Highlights

Optimality of Maximal-Effort Vaccination

Matthew J. Penn, Christl A. Donnelly

- It is proved that it is optimal to vaccinate at maximal effort
- A novel model of vaccination is explored with COVID-19 vaccine data
- Results are presented which hold for both the novel and classical vaccination models.

Optimality of Maximal-Effort Vaccination

Matthew J. Penn[a], Christl A. Donnelly[a,b]

[a]*Department of Statistics, University of Oxford, 29 St Giles', Oxford, OX1 3LB, United Kingdom*
[b]*MRC Centre for Global Infectious Disease Analysis, Imperial College London, School of Public Health, Medical School Building, Norfolk Place, Oxford, W2 1PG, United Kingdom*

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 are distributed randomly to the unvaccinated population and 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: Optimal Vaccination, SIR Models, Epidemic Control, Epidemiology

1. Introduction

The COVID-19 pandemic has illustrated the importance of quickly implementing vaccination policies which target particular groups within a population [1]. The difference in final infections between targeted policies and uniform distribution to the entire population can be significant [2], [3] 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 [4], [5], [6]. This model divides the population into different groups based on characteristics such as age or occupation. Each group is then further sub-divided 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 [7], as will be done in this paper.

While many other approaches have been developed either by adding compartments to the SIR framework [8] or using completely different models such as networks [9] or stochastic simulations [10], 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 [11] and [12], 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 [5] and [13] 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 [14].

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 [15].

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 [16] 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 [13], [17], [18] and [19] to a much wider class of vaccination policies, although the proof was still restricted to a single group and to perfect vaccination. Moreover, [13] 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 [20] 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 [21] and [22] for a two-group model with imperfect vaccination, in [23] for a general n-group model with perfect vaccination and in [24] 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, $\tilde{\boldsymbol{U}}$ vaccinates each

individual at least as early as another vaccination policy, $\boldsymbol{U}$, 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, and seemingly more accurate, 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 [13], [17] and [25]) assumes that the entire population is randomly vaccinated at a controllable rate. In practice, this is unrealistic, as once a person has been vaccinated once, that information is available to policy-makers so they can ensure that only the unvaccinated are given new doses (assuming that only one dose of the vaccine is required). There can be significant differences between the optimal vaccination policy according to the two models, and so it seems that the model introduced in this paper should be strongly considered for new optimal vaccination studies.

However, it is shown in this paper that, in the context of the theorems proved, modifying the model so that vaccines are assigned randomly to the unvaccinated population provides a strictly larger set of possible vaccination policies. In particular, this means that the results proved here also apply to the more commonly-used case, and so can also be used by those following the more commonly-used vaccination model.

Alongside proving that the final infected and recovered populations are non-increasing with increased and earlier vaccination effort, some cautionary contradictions 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 $i = 1, ..., n$:

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, 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 and were infected after being vaccinated.

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

$U_i(t)dt \times \mathrm{P}$ (A person in group i is in S_i | A person is in group i is unvaccinated)

which is equal to

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

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 "unvaccinated-only" model

This results in the following model, based on SIR principles

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

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

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

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

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

$$\frac{dR_i^V}{dt} = \mu_i^2 I_i^V, \tag{2}$$

where

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

Here, β_{ij}^α represents transmission from group j to group i and μ_i^α represents the infectious period (note that the superscript denotes different parameter values, so that β_{ij}^2 is not necessarily the square of β_{ij}^1!)

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

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

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 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), the following constraints are imposed on the parameters:

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

Note that there is no constraint on the ordering of β_{ij}^2 and β_{ij}^3. It is assumed for convenience that all variables except the S_i and I_i are initially zero. Moreover, it will be assumed that the populations and parameters have been scaled such that $N = 1$.

2.2. Fully Random Vaccination

A more common model of vaccination in the literature is the "fully random" vaccination model ([13], [17] and [25]), where vaccinations assigned to each subgroup are given randomly to all members of that subgroup. That is, the equation (1) becomes

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

where $U_i^*(t)$ is the vaccination rate in this model. Note that $U_i^*(t)$ is dimensionless, unlike $U_i(t)$ (which has units of population before scaling N to be 1) and so the number of vaccinations assigned by this model in the small time interval $[t, t+dt]$ is $N_i U_i^*(t)dt$. This model then corresponds to random vaccination as

$$U_i^*(t)S_i dt = N_i U_i^*(t)dt \times \frac{S_i}{N_i} = (\text{Doses given to group } i) \times (\text{Susceptible proportion of group } i).$$

As in [17], $U_i^*(t)$ can equivalently be interpreted as the percentage of susceptibles being vaccinated per unit time. However, even with this interpretation, [17] constrains $U_i^*(t)$ to be less than a fixed, constant value. This is still in line with the random vaccination interpretation, as this constraint could again be interpreted as simply the number of randomly-assigned vaccines that are "successful" in being given to susceptibles, given an overall constant rate of vaccination to the whole population.

2.3. Justification of the Unvaccinated-Only Model

A significant problem with the fully random model is that it assumes that the number of vaccines wasted when vaccinating a population grows very quickly as more people are vaccinated. For example, in a (mostly susceptible) population that has 50% vaccination coverage, then it assumed that it will take twice as much effort to vaccinate the same number of people as at the start of program. This seems unrealistic given the vaccination uptake for COVID-19, where demand remained high until a large proportion of the

population had been vaccinated, as shown in [26]. Moreover, [27] shows that in Canada, the number of wasted vaccines was 2.6% of the total given out when their adult population had been mostly vaccinated (at least once) in November 2021, and so it is not the case that the vaccine wastage scales in the way predicted by the fully random model.

Figure 1, made from the data used in [26] and created in [15] gives an explicit example of this. It compares the median time taken (in the 27 countries from the dataset that recorded their data in the same way) to give a first vaccine dose to different 10% percentiles of the population that were over 80 years old in each country. As a baseline, it uses the time taken to move from 20% vaccination coverage of this age group to 30% coverage (to avoid errors from the very small but rapidly growing initial vaccine supply). The median values shown in the figure are the median of the times

$$\frac{(\text{Time taken to move from } x\% \text{ to } (x+10)\% \text{ coverage})}{(\text{Time taken to move from } 20\% \text{ to } 30\% \text{ coverage})}$$

across the 27 different countries, where x denotes the value on the x axis. Assuming that vaccination effort was roughly constant in the period during which these vaccinations occurred (which is not a completely accurate assumption, but in general, those aged over 80 years were given their first doses in a short period of time), one would then expect that these proportions would be equal to 1 for all x under the unvaccinated-only model. However, under the fully random model, significantly more effort is required to vaccinate people once the total percentage vaccinated is higher. Indeed, if X is the unvaccinated percentage then one has (assuming $U_i^*(t)$ is a constant)

$$\frac{dX}{dt} = -U_i^* X,$$

and hence

$$X(t) = X(0)e^{-U_i^* t} \quad \Rightarrow \quad t = \frac{1}{U_i^*} \log\left(\frac{X(t)}{X(0)}\right).$$

These predicted times according to the random model are given by the red line in Figure 1.

