N_1 - W_1(t; \epsilon)) P(t) dt = \left(\int_0^\tau + \int_\tau^T \right) \left(X(t) \sigma(N_1 - W_1(t; \epsilon)) P(t) dt \right) \quad (\text{A126})$$

and that, as $W_1(t; \epsilon) \leq \tilde{W}_1(t; \epsilon)$,

$$\int_\tau^T X(t) \sigma(N_1 - W_1(t; \epsilon)) P(t) dt \geq \int_\tau^T X(t) \sigma(N_1 - \tilde{W}_1(t; \epsilon)) P(t) dt. \quad (\text{A127})$$

Now, define $z(\epsilon)$ to be

$$z(\epsilon) = \inf \left\{ t : \sum_{i=1}^n W_i(t) = \epsilon \right\}. \quad (\text{A128})$$

Note that, for $\epsilon < w$, z exists and is bounded above by τ as

$$\sum_{i=1}^n W_i(\tau) = w. \quad (\text{A129})$$

Now, define a fixed value

$$z_0 := z\left(\frac{w}{2}\right) \quad (\text{A130})$$

so that, by continuity of W , $z_0 < \tau$ (and is independent of ϵ). Suppose that $\epsilon < \frac{w}{2}$ (which will be assumed for the rest of the proof). Note that

$$\int_0^{z_0} X(t) \sigma(N_1 - W_1(t; \epsilon)) P(t) dt \geq \int_0^{z_0} X(t) \sigma(N_1 - \tilde{W}_1(t; \epsilon)) P(t) dt \quad (\text{A131})$$

and that

$$\int_{z_0}^{\tau} X(t) \sigma(N_1 - \tilde{W}_1(t; \epsilon)) P(t) dt = 0 \quad (\text{A132})$$

as $\tilde{W}_1(t; \epsilon) = N_1$ for all $t > z(\epsilon)$. Moreover, by (A2)

$$\int_{z_0}^{\tau} X(t) \sigma(N_1 - W_1(t; \epsilon)) P(t) dt \geq (1 - \alpha) \epsilon \sigma \int_{z_0}^{\tau} X(t) P(t) dt. \quad (\text{A133})$$

Now, note that $P(t)$ is strictly positive for $t > 0$ as it is an exponential, while, as $\beta_{1j} > 0$ for some $j \neq 1$,

$$X(t) \geq \beta_{ij} I_j(t; 0) > 0 \quad \text{as } j \in \Pi(0). \quad (\text{A134})$$

Thus,

$$(1 - \alpha) \int_{z_0}^{\tau} X(t) P(t) dt > 0 \quad (\text{A135})$$

and this is independent of ϵ . This means that

$$\begin{aligned} R_1(\infty; \epsilon) - \tilde{R}(\infty; \epsilon) &\geq \epsilon(1 - \alpha) \int_z^{\tau} X(t) P(t) dt + O(\epsilon \delta^{\frac{1}{3}}) \\ &= \epsilon(1 - \alpha) \int_z^{\tau} X(t) P(t) dt + o(\epsilon) \end{aligned} \quad (\text{A136})$$

and so the leading-order change in $R_1(\infty; \epsilon)$ is indeed of order exactly ϵ .

Now, it is important to check the leading-order change in $R_1^V(\infty; \epsilon)$. Note that, as $S_1(t; \epsilon)$ and $S_1^V(\epsilon)$ are at most ϵ ,

$$\frac{d}{dt} \left(S_1(t; \epsilon) + S_1^V(t; \epsilon) \right) = -X(t) S_1(t; \epsilon) - Y(t) S_1^V(t; \epsilon) + O(\epsilon \delta). \quad (\text{A137})$$

Using (A122), this can be written as:

$$\frac{d}{dt} \left(S_1(t; \epsilon) + S_1^V(t; \epsilon) \right) + Y(t) (S_1(t; \epsilon) + S_1^V(t; \epsilon)) = (Y(t) - X(t)) S_1(t; \epsilon) + O(\epsilon \delta). \quad (\text{A138})$$

This equation can be integrated by defining

$$\mathcal{Y}(t) := \int_0^t Y(s) ds \quad (\text{A139})$$

so that

$$\frac{d}{dt} \left(e^{\mathcal{Y}(t)} (S_1(t; \epsilon) + S_1^V(t; \epsilon)) \right) = e^{\mathcal{Y}(t)} (Y(t) - X(t)) S_1(t; \epsilon) + O(\epsilon \delta). \quad (\text{A140})$$

Thus, for any $t \leq T$

$$\begin{aligned} S_1(t; \epsilon) + S_1^V(t; \epsilon) &= e^{-\mathcal{Y}(t)} (S_1(0; \epsilon) + S_1^V(0; \epsilon)) \\ &\quad + \int_0^t e^{\mathcal{Y}(s) - \mathcal{Y}(t)} (Y(s) - X(s)) S_1(s; \epsilon) ds + O(\epsilon \delta^{\frac{2}{3}}) \end{aligned} \quad (\text{A141})$$

which means that

$$\begin{aligned} S_1(t; \epsilon) + S_1^V(t; \epsilon) - \tilde{S}_1(t; \epsilon) - \tilde{S}_1^V(t; \epsilon) \\ = \int_0^t e^{\mathcal{Y}(s) - \mathcal{Y}(t)} (Y(s) - X(s)) \left(S_1(s; \epsilon) - \tilde{S}_1(s; \epsilon) \right) ds + O(\epsilon \delta^{\frac{2}{3}}) \end{aligned} \quad (\text{A142})$$

Thus,

$$\int_0^t Y(s) \left[S_1(s; \epsilon) + S_1^V(s; \epsilon) - \tilde{S}_1(s; \epsilon) - \tilde{S}_1^V(s; \epsilon) \right] ds \quad (\text{A143})$$

$$= \int_0^t \int_0^s Y(s) e^{\mathcal{Y}(k) - \mathcal{Y}(s)} (Y(k) - X(k)) \left(S_1(k; \epsilon) - \tilde{S}_1(k; \epsilon) \right) dk ds + O(\epsilon \delta^{\frac{1}{3}}) \quad (\text{A144})$$

$$= \int_0^t \int_k^t \left[Y(s) e^{-\mathcal{Y}(s)} \right] e^{\mathcal{Y}(k)} (Y(k) - X(k)) \left(S_1(k; \epsilon) - \tilde{S}_1(k; \epsilon) \right) ds dk + O(\epsilon \delta^{\frac{1}{3}}) \quad (\text{A145})$$

$$= \int_0^t (e^{-\mathcal{Y}(k)} - e^{-\mathcal{Y}(t)}) e^{\mathcal{Y}(k)} (Y(k) - X(k)) \left(S_1(k; \epsilon) - \tilde{S}_1(k; \epsilon) \right) dk + O(\epsilon \delta^{\frac{1}{3}}) \quad (\text{A146})$$

$$= \int_0^t (1 - e^{\mathcal{Y}(k) - \mathcal{Y}(t)}) (Y(k) - X(k)) \left(S_1(k; \epsilon) - \tilde{S}_1(k; \epsilon) \right) dk + O(\epsilon \delta^{\frac{1}{3}}). \quad (\text{A147})$$

Now, note that, as \mathcal{Y} is non-decreasing, and non-negative

$$0 \leq 1 - e^{\mathcal{Y}(k) - \mathcal{Y}(t)} \leq 1 - e^{-\mathcal{Y}(t)}. \quad (\text{A148})$$

Moreover, one has

$$\mathcal{Y}(t) = \int_0^t \sum_{j=1}^n \left[\beta_{1j}^3 I_j(s; 0) + \beta_{1j}^4 I_j^V(s; 0) \right] ds \quad (\text{A149})$$

$$= \sum_{j=1}^n \left[\frac{\beta_{1j}^3 R_j(t; 0)}{\mu_j^1} + \frac{\beta_{1j}^4 R_j(t; 0)}{\mu_j^2} \right] \quad (\text{A150})$$

$$\leq \sum_{j=1}^n \left[\frac{\beta_{1j}^3 N_j(1)}{\mu_j^1} + \frac{\beta_{1j}^4 N_j(1)}{\mu_j^2} \right] \quad (\text{A151})$$

and so $\mathcal{Y}(t)$ is bounded above by some constant (for $\epsilon \leq 1$). This in turn means that there exists some \mathcal{Y}^* such that

$$1 - e^{-\mathcal{Y}(t)} \leq \mathcal{Y}^* < 1. \quad (\text{A152})$$

Thus, as $Y(t) - X(t) \leq 0$ and $S_1(k; \epsilon) \geq \tilde{S}_1(k; \epsilon) + O(\delta\epsilon)$, for any $k \leq t$

$$\begin{aligned} & \int_0^t Y(s) \left[S_1(s; \epsilon) + S_1^V(s; \epsilon) - \tilde{S}_1(s; \epsilon) - \tilde{S}_1^V(s; \epsilon) \right] \\ & \geq \mathcal{Y}^* \int_0^t (Y(k) - X(k)) \left(S_1(k; \epsilon) - \tilde{S}_1(k; \epsilon) \right) dk + O(\epsilon\delta^{\frac{1}{3}}). \end{aligned} \quad (\text{A153})$$

Now, adding inequalities (A115) and (A117) together gives

$$R_1(\infty; \epsilon) + R_1^V(\infty; \epsilon) = I_1(0; \epsilon) + \int_0^T X(t)S_1(t; \epsilon) + Y(t)S_1^V(t; \epsilon)dt + o(\epsilon). \quad (\text{A154})$$

Note that

$$X(t)S_1(t; \epsilon) + Y(t)S_1^V(t; \epsilon) = (X(t) - Y(t))S_1(t; \epsilon) + Y(t)(S_1(t; \epsilon) + S_1^V(t; \epsilon)) \quad (\text{A155})$$

and hence

$$\begin{aligned} & R_1(\infty; \epsilon) + R_1^V(\infty; \epsilon) = I_1(0; \epsilon) + \\ & \int_0^T (X(t) - Y(t))S_1(t; \epsilon) + Y(t)(S_1(t; \epsilon) + S_1^V(t; \epsilon))dt + o(\epsilon). \end{aligned} \quad (\text{A156})$$

This means that

$$\begin{aligned} & R_1(\infty; \epsilon) + R_1^V(\infty; \epsilon) - \tilde{R}_1(\infty; \epsilon) - \tilde{R}_1^V(\infty; \epsilon) \\ & \geq (1 - \mathcal{Y}^*) \int_0^T (X(t) - Y(t)) \left(S_1(t; \epsilon) - \tilde{S}_1(t; \epsilon) \right) dt + O(\epsilon\delta^{\frac{1}{3}}). \end{aligned} \quad (\text{A157})$$

Now, as there is some $i \neq 1$ such that

$$\beta_{1i}^1 > \beta_{1i}^3 \geq 0 \quad (\text{A158})$$

and (as $i \neq 1$), $i \in \Pi(0)$ which means that

$$\beta_{1i}^1 I_i(t) > \beta_{1i}^3 I_i(t) \quad \forall t > 0. \quad (\text{A159})$$

