



# Optimality of Maximal-Effort Vaccination

Matthew J. Penn<sup>1</sup>  · Christl A. Donnelly<sup>1,2,3</sup>

Received: 28 February 2023 / Accepted: 2 June 2023  
© The Author(s) 2023

## Abstract

It is widely acknowledged that vaccinating at maximal effort in the face of an ongoing epidemic is the best strategy to minimise infections and deaths from the disease. Despite this, no one has proved that this is guaranteed to be true if the disease follows multi-group SIR (Susceptible–Infected–Recovered) dynamics. This paper provides a novel proof of this principle for the existing SIR framework, showing that the total number of deaths or infections from an epidemic is decreasing in vaccination effort. Furthermore, it presents a novel model for vaccination which assumes that vaccines assigned to a subgroup are distributed randomly to the unvaccinated population of that subgroup. It suggests, using COVID-19 data, that this more accurately captures vaccination dynamics than the model commonly found in the literature. However, as the novel model provides a strictly larger set of possible vaccination policies, the results presented in this paper hold for both models.

**Keywords** Vaccination · Epidemiology · Epidemics · SIR Modelling

## 1 Introduction

The COVID-19 pandemic has illustrated the importance of quickly implementing vaccination policies which target particular groups within a population (Fitzpatrick and Galvani 2021). The difference in final infections between targeted policies and uniform distribution to the entire population can be significant (Castro and Singer

---

✉ Matthew J. Penn  
matthew.penn@st-annes.ox.ac.uk  
Christl A. Donnelly  
christl.donnelly@stats.ox.ac.uk

<sup>1</sup> Department of Statistics, University of Oxford, St Giles', Oxford OX1 3LB, UK

<sup>2</sup> Department of Infectious Disease Epidemiology, Imperial College London, St Mary's Campus, London W2 1PG, UK

<sup>3</sup> Pandemic Sciences Institute, University of Oxford, Roosevelt Drive, Oxford OX3 7DQ, UK

2021; Estadilla et al. 2021) and so it is important that the models underlying these decisions provide realistic predictions of the outcomes of different policies.

One of the most commonly used models to forecast epidemics is the multi-group SIR (Susceptible–Infected–Recovered) model (Acemoglu et al. 2021; Kuniya 2019; Ram and Schaposnik 2021). This model divides the population into different groups based on characteristics such as age or occupation. Each group is then further subdivided into categories of susceptible, infected and recovered. Where vaccination does not give perfect immunity, further sub-categorization based on vaccination status can also be used (Kuga and Tanimoto 2018), as will be done in this paper.

While many other approaches have been developed either by adding compartments to the SIR framework (Moore et al. 2021) or using completely different models such as networks (Chen and Sun 2014) or stochastic simulations (Ball and Lyne 2002), the multi-group SIR model remains popular because of its comparatively small number of parameters and its relatively simple construction and solution. In this paper, attention will thus be restricted to the multi-group SIR model, although it would be beneficial for future work to consider a wider range of disease models.

There are two general frameworks that are used to model optimal vaccination policies in a resource-limited setting. The first, used in papers such as Hill and Longini Jr (2003) and Becker and Starczak (1997), seeks to reduce the reproduction number,  $R_0$  of the epidemic as much as possible by vaccinating before infections arrive in a population. It is simple to show that in this case, one should use all of the vaccinations available, and so this problem will not be considered further in this paper.

The second framework, used in papers such as Acemoglu et al. (2021) and Hansen and Day (2011) aims to minimise the total cost of an epidemic. This is the framework that will be discussed in this paper. The “cost” of an epidemic is, in general, defined to be the number of deaths (or equivalently, infections), with many papers also considering the cost of vaccination alongside the cost of other control measures, such as isolation, lockdown or treatment (Fu et al. 2022).

One important principle which underlies all of these vaccination policies is the acceptance that giving people their first dose of vaccine as soon as possible reduces the number of infections. Of course, this only holds when the timescale considered is sufficiently short for effects such as waning immunity and disease seasonality to be negligible, and a more complicated framework would be needed to model these effects. However, the acceptance of at least short-term optimality of maximal vaccination effort has been highlighted in the COVID-19 pandemic response, as countries began their vaccination programs as soon as vaccines became available (Mathieu et al. 2021).

To the best of the authors’ knowledge, no one has provided a mathematical proof that in a general, multi-group SIR model with imperfect vaccination, it is always best to vaccinate people as early as possible. Of course, it is not difficult to create a conceptually sound justification—vaccinating more people means that fewer people will catch the disease which will reduce the total number of infections. However, the SIR model is an approximation of the process of a disease spreading, and so it is important that it obeys this principle for all physical parameter values and vaccination policies.

Some special cases of the theorem presented in this paper have been previously proved in the literature. In particular, a significant number of papers have considered

the optimal vaccination policy for a homogeneous population, with Abakuks (1972) first proving that, in this case, it is optimal to vaccinate at maximal effort (if one ignores the cost of vaccination). This proof held for vaccination policies that were finite sums of point mass “impulse” vaccinations, and has been generalised by papers such as Hansen and Day (2011), Zaman et al. (2008), Morton and Wickwire (1974) and Zhou et al. (2014) to a much wider class of vaccination policies, although the proof was still restricted to a single group and to perfect vaccination. Moreover, Hansen and Day (2011) notes that the case of imperfect vaccination (where vaccinated individuals can still get infected, although at a lower rate) remained a topic of open investigation, and so it can not easily be solved using the same methods presented in these papers. A slight extension is made in Duijzer et al. (2018) where it is shown that maximal effort is optimal in the case of perfect vaccination of any number of disconnected groups, but the full problem is still far from understood.

The general method of proof in the literature relies on Pontryagin’s Maximum Principle, which is difficult to apply to multi-group models due to the more complex structure of the equations. It is simple to characterise the solution in terms of the adjoint variables, as is done in Zhang et al. (2020) and Zavrakli et al. (2021) for a two-group model with imperfect vaccination, in Boutayeb et al. (2021) for a general  $n$ -group model with perfect vaccination and in Lee et al. (2012) for a six-group model with imperfect vaccination. However, determining whether this solution corresponds to the maximal effort solution in the case of zero vaccination cost requires the analysis of the adjoint ODE system, which is often just as complicated as the original disease model. In particular, the fact that vaccinated people need to be no more infectious, no more susceptible and be infected for no more time than unvaccinated people means that any analysis of the adjoint system would be complicated, as the properties of all the constituent parameters would need to be used.

Thus, in this paper, a novel approach is developed. Rather than attempting to use the general optimal control theory methodology, the specific structure of the SIR equation system is exploited. Using this, an inequality is derived which shows that if a given vaccination policy,  $\tilde{U}$  vaccinates each individual at least as early as another vaccination policy,  $U$ , then the latter policy will lead to at least as many deaths (or equivalently, infections) as the former. As well as providing a constraint on the optimal solution, this theorem also highlights important structural properties of the model, as it shows that the number of deaths is everywhere non-increasing in the vaccination rates, rather than this just holding near the optimal solutions.

Also introduced in this paper is a more general model of vaccination than the one normally used in the literature. The one that is typically used (in almost all papers cited in this work such as Hansen and Day (2011), Zaman et al. (2008) and Kar and Batabyal (2011)) models decreasing vaccination uptake by assuming that the total number rate of people being vaccinated is the product of a vaccination rate and the proportion of susceptible people in the population. The model introduced in this paper allows for more flexibility in modelling the demand. However, the standard vaccination model is a special case of the general model introduced here, and so the theoretical results proved in this paper can be used by those following the standard model.

Alongside proving that the final infected, recovered and dead populations are non-increasing with increased and earlier vaccination effort, some cautionary contra-

dictions to perhaps intuitive conjectures are also provided which show the importance of mathematical proof instead of simply intuition. In particular, it is shown that increased vaccination (under this model) can lead to, at a fixed finite time of the simulation, higher infection rates or a higher death count, despite the longer-term better performance of this policy. Indeed, it is results similar to these which make the proof of the optimality of maximal effort difficult, as it means that one must be very careful when constructing the inequalities that do hold for all models.

## 2 Modelling

### 2.1 Disease Transmission and Vaccination Model

Suppose that the population is divided into  $n$  subgroups, such that population of people in group  $i$  is  $N_i$  and define

$$N := \sum_{i=1}^n N_i.$$

Define the compartments of people as follows, for  $i = 1, \dots, n$ :

$S_i$  := Number of people that are in group  $i$ , are susceptible, and are unvaccinated,

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

$R_i$  := Number of people that are in group  $i$ , are recovered or dead, and were infected while unvaccinated,

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

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

$R_i^V$  := Number of people that are in group  $i$ , are recovered or dead and were infected after being vaccinated.

This paper introduces a more general and flexible framework for vaccination, which is motivated as follows. It is assumed that there is a record of people who have received a vaccination and that protection from vaccination does not decay over time, so that no one is vaccinated more than once. Thus, if a total number,  $U_i(t)dt$ , of people in group  $i$  are given vaccines in a small time interval  $(t, t + dt)$ , and these vaccines are distributed randomly to the unvaccinated population in group  $i$ , the total population of susceptibles given vaccines in group  $i$  is

$$U_i(t)dt \times \mathbb{P}\left(\text{A person in group } i \text{ is in } S_i \mid \text{A person in group } i \text{ is unvaccinated}\right)$$

which is equal to

$$\frac{U_i(t)dt S_i(t)}{N_i - \int_0^t U_i(s)ds},$$

as  $\int_0^t U_i(s)ds$  is the total population that are in group  $i$  and have been vaccinated before time  $t$ . For the remainder of this section, this vaccination model will be referred to as the “general” model

This results in the following model, based on SIR principles

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

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

$$\frac{dR_i}{dt} = \mu_i^1 I_i,$$

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

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

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

where

$$W_i(t) := \int_0^t U_i(s)ds.$$

Here,  $\beta_{ij}^1$  represents transmission from the unvaccinated members of group  $j$  to the unvaccinated members of group  $i$ ,  $\beta_{ij}^2$  represents transmission from vaccinated members to unvaccinated members,  $\beta_{ij}^3$  represents transmission from vaccinated members to unvaccinated members and  $\beta_{ij}^4$  represents transmission from vaccinated members to vaccinated members. Additionally,  $\mu_i^1$  represents the infectious period of unvaccinated infected members in group  $i$  while  $\mu_i^2$  represents the infectious period of vaccinated members. Note that the superscript denotes different parameter values, so that  $\beta_{ij}^2$  is not necessarily the square of  $\beta_{ij}^1$ .

To ensure that vaccination is “locally effective” (that is, a vaccinated individual is no more likely to transmit or be infected by the disease, and is infectious for no longer than an unvaccinated individual in the same subgroup), and that the parameters are epidemiologically feasible, the following constraints are imposed:

$$\beta_{ij}^1 \geq \beta_{ij}^2, \beta_{ij}^3 \geq \beta_{ij}^4 \geq 0 \quad \text{and} \quad \mu_i^2 \geq \mu_i^1 > 0$$

Note that there is no constraint on the ordering of  $\beta_{ij}^2$  and  $\beta_{ij}^3$ . It is assumed for convenience that all variables except the  $S_i$  and  $I_i$  are initially zero. Finally, we assume that all initial conditions are non-negative.

Ultimately, the objective of the vaccination program will be to minimize a weighted sum of the total infections in each group—that is

$$\sum_{i=1}^n p_i (R_i(\infty) + \kappa_i R_i^V(\infty)).$$

Here  $p_i$  is the weighting of a member of group  $i$  who is infected before being vaccinated, while  $p_i \kappa_i$  is the weighting of a member of group  $i$  who is infected after being vaccinated. These parameters could be chosen to capture one of a range of objectives, such as minimizing deaths, minimizing hospitalisations, or minimizing total infections. Again assuming “local effectiveness” of the vaccination, it is imposed that  $\kappa_i \leq 1$ , as vaccination should not increase the severity of the infection.

The equations (1)–(2) sum to zero on the right-hand side, and so for each  $i$ ,

$$S_i(t) + I_i(t) + R_i(t) + S_i^V(t) + I_i^V(t) + R_i^V(t) = N_i \quad \forall t \geq 0. \quad (3)$$

It will be assumed that the populations and parameters have been scaled such that  $N = 1$ . Finally, it is assumed that

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

to ensure feasibility of the vaccination policies.

## 2.2 Comparison to the Standard Vaccination Model

A more common model of vaccination in the literature is the “standard” vaccination model (Hansen and Day 2011; Zaman et al. 2008; Kar and Batabyal 2011), where Eq. (1) becomes

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

Here,  $U_i^*(t)$  is the vaccination rate in this model. In general,  $U_i^*(t)$  is constrained such that  $U_i^*(t) \leq \mathcal{U}_i(t)$  for some function  $\mathcal{U}_i(t)$

The  $U_i^*(t)S_i$  term seeks to capture the fact that vaccination uptake will decrease even if the vaccination “effort” (or, equivalently, the doses available) remains constant. However, the rate at which uptake decreases is fixed by the model. For example, if the

vaccination effort  $U_i^*(t)$  was equal to a constant  $\mathcal{U}_i$  and was much quicker than the rate of infection, then the leading order equation is

$$\frac{dS_i}{dt} = -\mathcal{U}_i S_i \Rightarrow S_i(t) = S_i(0)e^{-\mathcal{U}_i t}$$

and hence

$$\frac{dS_i}{dt} = -\mathcal{U}_i S_i(0)e^{-\mathcal{U}_i t}$$

which means that vaccination uptake decreases exponentially. Even for some human pandemics, such as COVID-19, where demand remained high until a large proportion of the population had been vaccinated, as shown in Ritchie et al. (2020), such a model may be inappropriate.

The general vaccination model provides substantially more flexibility. For example, it is possible for a group to be completely vaccinated in the general case, whereas this is impossible in the standard case (while one may never be able to fully vaccinate a human population, it would be possible, for example, in a group of animals on a farm). Moreover, by bounding the vaccination rate  $U_i(t)$  above by some function of vaccination demand  $K_i(W(t), t)$ , decreasing vaccination uptake can still be modelled.

### 2.3 Recovery of the Standard Model

The standard model is a special case of the general model, meaning that the results of this paper are applicable to both frameworks. To show this, one can solve the equation

$$\frac{U_i(t)}{N_i - W_i(t)} = U_i^*(t) \Rightarrow \frac{d}{dt} \left( \log(N_i - W_i(t)) + W_i^*(t) \right) = 0, \quad (4)$$

where

$$W_i^*(t) := \int_0^t U_i^*(s) ds.$$

Thus, by integrating (4), and noting that  $W_i^*(0) = W_i(0) = 0$

$$\log(N_i - W_i(t)) + W_i^*(t) = \log(N_i)$$

and so

$$W_i(t) = N_i(1 - e^{-W_i^*(t)}).$$

The constraint  $U_i^*(t) \leq \mathcal{U}_i$  is equivalent to  $U_i(t) \leq (N_i - W_i(t))\mathcal{U}_i$  and so this can also be represented in the general model. Thus, given any standard vaccination policy  $U^*$ , it can be replaced by a general policy  $U$  (although the converse does not hold as  $W_i(t) = N_i$  requires  $W_i^*(t) = \infty$ ).

Moreover, note that  $W_i^*(t)$  is increasing in  $W_i(t)$ . Thus, if a pair of general policies  $U$  and  $\tilde{U}$  satisfy  $W_i(t) \leq \tilde{W}_i(t)$  then this inequality is preserved by the corresponding standard policies as  $W_i^*(t) \leq \tilde{W}_i^*(t)$ . This property means that the theorems proved in this paper will hold for both models (as they will be proved using the general model).

### 3 Optimisation Problem

Now that the model has been formulated, it is possible to set up the optimisation problem that will be considered in the remainder of this paper.

#### 3.1 Constraints on $U_i(t)$

In order to assist the proof of the theorems, it is necessary to make some (unrestrictive) assumptions on the vaccination rates,  $U_i(t)$ .

Firstly, there are the physical constraints that for each  $i \in \{1, \dots, n\}$

$$U_i(t) \geq 0 \quad \text{and} \quad \int_0^t U_i(s) ds \leq N_i \quad \forall t \geq 0. \quad (5)$$

It is also necessary that  $U_i(t)$  is within the class of functions such that solutions to the model equations exist and are unique. Discussion of the exact conditions necessary for this to hold is outside the scope of this paper. However, from the Picard-Lindelöf Theorem (Collins 2006), a sufficient condition for this is that  $U_i(t)$  is a piecewise Lipschitz continuous function. While this is not a necessary condition, this illustrates that this assumption will hold for a large class of functions. However, it will be helpful throughout the course of the proof to explicitly assume two conditions on  $U_i(t)$  - namely, that it is bounded and that it is Lebesgue integrable on  $\mathfrak{R}$  for each  $i$ .

For the remainder of this paper, define the set of feasible vaccination policies,  $C$ , is the set of functions  $U$  satisfying (5) such that unique solutions to the model equations exist with these functions as the vaccination policy and such that each  $U_i(t)$  is bounded and Lebesgue integrable on  $\mathfrak{R}$ .

#### 3.2 Optimisation Problem

The aim is to choose the vaccination policy  $U \in C$  such that the total number of deaths (or any linear function of the infections in each subgroup) is minimised while meeting additional constraints on vaccine supply and vaccination rate. It is assumed that the maximal rate of vaccination at time  $t$  is  $A(t)$  and that there is a total (non-decreasing) supply of  $B(t)$  vaccinations that has arrived by time  $t$ . Thus, for each  $i$ ,  $U_i(t)$  is constrained to satisfy

$$\sum_{i=1}^n U_i(t) \leq A(t) \quad \text{and} \quad \sum_{i=1}^n W_i(t) \leq B(t).$$



As previously discussed, it is assumed that each infection of unvaccinated people in group  $i$  is weighted by some  $p_i$  and that the infection is no more serious for those that have been vaccinated, so that the weighting of an infection of a vaccinated person in group  $i$  is  $p_i \kappa_i$ , where  $\kappa_i \leq 1$ . Thus, the objective function is

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

where, for example

$$R_i(\infty) = \lim_{t \rightarrow \infty} (R_i(t)).$$

Note these limits exist as  $R_i$  is non-decreasing and bounded by Lemma C.3. Hence, the optimisation problem is

$$\min \left\{ \sum_{i=1}^n p_i \left( R_i(\infty) + \kappa_i R_i^V(\infty) \right) : \sum_{i=1}^n U_i(t) \leq A(t), \sum_{i=1}^n W_i(t) \leq B(t) \quad \forall i, t \dots \right. \\ \left. \text{and } U \in C \right\}. \quad (6)$$

## 4 Main Results

The main results of this paper are as follows. Firstly, it is shown that the objective function is non-increasing in vaccination effort.

**Theorem 1** *Suppose that  $U, \tilde{U} \in C$ . Suppose further that for each  $i \in \{1, \dots, n\}$  and  $t \geq 0$*

$$\int_0^t U_i(s) ds \leq \int_0^t \tilde{U}_i(s) ds$$

*Then*

$$H(U) \geq H(\tilde{U}).$$

Then, it is shown that if an optimal solution exists, there is an optimal maximal effort solution.

**Theorem 2** *Suppose that  $B$  is differentiable, and that there is an optimal solution  $U$  to (6). Then, define the function*

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

and suppose that  $\chi(t)$  exists and is bounded. Then, there exists an optimal solution  $\tilde{U}$  to the problem (6) such that

$$\sum_{i=1}^n \tilde{W}_i(t) = \min \left( \int_0^t \chi(s) ds, 1 \right). \quad (7)$$

Moreover, if  $\chi(t)$  is continuous almost everywhere, there exists an optimal solution  $\tilde{U}$  such that

$$\sum_{i=1}^n \tilde{U}_i(t) = \begin{cases} \chi(t) & \text{if } \int_0^t \chi(s) ds < 1 \\ 0 & \text{otherwise} \end{cases}$$

It is perhaps concerning to the reader that the existence of  $\chi$  is left as an assumption in this theorem. However, while the exact conditions on its existence are beyond the scope of this paper, it certainly exists for a wide class of functions  $A(t)$  and  $B(t)$ , as proved in Lemma B.11.