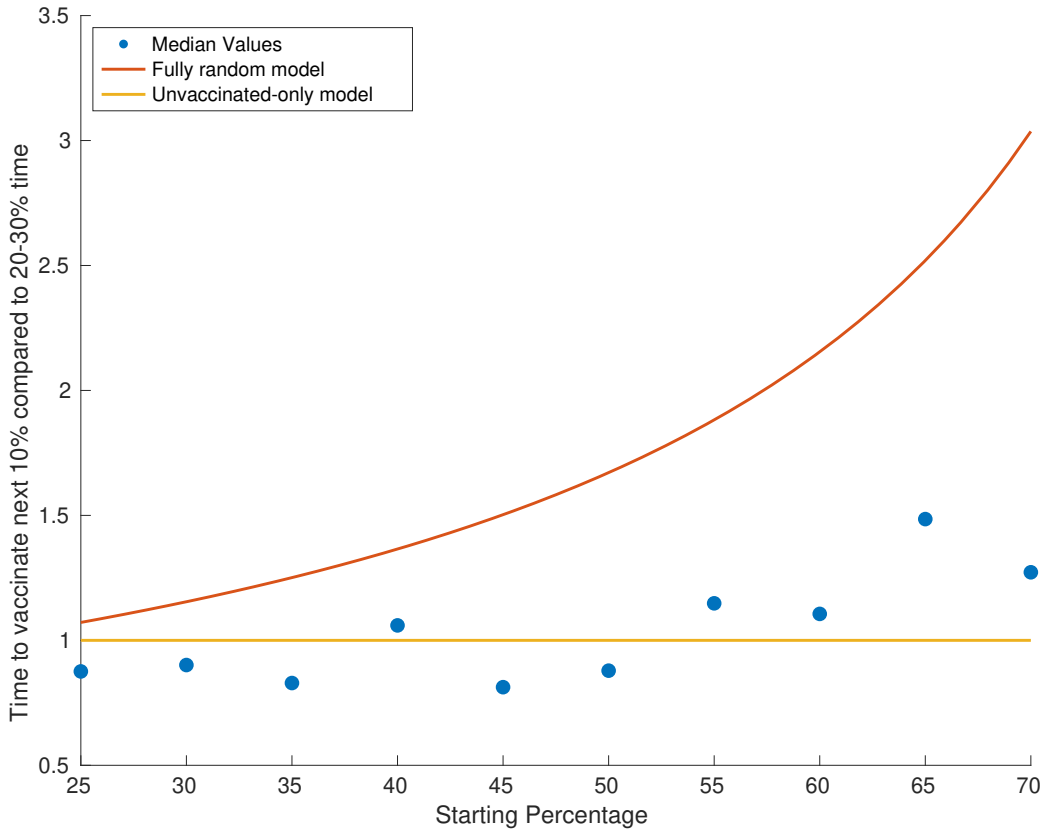

Figure 1: A comparison of the expected time to vaccinate the next 10% of those aged over 80 years according to the fully random model, unvaccinated-only model and median values from data taken from 27 countries.

It is clear from Figure 1 that the constant-time-interval model provides a much better prediction of the median time intervals from the data. Of course, it is important to model the fact that a significant proportion of each group in the population will remain unvaccinated, and that rates of uptake will decrease as the number of unvaccinated people decreases, evidence of which is apparent in Figure 1. However, both models could easily be modified to (more accurately) incorporate the fact that not everyone will get vaccinated by constraining $W_i(t) \leq \alpha_i N_i$ for some $\alpha_i < 1$. Furthermore, the demand could be modelled by introducing some general constraint $U_i(t) \leq K(W_i(t), t)$ which would again more accurately model this phenomenon.

Thus, it seems that the unvaccinated-only model does give a more accurate representation of a vaccination program than the more traditional

random model, although more detailed analysis would be needed to justify this further. Therefore, future studies should consider carefully their choice of vaccination function to ensure that their results are as practically applicable as possible.

2.4. The Importance of Choosing the Correct Model

The distinction between the two models is important, as there are many examples where the two models will give significantly different solutions to the optimal vaccination problem. Consider a two-group case with perfect vaccination (so $\beta_{ij}^{\alpha} = 0$ for $\alpha > 1$) with $N_1 = N_2 = 1$ (which means that U_i^* and U_i are constrained in the same way). Then, suppose the parameters are

$$\beta^1 = \begin{pmatrix} 4 & 0.5 \\ 0.5 & 0.25 \end{pmatrix} \quad \mu^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \boldsymbol{S}(0) = \begin{pmatrix} 1 - 10^{-4} \\ 1 - 10^{-4} \end{pmatrix} \quad \boldsymbol{I}(0) = \begin{pmatrix} 10^{-4} \\ 10^{-4} \end{pmatrix} \quad \boldsymbol{p} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

where $\boldsymbol{p}$ is the case fatality ratio in each group. Suppose that the objective is to minimise the final number of deaths, given by $R_2(\infty)$. Moreover, suppose that the overall vaccination rate is given by

$$U_1(t) + U_2(t) = U_1^*(t) + U_2^*(t) = \begin{cases} 10 & \text{if } t < 0.08 \\ 0 & \text{otherwise} \end{cases}$$

so that vaccination takes place almost instantly. Suppose that one has the choice between vaccinating exactly group 1 or group 2 at maximum rate until the supply runs out. Then, under the random vaccination model, it is not possible to reach herd immunity by vaccinating the infectious group 1, and so the optimal strategy is to vaccinate group 2. However, in the unvaccinated-only model it *is* possible to reach herd immunity and the optimal strategy is to vaccinate group 1. Indeed, if the random model is used to make the decision, but the unvaccinated-only model is the actual population response to vaccination, then the wrong decision is made and the number of deaths increases from 7×10^{-4} to 0.08. Thus, it is clearly important to ensure that the correct model is chosen.