This means that $X(t) > Y(t)$ for all $t > 0$ and hence

$$\int_0^T (X(t) - Y(t))dt > 0. \quad (\text{A160})$$

Thus, following the arguments from before, one can see that

$$\int_0^t (X(s) - Y(s)) \left(S_1(s; \epsilon) - \tilde{S}_1(s; \epsilon) \right) ds > \epsilon(1 - \mathcal{Y}^*) \int_{z_0}^{\tau} (X(t) - Y(t)) P(t) dt + o(\epsilon) \quad (\text{A161})$$

where the leading-order term is positive as required (as $P(t)$ is positive). Hence, from (A157)

$$\begin{aligned} R_1(\infty; \epsilon) + R_1^V(\infty; \epsilon) - (\tilde{R}_1(\infty; \epsilon) + \tilde{R}_1^V(\infty; \epsilon)) \\ \geq (1 - \mathcal{Y}^*)\epsilon(1 - \alpha) \int_{z_0}^{\tau} (X(t) - Y(t)) P(t) dt + o(\epsilon). \end{aligned} \quad (\text{A162})$$

Thus, for any $\kappa_1 \in [0, 1]$, combining (A136) and (A162)

$$\begin{aligned} R_1(\infty) + \kappa_1 R_1^V(\infty) &= \kappa_1 (R_1(\infty) + R_1^V(\infty)) + (1 - \kappa_1) R_1(\infty) \\ &\geq \epsilon \int_{z_0}^{\tau} (1 - \alpha) P(t) \left[(1 - \mathcal{Y}^*)\kappa_1 (X(t) - Y(t)) + (1 - \kappa_1) X(t) \right] dt \end{aligned} \quad (\text{A163})$$

$$+ \kappa_1 \tilde{R}_1^V(\infty) + \tilde{R}_1(\infty) + o(\epsilon). \quad (\text{A164})$$

Thus, recalling (A85) and that $p_1 = \frac{1}{\epsilon}$

$$H(U) \geq H(\tilde{U}) + \int_{z_0}^{\tau} (1 - \alpha) [\kappa_1 (X(t) - Y(t)) + (1 - \kappa_1) X(t)] dt + o(1) \quad (\text{A165})$$

for some constant K . Moreover, for sufficiently small ϵ ,

$$\int_{z_0}^{\tau} \alpha [\kappa_1 (X(t) - Y(t)) + (1 - \kappa_1) X(t)] dt + o(1) > 0 \quad (\text{A166})$$

and hence

$$H(U(t; \epsilon)) > H(\tilde{U}(t; \epsilon)), \quad (\text{A167})$$

as required.

Appendix B Proof of Theorem 2

Note that, throughout this section, the tilde will be removed from the rescaled p_i terms to reduce notation (and hence all p_i terms used here will be assumed to be rescaled).

Recall from the main text that, using the results in Penn and Donnelly (2022), if one defines

$$\chi(t) := \begin{cases} A(t) & \text{if } \int_0^t A(s)ds < B(t) \\ \min(A(t), B'(t)) & \text{if } \int_0^t A(s)ds \geq B(t) \end{cases}, \quad (\text{B168})$$

then (assuming that there is an optimal solution, and under mild smoothness conditions on U , A and B) there must be an optimal solution satisfying

$$\sum_{i=1}^n W_i(t) = \max \left(\int_0^t \chi(s)ds, 1 \right). \quad (\text{B169})$$

Theorem 2 *With the definitions of Theorem 1, suppose additionally that*

$$\sum_{j=2}^n (\beta_{1j}^1 - \beta_{1j}^3) I_j(0; \epsilon) > 0. \quad (\text{B170})$$

That is, the initial difference between the infective force on vaccinated and unvaccinated members of the population is positive. Suppose further that

$$\sigma = 1. \quad (\text{B171})$$

Suppose an optimal vaccination policy for each ϵ is given by $\bar{U}(t; \epsilon)$ and suppose that $\bar{U}(t; \epsilon)$ has uniformly bounded finite support. Then, there exists an η depending only on α , τ , w and the model parameters such that, for any U satisfying the condition (A2) as defined in Theorem 1

$$\epsilon \in (0, \eta) \Rightarrow H(U(t; \epsilon)) > H(\bar{U}(t; \epsilon)). \quad (\text{B172})$$

Moreover, there is a sequence of optimal vaccination policies $\bar{U}(t; \epsilon)$ satisfying

$$\lim_{\epsilon \rightarrow 0} \left(\frac{\bar{W}_1(t; \epsilon)}{\epsilon} \right) = 1 \quad \forall t \quad \text{s.t.} \quad \int_0^t \chi(s)ds > 0. \quad (\text{B173})$$

To make things clearer in the course of this proof, note that H will be written as

$$H(U; \epsilon) \quad (\text{B174})$$

where the ϵ refers to the size of the population N_1 under consideration.

Proposition 6

It remains to show that, for sufficiently small ϵ and fixed α , τ and w , there is no U satisfying the conditions (A2) that is the optimal vaccination policy.

To do this, it is necessary to prove the function $H(U(t; \epsilon); \epsilon)$ is non-increasing in ϵ . This uses the work of Penn and Donnelly (2022) as the main result of that paper gives a method of finding inequalities between the objective values of different vaccination policies. However, it is a meaningful extension, as here the population sizes are not identical when objective values are compared.

Proposition 6 Suppose that $I_1(0; \epsilon) = 0$ for all ϵ . Consider, for $\epsilon \leq 1$ any bounded vaccination policy $U(t; \epsilon)$ given by

$$U_1(t; \epsilon) = \begin{cases} U_1(t; 1) & \text{if } W_1(t; 1) < \epsilon \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad U_i(t; \epsilon) = U_i(t; 1) \quad \forall i \neq 1 \quad (\text{B175})$$

Then, if $H(U(t; \epsilon); \epsilon)$ is the value of the objective function for a given value of ϵ ,

$$\epsilon > \epsilon' \Rightarrow H(U(t; \epsilon); \epsilon) \geq H(U(t; \epsilon'); \epsilon'). \quad (\text{B176})$$

Proof Fix ϵ and ϵ' such that $\epsilon > \epsilon'$. Define the vaccination policy $U^*(t, \Delta; \epsilon)$ to be

$$U_1^*(t, \Delta; \epsilon) = \begin{cases} \frac{(\epsilon - \epsilon' + W_1(\Delta; \epsilon'))}{\Delta} & \text{if } t < \Delta \\ U_1(t; \epsilon') & \text{otherwise} \end{cases} \quad \text{and} \quad U_i^*(t, \Delta; \epsilon) = U_i(t; 1) \quad \forall i \neq 1. \quad (\text{B177})$$

Then, in particular

$$W_1^*(t, \Delta; \epsilon) = \frac{t(\epsilon - \epsilon' + W_1(\Delta; \epsilon'))}{\Delta} \quad \forall t < \Delta. \quad (\text{B178})$$

Now, as $U(t; \epsilon)$ is bounded by some M , it is necessary that $W_1(t; \epsilon)$ is bounded above by tM . Conversely, $W_1^*(t; \epsilon)$ is bounded below by $\frac{(\epsilon - \epsilon')t}{\Delta}$ for $t < \Delta$. Thus, taking Δ sufficiently small gives

$$W_1^*(t, \Delta; \epsilon) > W_1(t; \epsilon) \quad \forall t < \Delta. \quad (\text{B179})$$

Moreover, note that, assuming $\Delta < \epsilon'$, if $t > \Delta$ is chosen such that $W_1(t; 1) < \epsilon' < \epsilon$, then

$$U_1^*(t, \Delta; \epsilon) = U_1(t; \epsilon') = U_1(t; \epsilon) = U_1(t; 1) \quad (\text{B180})$$

and hence

$$W_1^*(t, \Delta; \epsilon) > W_1(t; \epsilon) \quad \forall t \quad \text{s.t.} \quad W_1(t; 1) < \epsilon'. \quad (\text{B181})$$

Finally, note that if $W_1(t; 1) \geq \epsilon'$ then $W_1(t; \epsilon') = \epsilon'$ and hence

$$W_1^*(t, \Delta; \epsilon) = W_1^*(\Delta; \Delta) + \int_{\Delta}^t U_1(s; \epsilon') ds \quad (\text{B182})$$

$$= \epsilon - \epsilon' + W_1(\Delta; \epsilon') + \int_{\Delta}^t U_1(s; \epsilon') ds \quad (\text{B183})$$

$$= \epsilon - \epsilon' + W_1(\Delta; \epsilon') + W_1(t; \epsilon') - W_1(\Delta; \epsilon') \quad (\text{B184})$$

$$= \epsilon \quad (\text{B185})$$

and so

$$W_1^*(t, \Delta; \epsilon) = \epsilon \geq W_1(t; \epsilon) \quad \forall t \geq 0. \quad (\text{B186})$$

Moreover,

$$W_i^*(t, \Delta; \epsilon) = W_i(t; \epsilon) \quad \forall t \geq 0 \quad \text{and} \quad \forall i \in \{2, \dots, n\}. \quad (\text{B187})$$

Thus, in particular, by Theorem 17, proved in Penn and Donnelly (2022), for each $i \in \{1, \dots, n\}$,

$$\begin{aligned} I_i^*(t, \Delta; \epsilon) + R_i^*(t, \Delta; \epsilon) + I_i^{V*}(t, \Delta; \epsilon) + R_i^{V*}(t, \Delta; \epsilon) \\ \leq I_i(t; \epsilon) + R_i(t; \epsilon) + I_i^V(t; \epsilon) + R_i^V(t; \epsilon) \end{aligned} \quad (\text{B188})$$

and

$$R_i^*(t, \Delta; \epsilon) \leq R_i(t; \epsilon) \quad (\text{B189})$$

where the $f_i^*(t, \Delta; \epsilon)$ are the values of the model variables under the $U^*(t, \Delta; \epsilon)$ vaccination policy and the $f_i(t; \epsilon)$ are their values under the $U(t; \epsilon)$ vaccination policy. \square

Now, for all $\Delta > 0$ and all f and i

$$f_i^*(0, \Delta; \epsilon) = f_i(0; \epsilon) \quad (\text{B190})$$

so, as all model variables except S_i and S_i^V have derivatives that are bounded independently, there exists some L such that, for all $f \in \{I_i(t, \Delta; \epsilon), I_i^V(t, \Delta; \epsilon), R_i(t, \Delta; \epsilon), R_i^V(t, \Delta; \epsilon)\}$,

$$|f_i^*(\Delta; \Delta; \epsilon) - f_i(0; \epsilon)| = |f_i^*(\Delta; \Delta) - f_i^*(0; \Delta)| < L\Delta. \quad (\text{B191})$$

Moreover, the initial conditions are the same for $f(0; \epsilon)$ and $f(0; \epsilon')$ except in the case of $S_1(0; \epsilon)$. Thus,

$$|f_i^*(\Delta; \Delta; \epsilon) - f_i(0; \epsilon')| < L\Delta \quad \forall f \in \{I_i(t, \Delta; \epsilon), I_i^V(t, \Delta; \epsilon), R_i(t, \Delta; \epsilon), R_i^V(t, \Delta; \epsilon)\}. \quad (\text{B192})$$