Finally, it is shown that this principle still holds if the cost of vaccination is considered.

**Theorem 3** *Under the assumptions of Theorem 2, consider a modified objective function  $\mathcal{H}$  given by*

$$\mathcal{H}(U) = H(U) + F(W(\infty))$$

for any function  $F$ . Then, with  $\chi$  defined to be the maximal vaccination effort as in Theorem 2, there exists an optimal solution  $\tilde{U}$  such that, for some  $\tau \geq 0$

$$\sum_{i=1}^n \tilde{W}_i(t) = \begin{cases} \int_0^t \chi(s) ds & \text{if } t \leq \tau \\ W_i(\tau) & \text{otherwise} \end{cases}.$$

Moreover, if  $\chi$  is continuous almost everywhere, then there is an optimal solution  $\tilde{U}$  such that

$$\sum_{i=1}^n U_i(t) = \begin{cases} \chi(t) & \text{if } t \leq \tau \\ 0 & \text{otherwise} \end{cases}.$$

## 5 Sketch Proof

The full proofs of Theorems 1, 2 and 3 can be found in Appendix A, with supplementary lemmas found in Appendix B and C. However, this section provides a high-level sketch of the main arguments.

## 5.1 Bounds on the Inter-Group Infectious Forces

Define

$$K_{ij}(t) = \frac{\beta_{ij}^1}{\mu_j^1} R_j(t) + \frac{\beta_{ij}^2}{\mu_j^2} R_j^V(t)$$

and

$$L_{ij}(t) := \frac{\beta_{ij}^3}{\mu_i^1} R_j(t) + \frac{\beta_{ij}^4}{\mu_i^2} R_j^V(t).$$

$K_{ij}(t)$  can be interpreted as the total infectious force up to time  $t$  from the members of group  $j$  on the unvaccinated members of group  $i$  as

$$K_{ij}(t) = \int_0^t (\beta_{ij}^1 I_j(\tilde{t}) + \beta_{ij}^2 I_j^V(\tilde{t})) d\tilde{t}.$$

Similarly,  $L_{ij}(t)$  can be interpreted as the total infectious force up to time  $t$  from the members of group  $j$  on the vaccinated members of group  $i$ .

The first part of the proof shows that increasing the vaccination effort will decrease these infectious forces. To facilitate the proof, some extra assumptions are made on the parameters (which will be removed in subsequent propositions).

**Proposition 1** *Suppose that  $U_i(t)$  and  $\tilde{U}_i(t)$  are right-continuous step functions. Moreover, suppose that*

$$\begin{aligned} \beta_{ij}^1 &> \beta_{ij}^3 > 0 \quad \forall i, j \in \{1, \dots, n\}, \\ S_i(0)I_i(0) &> 0 \quad \forall i \in \{1, \dots, n\}. \end{aligned}$$

and that

$$W_i(t) < N_i \quad \forall t \geq 0 \quad \text{and} \quad \forall i \in \{1, \dots, n\}$$

Then,

$$K_{ij}(t) \geq \tilde{K}_{ij}(t) \quad \text{and} \quad L_{ij}(t) \geq \tilde{L}_{ij}(t) \quad \forall t \geq 0. \quad (8)$$

This proposition is proved by contradiction in two parts. Firstly, a time  $T$  is introduced, which is the infimum of the times where at least one of  $K_{ij}(t) < \tilde{K}_{ij}(t)$  or  $L_{ij}(t) < \tilde{L}_{ij}(t)$  for some  $i$  and  $j$ . As the infectious forces do not satisfy this condition in  $[0, T]$ , one can show that, necessarily, they must all have been equal in  $[0, T]$ , which means that one must have  $W_i(t) = \tilde{W}_i(t)$  for all  $t \in [0, T]$ .

From here, the proof can proceed by a short-time linearisation, considering the small interval  $[T, T + \delta]$ . The condition on  $U_i$  and  $\tilde{U}_i$  being step functions allows for them to be considered constant in this interval. It can then be shown that (8) must hold in  $[T, T + \delta]$ , which contradicts the definition of  $T$  and completes the proof.

## 5.2 A Proof for a Restricted Parameter and Policy Set

Proposition 1 can be extended to prove the result of Theorem 1 under the more restrictive set of conditions it introduced.

**Proposition 2** *Under the conditions of Proposition 1, for any  $t \geq 0$  and  $i \in \{1, \dots, n\}$*

$$I_i(t) + I_i^V(t) + R_i(t) + R_i^V(t) \geq \tilde{I}_i(t) + \tilde{I}_i^V(t) + \tilde{R}_i(t) + \tilde{R}_i^V(t)$$

and

$$R_i(t) \geq \tilde{R}_i(t).$$

Moreover, for any  $\lambda \in [0, 1]$

$$R_i(\infty) + \lambda R_i^V(\infty) \geq \tilde{R}_i(\infty) + \lambda \tilde{R}_i^V(\infty)$$

and hence, the objective function is lower for  $\tilde{U}$ , provided the conditions of Proposition 1 are met.

This comes from finding  $S_i + S_i^V$  in terms of  $K_{ij}$ ,  $L_{ij}$  and  $W$ , and showing that  $S_i + S_i^V \leq \tilde{S}_i + \tilde{S}_i^V$ —that is, that more people were infected in the  $U_i$  case. Taking limits, and using a similar approach to consider the number of unvaccinated infections then shows the required result.

## 5.3 Generalisation

This result can be generalised to the original set of parameters and vaccination policies by using the continuous dependence of the number of infections on the parameters and the vaccination policy.

From here, it is simple to weaken the inequalities on the parameters introduced in Proposition 1. The treatment of the vaccination policies requires more care, as it is not necessarily true that a Lebesgue integrable  $U$  can be approximated by step functions. However, its integral,  $W$ , can be approximated by the integral of step functions, and this allows the result of Proposition 2 to be generalised to Theorem 1.

## 5.4 Theorem 2

Theorem 2 is proved as follows. Firstly, one can show that, for any vaccination policy  $U$  and  $t \geq 0$ ,

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

using the definition of  $\chi$  in terms of the constraints on  $U$ . This means that the total rate of vaccination given by  $\tilde{U}$  is at least as high as that given by  $U$ .

One can then show that  $\chi(t) \leq A(t)$

$$\int_0^t \chi(s) ds \leq B(t)$$

which means that  $\tilde{U}$  satisfies the vaccination constraints.

From here, one can transform any optimal vaccination policy  $U$  into suitable  $\tilde{U}$ . Initially, the quantities  $\tilde{W}_i(t)$  are constructed. The details of this are left to the appendix but the general principle is that the policy  $U$  is compressed in time so that the total number of vaccinations given out matches  $\min\left(\int_0^t \chi(s) ds, 1\right)$ . It may also be necessary to add additional vaccinations if the overall total differs—these can be assigned in proportion to the number of unvaccinated people in each group.

This construction ensures that the feasibility constraints  $\tilde{W}_i \leq N_i$  are satisfied. Moreover, one can show that  $\tilde{W}_i$  is Lipschitz continuous, which allows for the construction of a derivative  $\tilde{U}_i$  which integrates to  $\tilde{W}_i$ . Finally, one can show that  $\tilde{W}_i(t) \geq W_i(t)$ , meaning that, by Theorem 1,  $\tilde{U}$  must also be an optimal vaccination policy.

### 5.5 Theorem 3

The proof of Theorem 3 then follows from a similar construction to Theorem 2—the only difference is that no additional vaccinations are assigned by  $\tilde{U}$  compared to  $U$ .

## 6 Limitations of Theorem 1

It is helpful to consider the limitations of Theorem 1, as it does not prove that every conceivable cost function is non-increasing in vaccination effort. This will be illustrated through some examples based on theoretical COVID-19 outbreaks in the United Kingdom.

Using the work of Prem et al. (2017), one can split the UK into 16 age-groups (comprising five-year intervals from 0 to 75 and a group for those aged 75+) which mix heterogeneously. The contact matrices estimated in Prem et al. (2017) allow for the construction of a matrix  $\beta^*$ , which will be proportional to each of the matrices  $\beta^\alpha$  in the model.

As illustrated in Liu et al. (2020), estimation of the basic reproduction number  $R_0$  for COVID-19 is complicated, and a wide range of estimates have been produced. For the examples in this paper, a reproduction number of 4 will be used, meaning that  $\beta^1$  will be scaled so that the largest eigenvalue of the matrix given by

$$M_{ij} = \frac{\beta_{ij}^1 N_i}{\mu_i^1}$$

is equal to 4. Note that the population of each group  $N$ —normalised to have total sum 1—is taken from (UN 2019). Moreover, based on the estimates in Ram and Schaposnik (2021), the value of  $\mu_i^1$  and, in the first example,  $\mu_i^2$  will be set equal to  $\frac{1}{14}$ .

To model the effectiveness of vaccination, the estimates of Dean and Halloran (2022) will be used so that  $\beta^2 = 0.77\beta^1$ , (modelling the reduction in infectiousness),  $\beta^3 = 0.3\beta^1$  (modelling the reduction in susceptibility) and  $\beta^4 = 0.77 \times 0.3 \times \beta^1$  (assuming these effects are independent). Finally, the initial conditions used are  $S_i(0) = (1 - 10^{-4})N_i$  and  $I_i(0) = (10^{-4})N_i$  for each  $i$ , modelling a case where 0.01% of the population is initially infected. It should be emphasised however, that this model has purely been made for illustrative purposes and substantially more detailed fitting analysis would be required to use it for forecasting COVID-19 in the UK.

In both the subsequent examples, it will be assumed that 0.5% of the population is vaccinated homogeneously each day in the vaccination case. This will be compared to a case with no vaccination.

## 6.1 Infections Are Not Decreasing For All Time

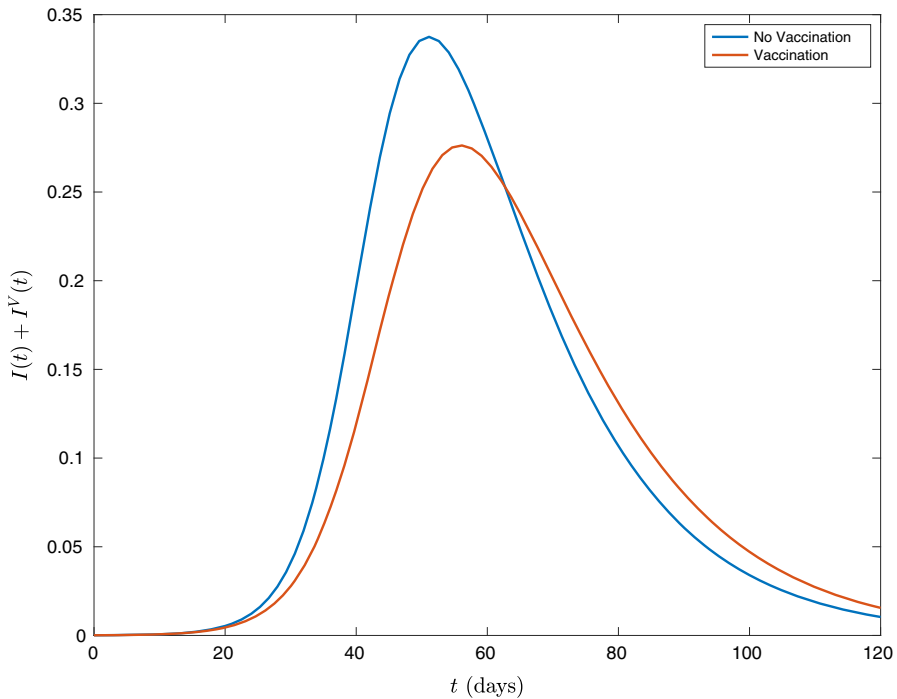
While the overall number of infections will decrease as vaccination effort increases, the infections at a particular point in time will not. Figure 1 shows that the effect of vaccination is both to reduce, but also delay the peak of the infections. This is an important consideration when deciding vaccination policy, as increasing infections at a time in the year when hospitals are under more pressure could have negative consequences, and so it is important not to simply assume that vaccination will reduce all infections at all times.

## 6.2 Deaths Are Not Decreasing For All Time

Perhaps most surprisingly, the total deaths in the epidemic may at some finite times (although not at  $t = \infty$ ) be higher when vaccination occurs, at least under the assumptions of the SIR model. This is a rarer phenomenon, but is possible if vaccination increases the recovery rate as well as decreasing infectiousness.

For illustrative purposes, suppose that vaccination doubles the recovery rate (so that  $\mu_i^2 = \frac{1}{7}$ ) and has no effect on mortality rates. Then, using Bonanad et al. (2020) to get age-dependent mortality rates for COVID-19, Fig. 2 shows that initially, the number of deaths is higher in the case of vaccination. This occurs because the higher value of  $\mu^2$  means that vaccinated people move more quickly to the  $R^V$  compartment than their unvaccinated counterparts and so, while they will infect fewer people, when the number of infections is comparable in the early epidemic, this means that more people will die. Indeed, this property can still hold if vaccination reduced mortality rates (although this reduces the already small difference between the two further—in this example, one needs  $\kappa_i \gtrsim 0.9999$  for deaths to ever be lower in the non-vaccinated case).

Of course, this is not a realistic reflection of the course of an epidemic—the reason for  $\mu^2$  being higher is that vaccinated people are likely to get *less* ill rather than dying more quickly—but it illustrates a potential limitation of the SIR framework. One



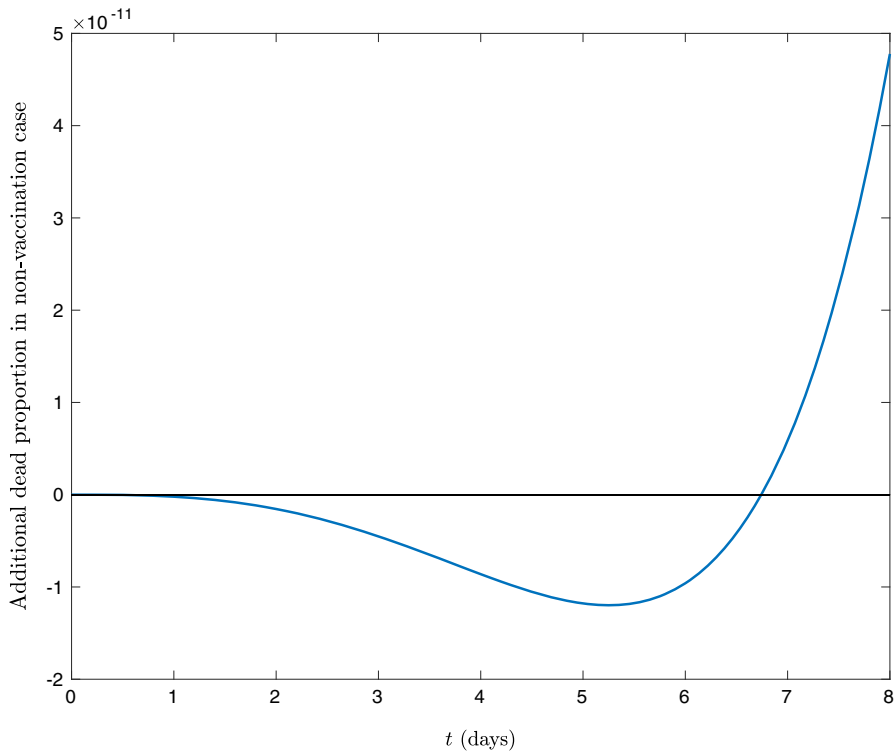
**Fig. 1** A comparison of the total infections over time for a simulated COVID-19 epidemic in the UK, depending on whether a uniform vaccination strategy of constant rate is used

possible way to avoid this problem would be to split the recovered compartment up into the truly recovered and dead subsections. Then, vaccination could increase the speed at which infected members of the population moved to the recovered compartment, but not the speed at which they moved to the dead compartment. This would remove the possibility of seeing the counter-intuitive behaviour of Fig. 2.

## 7 Discussion

It is comforting that the multi-group SIR model does indeed satisfy the condition that the final numbers of infections and deaths are non-increasing in vaccination effort. This shows the importance of ensuring that vaccinations are available as early as possible in a disease outbreak. To achieve this, it is important that good plans for vaccine roll-out and supply chains are available in advance of them being needed to ensure that maximum benefit from the vaccination program is obtained.

For  $n > 1$ , there are, of course, many possible maximal-effort vaccination policies. The results of this paper, in effect, reduce the dimension of the space of possible vaccination policies from  $n$  to  $n - 1$ , as one can assume that an optimal policy satisfies the condition (7) in Theorem 2. However, choosing the correct groups to prioritise is still of crucial importance and can have a substantial impact of the effectiveness of



**Fig. 2** The difference between proportion of the population that has died by each time  $t$  in the case of vaccination and non-vaccination. Positive values indicate that the deaths are higher in the non-vaccination case

the vaccination campaign (Fitzpatrick and Galvani 2021). Applying similarly rigorous techniques to finding the optimal vaccination policy is beyond the scope of this paper, although we extended the results of this paper to apply asymptotic techniques to understand the behaviour of the optimal solution under certain special cases in Penn and Donnelly (2023).

However, there are limitations to these results. Indeed, while the final numbers of infections and deaths are guaranteed to decrease, this is not necessarily true at a given finite time. In particular, vaccination can move the peak of the epidemic, and so it is important to consider the consequences of this, particularly if only a small number of lives are saved by vaccination.

Moreover, while this has not been discussed in this paper, it is also important to emphasise that these results only apply if vaccine efficacy does not decay over time. Indeed, if vaccination efficacy does decay significantly, then vaccinating the most vulnerable groups in a population very early may be worse than vaccinating them later, unless booster jabs are available. If the main epidemic occurs long after the vulnerable have been vaccinated, their immunity may have worn off significantly by the time that the majority of disease exposure occurs. Thus, in this case a more detailed analysis would be needed to determine the optimal vaccination rate.



The authors believe that future models for optimal vaccination should consider using the more general vaccination model introduced in this paper. This allows for greater flexibility in modelling the effect of decreasing demand. Of course, this modified model is slightly more complicated, and care needs to be taken to avoid numerical instabilities arising from the removable singularity in the  $\frac{U_i S_i}{N_i - W_i}$  term when  $W_i \rightarrow N_i$ . However, it has been shown that many of the standard properties of SIR models, and indeed the results of this paper, still hold for this model, and so these extra technical difficulties appear to be a small price to pay for the significantly increased accuracy and potentially large difference between the optimal solutions for the two models.

The results of this paper could be extended to cover a wider range of disease models that are currently being used in the literature. In particular, the next step could be to prove the results for SEIR (Susceptible-Exposed-Infected-Recovered) models, and indeed models with multiple exposed compartments for each subgroup. This would help to build a general mathematical theory of maximal-effort vaccination that would provide evidence for the reliability of contemporary epidemiological modelling.

## 8 Conclusion

The results of this paper are summarised below:

- Vaccinating at maximal effort is optimal for a multi-group SIR model with non-decaying vaccination efficacy.
- The general vaccination model introduced in this paper provides greater flexibility in modelling the effect of decreasing vaccination uptake.
- While vaccinating at maximal effort gives optimality, there can be finite times at which, according to the SIR model, infections or deaths are higher if vaccination has occurred.

**Acknowledgements** The authors would like to thank Cameron Simmons, Joseph Penn and Grace Penn for their invaluable proof-reading work.

**Funding** This work was funded by a DTP Studentship awarded to Matthew Penn by the Engineering and 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** As cited in the text, the data used to create Fig. 1 is available from <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005697> and <https://population.un.org/wpp/Download/Standard/Population/>. The additional data used to create Fig. 2 is available from <https://www.sciencedirect.com/science/article/pii/S1525861020304412>.

## 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/>.

## A Proofs of Theorems 1, 2 and 3

Before beginning the main proof, it is helpful to note some fundamental results about the SIR equations that will be used throughout. Namely, for each  $i \in \{1, \dots, n\}$  and  $t \geq 0$

$$0 \leq S_i(t), I_i(t), R_i(t), S_i^V(t), I_i^V(t), R_i^V(t) \leq N_i$$

and

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

These results are proved in Lemmas C.3 and C.4.