2.5. Globality of Results in This Paper

However, the principle of maximal vaccination effort will hold for both models of vaccination, which can be shown as follows. The two models are equivalent whenever

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

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

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

Thus, given any fully random vaccination policy $\boldsymbol{U}^*$, it can be replaced by a unvaccinated-only policy $\boldsymbol{U}$ (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 fully random vaccination policies $\boldsymbol{U}$ and $\tilde{\boldsymbol{U}}$ satisfy $W_i(t) \leq \tilde{W}_i(t)$ then this inequality is preserved by the corresponding unvaccinated-only 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 an "unvaccinated only" 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 $i \in \{1, ..., n\}$

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

Moreover, it is assumed that the $U_i(t)$ are sufficiently smooth so that the solutions to the model equations exist and are unique. From the Picard-Lindelöf Theorem [28], 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.

Finally, it is assumed that $U_i(t)$ is bounded and Lebesgue integrable on $\Re$ for each i.

For the remainder of this paper, define the set of feasible vaccination policies, C, is the set of functions $\boldsymbol{U}$ 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 $\Re$.

3.2. Optimisation Problem

The aim is to choose the vaccination policy $\boldsymbol{U} \in C$ 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

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

It is assumed that each infection of unvaccinated people in group i is weighted by some p_i (this could be interpreted as a probability of death, hospitalisation or, if $p_i = 1$, simply counting the number of infections). Moreover, it is assumed 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

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

where, for example

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

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

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

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

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

Then

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

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* $\boldsymbol{U}$ *to (6). Then, define the function*

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

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

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

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

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

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*

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

for any function F. Then, with χ defined to be the maximal vaccination effort as in Theorem 2, there exists an optimal solution $\tilde{\boldsymbol{U}}$ 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 χ is continuous almost everywhere, then there is an optimal solution $\tilde{\boldsymbol{U}}$ such that

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

5. Proof

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 $i \in \{1, ..., n\}$ 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

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

These results are proved in Lemmas B.3 and B.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.

5.1. 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 A. 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 $\Re$. 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 $\Re$).

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

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

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

and that

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

Then,

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

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

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

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

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

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

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

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

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

$$0 \leq \kappa \leq \eta \leq 1. \tag{12}$$

The derivations of inequalities (7) - (12) are found in Lemmas A.2 - A.5. Moreover,

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

which comes from (3). Note that (10) 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 A.6, the system (7) - (13) implies that

$$I_b(T) + R_b(T) = \tilde{I}_b(T) + \tilde{R}_b(T), \tag{14}$$

$$I_b^V(T) + R_b^V(T) = \tilde{I}_b^V(T) + \tilde{R}_b^V(T), \tag{15}$$

$$S_b(T) + S_b^V(T) = \tilde{S}_b(T) + \tilde{S}_b^V(T), \tag{16}$$

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

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

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

Define functions Δ_i^f to be

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

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 A.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 A.9,

$$\sum_{j=1}^{n} K_{ij}(t) \geq \sum_{j=1}^{n} \tilde{K}_{ij}(t) \quad \forall t \in [0, T + \delta] \tag{17}$$

and

$$\sum_{j=1}^{n} L_{ij}(t) \geq \sum_{j=1}^{n} \tilde{L}_{ij}(t) \quad \forall t \in [0, T + \delta] \tag{18}$$

for sufficiently small δ.

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 A.10, there exists some $t \in (T, T+\delta)$ such that

$$R_b(t) + \kappa R_b^V(t) < \tilde{R}_b(t) + \kappa \tilde{R}_b^V(t) \quad \text{and} \quad I_b(t) + \eta I_b^V(t) \leq \tilde{I}_b(t) + \eta \tilde{I}_b^V(t) \tag{19}$$

for some

$$0 \leq \kappa \leq \eta \leq 1. \tag{20}$$

Now, by Lemmas A.2 - A.4 (which only require the properties (17) and (18)), the system of inequalities (7)-(9) holds for group b at time t. These can be combined with (19), (20) and (13) to use Lemma A.6, showing

$$\eta I_b^V(t) + \kappa R_b^V(t) = \eta \tilde{I}_b^V(t) + \kappa \tilde{R}_b^V(t) \tag{21}$$

$$I_b(t) + R_b(t) = \tilde{I}_b(t) + \tilde{R}_b(t). \tag{22}$$

By adding the inequalities in (19) 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, (21) and (22) 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 1. □

It is now possible to prove Theorem 1 under the extra restrictions given Proposition 1.

5.2. A Proof for a Restricted Parameter and Policy Set

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

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

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.

Proof: Note that, by Proposition 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 A.2, for each $i \in \{1, ..., n\}$

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

Combining this with the conservation of population equation, (13), 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 $t \to \infty$ and noting that the infections tend to zero by Lemma B.4 gives

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

Moreover, by Lemma A.5, for any $t \geq 0$ and $i \in \{1, ..., n\}$

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

as required. Also, taking $t \to \infty$ 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. □

By summing the i inequalities at $t = \infty$ from Proposition 2 (and using $\lambda = \kappa_i$), Theorem 1 holds under the additional conditions given in Proposition 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.

5.3. Continuous Dependence

Proposition 3. *Define the set of functions*

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

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

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

initial conditions

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

and vaccination policy $\boldsymbol{U}(t;\epsilon)$. 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 $i \in \{1, ..., n\}$ and $\epsilon \geq 0$,

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

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

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

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

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

5.4. Theorem 1

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

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

Then

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

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 1 and 2. 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 1 and 2 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

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

and

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

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

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

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.

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,

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

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

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

and

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

Then, taking $\epsilon \to 0$ and using Proposition 3 (noting that the perturbations to the parameters, initial conditions and vaccination policies are all bounded by a constant multiple of ϵ) 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 2. □

5.5. Theorem 2

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

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

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

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

Moreover, if $\chi(t)$ is continuous almost everywhere, there exists an optimal solution $\tilde{\boldsymbol{U}}$ 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 $\boldsymbol{U}$ is an optimal vaccination policy. To begin, it will be shown that the total vaccination rate χ 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 $\chi(t)$).