As only the W_1 policy has an unbounded derivative in the $\Delta \rightarrow 0$ limit, it is also true that

$$|f_i^*(\Delta; \Delta; \epsilon) - f_i(0; \epsilon')| < L\Delta \quad \forall f \in \{S_i(t, \Delta; \epsilon), S_i^V(t, \Delta; \epsilon)\} \quad \text{and } i \neq 1. \quad (\text{B193})$$

Moreover, note that (here suppressing the dependence on ϵ)

$$S_1^*(\Delta; \Delta) = \frac{S_1^*(0; \Delta)}{N_1(\epsilon)} (N_1(\epsilon) - W_1^*(\Delta; \Delta)) e^{\sum_{j=1}^n \left[\frac{\beta_{1j}^1}{\mu_j^1} R_j^*(\Delta; \Delta) + \frac{\beta_{1j}^3}{\mu_j^3} R_j^{V*}(\Delta; \Delta) \right]} \quad (\text{B194})$$

$$= \sigma(\epsilon - (\epsilon - \epsilon') - W_1(\Delta; \epsilon')) e^{\sum_{j=1}^n \left[\frac{\beta_{1j}^1}{\mu_j^1} R_j(0; \epsilon) + \frac{\beta_{1j}^3}{\mu_j^3} R_j^V(0; \epsilon) \right]} + O(\Delta) \quad (\text{B195})$$

$$= \sigma(\epsilon' - W_1(0; \epsilon')) e^{\sum_{j=1}^n \left[\frac{\beta_{1j}^1}{\mu_j^1} R_j(0; \epsilon') + \frac{\beta_{1j}^3}{\mu_j^3} R_j^V(0; \epsilon') \right]} + O(\Delta) \quad (\text{B196})$$

$$= S_1(0; \epsilon') + O(\Delta) \quad (\text{B197})$$

and hence,

$$|S_1(\Delta; \Delta; \epsilon) - S_1(0; \epsilon')| < L' \Delta \quad (\text{B198})$$

for some $L' > 0$. Now, as (again suppressing the dependence on ϵ)

$$S_1(\Delta; \Delta) + I_1(\Delta; \Delta) + R_1(\Delta; \Delta) + S_1^V(\Delta; \Delta) + I_1^V(\Delta; \Delta) + R_1^V(\Delta; \Delta) = \epsilon, \quad (\text{B199})$$

it is necessary that

$$|S_1^V(\Delta; \Delta; \epsilon) - (\epsilon - \epsilon')| \leq L'' \Delta \quad (\text{B200})$$

for some $L'' > 0$. Thus, in particular, the values of the model variables f_i^* at time Δ converge to the initial conditions of the ϵ' case, except that

$$\lim_{\Delta \rightarrow 0} (S_1^{V*}(0; \Delta)) > S_1^V(0; \epsilon'). \quad (\text{B201})$$

Moreover, note that for any $t \geq 0$,

$$W_i^*(\Delta + t, \Delta; \epsilon) - W_i^*(\Delta; \Delta; \epsilon) = W_i(\Delta + t; \epsilon') - W_1(\Delta; \epsilon') \quad (\text{B202})$$

and so, as U^* is bounded in $[\Delta, \infty)$

$$\left| \left(W_i^*(\Delta + t, \Delta; \epsilon) - W_1^*(\Delta; \Delta; \epsilon) \right) - W_i(t; \epsilon') \right| < L''' \Delta \quad \forall t > 0 \quad (\text{B203})$$

for some L''' . Thus, define variables with a hat to denote those from the disease trajectory with initial conditions given by

$$\hat{f}_i(0; \epsilon') := \lim_{\Delta \rightarrow 0} \left(f_i^*(\Delta; \Delta; \epsilon) \right) \quad (\text{B204})$$

and with vaccination policy given by $W_i(t; \epsilon')$. Then, by considering the starred variables to come from an epidemic started at time $t = \Delta$, Lemma 14 shows that

$$\lim_{\Delta \rightarrow 0} (f_i^*(t; \Delta; \epsilon)) = \hat{f}_i(t; \epsilon'). \quad (\text{B205})$$

Thus, one can take the $\Delta \rightarrow 0$ limit in (B188) and (B189) to show

$$\hat{I}_i(t; \epsilon') + \hat{R}_i(t; \epsilon') + \hat{I}_i^V(t; \epsilon') + \hat{R}_i^V(t; \epsilon') \leq I_i(t; \epsilon) + R_i(t; \epsilon) + I_i^V(t; \epsilon) + R_i^V(t; \epsilon) \quad (\text{B206})$$

and

$$R_i(t; \epsilon') \leq R_i(t; \epsilon). \quad (\text{B207})$$

Taking $t \rightarrow \infty$ in these inequalities shows that

$$\hat{R}_i(\infty; \epsilon') + \hat{R}_i(\infty; \epsilon') \leq R_i(\infty; \epsilon) + R_i^V(\infty; \epsilon) \quad \text{and} \quad \hat{R}_i(\infty; \epsilon') \leq R_i(\infty; \epsilon) \quad (\text{B208})$$

and hence, for any $\kappa_i \in [0, 1]$,

$$\hat{R}_i(\infty; \epsilon') + \kappa_i \hat{R}_i(\infty; \epsilon') = (1 - \kappa_i) \hat{R}_i(\infty; \epsilon') + \kappa_i (\hat{R}_i(\infty; \epsilon') + \hat{R}_i(\infty; \epsilon')) \quad (\text{B209})$$

$$\leq (1 - \kappa_i) R_i(\infty; \epsilon) + \kappa_i (R_i(\infty; \epsilon) + R_i^V(\infty; \epsilon)) \quad (\text{B210})$$

$$= R_i(\infty; \epsilon) + \kappa_i R_i^V(\infty; \epsilon). \quad (\text{B211})$$

Summing these inequalities over i gives

$$\hat{H}(\hat{W}(t; \epsilon'); \epsilon') \leq H(W(t; \epsilon); \epsilon). \quad (\text{B212})$$

Finally, note that by Lemma 16, as the only change between cases \hat{H} and H is an increase in one of the values of S^V ,

$$H(W(t; \epsilon'); \epsilon') \leq \hat{H}(W(t; \epsilon'); \epsilon') \quad (\text{B213})$$

which, combined with (B212), completes the proof of this proposition.

Theorem 2

This allows the overall proof of Theorem 2. The proof will rely on Theorem 1, which allows the creation of an $O(1)$ decrease in the objective function by reducing ϵ . By comparing a sequence of policies satisfying (A2) with a sequence that does not satisfy (A2) and using Proposition 6, one can then create a sequence of optimal policies such that the associated sequence of objective values decreases by at least a fixed quantity at each step (and thus will eventually become negative, giving a contradiction). Suppose (for a contradiction) that Theorem 2 does not hold for some fixed α , τ and w . Thus, for any $\eta > 0$, there is an $\epsilon \in (0, \eta)$ such that, for some U satisfying (A2),

$$H(U(t; \epsilon); \epsilon) \leq H(\bar{U}(t; \epsilon); \epsilon). \quad (\text{B214})$$

By optimality of $\bar{U}(t; \epsilon)$, (B214) must in fact be an equality, and so it can be assumed that $U(t; \epsilon) = \bar{U}(t; \epsilon)$, which will be done in the remainder of this proof (that is, if for some ϵ there is an optimal solution satisfying (A2), then it will be assumed that \bar{U} satisfies (A2)). Thus, there is some ϵ_0 such that

$$H(\bar{U}(t; \epsilon_0); \epsilon_0) \leq H(\tilde{\bar{U}}(t; \epsilon_0); \epsilon_0) \quad (\text{B215})$$

where $\tilde{\bar{U}}$ is defined by (A4). Now, for $\epsilon < \epsilon_0$, define $U^0(t; \epsilon)$ by

$$U_1^0(t; \epsilon) = \begin{cases} \bar{U}_1(t; \epsilon_0) & \text{if } W_1(t; \epsilon_0) < \epsilon \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \bar{U}_i(t; \epsilon) = U_i^0(t; \epsilon_0) \quad \forall i \neq 1 \quad (\text{B216})$$

and note that this means that

$$U^0(t; \epsilon_0) = \bar{U}(t; \epsilon_0) \quad \forall t \geq 0. \quad (\text{B217})$$

By (A163) in the proof of Theorem 1, there exists some $\delta_1 > 0$ such that, for all $\epsilon < \delta_1$,

$$H(U^0(t; \epsilon); \epsilon) > H(\tilde{U}^0(t; \epsilon); \epsilon) + \frac{1}{2} \int_{z_0}^{\tau} (1 - \alpha) P(t) \left[(1 - \mathcal{Y}^*) \kappa_1 (X(t) - Y(t)) + (1 - \kappa_1) X(t) \right] dt. \quad (\text{B218})$$

where

$$X(t) = \sum_{j=1}^n \beta_{1j}^1 I_j(t; \epsilon) + \beta_{1j}^3 I_j^V(t; \epsilon), \quad (\text{B219})$$

$$Y(t) = \sum_{j=1}^n \beta_{1j}^2 I_j(t; \epsilon) + \beta_{1j}^4 I_j^V(t; \epsilon) \quad (\text{B220})$$

and

$$P(t) = \exp \left[- \sum_{j=1}^n \left(\frac{\beta_{1j}^1 R_j(t; 0)}{\mu_j^1} + \frac{\beta_{1j}^2 R_j^V(t; 0)}{\mu_j^2} \right) \right]. \quad (\text{B221})$$

Note that ρ_0 , τ and \mathcal{Y}^* are independent of U^0 , but $X(t)$, $Y(t)$ and $P(t)$ are not. However, note that

$$\frac{dI_i(t; \epsilon)}{dt} \geq -\mu_i^1 I_i(t; \epsilon) \quad (\text{B222})$$

and so

$$X(t) - Y(t) \geq \sum_{j=2}^n (\beta_{1j}^1 - \beta_{1j}^3) e^{-\mu_j^1 t} I_j(0; \epsilon) > 0, \quad (\text{B223})$$

by assumption (B170), giving a bound that is independent of U^0 . Moreover,

$$X(t) \geq X(t) - Y(t) > 0. \quad (\text{B224})$$

Finally, for $\epsilon \leq 1$,

$$\begin{aligned} P(t) &\geq \exp \left[- \sum_{j=1}^n \left(\frac{\beta_{1j}^1 N_j(\epsilon)}{\mu_j^1} + \frac{\beta_{1j}^2 N_j(\epsilon)}{\mu_j^2} \right) \right] \\ &\geq \exp \left[- \sum_{j=1}^n \left(\frac{\beta_{1j}^1 N_j(1)}{\mu_j^1} + \frac{\beta_{1j}^2 N_j(1)}{\mu_j^2} \right) \right] > 0 \end{aligned} \quad (\text{B225})$$

and this bound is again independent of U^0 . Thus,

$$H(U^0(t; \epsilon); \epsilon) > H(\tilde{U}^0(t; \epsilon); \epsilon) + K \quad \forall \epsilon < \delta_1 \quad (\text{B226})$$

for some constant $K > 0$ where this is now independent of U^0 . Now, by assumption, there must exist some $\epsilon_1 \in (0, \delta_1)$ such that $\bar{U}(t; \epsilon_1)$ meets the conditions (A2) so

$$H(\bar{U}(t; \epsilon_1); \epsilon_1) \leq H(\tilde{\bar{U}}(t; \epsilon_1); \epsilon_1) \quad (\text{B227})$$

while by optimality

$$H(\bar{U}(t; \epsilon_1); \epsilon_1) \leq H(\tilde{U}^0(t; \epsilon_1); \epsilon_1) < H(U^0(t; \epsilon_1); \epsilon_1) - K. \quad (\text{B228})$$