It is first useful to define

$$K_{ij}(t) = \frac{\beta_{ij}^1}{\mu_j^1} R_j(t) + \frac{\beta_{ij}^2}{\mu_j^2} R_j^V(t)$$

and

$$L_{ij}(t) := \frac{\beta_{ij}^3}{\mu_i^1} R_j(t) + \frac{\beta_{ij}^4}{\mu_i^2} R_j^V(t).$$

Then, the following propositions hold.

### A.1 An Inequality for $K_{ij}$ and $L_{ij}$

Note that the proof of this proposition requires a significant amount of algebra, and the majority of it has hence been left to lemmas which can be found in Appendix B. However, the key logic of the proof will be presented here.

Also, note that in this paper, a step function is defined to be a function that is piecewise constant on any *bounded* interval of  $\mathfrak{R}$ . Thus, it may have infinitely many discontinuities, but only finitely many in any bounded interval. This differs from the definition used in some other papers (which impose that a step function is piecewise constant on  $\mathfrak{R}$ ).

**Proposition A.1.1** *Suppose that  $U_i(t)$  and  $\tilde{U}_i(t)$  are right-continuous step functions. Moreover, suppose that*

$$\beta_{ij}^1 > \beta_{ij}^3 > 0 \quad \forall i, j \in \{1, \dots, n\},$$

$${}_i(0)I_i(0) > 0 \quad \forall i \in \{1, \dots, n\}.$$

and that

$$W_i(t) < N_i \quad \forall t \geq 0 \quad \text{and} \quad \forall i \in \{1, \dots, n\}$$

Then,

$$K_{ij}(t) \geq \tilde{K}_{ij}(t) \quad \text{and} \quad L_{ij}(t) \geq \tilde{L}_{ij}(t) \quad \forall t \geq 0.$$

**Proof** Suppose that the proposition does not hold. Hence, one can define

$$T := \inf \left\{ t : K_{ij}(t) < \tilde{K}_{ij}(t) \quad \text{or} \quad L_{ij}(t) < \tilde{L}_{ij}(t) \quad \text{for some } i, j \in \{1, \dots, n\} \right\}.$$

Then, there exists some  $b \in \{1, \dots, n\}$  and some real constants  $\kappa$  and  $\eta$  such that the following system of inequalities holds at time  $T$ :

$$S_b(T) + S_b^V(T) \leq \tilde{S}_b(T) + \tilde{S}_b^V(T), \quad (9)$$

$$I_b(T) + R_b(T) \geq \tilde{I}_b(T) + \tilde{R}_b(T) \quad (10)$$

$$R_b(T) \geq \tilde{R}_b(T), \quad (11)$$

$$R_b(T) + \kappa R_b^V(T) \leq \tilde{R}_b(T) + \kappa \tilde{R}_b^V(T), \quad (12)$$

$$I_b(T) + \eta I_b^V(T) \leq \tilde{I}_b(T) + \eta \tilde{I}_b^V(T), \quad (13)$$

$$0 \leq \kappa \leq \eta \leq 1. \quad (14)$$

The derivations of inequalities (9) - (14) are found in Lemmas B.2–B.5. Moreover,

$$\begin{aligned} S_b(T) + I_b(T) + R_b(T) + S_b^V(T) + I_b^V(T) + R_b^V(T) \\ = \tilde{S}_b(T) + \tilde{I}_b(T) + \tilde{R}_b(T) + \tilde{S}_b^V(T) + \tilde{I}_b^V(T) + \tilde{R}_b^V(T), \end{aligned} \quad (15)$$

which comes from (3). Note that (12) in fact holds to equality in this case, but this is not necessary for the proof (and later, the same system will be considered where such an equality is not guaranteed).

By Lemma B.6, the system (9)–(15) implies that

$$I_b(T) + R_b(T) = \tilde{I}_b(T) + \tilde{R}_b(T), \quad (16)$$

$$I_b^V(T) + R_b^V(T) = \tilde{I}_b^V(T) + \tilde{R}_b^V(T), \quad (17)$$

$$S_b(T) + S_b^V(T) = \tilde{S}_b(T) + \tilde{S}_b^V(T), \quad (18)$$

If  $T > 0$ , then Lemma B.7 can be used to show that

$$W_k(t) = \tilde{W}_k(t) \quad \forall t \in [0, T] \quad \text{and} \quad \forall k \in \{1, \dots, n\}$$

while if  $T = 0$  then this is immediate. Thus, the two ODE systems are the same up to time  $T$ , which means that all variables (in all groups) are equal at time  $T$ .

From this point, the proof of Proposition A.1.1 can be completed by considering the behaviour of the system at time  $T + \delta$  for small  $\delta$ . For sufficiently small  $\delta$ ,  $U_i(t)$  and  $\tilde{U}_i(t)$  are constant on  $[T, T + \delta]$  (as they are step functions) and this condition on  $\delta$  will be assumed for the remainder of this proof

Define functions  $\Delta_i^f$  to be

$$\Delta_i^f(t) := f_i(T + t) - \tilde{f}_i(T + t) \quad \text{for } f \in \{S, I, R, S^V, I^V, R^V, W\}$$

and note that

$$\Delta_i^f(0) = 0 \quad \forall f \in \{S, I, R, S^V, I^V, R^V, W\}.$$

Then, by Lemma B.8, for  $t \in [0, \delta]$  and any real numbers  $x$  and  $y$

$$\frac{x}{\mu_i^1} \Delta_i^R + \frac{y}{\mu_i^2} \Delta_i^{R^V} = \frac{t^3 S_i(T)(U_i(T) - \tilde{U}_i(T))}{6(N_i - W_i(T))} \left[ x \sum_{j=1}^n (K'_{ij}(T)) - y \sum_{j=1}^n (L'_{ij}(T)) \right] + O(\delta^4).$$

Hence, by Lemma B.9,

$$\sum_{j=1}^n K_{ij}(t) \geq \sum_{j=1}^n \tilde{K}_{ij}(t) \quad \forall t \in [0, T + \delta] \quad (19)$$

and

$$\sum_{j=1}^n L_{ij}(t) \geq \sum_{j=1}^n \tilde{L}_{ij}(t) \quad \forall t \in [0, T + \delta] \quad (20)$$

for sufficiently small  $\delta$ .

Now, by the definition of  $T$ , there exists some  $t$  in  $[T, T + \delta]$  such that, for some  $a, b$

$$K_{ab}(t) < \tilde{K}_{ab}(t) \quad \text{or} \quad L_{ab}(t) < \tilde{L}_{ab}(t).$$

Indeed, from Lemma B.10, there exists some  $t \in (T, T + \delta)$  such that

$$R_b(t) + \kappa R_b^V(t) < \tilde{R}_b(t) + \kappa \tilde{R}_b^V(t) \quad \text{and} \quad I_b(t) + \eta I_b^V(t) \leq \tilde{I}_b(t) + \eta \tilde{I}_b^V(t) \quad (21)$$

for some

$$0 \leq \kappa \leq \eta \leq 1. \quad (22)$$

Now, by Lemmas B.2–B.4 (which only require the properties (19) and (20)), the system of inequalities (9)–(11) holds for group  $b$  at time  $t$ . These can be combined with (21), (22) and (15) to use Lemma B.6, showing

$$\eta I_b^V(t) + \kappa R_b^V(t) = \eta \tilde{I}_b^V(t) + \kappa \tilde{R}_b^V(t) \quad (23)$$

$$I_b(t) + R_b(t) = \tilde{I}_b(t) + \tilde{R}_b(t). \quad (24)$$

By adding the inequalities in (21) together,

$$R_b(t) + \kappa R_b^V(t) + I_b(t) + \eta I_b^V(t) < \tilde{R}_b(t) + \kappa \tilde{R}_b^V(t) + \tilde{I}_b(t) + \eta \tilde{I}_b^V(t).$$

Then, (23) and (24) show that this must in fact be an equality which is a contradiction. Thus,  $t$  cannot exist. This provides a contradiction to the definition of  $T$ , and hence finishes the proof of Proposition A.1.1.

It is now possible to prove Theorem 1 under the extra restrictions given Proposition A.1.1.  $\square$

## A.2 A Proof for a Restricted Parameter and Policy Set

**Proposition A.2.1** *Under the conditions of Proposition A.1.1, for any  $t \geq 0$  and  $i \in \{1, \dots, n\}$*

$$I_i(t) + I_i^V(t) + R_i(t) + R_i^V(t) \geq \tilde{I}_i(t) + \tilde{I}_i^V(t) + \tilde{R}_i(t) + \tilde{R}_i^V(t)$$

and

$$R_i(t) \geq \tilde{R}_i(t).$$

Moreover, for any  $\lambda \in [0, 1]$

$$R_i(\infty) + \lambda R_i^V(\infty) \geq \tilde{R}_i(\infty) + \lambda \tilde{R}_i^V(\infty)$$

and hence, the objective function is lower for  $\tilde{U}$ , provided the conditions of Proposition A.1.1 are met.

**Proof** Note that, by Proposition A.1.1,

$$K_{ij}(t) \geq \tilde{K}_{ij}(t) \quad \text{and} \quad L_{ij}(t) \geq \tilde{L}_{ij}(t) \quad \forall t \geq 0$$

and hence, by Lemma B.2, for each  $i \in \{1, \dots, n\}$

$$S_i(t) + S_i^V(t) \leq \tilde{S}_i(t) + \tilde{S}_i^V(t).$$

Combining this with the conservation of population Eq. (15), shows that

$$I_i(t) + I_i^V(t) + R_i(t) + R_i^V(t) \geq \tilde{I}_i(t) + \tilde{I}_i^V(t) + \tilde{R}_i(t) + \tilde{R}_i^V(t)$$

as required. Now, taking  $t \rightarrow \infty$  and noting that the infections tend to zero by Lemma C.4 gives

$$R_i(\infty) + R_i^V(\infty) \geq \tilde{R}_i(\infty) + \tilde{R}_i^V(\infty).$$

Moreover, by Lemma B.4, for any  $t \geq 0$  and  $i \in \{1, \dots, n\}$

$$R_i(t) \geq \tilde{R}_i(t)$$

as required. Also, taking  $t \rightarrow \infty$  shows that

$$R_i(\infty) \geq \tilde{R}_i(\infty).$$

Thus, for any  $\lambda \in [0, 1]$

$$\begin{aligned} R_i(\infty) + \lambda R_i^V(\infty) &= (1 - \lambda)R_i(\infty) + \lambda(R_i(\infty) + R_i^V(\infty)) \\ &\geq (1 - \lambda)\tilde{R}_i(\infty) + \lambda(\tilde{R}_i(\infty) + \tilde{R}_i^V(\infty)) \\ &= \tilde{R}_i(\infty) + \lambda\tilde{R}_i^V(\infty) \end{aligned}$$

as required.  $\square$

By summing the  $i$  inequalities at  $t = \infty$  from Proposition A.2.1 (and using  $\lambda = \kappa_i$ ), Theorem 1 holds under the additional conditions given in Proposition A.1.1. Note that the closure of the set of parameters, initial conditions and vaccination policies which satisfy these conditions is the original set specified in Theorem 1. Thus, one can generalise the result with the help of the following proposition.

### A.3 Continuous Dependence

**Proposition A.3.1** *Define the set of functions*

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

where for each fixed  $\epsilon$ , these functions solve the model equations with parameters

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

initial conditions

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

and vaccination policy  $U(t; \epsilon)$ . Suppose that

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

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

and that

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

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

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

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

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

**Proof** The proof is simple but algebraically dense and so is left to Lemma C.8 in the appendices.

This now allows a proof of Theorem 1 to be formed.  $\square$

#### A.4 Theorem 1

**Theorem 1** Suppose that  $U, \tilde{U} \in C$ . Suppose further that for each  $i \in \{1, \dots, n\}$  and  $t \geq 0$

$$\int_0^t U_i(s) ds \leq \int_0^t \tilde{U}_i(s) ds.$$

Then

$$H(U) \geq H(\tilde{U}).$$

**Proof** Define the parameters  $\beta_{ij}^a(\epsilon)$  and  $\mu_i^a(\epsilon)$  by

$$\beta_{ij}^a(\epsilon) = \beta_{ij}^a + \frac{\epsilon}{a} \quad \text{and} \quad \mu_i^a(\epsilon) = \mu_i^a.$$

This means that, for any  $\epsilon > 0$ , these parameters satisfy the conditions of Propositions A.1.1 and A.2.1. Define, for  $\epsilon < 1$ , the initial conditions

$$S_i(0; \epsilon) = \begin{cases} S_i(0; 0) & \text{if } S_i(0; 0), I_i(0; 0) > 0 \\ S_i(0; 0) + \epsilon N_i & \text{if } S_i(0; 0) = 0 \\ S_i(0; 0) - \epsilon N_i & \text{if } I_i(0; 0) = 0 \end{cases}$$

and

$$I_i(0; \epsilon) = N_i - S_i(0; \epsilon).$$

Then, the conditions of Propositions A.1.1 and A.2.1 are met by these initial conditions for any  $\epsilon > 0$ .

Now, define the set of points

$$\sigma(\epsilon) := \left\{ n\epsilon : n \in \mathcal{N}_{\geq 0} \right\}.$$

Then, define  $W_i^*(t; \epsilon)$  to be the first order approximation to the function  $\mathcal{W}_i(t; \epsilon) := \max(W_i(t), N_i - \epsilon)$  using the points of  $\sigma(\epsilon)$ . That is, for each  $t$  define

$$K(t; \epsilon) := \inf \left\{ m : m \in \sigma(\epsilon) \text{ and } m \geq t \right\}$$

and

$$k(t; \epsilon) := \sup \left\{ m : m \in \sigma(\epsilon) \text{ and } m \leq t \right\}$$

Note that, as  $\sigma(\epsilon)$  is nowhere dense, one must have

$$k(t; \epsilon), K(t; \epsilon) \in \sigma(\epsilon) \text{ and } k(t; \epsilon) \leq t \leq K(t; \epsilon)$$

Then, define

$$W_i^*(t; \epsilon) = (t - k(t; \epsilon))\mathcal{W}_i(k(t; \epsilon); \epsilon) + (K(t; \epsilon) - t)\mathcal{W}_i(K(t; \epsilon); \epsilon).$$

Thus, as  $k$  and  $K$  are constant on any interval not containing a point in  $\sigma(\epsilon)$ ,  $W_i^*$  is linear on any interval not containing a point of  $\sigma(\epsilon)$  and so its derivative is a step function.  $\square$

Now, note that, for each  $t$

$$|\mathcal{W}_i(t; \epsilon) - W_i(t)| \leq \epsilon$$

and, moreover,

$$t \in S \Rightarrow W_i^*(t; \epsilon) = \mathcal{W}_i(t; \epsilon).$$

Also, as  $U_i$  is bounded, each  $W_i$  (and hence each  $\mathcal{W}_i$ ) are Lipschitz continuous with some Lipschitz constant  $L$ . Moreover, each  $W_i^*$  is continuous and is differentiable in each interval  $(k(t; \epsilon), K(t; \epsilon))$  with a maximal (uniformly bounded) gradient of  $U_i(t)$ , meaning that  $W_i^*$  is also Lipschitz continuous with Lipschitz constant  $L$ .

It can now be shown that  $|W_i(t) - W_i^*(t; \epsilon)|$  is uniformly bounded in  $t$ . For each  $t \geq 0$ , one can find an element  $s \in \sigma(\epsilon)$  such that  $|t - s| < \epsilon$ . Then,

$$|W_i(t) - W_i^*(t; \epsilon)| \leq |W_i(t) - W_i(s)| + |W_i(s) - W_i^*(s; \epsilon)| + |W_i^*(s; \epsilon) - W_i^*(t; \epsilon)|$$



$$\begin{aligned} &\leq L\epsilon + |W_i(s) - \mathcal{W}_i(s; \epsilon)| + L\epsilon \\ &\leq (2L + 1)\epsilon \end{aligned}$$

and so  $W_i^*$  converges uniformly to  $W_i$ . The same results hold for the analogously defined  $\tilde{W}_i^*$ . Then, note that, as  $\tilde{W}_i(t) \geq W_i(t)$ , it must be that  $\tilde{\mathcal{W}}_i(t; \epsilon) \geq \mathcal{W}_i(t; \epsilon)$ . Thus, it follows that  $\tilde{W}_i^*(t; \epsilon) \geq W_i^*(t; \epsilon)$ .

This means that Proposition A.2.1 can be used. Define using stars the variables that come from the  $U^*$  and  $\tilde{U}^*$  policies. Then, from Proposition 2, for each  $t \geq 0$ ,  $\epsilon > 0$  and  $i \in \{1, \dots, n\}$

$$\begin{aligned} I_i^*(t; \epsilon) + I_i^{V*}(t; \epsilon) + R_i^*(t; \epsilon) + R_i^{V*}(t; \epsilon) &\geq \tilde{I}_i^*(t; \epsilon) \\ + \tilde{I}_i^{V*}(t; \epsilon) + \tilde{R}_i^*(t; \epsilon) + \tilde{R}_i^{V*}(t; \epsilon) \end{aligned}$$

and

$$R_i^*(t; \epsilon) \geq \tilde{R}_i^*(t; \epsilon).$$

Then, taking  $\epsilon \rightarrow 0$  and using Proposition A.3.1 (noting that the perturbations to the parameters, initial conditions and vaccination policies are all bounded by a constant multiple of  $\epsilon$ ) shows that

$$I_i(t) + I_i^V(t) + R_i(t) + R_i^V(t) \geq \tilde{I}_i(t) + \tilde{I}_i^V(t) + \tilde{R}_i(t) + \tilde{R}_i^V(t)$$

and

$$R_i(t) \geq R_i^V(t).$$

Then, the result follows using the same logic as in the proof of Proposition A.2.1.

## A.5 Theorem 2

**Theorem 2** Suppose that  $B$  is differentiable, and that there is an optimal solution  $U$  to (6). Then, define the function

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

and suppose that  $\chi(t)$  exists and is bounded. Then, there exists an optimal solution  $\tilde{U}$  to the problem (6) such that

$$\sum_{i=1}^n \tilde{W}_i(t) = \min \left( \int_0^t \chi(s)ds, 1 \right).$$

Moreover, if  $\chi(t)$  is continuous almost everywhere, there exists an optimal solution  $\tilde{U}$  such that

$$\sum_{i=1}^n \tilde{U}_i(t) = \begin{cases} \chi(t) & \text{if } \int_0^t \chi(s) ds < 1 \\ 0 & \text{otherwise} \end{cases}$$

**Proof** Suppose that  $U$  is an optimal vaccination policy. To begin, it will be shown that the total vaccination rate  $\chi$  is indeed a maximal-effort vaccination policy (in the sense that, at each time  $t^*$ , it is impossible to have given out more vaccines than a policy with total overall rate  $\chi(t)$ ).  $\square$

**Claim:**  $\min \left( 1, \int_0^t \chi(s) ds \right) \geq \int_0^t \sum_{i=1}^n U_i(s) ds$  for all  $t > 0$

**Proof** Consider any time  $t \geq 0$  such that

$$\int_0^t \chi(s) ds < 1$$

and define the set

$$\mathcal{T} := \left\{ s \leq t : \int_0^s \chi(k) dk \geq B(s) \right\}.$$

Suppose that  $\mathcal{T} = \emptyset$ . Then,

$$\chi(s) = A(s) \quad \forall s \leq t$$

and so

$$\int_0^t \chi(s) ds = \int_0^t A(s) ds \geq \int_0^t \sum_{i=1}^n U_i(s) ds.$$

Moreover, suppose that  $\mathcal{T} \neq \emptyset$  and define

$$\tau := \sup(\mathcal{T}).$$

Then,

$$\int_0^\tau \chi(s) ds \geq B(\tau) \geq \int_0^\tau \sum_{i=1}^n U_i(s) ds$$

and

$$\int_\tau^t \chi(s) ds = \int_\tau^t A(s) ds \geq \int_\tau^t \sum_{i=1}^n U_i(s) ds$$

so that

$$\int_0^t \chi(s) ds \geq \int_0^t \sum_{i=1}^n U_i(s) ds.$$

Thus, this holds in all cases for  $\int_0^t \chi(s) ds < 1$ . Finally, suppose that

$$\int_0^t \chi(s) ds \geq 1.$$

Then, one has

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

and so the claim is proved.  $\square$

It is now important to show that  $\chi$  gives a feasible vaccination rate. Note that  $\chi(t) \leq A(t)$  by definition.