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

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

and so the claim is proved. $\square$

It is now important to show that χ 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

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

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 χ. 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 $\boldsymbol{U}$ 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

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

so that Q is the earliest time at which all of the vaccines given out by the $\boldsymbol{U}$ policy could have been administered. Note that both Q and q_∞ may be infinite. By taking the limit $t \to Q$, 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, $\tilde{\boldsymbol{W}}$ 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 χ 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 $i \in \{1, ..., n\}$.

Proof: If $t < Q$, then $W_i(q(t)) \leq N_i$ for all $t < Q$ by feasibility of $\boldsymbol{U}$.

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

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

Proof: Note that for $s, t < Q$, if M is a bound for χ (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 χ) 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 [29]. 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 $\tilde{\boldsymbol{W}}$ 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 χ 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 $\boldsymbol{W}$.

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

Proof: Note that, by maximality of χ, for $t < Q$,

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

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

$$\tilde{W}_i(t) \geq W_i(q_\infty) \tag{23}$$

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 $t \to Q$,

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

and so

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

Thus, using (23),

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

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

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

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

5.6. Theorem 3

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

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

for any function F. Then, with χ defined to be the maximal vaccination effort as in Theorem 2, there exists an optimal solution $\tilde{\boldsymbol{U}}$ 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 χ is continuous almost everywhere, then there is an optimal solution $\tilde{\boldsymbol{U}}$ 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 $\tilde{\boldsymbol{U}}$ in the interval $(0, Q)$ (where Q is defined in the proof of Theorem 2) such that

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

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,

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

and so

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

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

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 simulated examples.

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. Indeed, consider a one-group example (where subscripts will thus be ignored) and the case where the parameters are

$$\beta^1 = 5 \quad \beta^2 = \beta^3 = \beta^4 = 2 \quad \mu^1 = \mu^2 = 1 \quad \text{and} \quad N = 1$$

with non-zero initial conditions

$$S(0) = 1 - 10^{-4}, I(0) = 10^{-4}$$

and a vaccination policy of

$$U(t) = \begin{cases} 20 & \text{if } t < 0.05 \\ 0 & \text{otherwise} \end{cases}. \tag{24}$$

Then, Figure 2 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 moving the peak to 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.

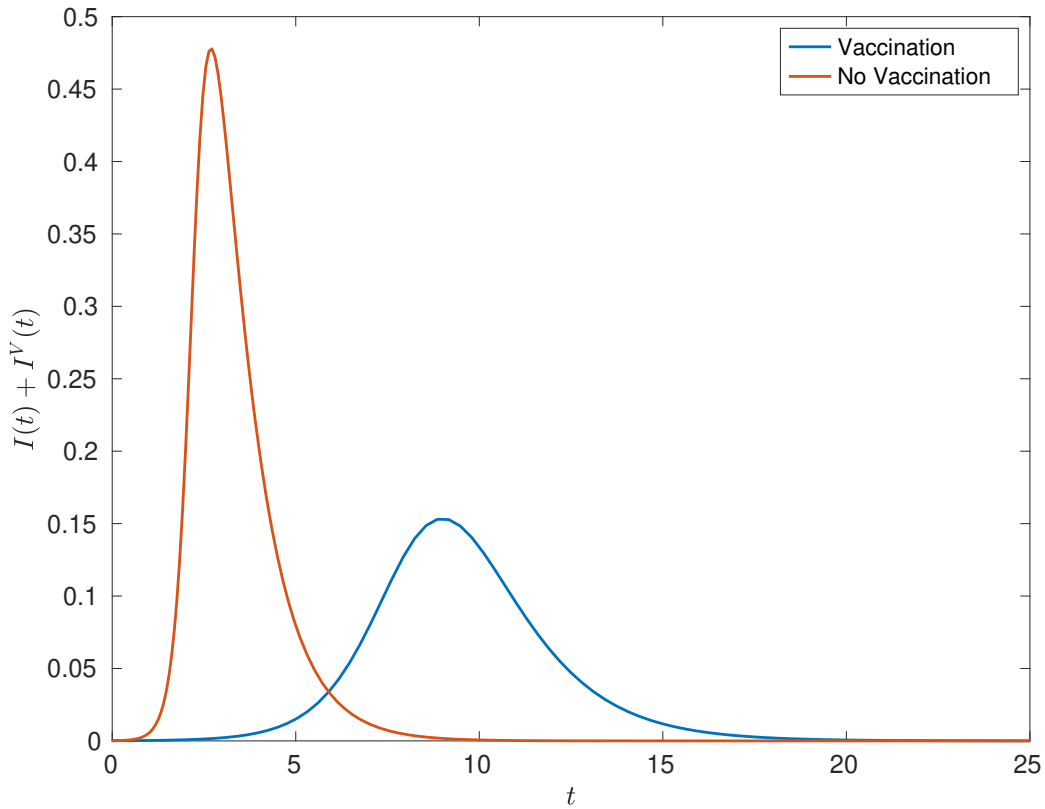

Figure 2: A comparison of the total infections over time for two, single group epidemics - one with vaccination and one without vaccination. The parameter values are $\beta^1 = 5 \quad \beta^2 = \beta^3 = \beta^4 = 2$, $\mu^1 = \mu^2 = N = 1$, $S(0) = 1 - 10^{-4}$, and $I(0) = 10^{-4}$. The vaccination policy (in the case of vaccination) is given by (24).

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. Indeed, consider the following one-group example where the parameters are

$$\beta^1 = \beta^2 = \beta^3 = \beta^4 = 3 \quad \mu^1 = 1 \quad \mu^2 = 2 \quad \text{and} \quad N = 1$$

and, as before, the non-zero initial conditions are

$$S(0) = 1 - 10^{-4}, I(0) = 10^{-4}$$

with a vaccination policy

$$U(t) = \begin{cases} 10 & \text{if } t < 0.05 \\ 0 & \text{otherwise} \end{cases}. \tag{25}$$

Suppose that this disease is always fatal, so that $R + R^V$ is the total number of deaths. Then Figure 3 shows that initially, the number of deaths is higher in the case of vaccination for some finite times. This occurs because the higher value of μ^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. Of course, this is not a realistic reflection of the course of an epidemic - the reason for μ^2 being higher is that vaccinated people are likely to get *less* ill rather than dying more quickly - but it illustrates that the model equations can have interesting, and unexpected properties.