Now, moreover, note that by Proposition 6,

$$H(U^0(t; \epsilon_1); \epsilon_1) \leq H(U^0(t; \epsilon_0); \epsilon_0) = H(\bar{U}(t; \epsilon^0); \epsilon^0) \quad (\text{B229})$$

and so

$$H(\overline{U}(t; \epsilon_1); \epsilon_1) \leq H(\overline{U}(t; \epsilon_0); \epsilon_0) - K. \quad (\text{B230})$$

Now, this can be continued iteratively so that, for any $n \geq 0$,

$$H(\overline{U}(t; \epsilon_n); \epsilon_n) \leq H(\overline{U}(t; \epsilon_0); \epsilon_0) - Kn \quad (\text{B231})$$

However, this means that eventually, one finds

$$H(\overline{U}(t; \epsilon_n); \epsilon_n) < 0 \quad (\text{B232})$$

which is a contradiction. Thus, for each fixed α , w and τ , there must exist some η such that for any $\epsilon \in (0, \eta)$, the optimal solution does not satisfy (A2).

Now, suppose that $\int_0^t \chi(s)ds > 0$ and suppose $\overline{U}(t; \epsilon)$ is an optimal solution for each value of ϵ such that, for each t

$$\sum_{i=1}^n \overline{W}_i(t; \epsilon) = \min \left(\int_0^t \chi(s)ds, 1 \right) \quad (\text{B233})$$

(note that this can be assumed by Theorem 2 in Penn and Donnelly (2022)). Now, suppose that, for some t

$$\lim_{\epsilon \rightarrow 0} \left(\frac{\overline{W}_1(t; \epsilon)}{\epsilon} \right) \neq 1 \quad \text{and} \quad \min \left(\int_0^t \chi(s)ds, 1 \right) > 0. \quad (\text{B234})$$

This means that there exists some $\delta > 0$ such that there is a subsequence ϵ_m satisfying

$$\frac{\overline{W}_1(t; \epsilon_m)}{\epsilon_m} < 1 - \delta < 1 \quad \text{and} \quad \lim_{m \rightarrow \infty} (\epsilon_m) = 0 \quad (\text{B235})$$

noting that

$$\frac{\overline{W}_1(t; \epsilon_m)}{\epsilon_m} \leq 1 \quad \forall \epsilon_m > 0. \quad (\text{B236})$$

However, this means that for each m , $\overline{U}(t; \epsilon_m)$ satisfies the condition (A2) with $\tau = t$, $\alpha = 1 - \delta$ and $w = \min \left(\int_0^t \chi(s)ds, 1 \right)$. This is a contradiction to the previous part of the proof (as $\lim_{m \rightarrow \infty} (\epsilon_m) = 0$) and hence

$$\lim_{\epsilon \rightarrow 0} \left(\frac{W_1^*(t; \epsilon)}{\epsilon} \right) = 1 \quad \forall t \quad \text{s.t.} \quad \min \left(\int_0^t \chi(s)ds, 1 \right) > 0, \quad (\text{B237})$$

as required.

Appendix C Proof of Theorem 3

Recall the definitions from the main text.

$$\beta'_{ij} = \begin{cases} \beta^1_{ij} & \text{if } i, j \leq n \\ \beta^2_{i(n-j)} & \text{if } i \leq n < j \leq 2n \\ \beta^3_{(n-i)j} & \text{if } j \leq n < i \leq 2n \\ \beta^4_{(n-i)(n-j)} & \text{if } n < i, j \leq 2n \end{cases}, \quad (\text{C239})$$

$$\mu'_i = \begin{cases} \mu^1_i & \text{if } i \leq n \\ \mu^2_{(i-n)} & \text{if } n < i \leq 2n \end{cases}, \quad (\text{C240})$$

$$p'_i = \begin{cases} p_i & \text{if } i \leq n \\ \kappa_{(i-n)} p_{(i-n)} & \text{if } n < i \leq 2n \end{cases}, \quad (\text{C241})$$

$$Q_{ij} = \frac{1}{1 - e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)}} \left[\delta_{ij} + \frac{S_i(0; 0) e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)}}{\mu'_j} \beta'_{ij} \right], \quad (\text{C242})$$

and

$$x = Q^{-T} p' \quad \text{and} \quad y_i = \frac{S_i(0; 0)}{N_i} (x_{i+n} - x_i) \quad \forall i \in \{1, \dots, n\}. \quad (\text{C243})$$

Theorem 3 Suppose that, for all $\epsilon > 0$

$$B(t; \epsilon) = \epsilon \quad \forall t \geq 0 \quad (\text{C244})$$

and that all other parameter values and initial conditions are independent of ϵ . Suppose that $A(t)$ is a continuous function with

$$A(0) > 0 \quad (\text{C245})$$

and that the matrix M is invertible. Assuming that ϵ is sufficiently small so that it exists, define

$$\tau(\epsilon) := \inf \left\{ t : \int_0^t A(s) ds = \epsilon \right\}. \quad (\text{C246})$$

Suppose that U satisfies the condition

$$\sum_{i=1}^n U_i(s) = \min \left(\int_0^t \chi(s) ds, 1 \right) \quad (\text{C247})$$

where χ is defined in (B169). Then, for sufficiently small ϵ , the objective function is given by

$$H(U(t; \epsilon)) = H(\mathbf{0}) + \mathbf{y}^T \mathbf{W}(\tau(\epsilon); \epsilon) + o(\epsilon). \quad (\text{C248})$$

Moreover, if there is a unique element of \mathbf{y} equal to the minimum of \mathbf{y} then the optimal vaccination policy (to leading order in ϵ) is uniquely given by

$$U_i(t; \epsilon) = \begin{cases} A(t) & \text{if } i = \min\{y_i : i \in \{1, \dots, n\}\} \text{ and } \int_0^t A(s)ds < \epsilon \\ 0 & \text{otherwise} \end{cases}. \quad (\text{C249})$$

Proposition 7

Note that the n -group model can be considered as a $2n$ -group model once vaccination has finished—an idea that is formalised in the below proposition. This moderately extends the previous work by incorporating the initial vaccination policy into the final size equation, but is not a major advancement on well-known results found in books such as Anderson and May (1992).

Proposition 7 Define for $i \in \{1, \dots, n\}$,

$$(S_{n+i}, I_{n+i}, R_{n+i}) := (S_i^V, I_i^V, R_i^V). \quad (\text{C250})$$

Define further

$$\sigma_i(\epsilon) = \begin{cases} -\frac{S_i(0;0)W_i(\tau(\epsilon))}{N_i} & \text{if } i \leq n \\ \frac{S_{i-n}(0;0)W_{i-n}(\tau(\epsilon))}{N_{i-n}} & \text{if } n < i \leq 2n \end{cases} \quad (\text{C251})$$

and

$$\rho_i(\epsilon) := R_i(\infty; \epsilon) - R_i(\infty; 0) \quad \forall i \in \{1, \dots, 2n\}. \quad (\text{C252})$$

Then, $\rho_i(\epsilon)$ is $o(1)$ as $\epsilon \rightarrow 0$ and

$$\sigma_i = \frac{\rho_i + S_i(0;0)e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty;0)} \left(\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \rho_j \right) + o(\sigma_i) + \sum_{j=1}^{2n} o(\rho_j) + O(\epsilon^2)}{1 - e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty;0)}}. \quad (\text{C253})$$

Proof As A is continuous, there is some region $(0, \delta)$ such that

$$\frac{A(0)}{2} < A(t) < 2A(0) \quad (\text{C254})$$

and hence

$$\int_0^\delta A(t)dt > \frac{\delta A(0)}{2}. \quad (\text{C255})$$

This lower bound is independent of ϵ , and hence, for sufficiently small ϵ ,

$$\int_0^\delta A(t)dt > \epsilon. \quad (\text{C256})$$

Now, by assumption,

$$\sum_{i=1}^n U_i(t; \epsilon) = \begin{cases} A(t) & \text{if } \int_0^t A(s)ds < \epsilon \\ 0 & \text{otherwise} \end{cases}. \quad (\text{C257})$$

By continuity and the definition of $\tau(\epsilon)$,

$$\int_0^{\tau(\epsilon)} A(t)dt = \epsilon \quad (\text{C258})$$

and note that it is necessary that $\tau(\epsilon) = O(\epsilon)$ as

$$\tau(\epsilon) \leq \frac{2\epsilon}{A(0)} \quad (\text{C259})$$

for sufficiently small ϵ .

Now, all of the variables are bounded independently of ϵ in the interval $[0, \tau(\epsilon)]$ (including U , which is bounded by $2A(0)$). Moreover, assuming $N_i > 0$ for each $i \in \{1, \dots, n\}$,

$$N_i - W_i > N_i - \epsilon > \frac{\min_i(N_i)}{2} \quad (\text{C260})$$

for sufficiently small ϵ . Thus, in particular, all of the derivatives of the model variables are bounded and so

$$S_i(\tau(\epsilon); \epsilon) = S_i(0; 0) + O(\epsilon) \quad (\text{C261})$$

with analogous results for the other model variables, noting that the initial conditions are identical in each case. Thus, in particular,

$$\frac{dS_i}{dt}(t; \epsilon) = \frac{dS_i}{dt}(0; \epsilon) - \frac{S_i(0; 0)(U_i(t; \epsilon) - U_i(0; \epsilon))}{N_i - W_i(0; \epsilon)} + O(\epsilon) \quad \forall t \in (0, \epsilon), \quad (\text{C262})$$

noting that the $U_i(t; \epsilon)$ are the only quantities that can change by an $O(1)$ amount in $O(\epsilon)$ time. Now, one can set $U_i(0; \epsilon) = 0$ to reduce notation (noting that the model depends only on the integral of U_i). Moreover, as $W_i(0; \epsilon) = 0$, the initial conditions are independent of ϵ and $\tau(\epsilon) = O(\epsilon)$, integrating gives

$$S_i(\tau(\epsilon); \epsilon) = S_i(0; 0) + \tau(\epsilon) \frac{dS_i}{dt}(0; 0) - \frac{S_i(0; \epsilon)W_i(\tau(\epsilon); \epsilon)}{N_i} + O(\epsilon^2). \quad (\text{C263})$$