**Claim:**  $\int_0^t \chi(s) ds \leq B(t)$  for all  $t \geq 0$ .

**Proof** Suppose, for a contradiction, that there exists a  $t$  such that

$$\int_0^t \chi(s) ds > B(t).$$

Then, define

$$\sigma := \sup \left\{ s \leq t : \int_0^s \chi(s) ds \leq B(t) \right\}$$

which must exist (as  $\int_0^0 \chi(s) ds \leq B(0)$ ) and satisfy  $\sigma < t$ , by continuity of  $\int_0^t \chi(s) ds$  and  $B(t)$ . Note that

$$s \in (\sigma, t) \Rightarrow \chi(s) \leq B'(s)$$

and so

$$\int_0^t \chi(s) ds \leq \int_0^\sigma \chi(s) ds + \int_\sigma^t B'(s) ds \leq B(\sigma) + (B(t) - B(\sigma)) = B(t),$$

which is a contradiction. Thus,

$$\int_0^t \chi(s) ds \leq B(t) \quad \forall t \geq 0$$

as required.  $\square$

Now, one can create a new optimal vaccination policy with total rate given by  $\chi$ . Define

$$q(t) = \begin{cases} \inf \left\{ s : \int_0^s \sum_{j=1}^n U_j(k) dk = \int_0^t \chi(k) dk \right\} & \text{if this exists} \\ \infty & \text{otherwise} \end{cases}$$

so that  $q(t)$  represents the earliest time at which  $\chi(t)$  vaccines were administered by the  $U$  policy. By continuity of the integral, this means that

$$\sum_{i=1}^n W_i(q(t)) = \int_0^{q(t)} \sum_{j=1}^n U_j(k) dk = \int_0^t \chi(s) ds.$$

Define further

$$Q := \sup\{t : q(t) < \infty\} \quad \text{and} \quad q_\infty := \lim_{t \rightarrow Q} (q(t))$$

so that  $Q$  is the earliest time at which all of the vaccines given out by the  $U$  policy could have been administered. Note that both  $Q$  and  $q_\infty$  may be infinite. By taking the limit  $t \rightarrow Q$ , and noting the left-hand side is bounded by 1,

$$\int_0^{q_\infty} \sum_{j=1}^n U_j(k) dk = \int_0^Q \chi(k) dk$$

Then, the integral of the new vaccination policy,  $\tilde{W}$  is given by

$$\tilde{W}_i(t) = \begin{cases} W_i(q(t)) & \text{if } t < Q \\ W_i(q_\infty) + \frac{(N_i - W_i(q_\infty)) \int_Q^t \chi(s) ds}{1 - \sum_{i=1}^n W_i(q_\infty)} & \text{if } \int_0^t \chi(s) ds < 1 \quad \text{and} \quad t \geq Q \\ N_i & \text{if } \int_0^t \chi(s) ds \geq 1 \quad \text{and} \quad t \geq Q \end{cases}$$

This is well-defined as

$$\sum_{i=1}^n W_i(q_\infty) = 1 \Rightarrow \int_0^Q \chi(s) ds = 1$$

and so, in this case, the second part of the definition of  $\chi$  is never used. It is important to establish for feasibility that each  $W_i$  is bounded by  $N_i$ .

**Claim:**  $\tilde{W}_i(t) \leq N_i$  for all  $t \geq 0$  and all  $i \in \{1, \dots, n\}$ .

**Proof** If  $t < Q$ , then  $W_i(q(t)) \leq N_i$  for all  $t < Q$  by feasibility of  $U$ . Otherwise, if  $t \geq Q$  and  $\int_0^t \chi(s)ds < 1$ , then one has

$$\begin{aligned} W_i(q_\infty) + \frac{(N_i - W_i(q_\infty)) \int_Q^t \chi(s)ds}{1 - \sum_{i=1}^n W_i(q_\infty)} &\leq W_i(q_\infty) + \frac{(N_i - W_i(q_\infty))(1 - \int_0^Q \chi(s)ds)}{1 - \sum_{i=1}^n W_i(q_\infty)} \\ &= W_i(q_\infty) + \frac{(N_i - W_i(q_\infty))(1 - \sum_{i=1}^n W_i(q_\infty))}{1 - \sum_{i=1}^n W_i(Q)} \\ &= N_i \end{aligned}$$

while if  $\int_0^t \chi(s)ds \geq 1$  then the result is immediate.

The optimisation problem is framed in terms of  $U$  rather than  $W$ , and so it is important to show that there is some  $\tilde{U}$  that integrates to  $\tilde{W}$ . One can do this by proving the Lipschitz continuity of  $\tilde{W}_i$  for each  $i$ .  $\square$

**Claim:**  $\tilde{W}_i(t)$  is Lipschitz continuous for each  $i \in \{1, \dots, n\}$

**Proof** Note that for  $s, t < Q$ , if  $M$  is a bound for  $\chi$  (which is assumed to exist)

$$\begin{aligned} |\tilde{W}_i(t) - \tilde{W}_i(s)| &= \left| \int_{q(s)}^{q(t)} U_i(k)dk \right| \\ &\leq \left| \int_{q(s)}^{q(t)} \sum_{j=1}^n U_j(k)dk \right| \\ &= \left| \int_s^t \chi(k)dk \right| \\ &\leq |t - s|M \end{aligned}$$

Moreover, if  $s, t > Q$  and  $\int_0^t \chi(k)dk, \int_0^s \chi(k)dk < 1$ , then

$$|\tilde{W}_i(t) - \tilde{W}_i(s)| \leq \left| \frac{(N_i - W_i(q_\infty)) \int_s^t \chi(s)ds}{1 - \sum_{i=1}^n W_i(q_\infty)} \right| \leq M \left| \frac{(N_i - W_i(q_\infty))}{1 - \sum_{i=1}^n W_i(q_\infty)} \right| |t - s|$$

and if  $s, t > Q$  and  $\int_0^t \chi(k)dk, \int_0^s \chi(k)dk \geq 1$ , then  $\tilde{W}_i(t) = \tilde{W}_i(s)$ . The intermediate cases (where  $s$  and  $t$  correspond to different cases in the definition of  $\chi$ ) can be proved by combining these bounds.

This means that (for each  $i$ ) there exists a Lebesgue integrable function  $\tilde{U}_i(t)$  such that

$$\frac{d\tilde{W}_i}{dt} = \tilde{U}_i(t) \quad \text{almost everywhere}$$

and, for all  $t \geq 0$

$$\int_0^t \tilde{U}_i(s)ds = \tilde{W}_i(t)$$

A proof of this (for the broader class of absolutely continuous functions) can be found in Bárcenas (2000). One can set  $\tilde{U}_i(t)$  to be zero for any  $t$  such that  $\tilde{W}_i(t)$  is not differentiable. Thus, noting that, where it is differentiable, the derivative of  $\tilde{W}_i$  is bounded by its Lipschitz constant,  $\tilde{U}_i(t)$  is bounded as required.

Note that, in all cases (as  $\sum_{i=1}^n N_i = 1$ )

$$\sum_{i=1}^n \tilde{W}_i(t) = \min \left( \int_0^t \chi(s) ds, 1 \right)$$

and so  $\tilde{W}$  does correspond to a maximal vaccination rate. If  $\chi(t)$  is continuous almost everywhere, then one can differentiate this relationship at  $t$  where each  $\tilde{W}_i$  is differentiable and  $\chi$  is continuous to show that  $\sum_{i=1}^n U_i(t) = \chi(t)$ . The complement of this set must have zero measure (as it is the finite union of zero measure sets), and so, in this case, one can change the values of each  $U_i(t)$  so that  $\sum_{i=1}^n U_i(t) = \chi(t)$  everywhere without changing the value of  $W$ .  $\square$

**Claim:**  $\tilde{W}_i(t) \geq W_i(t)$  for all  $i \in \{1, \dots, n\}$  and  $t \geq 0$

**Proof** Note that, by maximality of  $\chi$ , for  $t < Q$ ,

$$\sum_{j=1}^n \tilde{W}_j(t) = \sum_{j=1}^n W_j(q(t)) = \int_0^t \chi(s) ds \geq \sum_{j=1}^n W_j(t)$$

If  $q(t) \geq t$ , then  $W_i(q(t)) \geq W_i(t)$  for each  $i$ . If  $q(t) < t$ , then it is necessary that  $W_i(q(t)) = W_i(t)$  for each  $i$  as  $W_i$  is non-decreasing. Thus,  $W_i(q(t)) \geq W_i(t)$  for all  $i$  and for all  $t < Q$ .

If  $t > Q$  and  $\int_0^t \chi(s) ds < 1$ , then

$$\tilde{W}_i(t) \geq W_i(q_\infty) \quad (25)$$

Now, by definition of  $Q$ , it is necessary that

$$\int_0^t \chi(k) dk \geq \int_0^\infty \sum_{j=1}^n U_j(k) dk \quad \forall t > Q$$

as otherwise, there must exist some  $t > Q$  and some  $s < \infty$  such that

$$\int_0^t \chi(k) dk = \int_0^s \sum_{j=1}^n U_j(k) dk$$

which means that  $q(t) < \infty$ . Thus, by continuity, for all  $\tau \in (0, t)$ , there exists some  $s$  such that

$$\int_0^\tau \chi(k) dk = \int_0^s \sum_{j=1}^n U_j(k) dk$$

which means  $Q \geq t$ , which is a contradiction.

Thus, by taking  $t \rightarrow Q$ ,

$$\int_0^{q_\infty} \sum_{i=1}^n U_i(k) dk = \int_0^Q \chi(k) dk \geq \int_0^\infty \sum_{j=1}^n U_j(k) dk$$

and so

$$\int_{q_\infty}^\infty U_j(k) dk = 0 \Rightarrow W_i(t) = W_i(q_\infty) \quad \forall i \in \{1, \dots, n\} \quad \text{and} \quad \forall t \geq q_\infty$$

Thus, using (25),

$$\tilde{W}_i(t) \geq W_i(t).$$

Finally, if  $t > Q$  and  $\int_0^t \chi(s) ds \geq 1$ , then  $\tilde{W}_i(t) = N_i \geq W_i(t)$ . Thus, for all  $t$  and  $i$ ,

$$\tilde{W}_i(t) \geq W_i(t)$$

as required.

Thus, by Theorem 1, it is necessary that

$$H(U) \geq H(\tilde{U})$$

and hence, by the optimality of  $U$ ,  $\tilde{U}$  is optimal as required.  $\square$

## A.6 Theorem 3

**Theorem 3** *Under the assumptions of Theorem 2, consider a modified objective function  $\mathcal{H}$  given by*

$$\mathcal{H}(U) = H(U) + F(W(\infty))$$

*for any function  $F$ . Then, with  $\chi$  defined to be the maximal vaccination effort as in Theorem 2, there exists an optimal solution  $\tilde{U}$  such that, for some  $\tau \geq 0$*

$$\sum_{i=1}^n \tilde{W}_i(t) = \begin{cases} \int_0^t \chi(s) ds & \text{if } t \leq \tau \\ W_i(\tau) & \text{otherwise} \end{cases}.$$

*Moreover, if  $\chi$  is continuous almost everywhere, then there is an optimal solution  $\tilde{U}$  such that*

$$\sum_{i=1}^n U_i(t) = \begin{cases} \chi(t) & \text{if } t \leq \tau \\ 0 & \text{otherwise} \end{cases}.$$

**Proof** This follows directly from the proof of Theorem 2. One can again define  $\tilde{U}$  in the interval  $(0, Q)$  (where  $Q$  is defined in the proof of Theorem 2) such that

$$H(U) \geq H(\tilde{U}) \quad \text{and} \quad \int_0^t \sum_{i=1}^n \tilde{U}_i(s) ds = \int_0^t \chi(s) ds \quad \forall t < Q$$

with the only difference being that now

$$\tilde{U}_i(t) = 0 \quad \forall t \geq Q.$$

Thus, as shown in the proof of Theorem 2,

$$W(\infty) = W(q_\infty) = \tilde{W}(Q) = \tilde{W}(\infty)$$

and so

$$\mathcal{H}(U) \geq \mathcal{H}(\tilde{U}),$$

which means  $\tilde{U}$  is optimal as required.  $\square$

## B Supplementary Lemmas For Propositions A.1.1 and A.2.1 and Theorem 2

For the proofs of these lemmas, it is helpful to recall the following definitions of the following variables, which will be extensively used.

$$K_{ij}(t) = \frac{\beta_{ij}^1}{\mu_j^1} R_j + \frac{\beta_{ij}^2}{\mu_j^2} R_j^V,$$

$$L_{ij}(t) := \frac{\beta_{ij}^3}{\mu_i^1} R_j + \frac{\beta_{ij}^4}{\mu_i^2} R_j^V$$

and

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

Moreover, note that, under the assumptions of Proposition A.1.1 and A.2.1, each  $U_i(t)$  is a step function and is therefore piecewise smooth in each bounded interval. Thus, in particular, the derivatives of each of the model variables (and indeed, the derivative of  $W_i(t)$ ) are piecewise continuous in each bounded interval, meaning that each of the model variables is piecewise continuously differentiable in each bounded interval. This means that integration by parts can be performed (in a bounded interval), as will be done extensively throughout the proofs of these lemmas.



**B.1 Lemma B.1**

**Lemma B.1** Suppose that  $f(t)$  is a non-increasing, non-negative, continuous and piecewise continuously differentiable function and that the continuous and piecewise continuously differentiable functions  $g(t)$  and  $h(t)$  satisfy  $g(0) = h(0)$  and  $g(t) \leq h(t)$  for all  $t \geq 0$ . Then,

$$\int_0^t g'(s) f(s) ds \leq \int_0^t h'(s) f(s) ds.$$

**Proof** This follows from integrating by parts:

$$\begin{aligned} \int_0^t g'(s) f(s) ds &= g(t) f(t) - g(0) f(0) - \int_0^t g(s) f'(s) ds \\ &= g(t) f(t) - h(0) f(0) + \int_0^t g(s) |f'(s)| ds \\ &\leq h(t) f(t) - h(0) f(0) + \int_0^t h(s) |f'(s)| ds \\ &\leq h(t) f(t) - h(0) f(0) - \int_0^t h(s) f'(s) ds \\ &= \int_0^t h'(s) f(s) ds \end{aligned}$$

as required. □

**B.2 Lemma B.2**

**Lemma B.2** Suppose that

$$\begin{aligned} \sum_{j=1}^n K_{ij}(t) &\geq \sum_{j=1}^n \tilde{K}_{ij}(t) \quad \text{and} \quad \sum_{j=1}^n L_{ij}(t) \\ &\geq \sum_{j=1}^n \tilde{L}_{ij}(t) \quad \forall i \in \{1, \dots, n\} \quad \text{and} \quad t \in [0, T]. \end{aligned}$$

Then,

$$S_i(t) + S_i^V(t) \leq \tilde{S}_i(t) + \tilde{S}_i^V(t) \quad \forall t \in [0, T].$$

**Proof** To reduce notation in this proof, define

$$\mathcal{K}(t) := \sum_{j=1}^n K_{ij}(t) \quad \text{and} \quad \mathcal{L}(t) := \sum_{j=1}^n L_{ij}(t)$$

Note that

$$\begin{aligned} \frac{d}{dt}(S_i + S_i^V) &= - \sum_{j=1}^n \left( \beta_{ij}^1 I_j + \beta_{ij}^2 I_j^V \right) S_i - \sum_{j=1}^n \left( \beta_{ij}^3 I_j + \beta_{ij}^4 I_j^V \right) S_i^V \\ &= - \sum_{j=1}^n \left( \beta_{ij}^3 I_j + \beta_{ij}^4 I_j^V \right) (S_i + S_i^V) \dots \\ &\quad - \sum_{j=1}^n \left( (\beta_{ij}^1 - \beta_{ij}^3) I_j + (\beta_{ij}^2 - \beta_{ij}^4) I_j^V \right) S_i. \end{aligned}$$

Thus,

$$\begin{aligned} - \sum_{j=1}^n \left( (\beta_{ij}^1 - \beta_{ij}^3) I_j + (\beta_{ij}^2 - \beta_{ij}^4) I_j^V \right) S_i &= \frac{d}{dt}(S_i + S_i^V) \dots \\ &\quad + \sum_{j=1}^n \left( \beta_{ij}^3 I_j + \beta_{ij}^4 I_j^V \right) (S_i + S_i^V) \\ &= \frac{d}{dt} \left( (S_i + S_i^V) e^{\mathcal{L}(t)} \right) e^{-\mathcal{L}(t)}. \end{aligned}$$

This means that

$$\begin{aligned} S_i(t) + S_i^V(t) &= e^{-\mathcal{L}(t)} \left[ S_i(0) - \int_0^t e^{\mathcal{L}(s)} \sum_{j=1}^n \left( (\beta_{ij}^1 - \beta_{ij}^3) I_j + (\beta_{ij}^2 - \beta_{ij}^4) I_j^V \right) S_i ds \right] \\ &= S_i(0) \left[ e^{-\mathcal{L}(t)} - \int_0^t e^{\mathcal{L}(s) - \mathcal{K}(s) - \mathcal{L}(t)} (\mathcal{K}'(s) - \mathcal{L}'(s)) \left( \frac{N_i - W_i(s)}{N_i} \right) ds \right]. \end{aligned}$$

Now, one can see that, as  $0 \leq W_i(s) \leq N_i$ ,

$$0 \leq \frac{N_i - W_i(s)}{N_i} \leq 1 \quad \forall s \geq 0$$

and hence

$$e^{-\mathcal{L}(t)} = 1 - \int_0^t \mathcal{L}'(s) e^{-\mathcal{L}(s)} ds \leq 1 - \int_0^t \mathcal{L}'(s) e^{-\mathcal{L}(s)} \left( \frac{N_i - W_i(s)}{N_i} \right) ds.$$

Now, this means that

$$\begin{aligned} S_i(t) + S_i^V(t) &\leq S_i(0) - S_i(0) \int_0^t \left[ \mathcal{L}'(s) e^{-\mathcal{L}(s)} \right] \end{aligned}$$

$$+ e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)} \sum_{j=1}^n \left( \mathcal{K}'(s) - \mathcal{L}'(s) \right) \left] \left( \frac{N_i - W_i(s)}{N_i} \right) ds.$$

This allows the use of Lemma B.1. Firstly, note that, as  $\mathcal{K}'(s) \geq \mathcal{L}'(s) \geq 0$  and  $\tilde{W}_i(s) \geq W_i(s)$ , one has

$$\begin{aligned} S_i(t) + S_i^V(t) &\leq S_i(0) - S_i(0) \int_0^t \left[ \mathcal{L}'(s) e^{-\mathcal{L}(s)} \right. \\ &\quad \left. + e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)} \sum_{j=1}^n \left( \mathcal{K}'(s) - \mathcal{L}'(s) \right) \right] \left( \frac{N_i - \tilde{W}_i(s)}{N_i} \right) ds. \end{aligned}$$