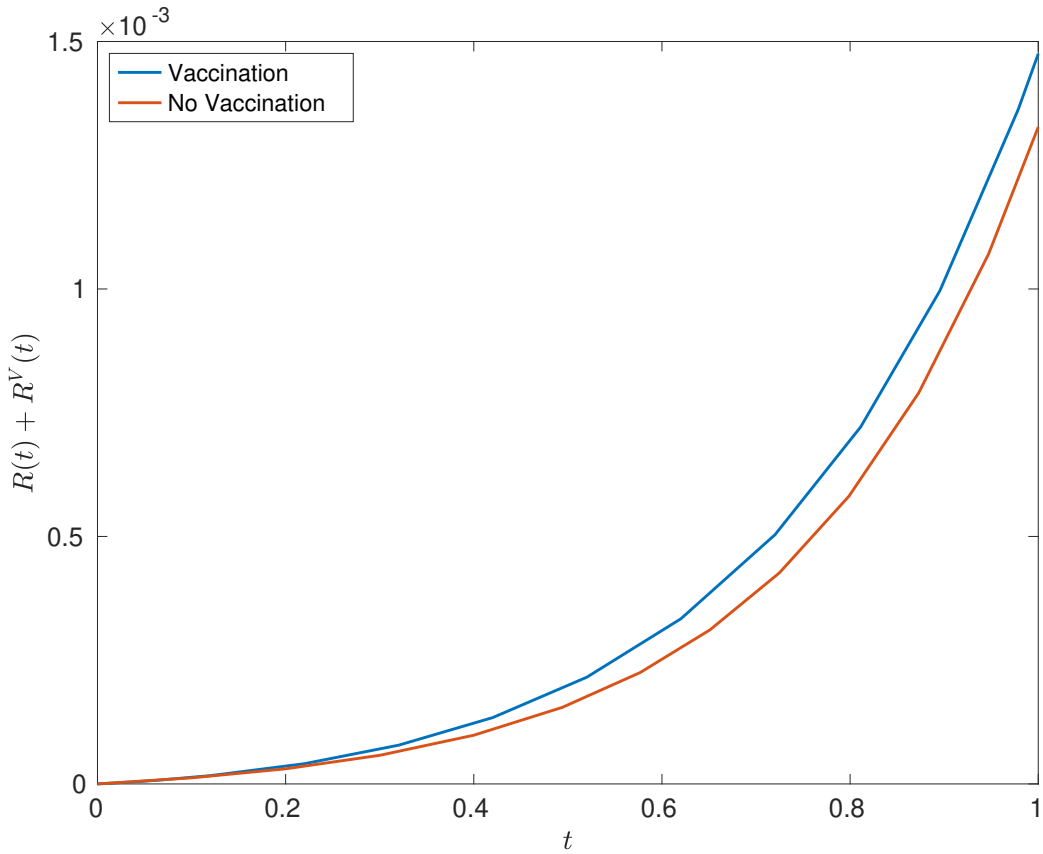

Figure 3: A comparison of the total deaths at early times for two single group epidemics - one with vaccination and one without vaccination. The parameter values are $\beta^1 == \beta^2 = \beta^3 = \beta^4 = 5$, $\mu^1 = 1$, $\mu^2 = 2$, $N = 1$, $S(0) = 1 - 10^{-4}$, and $I(0) = 10^{-4}$. The vaccination policy (in the case of vaccination) is given by (25).

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.

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.

It seems to the authors that future models for optimal vaccination should use the unvaccinated-only model used in in this paper, where vaccines are distributed randomly to the unvaccinated, rather than the entire population. While for small total vaccination numbers, the difference between the two models is small, when a significant proportion of the population is vaccinated, there can be large differences between the optimal policies, as was shown through the example in Section 2. Moreover, it appears to better reflect the vaccine uptake from the COVID-19 pandemic, as was illustrated in Figure 1.

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 $W_i \to N_i$. 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 "random vaccination model" found in the literature may overestimate the decay of vaccination uptake in a population and so the "unvaccinated-only model" presented here should be considered for finding optimal vaccination policies.
- 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.

References

[1] Meagan C Fitzpatrick and Alison P Galvani. "Optimizing age-specific vaccination". In: *Science* 371.6532 (2021), pp. 890–891.

[2] Marcia C Castro and Burton Singer. "Prioritizing COVID-19 vaccination by age". In: *Proceedings of the National Academy of Sciences* 118.15 (2021).

[3] Carlo Delfin S Estadilla et al. "Impact of vaccine supplies and delays on optimal control of the COVID-19 pandemic: mapping interventions for the Philippines". In: *Infectious Diseases of Poverty* 10.04 (2021), pp. 46–59.

[4] Vishaal Ram and Laura P Schaposnik. "A modified age-structured SIR model for COVID-19 type viruses". In: *Scientific Reports* 11.1 (2021), pp. 1–15.

[5] Daron Acemoglu et al. "Optimal targeted lockdowns in a multigroup SIR model". In: *American Economic Review: Insights* 3.4 (2021), pp. 487–502.

[6] Toshikazu Kuniya. "Global behavior of a multi-group SIR epidemic model with age structure and an application to the chlamydia epidemic in Japan". In: *SIAM Journal on Applied Mathematics* 79.1 (2019), pp. 321–340.

[7] Kazuki Kuga and Jun Tanimoto. "Impact of imperfect vaccination and defense against contagion on vaccination behavior in complex networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2018.11 (2018), p. 113402.

[8] Sam Moore et al. "Modelling optimal vaccination strategy for SARS-CoV-2 in the UK". In: *PLoS computational biology* 17.5 (2021), e1008849.

[9] Lijuan Chen and Jitao Sun. "Optimal vaccination and treatment of an epidemic network model". In: *Physics Letters A* 378.41 (2014), pp. 3028–3036.

[10] Frank G Ball and Owen D Lyne. "Optimal vaccination policies for stochastic epidemics among a population of households". In: *Mathematical Biosciences* 177 (2002), pp. 333–354.

[11] Andrew N Hill and Ira M Longini Jr. "The critical vaccination fraction for heterogeneous epidemic models". In: *Mathematical biosciences* 181.1 (2003), pp. 85–106.

[12] Niels G Becker and Dianna N Starczak. "Optimal vaccination strategies for a community of households". In: *Mathematical Biosciences* 139.2 (1997), pp. 117–132.

[13] Elsa Hansen and Troy Day. "Optimal control of epidemics with limited resources". In: *Journal of mathematical biology* 62.3 (2011), pp. 423–451.