Similarly,

$$S_i^V(\tau(\epsilon); \epsilon) = S_i^V(0; 0) + \tau(\epsilon) \frac{dS_i^V}{dt}(0; 0) + \frac{S_i(0; 0)W_i(\tau(\epsilon); \epsilon)}{N_i} + O(\epsilon^2) \quad (\text{C264})$$

while for the other model variables, f_i , there is no $O(1)$ change to the derivative so

$$f_i(\tau(\epsilon); \epsilon) = f(0; 0) + \tau(\epsilon) \frac{df_i}{dt}(0; 0) + O(\epsilon^2). \quad (\text{C265})$$

Now, for times $t \geq \tau(\epsilon)$, one has $U_i(t; \epsilon) = 0$ and so a standard multi-group SIR model (with initial conditions given by the model variables evaluated at time $\tau(\epsilon)$) is recovered. Thus in particular, the final number infected can be formulated in terms of a final size equation, following the work of Anderson and May (1992) among others. Define, for $i \in \{1, \dots, n\}$,

$$(S_{n+i}, I_{n+i}, R_{n+i}) = (S_i^V, I_i^V, R_i^V). \quad (\text{C266})$$

This new $2n$ group model has the same behaviour as the original model if the parameters are

$$\beta'_{ij} = \begin{cases} \beta_{ij}^1 & \text{if } i, j \leq n \\ \beta_{i(n-j)}^2 & \text{if } i \leq n < j \\ \beta_{(n-i)j}^3 & \text{if } j \leq n < i \\ \beta_{(n-i)(n-j)}^4 & \text{if } n < i, j \end{cases}, \quad \mu'_i = \begin{cases} \mu_i^1 & \text{if } i \leq n \\ \mu_{(i-n)}^2 & \text{if } i > n \end{cases} \quad (\text{C267})$$

and

$$p'_i = \begin{cases} p_i & \text{if } i \leq n \\ \kappa_{(i-n)} p_{(i-n)} & \text{if } i > n \end{cases}. \quad (\text{C268})$$

Thus, integrating the S_i equation between $\tau(\epsilon)$ and $t + \tau(\epsilon)$ gives:

$$\frac{d}{dt} (\log(S_i)) = - \sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \frac{dR_j}{dt} \quad (\text{C269})$$

$$\Rightarrow \ln(S_i(t + \tau(\epsilon); \epsilon)) = \ln(S_i(\tau(\epsilon); \epsilon)) - \sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \left[R_j(t + \tau(\epsilon); \epsilon) - R_j(\tau(\epsilon); \epsilon) \right] \quad (\text{C270})$$

$$\Rightarrow S_i(t + \tau(\epsilon); \epsilon) = S_i(\tau(\epsilon); \epsilon) e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \left[R_j(t + \tau(\epsilon); \epsilon) - \tau(\epsilon) \frac{dR_j}{dt}(0; 0) \right]} + O(\epsilon^2) \quad (\text{C271})$$

as $R_j(0; 0) = 0$ for each j . Now, note that for any $t \geq 0$,

$$\begin{aligned} S_i(\tau(\epsilon); \epsilon) + I_i(\tau(\epsilon); \epsilon) + R_i(\tau(\epsilon); \epsilon) \\ = S_i(t + \tau(\epsilon); \epsilon) + I_i(t + \tau(\epsilon); \epsilon) + R_i(t + \tau(\epsilon); \epsilon) \end{aligned} \quad (\text{C272})$$

and hence, taking $t \rightarrow \infty$ and using Lemma 11 shows that

$$S_i(\tau(\epsilon); \epsilon) + I_i(\tau(\epsilon); \epsilon) + R_i(\tau(\epsilon); \epsilon) = S_i(\infty; \epsilon) + R_i(\infty; \epsilon). \quad (\text{C273})$$

Hence, by (C265),

$$S_i(\tau(\epsilon); \epsilon) + I_i(0; 0) + \tau(\epsilon) \left[\frac{dI_i}{dt}(0; 0) + \frac{dR_i}{dt}(0; 0) \right] = S_i(\infty; \epsilon) + R_i(\infty; \epsilon) + O(\epsilon^2). \quad (\text{C274})$$

Now, substituting this into the limit of (C269) as $t \rightarrow \infty$ shows that

$$\begin{aligned} R_i(\infty; \epsilon) = S_i(\tau(\epsilon); \epsilon) + I_i(0; 0) + \tau(\epsilon) \left[\frac{dI_i}{dt}(0; 0) + \frac{dR_i}{dt}(0; \epsilon) \right] \\ - S_i(\tau(\epsilon); \epsilon) e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \left[R_j(\infty; \epsilon) - \tau(\epsilon) \frac{dR_j}{dt}(0; 0) \right]} + O(\epsilon^2). \end{aligned} \quad (\text{C275})$$

By treating this model as a model that has initial conditions given by the variable values at time $\tau(\epsilon)$, one sees that these initial conditions differ from the initial conditions of the $\epsilon = 0$ model by $O(\epsilon)$ (where no vaccination occurs in either case). This means that Proposition 5 can be used (as the vaccination policies U must have uniformly bounded finite support for sufficiently small ϵ) and so there exists some function $\delta(\epsilon)$ such that, for all sufficiently small ϵ ,

$$|R_j(\infty; \epsilon) - R_j(\infty; 0)| < \delta(\epsilon) \quad \forall j \quad \text{and} \quad \delta(\epsilon) = o(1). \quad (\text{C276})$$

Thus, in particular, one can define functions $\rho_j(\epsilon)$ such that

$$R_j(\infty; \epsilon) = R_j(\infty; 0) + \rho_j(\epsilon) \quad \forall j \in \{1, \dots, 2n\} \quad (\text{C277})$$

and

$$\rho_j(\epsilon) = o(1) \quad \text{as } \epsilon \rightarrow 0. \quad (\text{C278})$$

Furthermore, defining σ_i such that

$$\sigma_i(\epsilon) = \begin{cases} -\frac{S_i(0;0)W_i(\tau(\epsilon))}{N_i} & \text{if } i \leq n \\ \frac{S_{i-n}(0;0)W_{i-n}(\tau(\epsilon))}{N_{i-n}} & \text{if } n < i \leq 2n \end{cases} \quad (\text{C279})$$

gives

$$S_i(\tau(\epsilon); \epsilon) = S_i(0; 0) + \tau(\epsilon) \frac{dS_i}{dt}(0; 0) + \sigma_i(\epsilon) + O(\epsilon^2) \quad \forall i \in \{1, \dots, 2n\}. \quad (\text{C280})$$

Now, when $\sigma_i(\epsilon) = 0$ for all i , it must be the case that $\rho_i(\epsilon) = 0$ for all i as the final size is unchanged (as no vaccination has taken place). Thus, in this case, (C275) can be linearised to give

$$O(\epsilon^2) = \tau(\epsilon) \left[\frac{dS_i}{dt}(0; 0) + \frac{dI_i}{dt}(0; 0) + \frac{dR_i}{dt}(0; 0) e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)} \left(-\frac{dS_i}{dt}(0; 0) + S_i(0; 0) \sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \frac{dR_j}{dt}(0; 0) \right) \right]. \quad (\text{C281})$$

Note that this equality does indeed hold, as in the no vaccination case

$$\frac{dS_i}{dt}(0; 0) + \frac{dI_i}{dt}(0; 0) + \frac{dR_i}{dt}(0; 0) = 0 \quad (\text{C282})$$

is the conservation of population law, while

$$-\frac{dS_i}{dt}(0; 0) + S_i(0; 0) \sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \frac{dR_j}{dt}(0; 0) = -\frac{dS_i}{dt}(0; 0) + S_i(0; 0) \sum_{j=1}^{2n} \beta'_{ij} I_j(0; 0) = 0. \quad (\text{C283})$$

This means that, for nonzero σ_i , all terms not dependent on σ_i or ρ_i cancel and so the linearisation becomes

$$\begin{aligned} \rho_i = \sigma_i - \sigma_i e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)} - S_i(0; 0) e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)} \sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \rho_j \\ + o(\sigma_i) + \sum_{j=1}^{2n} o(\rho_j) + O(\epsilon^2) \end{aligned} \quad (\text{C284})$$

and so

$$\sigma_i = \frac{\rho_i + S_i(0; 0)e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)} \left(\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} \rho_j \right) + o(\sigma_i) + \sum_{j=1}^{2n} o(\rho_j) + O(\epsilon^2)}{1 - e^{-\sum_{j=1}^{2n} \frac{\beta'_{ij}}{\mu'_j} R_j(\infty; 0)}}. \quad (\text{C285})$$

as required.

Proposition 8

The result of Proposition 7 can be written as a system of equations for vectors σ and ρ

$$\sigma = Q\rho + o(\sigma) + \sum_{j=1}^{2n} o(\rho_j) + O(\epsilon^2) \quad (\text{C286})$$

for some matrix Q with nonzero determinant by assumption. However, it is important to establish the dominant balance in these equations, which is done through the following proposition, another result that the authors believe is novel to the literature.

Proposition 8

$$\rho_i(\epsilon) = O(\epsilon) \quad \forall i \in \{1, \dots, 2n\}. \quad (\text{C287})$$

Proof Suppose that this does not hold. Thus, there must be some sequence ϵ_m such that, for some i

$$\lim_{m \rightarrow \infty} \left(\frac{\rho_i(\epsilon_m)}{\epsilon_m} \right) = \infty \quad \text{and} \quad \lim_{m \rightarrow \infty} (\epsilon_m) = 0. \quad (\text{C288})$$

Define $J^*(\epsilon)$ such that

$$J^*(\epsilon) = \operatorname{argmax} \{ |\rho_j(\epsilon)| : j \in \{1, \dots, 2n\} \}. \quad (\text{C289})$$

Now, by the finiteness of $\{1, \dots, 2n\}$, there exists some subsequence ϵ_{m_k} and some fixed $J \in \{1, \dots, 2n\}$ such that

$$J^*(\epsilon_{m_k}) = J \quad \forall k. \quad (\text{C290})$$

For notational convenience, assume that the original sequence ϵ_m has this property. Note that

$$\lim_{m \rightarrow \infty} \left(\frac{\sigma_j(\epsilon_m)}{\rho_J(\epsilon_m)} \right) = \lim_{m \rightarrow \infty} \left(\frac{\sigma_j(\epsilon_m)}{\epsilon} \times \frac{\epsilon}{\rho_J(\epsilon_m)} \right) = 0. \quad (\text{C291})$$

as $\sigma_j(\epsilon) = O(\epsilon)$ and $\epsilon = o(\rho_i(\epsilon)) \leq o(\rho_J(\epsilon))$. Moreover,

$$\lim_{m \rightarrow \infty} \left(\frac{O(\epsilon_m^2)}{\rho_J(\epsilon_m)} \right) = \lim_{m \rightarrow \infty} \left(\epsilon_m \times \frac{O(\epsilon_m)}{\rho_J(\epsilon_m)} \right) = 0, \quad (\text{C292})$$

$$\lim_{m \rightarrow \infty} \left(\frac{o(\sigma_j(\epsilon_m))}{\rho_J(\epsilon_m)} \right) = \lim_{m \rightarrow \infty} \left(o(1) \times \frac{\sigma_j(\epsilon_m)}{\rho_J(\epsilon_m)} \right) = 0 \quad (\text{C293})$$

and

$$\left| \lim_{m \rightarrow \infty} \left(\frac{o(\rho_j(\epsilon_m))}{\rho_J(\epsilon_m)} \right) \right| \leq \lim_{m \rightarrow \infty} \left(\left| \frac{o(\rho_j(\epsilon_m))}{\rho_j(\epsilon_m)} \right| \right) = 0. \quad (\text{C294})$$