Moreover,

$$\begin{aligned} &\int_0^t \left[ \mathcal{L}'(s) e^{-\mathcal{L}(s)} + e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)} \sum_{j=1}^n \left( \mathcal{K}'(s) - \mathcal{L}'(s) \right) \right] ds \\ &= 1 - e^{-\mathcal{L}(t)} + e^{-\mathcal{L}(t)} - e^{-\mathcal{K}(t)} \\ &= 1 - e^{-\mathcal{K}(t)} \\ &\geq 1 - e^{-\tilde{\mathcal{K}}(t)} \\ &\geq \int_0^t \left[ \tilde{\mathcal{L}}'(s) e^{-\tilde{\mathcal{L}}(s)} + e^{\tilde{\mathcal{L}}(s)-\tilde{\mathcal{K}}(s)-\tilde{\mathcal{L}}(t)} \sum_{j=1}^n \left( \tilde{\mathcal{K}}'(s) - \tilde{\mathcal{L}}'(s) \right) \right] ds \end{aligned}$$

and  $N_i - W_i(s)$  is non-increasing in  $s$ . Thus, by Lemma B.1, with

$$g(s) = 1 - e^{-\mathcal{L}(s)} + e^{-\mathcal{L}(t)} - e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)},$$

$h(s)$  defined as the tilde version of  $g(s)$ , and  $f(s) := N_i - W_i(s)$ , one has

$$\begin{aligned} &\int_0^t \left[ \mathcal{L}'(s) e^{-\mathcal{L}(s)} + e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)} \sum_{j=1}^n \left( \mathcal{K}'(s) - \mathcal{L}'(s) \right) \right] \left( \frac{N_i - \tilde{W}_i(s)}{N_i} \right) ds \\ &\geq \int_0^t \left[ \tilde{\mathcal{L}}'(s) e^{-\tilde{\mathcal{L}}(s)} + e^{\tilde{\mathcal{L}}(s)-\tilde{\mathcal{K}}(s)-\tilde{\mathcal{L}}(t)} \sum_{j=1}^n \left( \tilde{\mathcal{K}}'(s) - \tilde{\mathcal{L}}'(s) \right) \right] \left( \frac{N_i - \tilde{W}_i(s)}{N_i} \right) ds. \end{aligned} \quad (26)$$

Thus, (as this integral is multiplied by -1 in (26)), combining this with (26) gives

$$S_i(t) + S_i^V(t) \leq \tilde{S}_i(t) + \tilde{S}_i^V(t) \quad \forall t \in [0, T]$$

as required  $\square$

### B.3 Lemma B.3

**Lemma B.3** *Suppose that*

$$\sum_{j=1}^n K_{ij}(t) \geq \sum_{j=1}^n \tilde{K}_{ij}(t) \quad \forall i \in \{1, \dots, n\} \quad \text{and} \quad t \in [0, T].$$

*Then*

$$I_i(t) + R_i(t) \geq \tilde{I}_i(t) + \tilde{R}_i(t) \quad \forall t \in [0, T].$$

To begin, one can write the equation for  $S_i$  as

$$\frac{1}{S_i} \frac{dS_i}{dt} = - \sum_{j=1}^n (K'_{ij}(t)) - \frac{U_i}{N_i - W_i}$$

and hence, integrating

$$\ln(S_i(t)) - \ln(S_i(0)) = - \sum_{j=1}^n K_{ij}(t) + \ln(N_i - W_i(t)) - \ln(N_i)$$

which implies

$$S_i(t) = \left( \frac{S_i(0)(N_i - W_i(t))}{N_i} \right) e^{-\sum_{j=1}^n K_{ij}(t)}$$

Using this result shows that

$$\begin{aligned} \frac{d}{dt}(I_i + R_i) &= \sum_{j=1}^n \left( \beta_{ij}^1 I_j + \beta_{ij}^2 I_j^V \right) S_i \\ &= \sum_{j=1}^n K'_{ij}(t) S_i \\ &= \left[ \sum_{j=1}^n K'_{ij}(t) \right] \left( \frac{S_i(0)(N_i - W_i(t))}{N_i} \right) e^{-\sum_{j=1}^n K_{ij}(t)}, \end{aligned}$$

Thus,

$$I_i(t) + R_i(t) = I_i(0) + \int_0^t \left[ \sum_{j=1}^n K'_{ij}(s) \right] \left( \frac{S_i(0)(N_i - W_i(s))}{N_i} \right) e^{-\sum_{j=1}^n K_{ij}(s)} ds$$

$$\geq \tilde{I}_i(0) + \int_0^t \left[ \sum_{j=1}^n K'_{ij}(s) \right] \left( \frac{S_i(0)(N_i - \tilde{W}_i(s))}{N_i} \right) e^{-\sum_{j=1}^n K_{ij}(s)} ds, \quad (27)$$

using the fact that the initial conditions are the same in both cases and that  $W_i \leq \tilde{W}_i$ . Now, one can use the results of Lemma B.1 with

$$g(t) = 1 - \exp \left( - \sum_{j=1}^n K_{ij}(t) \right), \quad h(t) = 1 - \exp \left( - \sum_{j=1}^n \tilde{K}_{ij}(t) \right)$$

and  $f(t) = (N_i - \tilde{W}_i(t))$ , noting that

$$\begin{aligned} \int_0^t \left[ \sum_{j=1}^n K'_{ij}(s) \right] e^{-\sum_{j=1}^n K_{ij}(s)} ds &= 1 - e^{-\sum_{j=1}^n K_{ij}(t)} \\ &\geq 1 - e^{-\sum_{j=1}^n \tilde{K}_{ij}(t)} \\ &= \int_0^t \left[ \sum_{j=1}^n \tilde{K}'_{ij}(s) \right] e^{-\sum_{j=1}^n \tilde{K}_{ij}(s)} ds \end{aligned}$$

and that  $N_i - \tilde{W}_i(t)$  is non-increasing. Thus,

$$I_i(t) + R_i(t) \geq \tilde{I}_i(t) + \tilde{R}_i(t) \quad \forall t \in [0, T]$$

as required.

#### B.4 Lemma B.4

**Lemma B.4** *Suppose that*

$$\sum_{j=1}^n K_{ij}(t) \geq \sum_{j=1}^n \tilde{K}_{ij}(t) \quad \forall i \in \{1, \dots, n\} \quad \text{and} \quad t \in [0, T].$$

*Then,*

$$R_i(t) \geq \tilde{R}_i(t) \quad \forall t \in [0, T]$$

**Proof** The result of Lemma B.3 can be written as

$$\frac{1}{\mu_i^1} \frac{dR_i}{dt} + R_i \geq \frac{1}{\mu_i^1} \frac{d\tilde{R}_i}{dt} + \tilde{R}_i \quad \forall t \in [0, T]$$

which implies

$$\frac{d}{dt} \left( R_i e^{\mu_i^1 t} \right) \geq \frac{d}{dt} \left( \tilde{R}_i e^{\mu_i^1 t} \right)$$

and hence, after integrating and cancelling exponentials, one finds

$$R_i(t) \geq \tilde{R}_i(t) \quad \forall t \in [0, T]$$

as required.  $\square$

### B.5 Lemma B.5

**Lemma B.5** *Suppose that*

$$T := \inf \left\{ t : K_{ij}(t) < \tilde{K}_{ij}(t) \text{ or } L_{ij}(t) < \tilde{L}_{ij}(t) \text{ for some } i, j \in \{1, \dots, n\} \right\}$$

*exists. Then, for some  $b \in \{1, \dots, n\}$ , and some real constants  $\kappa$  and  $\eta$ ,*

$$\begin{aligned} R_b(T) + \kappa R_b^V(T) &= \tilde{R}_b(T) + \kappa \tilde{R}_b^V(T), \\ I_b(T) + \eta I_b^V(T) &\leq \tilde{I}_b(T) + \eta \tilde{I}_b^V(T) \end{aligned}$$

*and*

$$0 \leq \kappa \leq \eta \leq 1$$

**Proof** Suppose that  $T$  exists. Then, by continuity, there exists some  $a$  and  $b$  such that  $K_{ab}(T) = \tilde{K}_{ab}(T)$  or  $L_{ab}(T) = \tilde{L}_{ab}(T)$ . These can be rearranged to give, respectively,

$$R_b(T) + \frac{\mu_b^1 \beta_{ab}^2}{\mu_b^2 \beta_{ab}^1} R_b^V(T) = \tilde{R}_b(T) + \frac{\mu_b^1 \beta_{ab}^2}{\mu_b^2 \beta_{ab}^1} \tilde{R}_b^V(T)$$

or

$$R_b(T) + \frac{\mu_b^1 \beta_{ab}^4}{\mu_b^2 \beta_{ab}^3} R_b^V(T) = \tilde{R}_b(T) + \frac{\mu_b^1 \beta_{ab}^4}{\mu_b^2 \beta_{ab}^3} \tilde{R}_b^V(T).$$

This can be written as

$$R_b(T) + \kappa R_b^V(T) = \tilde{R}_b(T) + \kappa \tilde{R}_b^V(T),$$

where, by the inequality constraints on the  $\beta_{ij}^\alpha$  and  $\mu_i^\alpha$

$$\kappa \leq \frac{\mu_b^1}{\mu_b^2}. \quad (28)$$

Moreover, note that

$$\frac{d}{dt} \left( R_b + \kappa R_b^V \right) = \mu_b^1 I_b + \frac{\beta_{ab}^2 \mu_b^1}{\beta_{ab}^1} I_b^V$$

is a continuous function. Thus, if

$$\left. \frac{d}{dt} \left( R_b + \kappa R_b^V \right) \right|_{t=T} > \left. \frac{d}{dt} \left( \tilde{R}_b + \kappa \tilde{R}_b^V \right) \right|_{t=T},$$

then there exists some  $\tau > 0$  such that

$$\int_T^{T+\tau} \frac{d}{dt} \left( R_b(s) + \kappa R_b^V(s) \right) ds > \int_T^{T+\tau} \frac{d}{dt} \left( \tilde{R}_b(s) + \kappa \tilde{R}_b^V(s) \right) ds \quad \forall t \in [0, \tau]$$

and hence, in particular

$$R_b(T+t) + \kappa R_b^V(T+t) > \tilde{R}_b(T+t) + \kappa \tilde{R}_b^V(T+t) \quad \forall t \in [0, \tau],$$

Thus, it is necessary that there is some  $b$  such that

$$\left. \frac{d}{dt} \left( R_b + \kappa R_b^V \right) \right|_{t=T} \leq \left. \frac{d}{dt} \left( \tilde{R}_b + \kappa \tilde{R}_b^V \right) \right|_{t=T}$$

so

$$I_b(T) + \frac{\kappa \mu_b^2}{\mu_b^1} I_b^V(T) \leq \tilde{I}_b(T) + \frac{\kappa \mu_b^2}{\mu_b^1} \tilde{I}_b^V(T).$$

This can be written as

$$I_b(t) + \eta I_b^V(t) \leq \tilde{I}_b(t) + \eta \tilde{I}_b^V(t),$$

where, by (28), the fact that  $\mu_b^2 \geq \mu_b^1$ , and the non-negativity of all parameters,

$$0 \leq \kappa \leq \eta \leq 1.$$

as required □

## B.6 Lemma B.6

For the purposes of this lemma, it is helpful to recall the inequality system (9)–(15).

$$S_b(T) + S_b^V(T) \leq \tilde{S}_b(T) + \tilde{S}_b^V(T), \quad (7)$$

$$I_b(T) + R_b(T) \geq \tilde{I}_b(T) + \tilde{R}_b(T) \quad (8)$$

$$R_b(T) \geq \tilde{R}_b(T), \quad (9)$$

$$R_b(T) + \kappa R_b^V(T) \leq \tilde{R}_b(T) + \kappa \tilde{R}_b^V(T), \quad (10)$$

$$I_b(T) + \eta I_b^V(T) \leq \tilde{I}_b(T) + \eta \tilde{I}_b^V(T), \quad (11)$$

$$0 \leq \kappa \leq \eta \leq 1. \quad (12)$$

and

$$\begin{aligned} S_b(T) + I_b(T) + R_b(T) + S_b^V(T) + I_b^V(T) + R_b^V(T) \\ = \tilde{S}_b(T) + \tilde{I}_b(T) + \tilde{R}_b(T) + \tilde{S}_b^V(T) + \tilde{I}_b^V(T) + \tilde{R}_b^V(T), \end{aligned} \quad (13)$$

**Lemma B.6** Suppose that the system (9) - (15) holds for some  $b \in \{1, \dots, n\}$  and some  $T \geq 0$ . Then,

$$\begin{aligned} \eta I_b^V(T) + \kappa R_b^V(T) &= \eta \tilde{I}_b^V(T) + \kappa \tilde{R}_b^V(T) \\ I_b(T) + R_b(T) &= \tilde{I}_b(T) + \tilde{R}_b(T) \\ I_b^V(T) + R_b^V(T) &= \tilde{I}_b^V(T) + \tilde{R}_b^V(T) \\ S_b(T) + S_b^V(T) &= \tilde{S}_b(T) + \tilde{S}_b^V(T). \end{aligned}$$

**Proof** To begin, note that adding inequalities (9), (12) and (13) gives

$$\begin{aligned} S_b(T) + S_b^V(T) + R_b(T) + \kappa R_b^V(T) + I_b(T) + \eta I_b^V(T) \\ \leq \tilde{S}_b(T) + \tilde{S}_b^V(T) + \tilde{R}_b(T) + \kappa \tilde{R}_b^V(T) + \tilde{I}_b(T) + \eta \tilde{I}_b^V(T) \end{aligned}$$

and then, using (15) shows that

$$(\kappa - 1)R_b^V(T) + (\eta - 1)I_b^V(T) \leq (\kappa - 1)\tilde{R}_b^V(T) + (\eta - 1)\tilde{I}_b^V(T). \quad (29)$$

Moreover, adding (12) and (13) shows that

$$I_b(T) + \eta I_b^V(T) + R_b(T) + \kappa R_b^V(T) \leq \tilde{I}_b(T) + \eta \tilde{I}_b^V(T) + \tilde{R}_b(T) + \kappa \tilde{R}_b^V(T)$$

and then, using (10) shows that

$$\eta I_b^V(T) + \kappa R_b^V(T) \leq \eta \tilde{I}_b^V(T) + \kappa \tilde{R}_b^V(T). \quad (30)$$



Now, from the inequality (12) combined with the inequality (11), it must be the case that

$$R_b^V(T) - \tilde{R}_b^V(T) \leq \frac{1}{\kappa}(\tilde{R}_b(T) - R_b(T)) \leq 0. \quad (31)$$

Define

$$x := R_b^V(T) - \tilde{R}_b^V(T) \quad \text{and} \quad y := I_b^V(T) - \tilde{I}_b^V(T)$$

so that the system given by (14), (29), (30) and (31) reduces to

$$(\kappa - 1)x + (\eta - 1)y \leq 0 \quad (32)$$

$$\kappa x + \eta y \leq 0$$

$$x \leq 0$$

$$0 \leq \kappa \leq \eta \leq 1. \quad (33)$$

Note first that  $x = 0$  implies that  $y = 0$  as  $\eta$  and  $(\eta - 1)$  have different signs. Thus, in this case, the inequalities (32) and (33) are in fact equalities.

Suppose instead that  $x \neq 0$  (so  $x < 0$ ). The first two of these inequalities can be rearranged (noting the signs of the denominators) to give

$$-\frac{(\kappa - 1)x}{(\eta - 1)} \leq y \leq \frac{-\kappa x}{\eta}$$

and so, as  $-x > 0$ ,

$$\frac{(\kappa - 1)}{(\eta - 1)} \leq -\frac{y}{x} \leq \frac{\kappa}{\eta}. \quad (34)$$

However, note that

$$\begin{aligned} \kappa < \eta &\Rightarrow \eta\kappa - \eta < \eta\kappa - \kappa \\ &\Rightarrow \eta(\kappa - 1) < \kappa(\eta - 1) \\ &\Rightarrow \frac{\kappa - 1}{\eta - 1} > \frac{\kappa}{\eta} \end{aligned}$$

and hence, as  $\kappa \leq \eta$ , for there to be solutions to the inequality (34), it is necessary that

$$\kappa = \eta \Rightarrow \frac{-y}{x} = 1 \Rightarrow y = -x.$$

This means that the inequalities (32) and (33) are satisfied to equality in this and hence, from before, all cases. Thus, it is necessary that

$$(\kappa - 1)R_b^V(T) + (\eta - 1)I_b^V(T) = (\kappa - 1)\tilde{R}_b^V(T) + (\eta - 1)\tilde{I}_b^V(T) \quad (35)$$

and

$$\eta I_b^V(T) + \kappa R_b^V(T) = \eta \tilde{I}_b^V(T) + \kappa \tilde{R}_b^V(T), \quad (36)$$

which is the first required equality. Thus, one can once again add the inequalities (12) and (13) to give

$$I_b(T) + R_b(T) + \left[ \eta I_b^V(T) + \kappa R_b^V(T) \right] \leq \tilde{I}_b(T) + \tilde{R}_b(T) + \left[ \eta \tilde{I}_b^V(T) + \kappa \tilde{R}_b^V(T) \right]$$

and so

$$I_b(T) + R_b(T) \leq \tilde{I}_b(T) + \tilde{R}_b(T), \quad (37)$$

which, combined with (10), shows that

$$I_b(T) + R_b(T) = \tilde{I}_b(T) + \tilde{R}_b(T). \quad (38)$$

Moreover, one can subtract (35) from (36) to get

$$I_b^V(T) + R_b^V(T) = \tilde{I}_b^V(T) + \tilde{R}_b^V(T)$$

and then, using (15) alongside (37) and (38) shows

$$S_b(T) + S_b^V(T) = \tilde{S}_b(T) + \tilde{S}_b^V(T)$$

as required.  $\square$

## B.7 Lemma B.7

Note that for this lemma, it will be assumed that each  $K_{ij}(t) \geq \tilde{K}_{ij}(t)$ , rather than the inequality simply holding for their sums as before.