[14] Yuting Fu et al. "Optimal lockdown policy for vaccination during COVID-19 pandemic". In: *Finance research letters* 45 (2022), p. 102123.

[15] Edouard Mathieu et al. "A global database of COVID-19 vaccinations". In: *Nature human behaviour* 5.7 (2021), pp. 947–953.

[16] Andris Abakuks. "Some optimal isolation and immunisation policies for epidemics". PhD thesis. University of Sussex, 1972.

[17] Gul Zaman, Yong Han Kang, and Il Hyo Jung. "Stability analysis and optimal vaccination of an SIR epidemic model". In: *BioSystems* 93.3 (2008), pp. 240–249.

[18] Richard Morton and Kenneth H Wickwire. "On the optimal control of a deterministic epidemic". In: *Advances in Applied Probability* 6.4 (1974), pp. 622–635.

[19] Yinggao Zhou et al. "Optimal vaccination policies for an SIR model with limited resources". In: *Acta biotheoretica* 62.2 (2014), pp. 171–181.

[20] Lotty E Duijzer et al. "Dose-optimal vaccine allocation over multiple populations". In: *Production and Operations Management* 27.1 (2018), pp. 143–159.

[21] Heting Zhang et al. "Optimal control strategies for a two-group epidemic model with vaccination-resource constraints". In: *Applied Mathematics and Computation* 371 (2020), p. 124956.

[22] Eleni Zavrakli et al. "Optimal age-specific vaccination control for COVID-19". In: *arXiv preprint arXiv:2104.15088* (2021).

[23] Hamza Boutayeb et al. "Automated optimal vaccination and travel-restriction controls with a discrete multi-region SIR epidemic model". In: *Commun. Math. Biol. Neurosci.* (2021).

[24] Sunmi Lee, Michael Golinski, and Gerardo Chowell. "Modeling optimal age-specific vaccination strategies against pandemic influenza". In: *Bulletin of mathematical biology* 74.4 (2012), pp. 958–980.

[25] Tapan Kumar Kar and Ashim Batabyal. "Stability analysis and optimal control of an SIR epidemic model with vaccination". In: *Biosystems* 104.2-3 (2011), pp. 127–135.

[26] Hannah Ritchie et al. "Coronavirus pandemic (COVID-19)". In: *Our world in data* (2020).

[27] Lauren Aubrey et al. "COVID-19 vaccine wastage in Canada, a reason for concern?" In: *Canadian Journal of Public Health* (2022), pp. 1–2.

[28] Peter J Collins. *Differential and integral equations*. OUP Oxford, 2006.

[29] Diómedes Bárcenas. "The fundamental theorem of calculus for Lebesgue integral." In: *Divulgaciones Matemáticas* 8.1 (2000), pp. 75–85.

[30] Jagdish Chandra and Paul W Davis. "Linear generalizations of Gronwall's inequality". In: *Proceedings of the American mathematical society* 60.1 (1976), pp. 157–160.

[31] DinosaurEgg. *Gronwall lemma for system of linear differential inequalities*. StackExchange. [Accessed: 2022 February 18] Available from: `https://math.stackexchange.com/questions/4090462/gronwall-lemma-for-system-of-linear-differential-inequalities`.

Acknowledgements

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

Funding

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC).

Appendix A. Supplementary Lemmas For Proposition 1 and 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

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

Moreover, note that, under the assumptions of Proposition 1 and 2, 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.

Appendix A.1. Lemma A.1

Lemma A.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.

Appendix A.2. Lemma A.2

Lemma A.2. *Suppose that*

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

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

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

Thus,

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

This means that

$$S_i(t)+S_i^V(t)=e^{-\mathcal{L}(t)}\left[S_i(0)-\int_0^t e^{\mathcal{L}(s)}\sum_{j=1}^{n}\left((\beta_{ij}^1-\beta_{ij}^3)I_j+(\beta_{ij}^2-\beta_{ij}^4)I_j^V\right)S_ids\right]$$

$$=S_i(0)\left[e^{-\mathcal{L}(t)}-\int_0^t e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)}(\mathcal{K}'(s)-\mathcal{L}'(s))\left(\frac{N_i-W_i(s)}{N_i}\right)\right]ds.$$

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

$$S_i(t)+S_i^V(t)\leq$$
$$S_i(0)-S_i(0)\int_0^t\left[\mathcal{L}'(s)e^{-\mathcal{L}(s)}+e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)}\sum_{j=1}^{n}\left(\mathcal{K}'(s)-\mathcal{L}'(s)\right)\right]\left(\frac{N_i-W_i(s)}{N_i}\right)ds.$$

This allows the use of Lemma A.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

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

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 A.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

$$\int_0^t \left[\mathcal{L}'(s)e^{-\mathcal{L}(s)} + e^{\mathcal{L}(s)-\mathcal{K}(s)-\mathcal{L}(t)} \sum_{j=1}^{n} \left(\mathcal{K}'(s) - \mathcal{L}'(s)\right)\right] \left(\frac{N_i - \tilde{W}_i(s)}{N_i}\right) ds \tag{A.1}$$

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

Thus, (as this integral is multiplied by -1 in (A.1)), combining this with (A.1) 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

Appendix A.3. Lemma A.3

Lemma A.3. *Suppose that*

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

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,

$$\begin{aligned}
I_i(t) + R_i(t) &= I_i(0) + \int_0^t \left[\sum_{j=1}^{n} K'_{ij}(s)\right]\left(\frac{S_i(0)(N_i - W_i(s))}{N_i}\right) e^{-\sum_{j=1}^{n} K_{ij}(s)} ds \\
&\geq \tilde{I}_i(0) + \int_0^t \left[\sum_{j=1}^{n} K'_{ij}(s)\right]\left(\frac{S_i(0)(N_i - \tilde{W}_i(s))}{N_i}\right) e^{-\sum_{j=1}^{n} K_{ij}(s)} ds,
\end{aligned} \tag{A.2}$$

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 A.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.

Appendix A.4. Lemma A.4

Lemma A.4. *Suppose that*

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

Then,

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

Proof: The result of Lemma A.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.

Appendix A.5. Lemma A.5

Lemma A.5. *Suppose that*

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

exists. Then, for some $b \in \{1,...,n\}$, and some real constants κ and η,

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

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

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 β_{ij}^α and μ_i^α

$$\kappa \leq \frac{\mu_b^1}{\mu_b^2}. \tag{A.3}$$

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

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

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

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

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

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

so

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

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 (A.3), 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

Appendix A.6. Lemma A.6

For the purposes of this lemma, it is helpful to recall the inequality system (7)-(13).