Note that there is some abuse of notation in these calculations, but, for example, an $O(\epsilon^2)$ term in the limit represents any function which is $O(\epsilon^2)$. Thus, dividing (C286) by $\rho_J(\epsilon_m)$ and taking m to ∞ shows that

$$\lim_{m \rightarrow \infty} \left(\frac{\mathbf{Q}\rho}{\rho_J(\epsilon_m)} \right) = \mathbf{0}. \quad (\text{C295})$$

Define

$$\hat{\rho}(\epsilon) := \frac{\rho(\epsilon)}{\sum_{j=1}^{2n} |\rho_j(\epsilon)|} \quad (\text{C296})$$

and note that

$$\left| \left(\frac{\sum_{j=1}^{2n} |\rho_j(\epsilon_m)|}{\rho_J(\epsilon_m)} \right) \right| \in [1, 2n] \quad (\text{C297})$$

and thus remains finite and nonzero. Thus,

$$\mathbf{0} = \lim_{m \rightarrow \infty} \left(\frac{\mathbf{Q}\rho}{\rho_J(\epsilon_m)} \right) \quad (\text{C298})$$

$$= \lim_{m \rightarrow \infty} \left(\frac{\sum_{j=1}^{2n} |\rho_j(\epsilon_m)|}{\rho_J(\epsilon_m)} \times \mathbf{Q}\hat{\rho}(\epsilon_m) \right), \quad (\text{C299})$$

which means

$$\mathbf{0} = \lim_{m \rightarrow \infty} \left(\mathbf{Q}\hat{\rho}(\epsilon_m) \right). \quad (\text{C300})$$

However, note that

$$\sum_{j=1}^{2n} |\hat{\rho}_i(\epsilon)| = 1 \quad (\text{C301})$$

and hence the sequence $\hat{\rho}$ is bounded. Thus, by the Bolzano–Weierstrass theorem, there must be some subsequence m_k such that $\lim_{k \rightarrow \infty} (\hat{\rho}(\epsilon_{m_k}))$ exists and is equal to some ρ^* where

$$\sum_{j=1}^{2n} |\rho_j^*| = 1. \quad (\text{C302})$$

However, then, by continuity and the fact that Q is invertible,

$$Q\rho^* = \mathbf{0} \Rightarrow \rho^* = \mathbf{0} \quad (\text{C303})$$

which is a contradiction to (C302) as required. Thus, it must be the case that $\rho(\epsilon) = O(\epsilon)$ \square

Theorem 3

Combining Proposition 8 with the fact that $\sigma_i = O(\epsilon)$ means that (C286) can be written as

$$\sigma = Q\rho + o(\epsilon). \quad (\text{C304})$$

Thus, one can multiply the equation by M^{-1} to get

$$\rho = Q^{-1}\sigma + o(\epsilon). \quad (\text{C305})$$

Hence, given vectors p and q where

$$p_i := p_i \quad \text{and} \quad q_i = p_i \kappa_i \quad \forall i \in \{1, \dots, n\}, \quad (\text{C306})$$

the change to the objective function is given by:

$$(p, q)^T \rho = (p, q)^T \left[Q^{-1}\sigma + o(\epsilon) \right] \quad (\text{C307})$$

$$:= x^T \sigma + o(\epsilon). \quad (\text{C308})$$

Now, note that, for $i \in \{1, \dots, n\}$,

$$\sigma_i = -\frac{S_i(0; 0)W_i(\tau(\epsilon); \epsilon)}{N_i} \quad (\text{C309})$$

while, for $i \in \{n+1, \dots, 2n\}$

$$\sigma_i = -\sigma_{i-n}. \quad (\text{C310})$$

Hence, one can write (C308) as

$$(\mathbf{p}^T, \mathbf{q}^T) \boldsymbol{\rho} = \mathbf{y}^T \mathbf{W}(\tau(\epsilon); \epsilon) + o(\epsilon), \quad (\text{C311})$$

where

$$\mathbf{y} = \frac{S_i(0; 0)}{N_i} \left[-(x_1, \dots, x_n)^T + (x_{n+1}, \dots, x_{2n})^T \right], \quad (\text{C312})$$

as required by Theorem 3. The only restriction is that all the W_i are non-negative and that

$$\sum_{i=1}^n W_i(\tau(\epsilon); \epsilon) = \epsilon \quad (\text{C313})$$

and so the optimisation problem becomes

$$\min\{\mathbf{y}^T \mathbf{w} : \mathbf{w} \geq \mathbf{0} \text{ and } \sum_{i=1}^n w_i = \epsilon\}. \quad (\text{C314})$$

Now, by Theorem 17, proved in Penn and Donnelly (2022) and stated in the appendices, it must be the case that the objective function is non-increasing in \mathbf{w} . Thus, in particular, one must have

$$\mathbf{y} \leq \mathbf{0} \quad (\text{C315})$$

as otherwise, if $y_i > 0$ then setting $\mathbf{w} = \epsilon \mathbf{e}_i$ (where \mathbf{e}_i is the i th canonical basis vector) means that

$$H(\mathbf{U}(t; \epsilon)) = H(\mathbf{U}(t; 0)) + y_i \epsilon + o(\epsilon) \quad (\text{C316})$$

and so, for sufficiently small ϵ ,

$$H(\mathbf{U}(t; \epsilon)) > H(\mathbf{U}(t; 0)) \quad (\text{C317})$$

which is a contradiction. Hence, $\mathbf{y} \leq \mathbf{0}$ which means that the optimisation problem is an example of a continuous knapsack problem and one can readily see that a solution is given by

$$w_i^* = \begin{cases} \epsilon & \text{if } i = \min\{y_i\} \\ 0 & \text{otherwise} \end{cases}. \quad (\text{C318})$$

As this minimum is unique by assumption, this is the unique leading-order optimal solution to the optimisation problem.

A technical note is that this only proves the form of the optimal solution to leading order. Indeed, if

$$w_i = w_i^* + o(\epsilon), \quad (\text{C319})$$

then the optimal objective value is unchanged to leading order. Hence, this restriction is given in the statement of the theorem (although in practice is unimportant).

Appendix D Supplementary Lemmas

This section contains the supplementary lemmas that have been used in the proofs of Theorems 1–3. All but two of these lemmas were proved in Penn and Donnelly (2022) and so their proofs will not be reproduced here, but they have been included for completeness and for ease of access. The exceptions are Lemma 15 and 16.

Lemma 9

Lemma 9 *Consider a continuous, time-dependent, matrix $A(t)$ which satisfies*

$$A(t)_{ij} \geq 0 \quad \forall t \geq 0 \quad \text{and} \quad \forall i \neq j \quad (\text{D320})$$

and a constant matrix B that satisfies

$$B_{ij} \geq 0 \quad \forall t \geq 0 \quad \text{and} \quad \forall i \neq j. \quad (\text{D321})$$

Then, suppose that each element of $A(t)$ is non-increasing with t and that

$$A(t)_{ij} \geq B_{ij} \quad \forall t \geq 0 \quad \text{and} \quad \forall i \neq j. \quad (\text{D322})$$

Moreover, define an initial condition \mathbf{v} and suppose that \mathbf{y} and \mathbf{z} solve the systems

$$\frac{d\mathbf{y}}{dt} = A(t)\mathbf{y} \quad \text{and} \quad \frac{d\mathbf{z}}{dt} = B\mathbf{z} \quad (\text{D323})$$

with

$$\mathbf{y}(0) = \mathbf{z}(0) = \mathbf{v} \geq \mathbf{0}. \quad (\text{D324})$$

Then,

$$\mathbf{y}(t) \geq \mathbf{z}(t) \geq \mathbf{0} \quad \forall t \geq 0. \quad (\text{D325})$$

Proof This was proved as Lemma B.2 in Penn and Donnelly (2022)

Lemma 10

Lemma 10 *Define the set of functions*

$$\mathcal{F}_i(t) := \left\{ S_i(t), I_i(t), R_i(t), S_i^V(t), I_i^V(t), R_i^V(t) \right\}. \quad (\text{D326})$$

Then, for all $t \geq 0$ and $i \in \{1, \dots, n\}$,

$$0 \leq f \leq N_i \quad \forall f \in \mathcal{F}_i(t). \quad (\text{D327})$$

Proof This was proved as Lemma B.3 in Penn and Donnelly (2022).

Lemma 11

Lemma 11 *For each i ,*

$$\lim_{t \rightarrow \infty} (I_i(t)) = \lim_{t \rightarrow \infty} (I_i^V(t)) = 0. \quad (\text{D328})$$

Proof This was proved as Lemma B.4 in Penn and Donnelly (2022). \square

Lemma 12

Lemma 12 *Suppose that $I_i(t) > 0$ for some $t \geq 0$ and some $i \in \{1, \dots, n\}$. Then*

$$I_i(s) > 0 \quad \forall s > t. \quad (\text{D329})$$

An analogous result holds for $I_i^V(t)$.

Proof This was proved as Lemma B.5 in Penn and Donnelly (2022).

Lemma 13

Lemma 13 *Define*

$$\Pi := \left\{ i : \exists t \geq 0 \text{ s.t. } I_i(t) > 0 \text{ or } I_i^V(t) > 0 \right\}. \quad (\text{D330})$$

Moreover, define

$$\Pi^0 := \{ i : I_i(0) > 0 \} \quad (\text{D331})$$

and the n by n matrix M by

$$M_{ij} = S_i(0)\beta_{ij}^1. \quad (\text{D332})$$

Then, define the connected component C of Π^0 in M as follows. The index $i \in \{1, \dots, n\}$ belongs to C if and only if there is some sequence a_1, \dots, a_k such that

$$a_j \in \{1, \dots, n\} \quad \forall j \in \{1, \dots, k\}, \quad (\text{D333})$$

$$M_{a_1, a_2} M_{a_2, a_3} \dots, M_{a_{k-1}, a_k} > 0 \quad (\text{D334})$$

and

$$a_1 = i \quad \text{and} \quad a_k \in \Pi^0. \quad (\text{D335})$$

Then,

$$(a) \quad i \in C \Rightarrow I_i(t) > 0 \quad \forall t > 0.$$

$$(b) \quad \Pi = C \cup \Pi^0.$$

Thus, in particular,

$$i \in C \cup \Pi^0 = \Pi \Leftrightarrow I(t) > 0 \quad \forall t > 0. \quad (\text{D336})$$

Proof This was proved as Lemma B.6 in Penn and Donnelly (2022).