**Lemma B.7** *Under the assumptions of Proposition A.1.1, suppose that the system of inequalities (9)–(15) holds for some  $b \in \{1, \dots, n\}$  and some  $T > 0$ . Suppose further that*

$$K_{ij}(t) \geq \tilde{K}_{ij}(t) \quad \forall i, j \in \{1, \dots, n\}.$$

*Then,*

$$W_i(t) = \tilde{W}_i(t) \quad \forall i \in \{1, \dots, n\} \quad \text{and} \quad \forall t \in [0, T].$$

**Proof** By Lemma B.6, the system (16)–(18) must hold for  $b$ . Now, Equation (27) in the proof of Lemma B.3 shows that

$$I_b(T) + R_b(T) = \frac{S_b(0)}{N_b} \int_0^T \left[ \sum_{k=1}^n K'_{bk}(s) \right] (N_b - W_b(s)) e^{-\sum_{j=1}^n K_{bj}(s)} ds. \quad (39)$$

Now, the equality (16) shows

$$I_b(T) + R_b(T) = \tilde{I}_b(T) + \tilde{R}_b(T)$$

and hence, after cancelling the non-zero  $S_b(0)$  and  $N_b$  terms, (39) (and its tilde equivalent) shows that

$$\begin{aligned} & \int_0^T \left[ \sum_{k=1}^n K'_{bk}(s) \right] (N_b - W_b(s)) e^{-\sum_{k=1}^n K_{bk}(s)} ds \\ &= \int_0^T \left[ \sum_{k=1}^n \tilde{K}'_{bk}(s) \right] (N_b - \tilde{W}_b(s)) e^{-\sum_{k=1}^n \tilde{K}_{bk}(s)} ds. \end{aligned} \quad (40)$$

Note that, from Lemma C.6, as  $\Pi = \{1, \dots, n\}$

$$\tilde{I}_k(s), I_k(s) > 0 \quad \forall k \in \{1, \dots, n\} \quad \text{and} \quad s > 0.$$

Thus,

$$K'_{bk}(t) \geq \beta_{bk}^1 I_j(t) > 0 \quad \forall t > 0.$$

In particular,

$$\left[ \sum_{j=1}^n K'_{bk}(s) \right] e^{-\sum_{j=1}^n K_{bj}(s)} > 0 \quad \forall s \in [0, T].$$

Moreover, by continuity of  $K'_{ik}$  (as continuous functions attain their bounds on closed intervals), there exists some  $m > 0$  such that

$$\left[ \sum_{k=1}^n K'_{bk}(s) \right] e^{-\sum_{k=1}^n K_{bk}(s)} > m \quad \forall s \in [0, T].$$

Hence, as  $W_b \leq \tilde{W}_b$

$$\int_0^T \left[ \sum_{k=1}^n K'_{bk}(s) \right] (N_b - W_b(s)) e^{-\sum_{k=1}^n K_{bk}(s)} ds$$

$$\begin{aligned}
&= \int_0^T \left[ \sum_{k=1}^n K'_{bk}(s) \right] (N_b - \tilde{W}_b(s) + (\tilde{W}_b(s) - W_b(s)) e^{-\sum_{k=1}^n K_{bk}(s)}) ds \\
&\geq \int_0^T \left[ \sum_{k=1}^n K'_{bk}(s) \right] (N_b - \tilde{W}_b(s)) e^{-\sum_{k=1}^n K_{bk}(s)} ds + m \int_0^T \tilde{W}_b(s) - W_b(s) ds.
\end{aligned} \tag{41}$$

Finally, as  $N - \tilde{W}_b$  is decreasing and for any  $t \in [0, T]$ ,

$$\int_0^t \left[ \sum_{k=1}^n K'_{bk}(s) \right] e^{-\sum_{k=1}^n K_{bk}(s)} ds \geq \int_0^t \left[ \sum_{k=1}^n \tilde{K}'_{bk}(s) \right] e^{-\sum_{k=1}^n \tilde{K}_{bk}(s)} ds$$

one has, by Lemma B.1, setting

$$g(t) = 1 - e^{-\sum_{k=1}^n K_{bk}(t)}, \quad h(t) = e^{-\sum_{k=1}^n \tilde{K}_{bk}(t)}$$

and  $f(t) = N_b - \tilde{W}_b(s)$ ,

$$\begin{aligned}
&\int_0^T \left[ \sum_{k=1}^n K'_{bk}(s) \right] (N_b - \tilde{W}_b(s)) e^{-\sum_{k=1}^n K_{bk}(s)} ds \\
&\geq \int_0^T \left[ \sum_{k=1}^n \tilde{K}'_{bk}(s) \right] (N_b - \tilde{W}_b(s)) e^{-\sum_{k=1}^n \tilde{K}_{bk}(s)} ds \\
&= \tilde{I}_b(T) + \tilde{R}_b(T)
\end{aligned}$$

and so, combining this with (41),

$$\begin{aligned}
I_b(T) + R_b(T) &\geq \tilde{I}_b(T) + \tilde{R}_b(T) + m \int_0^T \tilde{W}_b(s) - W_b(s) ds \\
&\geq \tilde{I}_b(T) + \tilde{R}_b(T) = I_b(T) + R_b(T).
\end{aligned}$$

Hence,

$$\int_0^T \tilde{W}_b(s) - W_b(s) ds = 0,$$

which by continuity means

$$W_b(t) = \tilde{W}_b(t) \quad \forall t \in [0, T]$$

Now, moreover, substituting this back into the equality given in (40) shows that

$$\int_0^T \left[ \sum_{k=1}^n K'_{bk}(s) \right] (N_b - W_b(s)) e^{-\sum_{k=1}^n K_{bk}(s)} ds$$

$$= \int_0^T \left[ \sum_{k=1}^n \tilde{K}'_{bk}(s) \right] (N_b - W_b(s)) e^{-\sum_{k=1}^n \tilde{K}_{bk}(s)} ds.$$

Hence, integrating by parts, this shows that

$$\begin{aligned} 0 &= (N_b - W_b(T)) (e^{-\sum_{k=1}^n K_{bk}(T)} - e^{-\sum_{k=1}^n \tilde{K}_{bk}(T)}) \dots \\ &\quad + \int_0^T U_b(s) (e^{-\sum_{k=1}^n K_{bk}(s)} - e^{-\sum_{k=1}^n \tilde{K}_{bk}(s)}) ds \end{aligned}$$

Now,

$$\sum_{k=1}^n \tilde{K}_{bk}(s) \geq \sum_{k=1}^n K_{bk}(s) \quad \forall s \in [0, T]$$

and so, for equality, it is necessary that

$$(N_b - W_b(T)) (e^{-\sum_{k=1}^n K_{bk}(T)} - e^{-\sum_{k=1}^n \tilde{K}_{bk}(T)}) = 0$$

Thus, as it is assumed that  $W_b(t) < N_b$  for all  $t \geq 0$ ,

$$e^{-\sum_{k=1}^n K_{bk}(T)} - e^{-\sum_{k=1}^n \tilde{K}_{bk}(T)} = 0$$

and hence, as  $K_{bk}(T) \geq \tilde{K}_{bk}(T)$  for all  $k \in \{1, \dots, n\}$ ,

$$K_{bk}(T) = \tilde{K}_{bk}(T) \quad \forall k \in \{1, \dots, n\} \quad (42)$$

Now, suppose that  $K'_{bk}(T) > \tilde{K}'_{bk}$  for some  $k$ . Then, by continuity and the fact that  $T > 0$ , it is necessary that there is some  $\tau \in (0, T)$  such that

$$\int_{T-\tau}^T K'_{bk}(s) ds > \int_{T-\tau}^T \tilde{K}'_{bk}(s) ds$$

which means that

$$K_{bk}(T - \tau) < \tilde{K}_{bk}(T - \tau)$$

which is a contradiction to the definition of  $T$ . Thus, it is necessary that

$$K'_{bk}(T) \leq \tilde{K}'_{bk}(T) \quad \forall k \in \{1, \dots, n\}. \quad (43)$$

Dividing (42) by  $\beta_{bk}^1/\mu_k^1$  and (43) by  $\beta_{bk}^1$  shows that the inequality system (9)–(15) holds for each  $k$  (as Lemmas B.2–B.4 hold for any group) and so, following Lemma B.6 and the previous work of this proof, it is necessary that

$$W_k(t) = \tilde{W}_k(t) \quad \forall t \in [0, T]$$

This holds for each  $k$  and hence the proof is complete.  $\square$

### B.8 Lemma B.8

**Lemma B.8** Define functions  $\Delta_i^f$  to be

$$\Delta_i^f(t) := f_i(T+t) - \tilde{f}_i(T+t) \text{ for } f \in \{S, I, R, S^V, I^V, R^V, W\}$$

and suppose that

$$\Delta_i^f(0) = 0 \quad \forall f \in \{S, I, R, S^V, I^V, R^V, W\}. \quad (44)$$

Suppose further that the  $U_i(t)$  are right-continuous step functions. Then, for  $t \in [0, \delta]$  in the limit  $\delta \rightarrow 0$ , and for any  $x, y \in \mathbb{R}$

$$\begin{aligned} \frac{x}{\mu_i^1} \Delta_i^R + \frac{y}{\mu_i^2} \Delta_i^{R^V} &= \frac{t^3 S_i(T)(U_i(T) - \tilde{U}_i(T))}{6(N_i - W_i(T))} \\ &\quad \left[ x \sum_{j=1}^n (K'_{ij}(T)) - y \sum_{j=1}^n (L'_{ij}(T)) \right] + O(\delta^4). \end{aligned}$$

**Proof** As the  $U_i(t)$  are step functions, for sufficiently small  $\delta$ , they are constant on the interval  $[T, T + \delta]$ , so this will be assumed. Note that, for any  $i \in \{1, \dots, n\}$  and any  $t \geq 0$

$$\begin{aligned} \left| \frac{dS_i}{dt}(t) \right| &\leq \left| \sum_{j=1}^n S_i(t) \beta_{ij}^1 I_j(t) \right| + \left| \frac{S_i(t)}{N_i - W_i(t)} U_i(t) \right| \\ &\leq \left| \sum_{j=1}^n N_i \beta_{ij}^1 N_j \right| + |1 \times U_i(t)| \\ &\leq U_i(T) + C, \end{aligned}$$

where the constant term,  $C$ , is independent of  $t$  and the vaccination policy. Note the second line follows from the fact that, as  $W_i(t) < N_i$ ,

$$\frac{S_i(t)}{N_i - W_i(t)} = \frac{S_i(0)}{N_i} \exp \left[ - \sum_{j=1}^n K_{ij}(t) \right] \leq 1.$$

Similarly, one can show (by increasing the constant  $C$  if necessary) that

$$\left| \frac{dS_i^V}{dt}(t) \right| \leq U_i(T) + C$$

$$\left| \frac{dI_i^V}{dt}(t) \right|, \left| \frac{dR_i^V}{dt}(t) \right|, \left| \frac{dI_i}{dt}(t) \right|, \left| \frac{dR_i}{dt}(t) \right| \leq C$$

$$\left| \frac{dW_i}{dt}(t) \right| \leq U_i(T).$$

Then, for  $t \in (0, \delta)$  and  $f \in \{S, I, R, S^V, I^V, R^V, W\}$

$$|f_i(T+t) - f_i(T)| = \left| \int_T^{T+t} \frac{df_i}{dt}(s) ds \right| \leq (C + U_i(T))\delta$$

so that, in particular

$$f_i(T+t) = f_i(T) + O(\delta) \quad \forall f \in \{S, I, R, S^V, I^V, R^V, W\}. \quad (45)$$

Now,

$$\frac{d\Delta_i^S}{dt} = - \sum_{j=1}^n (K_{ij}S_i - \tilde{K}_{ij}\tilde{S}_j) + \frac{S_i U_i}{N_i - W_i} - \frac{\tilde{S}_i \tilde{U}_i}{N_i - \tilde{W}_i}.$$

Using (44) and (45), this equation linearises to

$$\frac{d\Delta_i^S}{dt}(t) = \frac{S_i(T)(U_i(t+T) - \tilde{U}_i(t+T))}{N_i - W_i(T)} + O(\delta).$$

Noting that

$$U_i(t+T) - \tilde{U}_i(t+T) = U_i(T) - \tilde{U}_i(T) \quad \forall t \in [0, \delta]$$

this means that

$$\frac{d\Delta_i^S}{dt} = \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} + O(\delta)$$

and so (for  $t < \delta$ )

$$\Delta_i^S(t) = t \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} + O(\delta^2).$$

Now, one can linearise the equation for  $\Delta_i^I$ . Note that

$$\frac{d\Delta_i^I}{dt} = \sum_{j=1}^n (K'_{ij}S_i - \tilde{K}'_{ij}\tilde{S}_i) + \mu_i^1(I_i - \tilde{I}_i)$$

and so, with

$$I_i(t + T) = I_i(T) + O(\delta)$$

and similar expressions for other variables,

$$\frac{d\Delta_i^I}{dt} = O(\delta) \Rightarrow \Delta_i^I(t) = O(\delta^2) \quad \text{for } t < \delta$$

Now, one can linearise in a different way. Note that

$$\tilde{I}_i(T + t) = I_i(T + t) + O(\delta^2) \quad \text{and} \quad \tilde{I}_i^V(T + t) = I_i^V(T + t) + O(\delta^2)$$

so

$$\tilde{K}_{ij}'(T + t) = K_{ij}'(T + t) + O(\delta^2).$$

Thus,

$$\begin{aligned} \frac{d\Delta_i^I}{dt}(T + t) &= \sum_{j=1}^n (K_{ij}'(T + t)S_i(T + t) - \tilde{K}_{ij}'(T + t)\tilde{S}_i(T + t)) + \mu_i^1 \Delta_i^I(T + t) + O(\delta^2) \\ &= \Delta_i^S(t) \sum_{j=1}^n (K_{ij}'(T + t)) + O(\delta^2) \\ &= t \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} \sum_{j=1}^n (K_{ij}'(T) + O(\delta)) + O(\delta^2) \\ &= t \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} \sum_{j=1}^n (K_{ij}'(T)) + O(\delta^2) \end{aligned}$$

and hence

$$\Delta_i^I = \frac{t^2}{2} \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} \sum_{j=1}^n (K_{ij}'(T)) + O(\delta^3).$$

Thus,

$$\frac{d\Delta_i^R}{dt} = \Delta_i^I \mu_i^1 \Rightarrow \Delta_i^R(t) = \frac{\mu_i^1 t^3}{6} \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} \sum_{j=1}^n (K_{ij}'(T)) + O(\delta^4).$$

Now, note that

$$\frac{d(\Delta_i^S + \Delta_i^{SV})}{dt} = O(\delta)$$



as this derivative has no explicit dependence on  $U$ . Thus, in particular,

$$\Delta_i^S + \Delta_i^{S^V} = O(\delta^2)$$

and so

$$\Delta_i^{S^V} = -t \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} + O(\delta^2).$$

Then, as before (as the equation for  $\frac{dI_i}{dt}$  is the same as that for  $\frac{dI_i^V}{dt}$ , but with  $S_i^V$  instead of  $S_i$ ,  $\mu_i^1$  instead of  $\mu_i^2$  and  $K_{ij}$  instead of  $L_{ij}$ )

$$\frac{d\Delta_i^{I^V}}{dt}(T+t) = -t \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} \sum_{j=1}^n (L'_{ij}(T)) + O(\delta^2),$$

which means

$$\Delta_i^{I^V} = -\frac{t^2}{2} \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} \sum_{j=1}^n (L'_{ij}(T)) + O(\delta^3)$$

and hence

$$\Delta_i^{R^V}(t) = -\frac{\mu_i^2 t^3}{6} \frac{S_i(T)(U_i(T) - \tilde{U}_i(T))}{N_i - W_i(T)} \sum_{j=1}^n (L'_{ij}(T)) + O(\delta^4).$$

Thus,

$$\begin{aligned} \frac{x}{\mu_i^1} \Delta_i^R + \frac{y}{\mu_i^2} \Delta_i^{R^V} &= \frac{t^3 S_i(T)(U_i(T) - \tilde{U}_i(T))}{6(N_i - W_i(T))} \\ &\quad \left[ x \sum_{j=1}^n (K'_{ij}(T)) - y \sum_{j=1}^n (L'_{ij}(T)) \right] + O(\delta^4) \end{aligned}$$

as required.  $\square$

## B.9 Lemma B.9

**Lemma B.9** *Suppose that*

$$T := \inf \left\{ t : K_{ij}(t) \geq \tilde{K}_{ij} \text{ or } L_{ij}(t) \geq \tilde{L}_{ij}(t) \text{ for some } i, j \in \{1, \dots, n\} \right\}$$

exists. Define functions  $\Delta_i^f$  to be

$$\Delta_i^f(t) := f_i(T+t) - \tilde{f}_i(T+t) \quad \text{for } f \in \{S, I, R, S^V, I^V, R^V, W\}$$

and suppose that

$$\Delta_i^f(0) = 0 \quad \forall f \in \{S, I, R, S^V, I^V, R^V, W\}.$$

Suppose further that the  $U_i(t)$  are right-continuous step functions,  $\Pi = \{1, \dots, n\}$  and that

$$\beta_{ij}^1 > \beta_{ij}^3 > 0 \quad \text{and} \quad I_i(0) > 0 \quad \forall i, j \in \{1, \dots, n\}.$$

Then,

$$\sum_{j=1}^n K_{ij}(t) \geq \sum_{j=1}^n \tilde{K}_{ij}(t) \quad \forall t \in [0, T + \delta]$$

and

$$\sum_{j=1}^n L_{ij}(t) \geq \sum_{j=1}^n \tilde{L}_{ij}(t) \quad \forall t \in [0, T + \delta],$$

for sufficiently small  $\delta$ .

**Proof** By Lemma B.8, with  $x = \beta_{li}^1$  and  $y = \beta_{li}^2$  for some  $l \in \{1, \dots, n\}$

$$\begin{aligned} \frac{\beta_{li}^1}{\mu_i^1} \Delta_i^R + \frac{\beta_{li}^2}{\mu_i^2} \Delta_i^{R^V} &= \frac{t^3 S_i(T)(U_i(T) - \tilde{U}_i(T))}{6(N_i - W_i(T))} \\ &\quad \left[ \beta_{li}^1 \sum_{j=1}^n (K'_{ij}(T)) - \beta_{li}^2 \sum_{j=1}^n (L'_{ij}(T)) \right] + O(\delta^4). \end{aligned}$$

Now, as  $\beta_{li}^1 \geq \beta_{li}^2$ ,  $\beta_{li}^1 > 0$  and  $K'_{ij}(t)$  and  $L'_{ij}(t)$  are non-negative

$$\beta_{li}^1 \sum_{j=1}^n (K'_{ij}(T)) - \beta_{li}^2 \sum_{j=1}^n (L'_{ij}(T)) \leq 0 \Rightarrow \sum_{j=1}^n (K'_{ij}(T)) \leq \sum_{j=1}^n (L'_{ij}(T)). \quad (46)$$

Noting that

$$K'_{ij}(T) \geq L'_{ij}(T) \quad \forall j \in \{1, \dots, n\},$$

(46) requires

$$K'_{ij}(T) = L'_{ij}(T) \quad \forall j \in \{1, \dots, n\}$$

which, from the definitions of  $K'$  and  $L'$  requires

$$\beta_{ij}^1 I_j(T) + \beta_{ij}^2 I_j^V(T) = \beta_{ij}^3 I_j(T) + \beta_{ij}^4 I_j^V(T).$$

Thus, as  $I_j(T) > 0$  (as  $\Pi \in \{1, \dots, n\}$ ) and  $\beta_{ij}^2 I_j^V(T) \geq \beta_{ij}^4 I_j^V(T)$ , it is necessary that

$$\beta_{ij}^1 \leq \beta_{ij}^3,$$

which is a contradiction. Thus,

$$\beta_{ij}^1 \sum_{j=1}^n (K'_{ij}(T)) - \beta_{ij}^2 \sum_{j=1}^n (L'_{ij}(T)) > 0$$

which means

$$S_i(T)U_i(T) < S_i(T)\tilde{U}_i(T) \Rightarrow \frac{\beta_{ij}^1}{\mu_i^1} \Delta_i^R + \frac{\beta_{ij}^2}{\mu_i^2} \Delta_i^{R^V} = -C\delta^3 + O(\delta^4)$$

for some positive constant  $C$ . Now, if

$$S_i(T)U_i(T) > S_i(T)\tilde{U}_i(T)$$

then, necessarily,  $U_i(T) > \tilde{U}_i(T)$ . Thus, as  $\Delta_i^W(0) = 0$ , one will have

$$W_i(T+t) > \tilde{W}_i(T+t)$$

for sufficiently small  $t$ , which is a contradiction. Moreover, if

$$S_i(T)U_i(T) = S_i(T)\tilde{U}_i(T) \quad \forall i \in \{1, \dots, n\}$$

then the vaccination policies are the same in the interval  $[T, T+\delta]$ , as for each  $i$ , either  $S_i(T) = 0$  (in which case there is no more vaccination in group  $i$  so  $U_i(T) = \tilde{U}_i(T) = 0$ ) or  $U_i(T) = \tilde{U}_i(T)$ . Thus, the disease trajectories are the same, which contradicts the definition of  $T$ , as then  $K_{ij}(T+t) = \tilde{K}_{ij}(T+t)$  and  $L_{ij}(T+t) = \tilde{L}_{ij}(T+t)$  for all  $t \in [0, \delta]$ .