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

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

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

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

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

$$0 \leq \kappa \leq \eta \leq 1. \tag{12}$$

and

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

Lemma A.6. *Suppose that the system (7) - (13) holds for some $b \in \{1, ..., n\}$ 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 (7), (10) and (11) 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 (13) shows that

$$(\kappa - 1)R_b^V(T) + (\eta - 1)I_b^V(T) \leq (\kappa - 1)\tilde{R}_b^V(T) + (\eta - 1)\tilde{I}_b^V(T). \tag{A.4}$$

Moreover, adding (10) and (11) 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 (8) shows that

$$\eta I_b^V(T) + \kappa R_b^V(T) \leq \eta \tilde{I}_b^V(T) + \kappa \tilde{R}_b^V(T). \tag{A.5}$$

Now, from the inequality (10) combined with the inequality (9), it must be the case that

$$R_b^V(T) - \tilde{R}_b^V(T) \leq \frac{1}{\kappa}(\tilde{R}_b(T) - R_b(T)) \leq 0. \tag{A.6}$$

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 (12), (A.4), (A.5) and (A.6) reduces to

$$\begin{aligned}
(\kappa - 1)x + (\eta - 1)y &\leq 0 && \text{(A.7)} \\
\kappa x + \eta y &\leq 0 && \text{(A.8)} \\
x &\leq 0 \\
0 \leq \kappa \leq \eta &\leq 1.
\end{aligned}$$

Note first that $x = 0$ implies that $y = 0$ as η and $(\eta - 1)$ have different signs. Thus, in this case, the inequalities (A.7) and (A.8) 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 $-x > 0$,

$$\frac{(\kappa - 1)}{(\eta - 1)} \leq -\frac{y}{x} \leq \frac{\kappa}{\eta}. \tag{A.9}$$

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 (A.9), it is necessary that

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

This means that the inequalities (A.7) and (A.8) are satisfied to equality in this and hence, from before, all cases. Thus, it is necessary that

$$(\kappa - 1)R_b^V(T) + (\eta - 1)I_b^V(T) = (\kappa - 1)\tilde{R}_b^V(T) + (\eta - 1)\tilde{I}_b^V(T) \tag{A.10}$$

and

$$\eta I_b^V(T) + \kappa R_b^V(T) = \eta \tilde{I}_b^V(T) + \kappa \tilde{R}_b^V(T), \tag{A.11}$$

which is the first required equality. Thus, one can once again add the inequalities (10) and (11) 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

$$I_b(T) + R_b(T) \leq \tilde{I}_b(T) + \tilde{R}_b(T), \tag{A.12}$$

which, combined with (8), shows that

$$I_b(T) + R_b(T) = \tilde{I}_b(T) + \tilde{R}_b(T). \tag{A.13}$$

Moreover, one can subtract (A.10) from (A.11) to get

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

and then, using (13) alongside (A.12) and (A.13) shows

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

as required.

Appendix A.7. Lemma A.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 A.7. *Under the assumptions of Proposition 1, suppose that the system of inequalities (7) - (13) holds for some $b \in \{1, ..., n\}$ and some $T > 0$. Suppose further that*

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

Then,

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

Proof: By Lemma A.6, the system (14) - (16) must hold for b. Now, Equation (A.2) in the proof of Lemma A.3 shows that

$$I_b(T) + R_b(T) = \frac{S_b(0)}{N_b} \int_0^T \left[\sum_{k=1}^n K'_{bk}(s) \right] (N_b - W_b(s)) e^{-\sum_{j=1}^n K_{bk}(s)} ds. \tag{A.14}$$

Now, the equality (14) 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, (A.14) (and its tilde equivalent) shows that

$$\int_0^T \left[\sum_{k=1}^n K'_{bk}(s) \right] (N_b - W_b(s)) e^{-\sum_{k=1}^n K_{bk}(s)} ds = \int_0^T \left[\sum_{k=1}^n \tilde{K}'_{bk}(s) \right] (N_b - \tilde{W}_b(s)) e^{-\sum_{k=1}^n \tilde{K}_{bk}(s)} ds. \tag{A.15}$$

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

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

Thus,

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

In particular,

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

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 $W_b \leq \tilde{W}_b$

$$\begin{aligned}
&\int_0^T \left[\sum_{k=1}^{n} K'_{bk}(s)\right] (N_b - W_b(s)) e^{-\sum_{k=1}^{n} K_{bk}(s)} ds \\
&= \int_0^T \left[\sum_{k=1}^{n} K'_{bk}(s)\right] (N_b - \tilde{W}_b(s) + (\tilde{W}_b(s) - W_b(s)) e^{-\sum_{k=1}^{n} K_{bk}(s)} ds \\
&\geq \int_0^T \left[\sum_{k=1}^{n} K'_{bk}(s)\right] (N_b - \tilde{W}_b(s)) e^{-\sum_{k=1}^{n} K_{bk}(s)} ds + m \int_0^T \tilde{W}_b(s) - W_b(s) ds.
\end{aligned} \tag{A.16}$$

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 A.1, setting

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

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

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

and so, combining this with (A.16),

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

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 (A.15) 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

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

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 $k \in \{1, ..., n\}$,

$$K_{bk}(T) = \tilde{K}_{bk}(T) \quad \forall k \in \{1, ..., n\} \tag{A.17}$$

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

$$K'_{bk}(T) \leq \tilde{K}'_{bk}(T) \quad \forall k \in \{1, ..., n\}. \tag{A.18}$$

Dividing (A.17) by β^1_{bk}/μ^1_k and (A.18) by β^1_{bk} shows that the inequality system (7) - (13) holds for each k (as Lemmas A.2 - A.4 hold for any group) and so, following Lemma A.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.