Lemma 14

Lemma 14 Define the set of functions

$$\mathcal{F} := \left\{ S_i(t; \epsilon), I_i(t; \epsilon), R_i(t; \epsilon), S_i^V(t; \epsilon), I_i^V(t; \epsilon), R_i^V(t; \epsilon) : i \in \{1, \dots, n\}, \quad \epsilon, t \geq 0 \right\}, \quad (\text{D337})$$

where for each fixed ϵ , these functions solve the model equations with parameters

$$\mathcal{P} = \left\{ \beta_{ij}^\alpha(\epsilon), \mu_i^\gamma(\epsilon) : i, j \in \{1, \dots, n\}, \quad \alpha \in \{1, 2, 3, 4\}, \quad \gamma \in \{1, 2\} \quad \text{and} \quad \epsilon \geq 0 \right\}, \quad (\text{D338})$$

initial conditions

$$\mathcal{I} = \left\{ f(0; \epsilon) : i \in \{1, \dots, n\}, \quad f \in \mathcal{F} \quad \text{and} \quad \epsilon \geq 0 \right\} \quad (\text{D339})$$

and vaccination policy $U(t; \epsilon)$. Suppose further that the population sizes are independent of ϵ , except in group 1 where $N_1(\epsilon)$ satisfies

$$|N_1(\epsilon) - N_1(0)| \leq \epsilon \quad \text{and} \quad \frac{S_1(0; \epsilon)}{N_1} = \sigma \quad (\text{D340})$$

for some constant σ .

Suppose that

$$|p(\epsilon) - p(0)| \leq \epsilon \quad \forall p \in \mathcal{P}, \quad (\text{D341})$$

$$|f_i(0; \epsilon) - f_i(0; 0)| \leq \epsilon \quad \forall f \in \mathcal{F} \quad (\text{D342})$$

and that

$$|W_i(t, \epsilon) - W_i(t, 0)| < \epsilon \quad \forall t \geq 0. \quad (\text{D343})$$

Moreover, suppose that for each $i \in \{1, \dots, n\}$ and $\epsilon \geq 0$,

$$U_i(s; \epsilon) \geq 0 \quad \text{and} \quad \int_0^t U_i(s; \epsilon) ds \leq N_i \quad \forall t \geq 0. \quad (\text{D344})$$

Then, for each $\delta > 0$ and each $T > 0$ there exists some $\eta > 0$ (that may depend on T and δ) such that

$$\epsilon \in (0, \eta) \Rightarrow |f(t; \epsilon) - f(t; 0)| < \delta \quad \forall f \in \mathcal{F} \quad \text{and} \quad \forall t \in [0, T]. \quad (\text{D345})$$

Proof An almost identical result is proved in Lemma B.8 from Penn and Donnelly (2022), with the only exception being that N_i can vary in this example. However, note that by replacing $\frac{S_i(0; \epsilon)}{N_i(\epsilon)}$ with σ , this lemma can be proved identically.

Lemma 15

The following lemma is a new result, proved using similar techniques to results in Penn and Donnelly (2022) such as Lemma 13.

Lemma 15 Suppose that $i \in \Pi$, with Π defined as in Lemma 13. Then, for $t > 0$,

$$I_i^V(t) = 0 \Rightarrow S_i^V(t)\beta_{ji}^3 = S_i^V(t)\beta_{ji}^4 = 0 \quad \forall j \in \Pi. \quad (\text{D346})$$

Proof Suppose that there exists some t and some $i, j \in \Pi$ such that

$$S_i^V(t)\beta_{ji}^3 > 0 \quad \text{and} \quad I_i^V(t) = 0. \quad (\text{D347})$$

Then, by continuity, there exists some $a < t$ such that

$$S_i^V(s)\beta_{ji}^3 > 0 \quad \forall s \in (a, t). \quad (\text{D348})$$

Moreover, by Lemma 12, it is necessary that

$$I_i^V(s) = 0 \quad \forall s \in (a, t), \quad (\text{D349})$$

while, by Lemma 13

$$I_j(t) > 0 \quad \forall s \in (a, t) \quad (\text{D350})$$

and hence (using the fact that $I_i^V(s) = 0 \quad \forall s \in (a, t)$)

$$\frac{dI_i^V}{dt} \geq S_i^V(s) \beta_{ji}^3 I_j(t) > 0 \quad \forall s \in (a, t) \quad (\text{D351})$$

and so

$$I_i^V(t) > I_i^V(a) = 0, \quad (\text{D352})$$

which is a contradiction as required. The final equality then follows as $\beta_{ji}^3 \geq \beta_{ji}^4 \geq 0$. \square

Lemma 16

The following result extends the main theorem from Penn and Donnelly (2022) in a similar way to Proposition 6 to provide an additional inequality on the objective values from the optimal vaccination problem.

Lemma 16 *Suppose that the disease trajectories S and \tilde{S} are given by the same model equations, parameters, vaccination policy U and initial conditions except for the fact that*

$$S_1^V(0) > \tilde{S}_1^V(0). \quad (\text{D353})$$

Then, if the objective functions are denoted by H and \tilde{H} for the two policies,

$$H(U) \geq \tilde{H}(U). \quad (\text{D354})$$

Proof Define a new disease model, denoted by hats where a new group $(n+1)$ is added in such that its unvaccinated compartments behave like the vaccinated compartments of group 1 and its vaccinated compartments are perfectly immune from the disease. That is,

$$\hat{\beta}_{(n+1)j}^1 = \beta_{1j}^3, \quad \hat{\beta}_{(n+1)j}^2 = \beta_{1j}^4, \quad \text{and} \quad \hat{\beta}_{(n+1)j}^3 = \hat{\beta}_{(n+1)j}^4 = 0 \quad \forall j \in \{1, \dots, n\}, \quad (\text{D355})$$

$$\hat{\beta}_{j(n+1)}^1 = \beta_{j1}^3, \quad \hat{\beta}_{j(n+1)}^2 = \beta_{j1}^4, \quad \text{and} \quad \hat{\beta}_{j(n+1)}^3 = \hat{\beta}_{j(n+1)}^4 = 0 \quad \forall j \in \{1, \dots, n\}, \quad (\text{D356})$$

$$\beta_{(n+1)(n+1)}^\alpha = 0 \quad \forall \alpha \in \{1, 2, 3, 4\} \quad (\text{D357})$$

and

$$\hat{\mu}_{n+1}^1 = \mu_1^2 \quad \text{and} \quad \hat{\mu}_{n+1}^2 = 1. \quad (\text{D358})$$

Suppose further that all other parameter values are identical and that the only differences in the initial conditions are that

$$\hat{S}_1^V(0) = \tilde{S}_1^V(0) \quad \text{and} \quad S_{n+1}(0) = S_1^V(0) - \tilde{S}_1^V(0) > 0. \quad (\text{D359})$$

Then, note that

$$\begin{aligned} \frac{d(\hat{S}_1^V + \hat{S}_{n+1})}{dt} &= - \sum_{j=1}^{n+1} \left[\hat{S}_1^V (\hat{\beta}_{1j}^3 \hat{I}_j + \hat{\beta}_{1j}^4 \hat{I}_j^V) + \hat{S}_{n+1} (\hat{\beta}_{(n+1)j}^1 \hat{I}_j + \hat{\beta}_{(n+1)j}^2 \hat{I}_j^V) \right] \dots \\ &\quad \dots - \frac{\hat{S}_{n+1} \hat{U}_{n+1}}{\hat{N}_{n+1} - \hat{W}_{n+1}} \\ &= -(\hat{S}_1^V + \hat{S}_{n+1}) \sum_{j=1}^{n+1} \left[\hat{\beta}_{1j}^3 \hat{I}_j + \hat{\beta}_{1j}^4 \hat{I}_j^V \right] - \frac{\hat{S}_{n+1} \hat{U}_{n+1}}{\hat{N}_{n+1} - \hat{W}_{n+1}}. \end{aligned} \quad (\text{D360})$$

Moreover, for $i \neq 1$

$$\begin{aligned} \frac{d}{dt}(\hat{S}_i) &= -\hat{S}_i \sum_{j=1}^{n+1} \left[\beta_{ij}^1 \hat{I}_i + \beta_{ij}^2 \hat{I}_i^V \right] - \frac{\hat{S}_i \hat{U}_i}{\hat{N}_i - \hat{W}_i} \\ &= -\hat{S}_i \left(\sum_{j=2}^n \left[\beta_{ij}^1 \hat{I}_i + \beta_{ij}^2 \hat{I}_i^V \right] + \beta_{ij}^1 \hat{I}_1 + \beta_{ij}^2 (\hat{I}_1^V + \hat{I}_{n+1}) \right) - \frac{\hat{S}_i \hat{U}_i}{\hat{N}_i - \hat{W}_i}. \end{aligned} \quad (\text{D361})$$

$$(\text{D362})$$

Thus, with similar calculations for \hat{I} , \hat{I}^V , \hat{R} and \hat{R}^V , by the initial conditions and by the uniqueness of solution, in the case that $\hat{U}_{n+1} = 0$,

$$\hat{S}_1^V + \hat{S}_{n+1} = S_1^V \quad \hat{I}_1^V + \hat{I}_{n+1} = I_1^V \quad \text{and} \quad \hat{R}_1^V + \hat{R}_{n+1} = R_1^V. \quad (\text{D363})$$

Thus, setting

$$p_{n+1} = p_1 \kappa_1, \quad (\text{D364})$$

this means that

$$\hat{H}(\hat{U}) = H(U) \quad (\text{D365})$$

for any \hat{U} such that $\hat{U}_{n+1} = 0$ and $\hat{U}_i = U_i$ for any $i \neq n+1$.