Now, note that

$$\sum_{i=1}^n K_{li}(t) - \sum_{i=1}^n \tilde{K}_{li}(t) = \sum_{i=1}^n \left( \frac{\beta_{li}^1}{\mu_i^1} \Delta_i^R + \frac{\beta_{li}^2}{\mu_i^2} \Delta_i^{R^V} \right) = - \sum_{i=1}^n E_i \delta^3 + O(\delta^4),$$

where  $E_i > 0$  if  $U_i(T) < \tilde{U}_i(T)$  and  $E_i = 0$  otherwise. Thus, in particular

$$\sum_{i=1}^n E_i > 0$$

and hence

$$\sum_{i=1}^n K_{li}(t) - \sum_{i=1}^n \tilde{K}_{li}(t) = -\sum_{i=1}^n E_i \delta^3 + O(\delta^4) < 0$$

for sufficiently small  $\delta$ . Thus,

$$\sum_{j=1}^n K_{ij}(t) \geq \sum_{j=1}^n \tilde{K}_{ij}(t) \quad \forall t \in [0, T + \delta]$$

and, by identical arguments (using  $x = \beta_{li}^3$  and  $y = \beta_{li}^4$  in Lemma B.8)

$$\sum_{j=1}^n L_{ij}(t) \geq \sum_{j=1}^n \tilde{L}_{ij}(t) \quad \forall t \in [0, T + \delta]$$

as required. □

### B.10 Lemma B.10

**Lemma B.10** *Suppose that*

$$T := \inf \left\{ t : K_{ij}(t) < \tilde{K}_{ij}(t) \text{ or } L_{ij}(t) < \tilde{L}_{ij}(t) \text{ for some } i, j \in \{1, \dots, n\} \right\}.$$

*Then, for any  $\delta > 0$ , there exists some  $t \in (T, T + \delta)$  and some real parameters  $0 \leq \kappa \leq \eta \leq 1$  such that*

$$R_b(t) + \kappa R_b^V(t) < \tilde{R}_b(t) + \kappa \tilde{R}_b^V(t) \quad \text{and} \quad I_b(t) + \eta I_b^V(t) \leq \tilde{I}_b(t) + \eta \tilde{I}_b^V(t).$$

**Proof** Firstly, note that by the definition of  $T$ , for each  $\delta > 0$ , there must exist  $i, j \in \{1, \dots, n\}$  and  $t \in (0, \delta)$  such that

$$K_{ij}(T + t) < \tilde{K}_{ij}(T + t) \quad \text{or} \quad L_{ij}(T + t) < \tilde{L}_{ij}(T + t).$$

That is, there is some  $b \in \{1, \dots, n\}$  such that

$$R_b(T + t) + \kappa R_b^V(T + t) < \tilde{R}_b(T + t) + \kappa \tilde{R}_b^V(T + t) \quad (47)$$

where

$$\kappa \leq \frac{\mu_b^1}{\mu_b^2}.$$

Note that

$$\mu_b^1 I_b(t) + \kappa \mu_b^2 I_b^V(t) = \frac{d}{dt} \left( R_b(t) + \kappa R_b^V(t) \right).$$

Now, define

$$\Delta_i^f(t) := f_i(T+t) - \tilde{f}_i(T+t) \quad \forall f \in \{I, I^V, R, R^V\}$$

and

$$\tau := \sup\{s \in [0, t] : \Delta_b^R(s) + \kappa \Delta_b^{R^V}(s) \geq 0\}$$

which exists as  $\Delta_b^R(0) + \kappa \Delta_b^{R^V}(0) = 0$ . Note that  $\tau < t$  by (47). Note also that by continuity, it is necessary that

$$\Delta_b^R(\tau) + \kappa \Delta_b^{R^V}(\tau) = 0.$$

Now, by the mean value theorem (as  $\Delta_b^R + \kappa \Delta_b^{R^V}$  is continuously differentiable), there exists an  $s$  in the non empty interval  $(\tau, t)$  such that

$$\begin{aligned} \mu_b^1 \Delta_b^I(s) + \kappa \mu_b^2 \Delta_b^{I^V}(s) &= \frac{1}{t - \tau} \left[ (\Delta_b^R(t) + \kappa \Delta_b^{R^V}(t)) - (\Delta_b^R(\tau) + \kappa \Delta_b^{R^V}(\tau)) \right] \\ &= \frac{1}{t - \tau} \left[ \Delta_b^R(t) + \kappa \Delta_b^{R^V}(t) \right] \\ &< 0 \end{aligned}$$

while also

$$\Delta_b^R(s) + \kappa \Delta_b^{R^V}(s) < 0,$$

by definition of  $\tau$ . Thus, defining  $\eta := \kappa \frac{\mu_b^2}{\mu_b^1} \leq 1$ ,

$$\Delta_b^R(s) + \kappa \Delta_b^{R^V}(s) < 0 \quad \Delta_b^I(s) + \eta \Delta_b^{I^V}(s) \geq 0 \quad \text{and} \quad 0 \leq \kappa \leq \eta \leq 1$$

as required.  $\square$

**B.11 Lemma B.11**

**Lemma B.11** Consider two non-negative functions  $A(t)$  and  $B(t)$  such that  $B(t)$  is non-decreasing and differentiable with a Lebesgue integrable derivative  $B'(t)$  satisfying

$$\int_0^t B'(s)ds = B(t) - B(0) \quad \forall t \geq 0.$$

Suppose further that for each  $T \geq 0$ , one can partition the interval  $[0, T]$  into a finite number of subintervals  $S_1^A, \dots, S_m^A$  and  $S_1^B, \dots, S_k^B$  such that

$$s \in \bigcup_{i=1}^m S_i^A \Leftrightarrow A(s) > B'(s)$$

Then, there exists a unique function  $\chi(t)$  for  $t \geq 0$  such that

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

**Proof**  $\chi$  can be constructed for each of the subintervals  $S_i^A$  and  $S_i^B$ . Note first that,

$$t \in S_i^B \Rightarrow B'(t) \geq A(t) \Rightarrow \chi(t) = A(t)$$

Now, suppose that  $t \in S_i^A$  for some  $i$ . Then, as  $S_i^A$  is an interval, one can suppose  $S_i^A = [c_i, d_i]$ . Define

$$\tau := \inf \left( \left\{ s \in S_i^A : B(s) \leq \int_0^{c_i} \chi(u)du + \int_{c_i}^s A(u)du \right\} \cup \{d_i\} \right).$$

If  $\tau = d_i$ , then one has (uniquely)  $\chi(t) = A(t)$  in  $S_i^A$ . Otherwise, one has (again uniquely)

$$\chi(t) = A(t) \quad \forall t \in [c_i, \tau] \quad \text{and} \quad \chi(t) = B'(t) \quad \forall t \in [\tau, d_i]$$

Uniqueness can be demonstrated as follows. If  $\chi(t) = B'(t)$  for some  $t \in [c_i, \tau]$ , then it is necessary (as  $A(t) > B'(t)$  so  $\chi(t) \neq A(t)$  in this case)

$$\int_0^t \chi(s)ds \geq B(t)$$

As  $A(t) \geq B'(t)$  in  $S_i^A$ , so  $\chi(t)$  is bounded by  $A(t)$ , the previous inequality can be extended to give

$$B(t) \leq \int_0^t \chi(s)ds \leq \int_0^{c_i} \chi(u)du + \int_{c_i}^s A(u)du$$

which contradicts the definition of  $\tau$ . A similar argument stands to prove uniqueness in  $[\tau, d_i]$ .

Thus,  $\chi$  is uniquely defined in each of the finite number of intervals and hence in  $[0, T]$  for each  $T$  and hence, it is uniquely defined for all  $t$  as required.  $\square$

## C Results on the SIR Equations

This section presents a variety of results on the SIR equations which are used in the proofs of the theorems in this paper. Many of them are well-known and widely used in the literature, but this appendix aims to provide a source of formal definitions and proofs of these results.

Before the results can be proved, it is necessary to establish two lemmas on differential equations.

### C.1 Lemma C.1

**Lemma C.1** *Suppose that  $H(t)$  is a continuous non-negative  $n \times n$  matrix for  $t \geq 0$  and that  $\mathbf{a} \in \mathbb{R}^n$ . Then, suppose that a function  $\mathbf{u} : \mathbb{R} \rightarrow \mathbb{R}^n$  satisfies*

$$\mathbf{u}(t) \leq \mathbf{a} + \int_0^t H(s)\mathbf{u}(s)ds \quad \forall t \geq 0.$$

Then,

$$\mathbf{u}(t) \leq \left(1 + \int_0^t V(t, s)H(s)ds\right)\mathbf{a},$$

where the matrix  $V(t, s)$  satisfies

$$V(t, s) = I_n + \int_s^t H(k)V(k, s)dk$$

and  $I_n$  is the  $n \times n$  identity matrix.

**Proof** This theorem is a special case of the theorem proved in Chandra and Davis (1976) where (in the notation of Chandra and Davis 1976),  $x$ ,  $y$  and  $z$  have been replaced by  $t$ ,  $s$  and  $k$  respectively,  $G(t)$  has been set to be the identity matrix and  $x^0$  has been set to zero.  $\square$

### C.2 Lemma C.2

**Lemma C.2** *Consider a continuous, time-dependent, matrix  $A(t)$  which satisfies*

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

and a constant matrix  $B$  that satisfies

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

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

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

Moreover, define a non-negative initial condition  $\mathbf{v}$  and suppose that  $\mathbf{y}$  and  $\mathbf{z}$  solve the systems

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

with

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

Then,

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

**Proof** To begin, define

$$\mu := \min_i (B_{ii})$$

so that, defining

$$A^*(t) := A(t) + \mu I \quad \text{and} \quad B^* := B + \mu I,$$

where  $I$  is the identity matrix,  $A^*$  and  $B^*$  are non-negative matrices. Moreover, note that

$$\frac{d\mathbf{y}}{dt} + \mu\mathbf{y} = A^*(t)\mathbf{y}$$

and so

$$e^{-\mu t} \frac{d}{dt} (e^{\mu t} \mathbf{y}) = A^*(t)\mathbf{y}.$$

Thus, define

$$\mathbf{y}^*(t) := e^{\mu t} \mathbf{y}(t)$$



so

$$\frac{dy^*}{dt} = A^*(t)y^*.$$

Similarly, defining

$$z^*(t) := e^{\mu t} z(t)$$

gives

$$\frac{dz^*}{dt} = Bz^*$$

while, moreover,

$$y^* \geq z^* \Leftrightarrow y \geq z \quad \text{and} \quad z^* \geq \mathbf{0} \Leftrightarrow z \geq \mathbf{0}.$$

Thus, it is simply necessary to prove that the results of this lemma hold when  $A(t)$  and  $B$  are non-negative matrices.

Now, it is helpful to note that, as the off-diagonal entries of  $A(t)$  and  $B$  are non-negative, the two differential systems are totally positive (Schwarz 1970). Thus, in particular, as  $v$  is non-negative,

$$y(t), z(t) \geq \mathbf{0} \quad \forall t \geq 0,$$

which proves one of the required inequalities. Now, one can also note that

$$\frac{d}{dt}(y - z) = A(t)y - Bz.$$

As  $A(t)$  is assumed to be non-negative, and  $y$  is non-negative,

$$\frac{d}{dt}(y - z) \geq B(y - z).$$

Defining  $\zeta := z - y$  and integrating gives

$$\zeta(t) \leq \int_0^t B(s)\zeta(s)ds,$$

noting that  $\zeta = \mathbf{0}$ . Hence, by Lemma C.1, one has

$$\zeta(t) \leq \mathbf{0} \Rightarrow y \geq z$$

as required. □

**C.3 Lemma C.3**

**Lemma C.3** Define the set of functions

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

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

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

**Proof** Noting that

$$\sum_{f \in \mathcal{F}_i(t)} f = N_i,$$

it is simply necessary to show that (for each  $t$  and  $i$ )

$$f(t) \geq 0 \quad \forall f(t) \in \mathcal{F}_i(t).$$

Now, note that

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

which means

$$\frac{d}{dt} \left( S_i \exp \left[ - \sum_{j=1}^n \left( \frac{\beta_{ij}^1}{\mu_j^1} R_j + \frac{\beta_{ij}^2}{\mu_j^2} R_j^V \right) - \ln(N_i - W_i) \right] \right) = 0$$

and hence (using the initial conditions)

$$S_i(t) = \frac{S_i(0)(N_i - W_i(t))}{N_i} \exp \left( - \sum_{j=1}^n \left[ \frac{\beta_{ij}^1}{\mu_j^1} R_j + \frac{\beta_{ij}^2}{\mu_j^2} R_j^V \right] \right).$$

As  $W_i(t) \leq N_i$  by construction, this means that

$$S_i(t) \geq 0 \quad \text{as required.}$$

Now, note that

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

so that

$$\frac{d}{dt} \left( S_i^V \exp \left[ \sum_{j=1}^n \left( \frac{\beta_{ij}^3}{\mu_j^1} R_j + \frac{\beta_{ij}^4}{\mu_j^2} R_j^V \right) \right] \right) \geq 0,$$

which means (as  $S_i^V(0) = 0$ )

$$S_i^V(t) \exp \left[ \sum_{j=1}^n \left( \frac{\beta_{ij}^3}{\mu_j^1} R_j(t) + \frac{\beta_{ij}^4}{\mu_j^2} R_j^V(t) \right) \right] \geq 0$$

and hence

$$S_i^V(t) \geq 0 \quad \text{as required.}$$

Now, define the vector

$$y := \begin{pmatrix} I \\ I^V \end{pmatrix}$$

Then, one can rewrite the equations for  $I_i$  and  $I_i^V$  in the form

$$\frac{dy}{dt} = M(S(t), S^V(t))y$$

for some matrix  $M$ , where, from the previous results

$$M_{ij} \geq 0 \quad \forall i \neq j.$$

Thus, from Lemma C.2,

$$y(t) \geq 0 \quad \forall t \geq 0.$$

Then,

$$\frac{dR_i}{dt} = \mu_i^1 I_i \geq 0 \quad \text{so} \quad R_i(t) \geq 0$$

and similarly,

$$R_i^V(t) \geq 0$$

and so the proof is complete.  $\square$

### C.4 Lemma C.4

**Lemma C.4** *For each  $i$ ,*

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

**Proof** Firstly, suppose

$$\lim_{t \rightarrow \infty} (\inf \{I_i(s) : s \geq t\}) = Q,$$

noting this infimum exists as  $I_i$  is bounded below by 0, and the limit exists as the sequence of infima given  $s \leq t$  is non-decreasing and bounded above by  $N_i$ . If  $Q \neq 0$ , there exists some  $m > 0$  and some  $t$  such that for all  $s \geq t$

$$I_i(s) \geq m \Rightarrow \frac{dR_i}{dt}(s) \geq m\mu_i^1 \Rightarrow R_i\left(t + \frac{2N_i}{m\mu_i^1}\right) > N_i$$

which contradicts Lemma C.3. Thus,  $Q = 0$  and so there exists some sequence  $t_n$  such that

$$\lim_{n \rightarrow \infty} (t_n) = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} (I(t_n)) = 0. \quad (48)$$

Now note that  $S_i(t)$  is non-increasing and bounded and that  $R_i(t)$  and  $(S_i^V(t) + I_i^V(t) + R_i^V(t))$  are non-decreasing and bounded. Thus, their limits as  $t \rightarrow \infty$  must exist and be finite, so in particular

$$\lim_{t \rightarrow \infty} (I_i(t)) = \lim_{t \rightarrow \infty} (N_i - S_i(t) - R_i(t) - S_i^V(t) - I_i^V(t) - R_i^V(t))$$

must exist. Thus, by (48), the only possible limit is 0 so

$$\lim_{t \rightarrow \infty} (I_i(t)) = Q = 0$$

as required. By noting that  $S_i(t) + S_i^V(t)$  is non-increasing and that  $I_i(t) + R_i(t)$  and  $R_i^V(t)$  are non-decreasing, an identical argument shows that

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

### C.5 Lemma C.5

**Lemma C.5** *Suppose that  $I_i(t) > 0$  for some  $t \geq 0$  and some  $i \in \{1, \dots, n\}$ . Then,*

$$I_i(s) > 0 \quad \forall s > t.$$

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

**Proof** Note that

$$\frac{dI_i}{dt} \geq -\mu_i^1 I_i$$

and so

$$\frac{d}{dt} \left( e^{\mu_i^1 t} I_i(t) \right) \geq 0$$

which means, for any  $s > t$

$$e^{\mu_i^1 s} I_i(s) \geq e^{\mu_i^1 t} I_i(t)$$

and hence

$$I_i(s) > 0$$

as required. The same argument then works for  $I_i^V(t)$  as well (with a  $\mu_i^2$  instead of a  $\mu_i^1$ ).  $\square$

## C.6 Lemma C.6

**Lemma C.6** *Define*

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

*Moreover, define*

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

*and the  $n$  by  $n$  matrix  $M$  by*

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

*Then, define the connected component  $C$  of  $\Pi^0$  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*

$$a_j \in \{1, \dots, n\} \quad \forall j \in \{1, \dots, k\}, \\ M_{a_1, a_2} M_{a_2, a_3} \dots M_{a_{k-1}, a_k} > 0$$

*and*

$$a_1 = i \text{ and } a_k \in \Pi^0.$$

Then,

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

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

Thus, in particular,

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

**Proof (a):** The proof will proceed by induction. For  $k \geq 1$ , define  $P^k$  is the set of elements of  $C$  that are connected to an element of  $\Pi^0$  by a sequence of length at most  $k$ . Then, note that

$$P^k \subseteq P^{k+1} \quad \forall k \geq 1$$

and

$$P^{n^2} = C$$

as there are  $n^2$  elements in  $M$ . (Thus, if  $i \in C$  then there must be a sequence of length at most  $n^2$  connecting  $i$  with an element in  $\Pi^0$  as any loops can be ignored.)

The inductive hypothesis is that

$$i \in P^k \Rightarrow I_i(t) > 0 \quad \forall t > 0.$$

The explanation of the base case will be left until the end of the proof. Suppose that this claim holds for some  $k \geq 0$ . If  $P^{k+1} = P^k$ , then

$$i \in P^{k+1} \Rightarrow i \in P^k \Rightarrow I_i(t) > 0 \quad \forall t > 0$$

and so the inductive step is complete. Otherwise, consider any  $i \in P^{k+1} \setminus P^k$ . Then, there exists some  $j$  such that

$$M_{ij} > 0 \quad \text{and} \quad j \in P^k.$$

Thus, by continuity, for sufficiently small  $\tau$ ,

$$t < \tau \Rightarrow S_i(t)\beta_{ij}^1 > 0$$

and indeed, by Boundedness Theorem, there exists some  $\chi > 0$  such that

$$S_i(t)\beta_{ij}^1 > \chi \quad \forall t \in [0, \tau].$$

Now, choose any  $\epsilon \in [0, \tau]$ . By Boundedness Theorem,  $I_i(t)$  achieves is bounded and achieves its maximum,  $\theta_\epsilon$  in the interval  $[0, \epsilon]$ . Moreover,  $\theta_\epsilon > 0$  as  $I_i(t) > 0$  in  $(0, \epsilon)$

by assumption. Thus, by continuity, there exists some non-empty region  $(\delta_\epsilon, \Delta_\epsilon)$  such that

$$t \in (\delta_\epsilon, \Delta_\epsilon) \Rightarrow I_i(t) > \frac{\theta_\epsilon}{2}.$$

Thus, in particular

$$\int_0^\epsilon S_i(t) \beta_{ij}^1 I_j(t) dt \geq \chi \int_{\delta_\epsilon}^{\Delta_\epsilon} I_j(t) dt \geq \frac{\chi \theta_\epsilon}{2} (\Delta_\epsilon - \delta_\epsilon) > 0.$$

Now, note that

$$\frac{dI_i}{dt} \geq S_i(t) \beta_{ij}^1 I_j(t) - \mu_i^1 I_i(t).$$

Suppose for a contradiction that  $I_i(t) = 0$  for all  $t \in [0, \epsilon]$ . Then,

$$\frac{dI_i}{dt} \geq S_i(t) \beta_{ij}^1 I_j(t) \Rightarrow I_i(\epsilon) \geq I_i(0) + \frac{\chi M_\epsilon}{2} (\Delta_\epsilon - \delta_\epsilon)$$

and hence,

$$I_i(\epsilon) > 0,$$

which is a contradiction. Thus, there exists a  $t \in [0, \epsilon]$  such that  $I_i(t) > 0$  and hence, by Lemma C.5,

$$I_i(t) > 0 \quad \forall t \in [\epsilon, \infty).$$

Thus, as  $\epsilon$  was any constant in the region  $(0, \tau)$ , and  $\tau > 0$ , this means that

$$I_i(t) > 0 \quad \forall t > 0$$

as required.

Finally, note that the base case  $k = 1$  can be proved in exactly the same way, except now  $j \in \Pi^0$  (but this still means that  $I_j(t) > 0$  for all  $t > 0$  by Lemma C.5), and so **(a)** has been proved.

**(b):** The previous work has shown that

$$C \subseteq \Pi.$$

Hence, as clearly  $\Pi^0 \subseteq \Pi$ , this means that

$$C \cup \Pi^0 \subseteq \Pi$$

and so it suffices to prove that

$$\Pi \subseteq C \cup \Pi^0.$$

That is, it suffices to prove

$$i \notin C \cup \Pi^0 \Rightarrow I_i(t) = I_i^V(t) = 0 \quad \forall t \geq 0.$$

To check that this solution satisfies the equations, one notes that, in this case, if  $i \notin C \cup \Pi^0$ , then

$$\begin{aligned} \frac{dI_i}{dt} &= \sum_{j=1}^n S_i(t) \beta_{ij}^1 I_j(t) + \sum_{j=1}^n S_i(t) \beta_{ij}^2 I_j^V(t) - \mu I_i(t) \\ &= \sum_{j \in C \cup \Pi^0} S_i(t) \beta_{ij}^1 I_j(t) + \sum_{j \in C \cup \Pi^0} S_i(t) \beta_{ij}^2 I_j^V(t) \end{aligned}$$

and, similarly,

$$\frac{dI_i^V}{dt} = \sum_{j \in C \cup \Pi^0} S_i^V(t) \beta_{ij}^3 I_j(t) + \sum_{j \in C \cup \Pi^0} S_i(t) \beta_{ij}^4 I_j^V(t),$$

as  $I_j(t) = I_j^V(t) = 0$  for all  $j \notin C \cup \Pi^0$ .

Now, suppose that  $i \notin C \cup \Pi^0$  and  $j \in C \cup \Pi^0$ . Then, by definition of  $C$ , this means that

$$M_{ij} = S_i(0) \beta_{ij}^1 = 0 \quad \forall j \in C \cup \Pi^0$$

and hence, as  $S_i$  is non-increasing and non-negative

$$S_i(t) \beta_{ij}^1 = 0 \quad \forall j \in C \cup \Pi^0.$$

Now, as  $\beta_{ij}^1 \geq \beta_{ij}^2 \geq 0$ , this means that

$$S_i(t) \beta_{ij}^2 = 0 \quad \forall j \in C \cup \Pi^0$$

so that

$$\sum_{j \in C \cup \Pi^0} S_i(t) \beta_{ij}^1 I_j(t) + \sum_{j \in C \cup \Pi^0} S_i(t) \beta_{ij}^2 I_j^V(t) = 0,$$

which means

$$\frac{dI_i}{dt} = 0 \quad \text{as required.}$$



Moreover, as  $S_i^V(0) = 0$ , it is necessary that

$$(S_i(0) + S_i^V(0))\beta_{ij}^1 = 0 \quad \forall j \in C \cup \Pi^0$$

so, as  $(S_i + S_i^V)\beta_{ij}^1$  is non-increasing and non-negative

$$(S_i(t) + S_i^V(t))\beta_{ij}^1 = 0 \quad \forall j \in C \cup \Pi^0$$

and hence, as  $S_i(t)$  is non-negative

$$(S_i^V(t))\beta_{ij}^1 = 0 \quad \forall j \in C \cup \Pi^0.$$

Thus, as  $\beta_{ij}^1 \geq \beta_{ij}^3 \geq \beta_{ij}^4 \geq 0$ , one has

$$\sum_{j \in C \cup \Pi^0} S_i^V(t)\beta_{ij}^3 I_j(t) + \sum_{j \in C \cup \Pi^0} S_i(t)\beta_{ij}^4 I_j^V(t) = 0$$

and hence

$$\frac{dI_i^V}{dt} = 0 \quad \text{as required.}$$

Then, one can separately solve the system for all  $j \in C \cup \Pi^0$  as the equations will now be independent of any indices  $i \notin C \cup \Pi^0$  (as they only depend on these indices via the  $I_i$  and  $I_i^V$  terms, which are identically zero). Thus, by the uniqueness of solution, one must have

$$i \in C \cup \Pi^0 \Rightarrow I_i(t) = I_i^V(t) = 0 \quad \forall t \geq 0$$

and hence part (b) is proved. Thus, the lemma has been proved.  $\square$

### C.7 Lemma C.7

**Lemma C.7** Consider a set  $C = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$  that is a Cartesian product of real intervals. Suppose that  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is differentiable with bounded derivatives in  $C$ . Then,  $f$  is Lipschitz continuous on  $C$  - that is, there exists some  $L > 0$  such that

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq L \sum_{i=1}^n |x_i - y_i| \quad \forall \mathbf{x}, \mathbf{y} \in C.$$

**Proof** Note that, by assumption, for each  $i$ ,

$$\frac{\partial f}{\partial x_i} \quad \text{is bounded in } C,$$

so define the global bound for all  $i$  to be  $M$ . Choose some  $\mathbf{x}, \mathbf{y} \in C$ . Define the points  $\mathbf{p}^k \in C$  for  $k = 0, 1, \dots, n$  by

$$\mathbf{p}_i^k = \begin{cases} \mathbf{y}_i & \text{if } i \leq k \\ \mathbf{x}_i & \text{otherwise} \end{cases}$$

and define the curve  $\gamma_i$  to be the straight line joining the point  $\mathbf{p}^{i-1}$  to the point  $\mathbf{p}^i$ . As  $C$  is a product of intervals, the  $\gamma_i$  lie entirely in  $C$ .

Define  $\Gamma$  to be the union of the curves  $\gamma_i$ , so that  $\Gamma$  joins  $\mathbf{p}^0 = \mathbf{x}$  to  $\mathbf{p}^n = \mathbf{y}$ . Then

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{y})| &= \left| \int_{\Gamma} \nabla f \cdot d\mathbf{x} \right| \\ &= \left| \sum_{i=1}^n \int_{\gamma_i} \nabla f \cdot d\mathbf{x} \right| \\ &= \left| \sum_{i=1}^n \int_{s=x_i}^{s=y_i} \frac{\partial f}{\partial x_i}(\mathbf{p}^{i-1} + (s - x_i)\mathbf{e}_i) ds \right| \\ &\leq \sum_{i=1}^n \left| \int_{s=x_i}^{s=y_i} \frac{\partial f}{\partial x_i}(\mathbf{p}^{i-1} + (s - x_i)\mathbf{e}_i) ds \right| \\ &\leq \sum_{i=1}^n \sup_{s \in C} \left| \frac{\partial f}{\partial x_i}(s) \right| |y_i - x_i| \\ &\leq M \sum_{i=1}^n |y_i - x_i| \end{aligned}$$

where  $\mathbf{e}_i$  is the  $i$ th canonical basis vector. Hence, the required Lipschitz continuity holds with  $M = L$ .  $\square$

### C.8 Lemma C.8

**Lemma C.8** *Define the set of functions*

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

where for each fixed  $\epsilon$ , these functions solve the model equations with parameters

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

initial conditions

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

and vaccination policy  $U(t; \epsilon)$ . Suppose that

$$\begin{aligned} |p(\epsilon) - p(0)| &\leq \epsilon \quad \forall p \in \mathcal{P}, \\ |f_i(0; \epsilon) - f_i(0; 0)| &\leq \epsilon \quad \forall f \in \mathcal{F} \end{aligned}$$

and that

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

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

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

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

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

**Proof** To begin, it is helpful to note that, by Lemma C.3,

$$f(t; \epsilon) \in [0, \max(N_i)] \quad \forall f \in \mathcal{F} \quad \text{and} \quad t \geq 0$$

and that, by assumption on the feasibility of  $U_i$

$$W(t; \epsilon) \in [0, \max(N_i)] \quad \forall t \geq 0.$$

Moreover, as the parameter values converge, it can be assumed that

$$p(\epsilon) \in [\alpha, \beta] \quad \forall \epsilon \geq 0 \quad \text{and} \quad p \in \mathcal{P}$$

for some  $\alpha, \beta \geq 0$ . Moreover, it can be assumed that, as each  $\mu_i^a > 0$ , there is some  $\gamma > 0$  such that  $\mu_i^a(\epsilon) > \gamma$  for all  $\epsilon \geq 0$ .

However, there is no condition on the maximal difference (at a point) between  $U_i(t; \epsilon)$  and  $U_i(t; 0)$ . To avoid this problem, it is helpful to consider the variable  $S_i^O := S_i + S_i^V$  instead of  $S_i^V$ . Then, the equations for  $S_i$  and  $S_i^O$  can be written as

$$\begin{aligned} S_i(t; \epsilon) &= \frac{S_i(0)(N_i - W_i(t; \epsilon))}{N_i} \exp \left[ - \sum_{j=1}^n \left( \frac{\beta_{ij}^1(\epsilon) R_j(t; \epsilon)}{\mu_j^1(\epsilon)} + \frac{\beta_{ij}^2(\epsilon) R_j^V(t; \epsilon)}{\mu_j^2(\epsilon)} \right) \right] \\ \frac{dS_i^O(t; \epsilon)}{dt} &= - \sum_{j=1}^n \left[ \left( \beta_{ij}^1(\epsilon) I_j(t; \epsilon) + \beta_{ij}^2(\epsilon) I_j^V(t; \epsilon) \right) S_i(t; \epsilon) \right] \\ &\quad - \sum_{j=1}^n \left[ \left( \beta_{ij}^3(\epsilon) I_j(t; \epsilon) + \beta_{ij}^4(\epsilon) I_j^V(t; \epsilon) \right) (S_i^O(t; \epsilon) - S_i(t; \epsilon)) \right]. \end{aligned}$$

Then, one can define

$$\mathbf{v} := (S^O, \mathbf{I}, \mathbf{I}^V, \mathbf{R}, \mathbf{R}^V)^T$$

and  $\mathbf{p}(\epsilon)$  to be a vector of the elements of  $\mathcal{P}$  at some  $\epsilon \geq 0$ . Then, (substituting for  $S$ ), the model equations can be written in the form

$$\frac{d\mathbf{v}(t; \epsilon)}{dt} = \Phi(\mathbf{v}(t; \epsilon), \mathbf{W}(t; \epsilon), \mathbf{p}(\epsilon))$$

where  $\Phi$  is a smooth function. Thus, from Lemma C.7, there exists some constant  $L$  such that, for  $\mathbf{v}$ ,  $\mathbf{W}$  and  $\mathbf{p}$  within the closed bounded feasible set of values and any  $j \in \{1, \dots, 5n\}$ ,

$$\begin{aligned} & |\Phi(\mathbf{v}, \mathbf{W}, \mathbf{p})_j - \Phi(\mathbf{v}^*, \mathbf{W}^*, \mathbf{p}^*)_j| \\ & \leq L \left( \sum_{i=1}^{5n} |v_i - v_i^*| + \sum_{i=1}^n |W_i - W_i^*| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right). \end{aligned}$$

Thus, in particular, this means that

$$\begin{aligned} \frac{d}{dt} \left( |v_j(t; \epsilon) - v_j(t; 0)| \right) & \leq \left| \frac{d}{dt} \left( v_j(t; \epsilon) - v_j(t; 0) \right) \right| \\ & \leq \left| \Phi(\mathbf{v}(t; \epsilon), \mathbf{W}(t; \epsilon), \mathbf{p}(\epsilon))_j - \Phi(\mathbf{v}(t; 0), \mathbf{W}(t; 0), \mathbf{p}(0))_j \right| \\ & \leq L \left( \sum_{i=1}^{5n} |v_i(t; \epsilon) - v_i(t; 0)| + \sum_{i=1}^n |W_i(t; \epsilon) - W_i(t; 0)| + \right. \\ & \quad \left. + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right). \end{aligned}$$

Now, adding these  $5n$  inequalities together, one seems that

$$\begin{aligned} & \frac{d}{dt} \left( \sum_{i=1}^{5n} |v_i(t; \epsilon) - v_i(t; 0)| \right) \\ & \leq 5nL \left( \sum_{i=1}^{5n} |v_i(t; \epsilon) - v_i(t; 0)| + \sum_{i=1}^n |W_i(t; \epsilon) - W_i(t; 0)| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right) \end{aligned}$$

and hence

$$\sum_{i=1}^{5n} \left[ \frac{d}{dt} \left( e^{-5nLt} |v_i(t; \epsilon) - v_i(t; 0)| \right) \right]$$

$$\begin{aligned} &\leq 5nLe^{-5nLt} \left( \sum_{i=1}^n |W_i(t; \epsilon) - W_i(t; 0)| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right) \\ &\leq (15n^2 + 20n^3)L\epsilon e^{-5nLt}. \end{aligned}$$

Thus, integrating (and using the fact that the initial conditions differ by at most  $\epsilon$ )

$$\begin{aligned} e^{-5nLt} \sum_{i=1}^{5n} |v_i(t; \epsilon) - v_i(t; 0)| &\leq \sum_{i=1}^{5n} |v_i(0; \epsilon) - v_i(0; 0)| + (3n + 4n^2)\epsilon(1 - e^{-5nLt}) \\ &\leq 5n\epsilon + (3n + 4n^2)\epsilon(1 - e^{-5nLt}) \end{aligned}$$

which means

$$\sum_{i=1}^{5n} |v_i(t; \epsilon) - v_i(t; 0)| \leq 5n\epsilon e^{5nLt} + (3n + 4n^2)\epsilon(e^{5nLt} - 1)$$

and hence, for each  $i \in \{1, \dots, 5n\}$

$$|v_i(t; \epsilon) - v_i(t; 0)| \leq 5n\epsilon e^{5nLt} + (3n + 4n^2)\epsilon(e^{5nLt} - 1).$$

The right-hand side is non-decreasing in  $t$  (as  $L > 0$ ) so, taking

$$\epsilon < \frac{\delta}{5ne^{5nLt} + (3n + 4n^2)(e^{5nLt} - 1)}$$

ensures that the required inequalities hold for  $\mathbf{I}$ ,  $\mathbf{I}^V$ ,  $\mathbf{R}$  and  $\mathbf{R}^V$  for all  $s \leq t$ . Now, note also that  $S_i(t; \epsilon)$  is a smooth function of  $W_i(t; \epsilon)$ ,  $\mathbf{v}(\epsilon)$ ,  $S_i(0; \epsilon)$  and  $\mathbf{p}$  so that there exists an  $L'$  such that

$$\begin{aligned} |S_i(t; \epsilon) - S_i(0; \epsilon)| &< L' \left( \sum_{i=1}^{5n} |v_i - v_i^*| + \sum_{i=1}^n |W_i - W_i^*| \dots \right. \\ &\quad \left. + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| + |S_i(0; \epsilon) - S_i(0; 0)| \right) \\ &< L'\epsilon \left[ 5ne^{5nLt} + (3n + 4n^2)(e^{5nLt} - 1) + (3n + 4n^2) + 1 \right] \\ &:= \chi(t)\epsilon \end{aligned}$$

and so, as  $\chi(t)$  is non-decreasing in  $t$ , taking

$$\epsilon < \frac{\delta}{\chi(t)}$$

gives the required inequalities for  $S$  for all times  $s \leq t$ . Finally, note that

$$\begin{aligned} |S_i^V(t; \epsilon) - S_i^V(t; 0)| &= |S_i^O(t; \epsilon) - S_i^O(t; 0) - S_i(t; \epsilon) + S_i(t; 0)| \\ &\leq |S_i^O(t; \epsilon) - S_i^O(t; 0)| + |S_i(t; \epsilon) - S_i(t; 0)| \\ &\leq +5n\epsilon e^{5nLt} + (3n + 4n^2)\epsilon(e^{5nLt} - 1) + \epsilon\chi(t) \end{aligned}$$

and so, as the right-hand side is increasing in  $t$ , taking

$$\epsilon < \frac{\delta}{5ne^{5nLt} + (3n + 4n^2)(e^{5nLt} - 1) + \chi(t)}$$

gives the required inequalities for  $S^V$  for all times  $s \leq t$  and hence completes the proof.  $\square$

## References

- Abakuks A (1972) Some optimal isolation and immunisation policies for epidemics. PhD thesis, University of Sussex
- 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
- Ball FG, Lyne OD (2002) Optimal vaccination policies for stochastic epidemics among a population of households. *Math Biosci* 177:333–354
- Bárcenas D (2000) The fundamental theorem of calculus for Lebesgue integral. *Divulgaciones Matemáticas* 8(1):75–85
- Becker NG, Starczak DN (1997) Optimal vaccination strategies for a community of households. *Math Biosci* 139(2):117–132
- Bonana C, García-Blas S, Tarazona-Santabalbina F et al (2020) The effect of age on mortality in patients with COVID-19: a meta-analysis with 611,583 subjects. *J Am Med Dir Assoc* 21(7):915–918
- 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*
- Castro MC, Singer B (2021) Prioritizing COVID-19 vaccination by age. *Proc Natl Acad Sci* 118(15):e2103700118
- Chandra J, Davis PW (1976) Linear generalizations of Gronwall's inequality. *Proc Am Math Soc* 60(1):157–160
- Chen L, Sun J (2014) Optimal vaccination and treatment of an epidemic network model. *Phys Lett A* 378(41):3028–3036
- Collins PJ (2006) *Differential and integral equations*. OUP, Oxford
- Dean NE, Halloran ME (2022) Protecting the herd with vaccination. *Science* 375(6585):1088–1089
- Duijzer LE, van Jaarsveld WL, Wallinga J et al (2018) Dose-optimal vaccine allocation over multiple populations. *Prod Oper Manag* 27(1):143–159
- Estadilla CDS, Uyheng J, de Lara-Tuprio EP et al (2021) Impact of vaccine supplies and delays on optimal control of the COVID-19 pandemic: mapping interventions for the Philippines. *Infect Dis Poverty* 10(04):46–59
- Fitzpatrick MC, Galvani AP (2021) Optimizing age-specific vaccination. *Science* 371(6532):890–891
- Fu Y, Jin H, Xiang H et al (2022) Optimal lockdown policy for vaccination during COVID-19 pandemic. *Financ Res Lett* 45(102):123
- Hansen E, Day T (2011) Optimal control of epidemics with limited resources. *J Math Biol* 62(3):423–451
- Hill AN, Longini IM Jr (2003) The critical vaccination fraction for heterogeneous epidemic models. *Math Biosci* 181(1):85–106
- Kar TK, Batabyal A (2011) Stability analysis and optimal control of an SIR epidemic model with vaccination. *Biosystems* 104(2–3):127–135

- Kuga K, Tanimoto J (2018) Impact of imperfect vaccination and defense against contagion on vaccination behavior in complex networks. *J Stat Mech: Theory Exp* 11:113402
- 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
- Liu Y, Gayle AA, Wilder-Smith A, et al (2020) The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med*
- Mathieu E, Ritchie H, Ortiz-Ospina E et al (2021) A global database of COVID-19 vaccinations. *Nat Hum Behav* 5(7):947–953
- 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
- Morton R, Wickwire KH (1974) On the optimal control of a deterministic epidemic. *Adv Appl Probab* 6(4):622–635
- Penn MJ, Donnelly CA (2023) Asymptotic analysis of optimal vaccination policies. *Bull Math Biol* 85(3):15
- 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
- Ritchie H, Mathieu E, Rod  s-Guirao L, et al (2020) Coronavirus pandemic (covid-19). Our world in data Schwarz B (1970) Totally positive differential systems. *Pac J Math* 32(1):203–229
- UN (2019) World population prospects - population division. <https://population.un.org/wpp/Download/Standard/Population/>
- Zaman G, Kang YH, Jung IH (2008) Stability analysis and optimal vaccination of an SIR epidemic model. *Biosystems* 93(3):240–249
- Zavrakli E, Parnell A, Malone D, et al (2021) Optimal age-specific vaccination control for COVID-19. arXiv preprint [arXiv:2104.15088](https://arxiv.org/abs/2104.15088)
- Zhang H, Yang Z, Pawelek KA et al (2020) Optimal control strategies for a two-group epidemic model with vaccination-resource constraints. *Appl Math Comput* 371(124):956
- Zhou Y, Yang K, Zhou K et al (2014) Optimal vaccination policies for an SIR model with limited resources. *Acta Biotheor* 62(2):171–181