Now, define a vaccination policy $\hat{U}^*(t; \Delta)$ such that

$$\hat{U}_i^*(t; \Delta) = \hat{U}_i(t) \quad \forall t \geq 0 \quad \text{and} \quad i \neq n+1 \quad (\text{D366})$$

and

$$\hat{U}_{n+1}^*(t; \Delta) = \begin{cases} \frac{1}{\Delta} \left(S_1^V(0) - \tilde{S}_1^V(0) \right) & \text{if } t \leq \Delta \\ 0 & \text{otherwise} \end{cases}. \quad (\text{D367})$$

Then, this means that

$$\hat{S}_{n+1}(\Delta; \Delta) = 0 \quad \text{and} \quad \hat{S}_{n+1}^V(\Delta; \Delta) = S_1^V(0) - \tilde{S}_1^V(0) + O(\Delta) \quad (\text{D368})$$

while all other variable values at time Δ differ by at most $O(\Delta)$ from their initial values. Thus, define by an overbar the model given by the initial conditions which are the same as those in the hat model, but with

$$\bar{S}_{n+1}(0) = 0 \quad \text{and} \quad \bar{S}_{n+1}^V = S_1^V(0) - \tilde{S}_1^V(0). \quad (\text{D369})$$

Suppose also that the vaccination policy in this case is equal to U , which is the point-wise limit of the vaccination policy $\hat{U}^*(t; \Delta)$ (for $t > 0$). Then, using Proposition 5, by considering the values of the variables \hat{f} at time Δ to be the initial conditions, one finds that for any finite time t ,

$$\lim_{\Delta \rightarrow 0} (\hat{H}(U^*(t; \Delta))) = \bar{H}(U). \quad (\text{D370})$$

Note this holds as it is assumed that U is bounded and so

$$|W_i(t + \Delta; \Delta) - W_i(\Delta; \Delta) - W_i(t)| = O(\Delta). \quad (\text{D371})$$

Moreover, note that the only difference between the bar model and the tilde model is in group $(n + 1)$. However, by the fact that $\beta_{ij}^3 = \beta_{ij}^4 = 0$ if $(n + 1) \in \{i, j\}$, the value of \bar{S}_{n+1}^V is constant and the other variable values are independent of it. Thus, by the uniqueness of solution, this means that

$$\bar{H}(U) = \tilde{H}(U). \quad (\text{D372})$$

Finally, note that by Theorem 17, it must be necessary that for any $\Delta > 0$

$$\hat{H}(U(t; \Delta)) \leq \hat{H}(U(t; \infty)) = H(U), \quad (\text{D373})$$

where $\Delta = \infty$ corresponds to no vaccination taking place in group $(n + 1)$ (and hence the original objective function H is recovered). Thus,

$$\tilde{H}(U) \leq H(U), \quad (\text{D374})$$

as required.

Theorem 17

Theorem 17 Suppose that U and \tilde{U} are feasible, bounded, Lebesgue-integrable vaccination policies. Suppose further that for each $i \in \{1, \dots, n\}$ and $t \geq 0$

$$\int_0^t U_i(s) ds \leq \int_0^t \tilde{U}_i(s) ds. \quad (D375)$$

Then, for each $t \geq 0$ and $i \in \{1, \dots, n\}$

$$I_i(t) + R_i(t) + I_i^V(t) + R_i^V(t) \geq \tilde{I}_i(t) + \tilde{R}_i(t) + \tilde{I}_i^V(t) + \tilde{R}_i^V(t) \quad (D376)$$

and

$$R_i(t) \geq \tilde{R}_i(t). \quad (D377)$$

Moreover,

$$H(U) \geq H(\tilde{U}). \quad (D378)$$

Proof A proof of this theorem is given in Penn and Donnelly (2022), where it is Theorem 1. Note that the first two results are not in the statement of Theorem 1 in Penn and Donnelly (2022), but can be found at the end of the proof.

Theorem 18

Theorem 18 Suppose that B is differentiable and that there is an optimal solution U . Then, define the function

$$\chi(t) := \begin{cases} A(t) & \text{if } \int_0^t \chi(s) ds < B(t) \\ \min(A(t), B'(t)) & \text{if } \int_0^t \chi(s) ds \geq B(t) \end{cases} \quad (D379)$$

and suppose that $\chi(t)$ exists and is bounded. Then, there exists an optimal solution \tilde{U} such that

$$\sum_{i=1}^n \tilde{W}_i(t) = \max \left(\int_0^t \chi(s) ds, 1 \right). \quad (D380)$$

Moreover, if $\chi(t)$ is continuous almost everywhere, there exists an optimal solution \tilde{U} such that

$$\sum_{i=1}^n \tilde{U}_i(t) = \begin{cases} \chi(t) & \text{if } \int_0^t \chi(s) ds < 1 \\ 0 & \text{otherwise} \end{cases} \quad (D381)$$

Proof A proof of this theorem is given in Penn and Donnelly (2022) where it is Theorem 2.

References

- Abou-Ismaïl A (2020) Compartmental models of the covid-19 pandemic for physicians and physician-scientists. *SN Compr Clin Med* 2(7):852–858
- Acemoglu D, Chernozhukov V, Werning I et al (2021) Optimal targeted lockdowns in a multigroup sir model. *Am Econ Rev Insights* 3(4):487–502
- Adhikari R, Bolitho A, Caballero F, et al (2020) Inference, prediction and optimization of non-pharmaceutical interventions using compartment models: the pyross library. [arXiv:2005.09625](https://arxiv.org/abs/2005.09625)
- Anderson RM, May RM (1992) *Infectious diseases of humans: dynamics and control*. Oxford University Press, Oxford
- Ball F, Neal P (2002) A general model for stochastic sir epidemics with two levels of mixing. *Math Biosci* 180(1–2):73–102
- Bedson J, Skrip LA, Pedi D et al (2021) A review and agenda for integrated disease models including social and behavioural factors. *Nat Hum Behav* 5(7):834–846
- Berman A, Plemmons RJ (1994) *Nonnegative matrices in the mathematical sciences*. SIAM, New Delhi
- Bloom DE, Cadarette D, Ferranna M (2021) The societal value of vaccination in the age of covid-19. *Am J Public Health* 111(6):1049–1054
- Blyth TS, Robertson EF (2002) *Basic linear algebra*. Springer, Berlin
- Boutayeb H, Bidah S, Zakary O, et al (2021) Automated optimal vaccination and travel-restriction controls with a discrete multi-region sir epidemic model. *Commun Math Biol Neurosci* 2021
- Brauer F (2008) Epidemic models with heterogeneous mixing and treatment. *Bull Math Biol* 70(7):1869–1885
- Brauer F, Castillo-Chavez C, Feng Z (2019) *Mathematical models in epidemiology*, vol 32. Springer, Berlin
- de Camino-Beck T (2020) A modified seir model with confinement and lockdown of covid-19 for costa rica. [medRxiv](https://medrxiv.org/content/2020.09.15.20191111v1)
- Delmas JF, Dronnier D, Zitt PA (2021) Optimal vaccination: various (counter) intuitive examples. [arXiv:2112.08756](https://arxiv.org/abs/2112.08756)
- Duijzer E, van Jaarsveld W, Wallinga J et al (2016) The most efficient critical vaccination coverage and its equivalence with maximizing the herd effect. *Math Biosci* 282:68–81
- Duijzer LE, van Jaarsveld WL, Wallinga J et al (2018) Dose-optimal vaccine allocation over multiple populations. *Prod Oper Manag* 27(1):143–159
- Dushoff J, Plotkin JB, Viboud C et al (2007) Vaccinating to protect a vulnerable subpopulation. *PLoS Med* 4(5):e174
- Dyer O (2021) Covid-19: unvaccinated face 11 times risk of death from delta variant, cdc data show
- Eichner M, Schwehm M, Eichner L et al (2017) Direct and indirect effects of influenza vaccination. *BMC Infect Dis* 17(1):1–8
- Ellison G (2020) *Implications of heterogeneous sir models for analyses of covid-19*. Technical report, National Bureau of Economic Research
- Fitzpatrick MC, Galvani AP (2021) Optimizing age-specific vaccination. *Science* 371(6532):890–891
- Gavish N, Katriel G (2022) Optimal vaccination at high reproductive numbers: sharp transitions and counter-intuitive allocations. [arXiv:2202.03909](https://arxiv.org/abs/2202.03909)
- Hansen E, Day T (2011) Optimal control of epidemics with limited resources. *J Math Biol* 62(3):423–451
- Kar TK, Batabyal A (2011) Stability analysis and optimal control of an sir epidemic model with vaccination. *Biosystems* 104(2–3):127–135
- Keeling MJ, White PJ (2011) Targeting vaccination against novel infections: risk, age and spatial structure for pandemic influenza in Great Britain. *J R Soc Interface* 8(58):661–670
- Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc Lond Ser A Contain Pap Math Phys Char* 115(772):700–721
- Kong L, Duan M, Shi J et al (2022) Compartmental structures used in modeling covid-19: a scoping review. *Infect Dis Poverty* 11(1):1–9
- Kuniya T (2019) Global behavior of a multi-group sir epidemic model with age structure and an application to the chlamydia epidemic in Japan. *SIAM J Appl Math* 79(1):321–340

- Lee S, Golinski M, Chowell G (2012) Modeling optimal age-specific vaccination strategies against pandemic influenza. *Bull Math Biol* 74(4):958–980
- Longini IM Jr, Ackerman E, Elveback LR (1978) An optimization model for influenza a epidemics. *Math Biosci* 38(1–2):141–157
- Mahase E (2022) Monkeypox: healthcare workers will be offered smallpox vaccine as UK buys 20,000 doses. *bmj* 377:o1379
- Medlock J, Meyers LA (2009) Optimizing allocation for a delayed influenza vaccination campaign. *PLoS Curr* 1
- Miura F, Leung KY, Klinkenberg D et al (2021) Optimal vaccine allocation for covid-19 in the Netherlands: a data-driven prioritization. *PLoS Comput Biol* 17(12):e1009697
- Molinari NAM, Ortega-Sanchez IR, Messonnier ML et al (2007) The annual impact of seasonal influenza in the US: measuring disease burden and costs. *Vaccine* 25(27):5086–5096
- Moore S, Hill EM, Dyson L et al (2021) Modelling optimal vaccination strategy for sars-cov-2 in the UK. *PLoS Comput Biol* 17(5):e1008849
- Muhammad L, Algehyne EA, Usman SS et al (2021) Supervised machine learning models for prediction of covid-19 infection using epidemiology dataset. *SN Comput Sci* 2(1):1–13
- Pakkanen MS, Miskouridou X, Berah T, et al (2021) Unifying incidence and prevalence under a time-varying general branching process. [arXiv:2107.05579](https://arxiv.org/abs/2107.05579)
- Penn MJ, Donnelly CA (2022) Optimality of maximal-effort vaccination. *medRxiv*
- Prem K, Cook AR, Jit M (2017) Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput Biol* 13(9):e1005697
- Ram V, Schaposnik LP (2021) A modified age-structured sir model for covid-19 type viruses. *Sci Rep* 11(1):1–15
- Rao IJ, Brandeau ML (2021) Optimal allocation of limited vaccine to control an infectious disease: simple analytical conditions. *Math Biosci* 337(108):621
- Saadi N, Chi Y, Ghosh S et al (2021) Models of covid-19 vaccine prioritisation: a systematic literature search and narrative review. *BMC Med* 19(1):1–11
- Sattenspiel L, Dietz K (1995) A structured epidemic model incorporating geographic mobility among regions. *Math Biosci* 128(1–2):71–91
- Shim E (2011) Prioritization of delayed vaccination for pandemic influenza. *Math Biosci Eng MBE* 8(1):95
- Shim E (2021) Optimal allocation of the limited covid-19 vaccine supply in South Korea. *J Clin Med* 10(4):591
- UN (2019) World population prospects-population division. <https://population.un.org/wpp/Download/Standard/Population/>
- Vardavas R, de Lima PN, Baker L (2021) Modeling covid-19 nonpharmaceutical interventions: exploring periodic npi strategies. *medRxiv*
- Zaman G, Kang YH, Jung IH (2008) Stability analysis and optimal vaccination of an sir epidemic model. *Biosystems* 93(3):240–249
- Zaric GS, Brandeau ML (2001) Resource allocation for epidemic control over short time horizons. *Math Biosci* 171(1):33–58

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.