Appendix A.8. Lemma A.8

Lemma A.8. *Define functions Δ^f_i to be*

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

and suppose that

$$\Delta^f_i(0) = 0 \quad \forall f \in \{S, I, R, S^V, I^V, R^V, W\}. \tag{A.19}$$

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

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

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

$$\left|\frac{dS_i^V}{dt}(t)\right| \leq U_i(T) + C, \quad \left|\frac{dI_i^V}{dt}(t)\right|, \left|\frac{dR_i^V}{dt}(t)\right|, \left|\frac{dI_i}{dt}(t)\right|, \left|\frac{dR_i}{dt}(t)\right| \leq C \quad \text{and} \quad \left|\frac{dW_i}{dt}(t)\right| \leq U_i(T).$$

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

$$f_i(T+t) = f_i(T) + O(\delta) \quad \forall f \in \{S, I, R, S^V, I^V, R^V, W\}. \tag{A.20}$$

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 (A.19) and (A.20), 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 Δ_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

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

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, μ_i^1 instead of μ_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,

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

as required.

Appendix A.9. Lemma A.9

Lemma A.9. *Suppose that*

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

exists. Define functions Δ_i^f *to be*

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

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,* $\Pi = \{1,...,n\}$ *and that*

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

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 δ.

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

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

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

$$\beta_{li}^1 \sum_{j=1}^n (K'_{ij}(T)) - \beta_{li}^2 \sum_{j=1}^n (L'_{ij}(T)) \leq 0 \Rightarrow \sum_{j=1}^n (K'_{ij}(T)) \leq \sum_{j=1}^n (L'_{ij}(T)). \tag{A.21}$$

Noting that

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

(A.21) requires

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

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

$$\beta^1_{ij} I_j(T) + \beta^2_{ij} I^V_j(T) = \beta^3_{ij} I_j(T) + \beta^4_{ij} I^V_j(T).$$

Thus, as $I_j(T) > 0$ (as $\Pi \in \{1, ..., n\}$) and $\beta^2_{ij} I^V_j(T) \geq \beta^4_{ij} I^V_j(T)$, it is necessary that

$$\beta^1_{ij} \leq \beta^3_{ij},$$

which is a contradiction. Thus,

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

which means

$$S_i(T)U_i(T) < S_i(T)\tilde{U}_i(T) \Rightarrow \frac{\beta^1_{ij}}{\mu^1_i} \Delta^R_i + \frac{\beta^2_{ij}}{\mu^2_i} \Delta^{R^V}_i = -C\delta^3 + O(\delta^4)$$

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 $\Delta^W_i(0) = 0$, one will have

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

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

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

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

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

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 δ. 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 A.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.

Appendix A.10. Lemma A.10

Lemma A.10. *Suppose that*

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

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

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

Proof: Firstly, note that by the definition of T, for each $\delta > 0$, there must exist $i, j \in \{1, ..., n\}$ 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 $b \in \{1, .., n\}$ such that

$$R_b(T+t) + \kappa R_b^V(T+t) < \tilde{R}_b(T+t) + \kappa \tilde{R}_b^V(T+t) \tag{A.22}$$

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 (A.22). 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 (τ, t) such that

$$\begin{aligned}
\mu_1^b \Delta_b^I(s) + \kappa \mu_b^2 \Delta_b^{I^V}(s) &= \frac{1}{t-\tau}\left[(\Delta_b^R(t) + \kappa \Delta_b^{R^V}(t)) - (\Delta_b^R(\tau) + \kappa \Delta_b^{R^V}(\tau))\right] \\
&= \frac{1}{t-\tau}\left[\Delta_b^R(t) + \kappa \Delta_b^{R^V}(t)\right] \\
&< 0
\end{aligned}$$

while also

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

by definition of τ. 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.

Appendix B. 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.

Appendix B.1. Lemma B.1

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

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

Then,

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

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 [30] where (in the notation of [30]), 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.

Appendix B.2. Lemma B.2

Lemma B.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 $\boldsymbol{v}$ and suppose that $\boldsymbol{y}$ and $\boldsymbol{z}$ solve the systems

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

with

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

Then,

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

Proof: Note that the general method for this proof was taken from [31]. To begin, define

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

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

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

and so

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

Thus, define

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

so

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

Similarly, defining

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

gives

$$\frac{d\boldsymbol{z}^*}{dt} = B(t)\boldsymbol{z}^*$$

while, moreover,

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

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

Now,

$$\frac{d\boldsymbol{y}}{dt} \geq B\boldsymbol{y} \Rightarrow \frac{d(-\boldsymbol{y})}{dt} \leq B(-\boldsymbol{y}),$$

which means (after integrating)

$$-\boldsymbol{y}(t) \leq -\boldsymbol{v} + \int_0^t B(-\boldsymbol{y}(s))ds$$

and so, from Lemma B.1, with $H(t) = B$, which is non-negative by assumption, $\boldsymbol{u} = -\boldsymbol{y}$ and $\boldsymbol{a} = -\boldsymbol{v}$,

$$-\boldsymbol{y}(t) \leq \left(I_n + \int_0^t V(t,s)Bds\right)(-\boldsymbol{v}).$$

Thus

$$\boldsymbol{y}(t) \geq \left(I_n + \int_0^t V(t,s)Bds\right)(\boldsymbol{v})$$

where

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

In this case, one can find an exact solution for $V(t,s)$. Consider

$$V(t,s) = e^{(t-s)B}$$

Then,

$$\begin{aligned} I_n + \int_s^t BV(k,s)ds &= I_n + \int_s^t Be^{(k-s)B}dk \\ &= I_n + \sum_{m=0}^{\infty} \int_s^t \frac{B^{m+1}(k-s)^m}{m!}dk \\ &= I_n + \sum_{m=0}^{\infty} \frac{B^{m+1}(t-s)^{m+1}}{(m+1)!} \\ &= e^{(t-s)B} \end{aligned}$$

as required. Thus, in particular, from the same logic

$$\int_0^t V(t,s)Bds = e^{(t-s)B} - I_n$$

and so

$$\boldsymbol{y}(t) \geq \left(I_n + e^{(t-s)B} - I_n\right)(\boldsymbol{v}) = e^{(t-s)B}\boldsymbol{v}.$$

Now, note that one can solve for $\boldsymbol{z}$ exactly as

$$\boldsymbol{z}(t) = e^{(t-s)B}\boldsymbol{v} \leq \boldsymbol{y}(t)$$

as required.

The final inequality follows from comparing the system for $\boldsymbol{z}$ to the system

$$\frac{d\boldsymbol{w}}{dt} = 0_{n\times n}\boldsymbol{w} \Rightarrow \boldsymbol{w} = \boldsymbol{v}$$

where here $0_{n\times n}$ denotes the $n \times n$ zero matrix. Thus, in particular (by now setting $A(t) = B$), one has

$$\boldsymbol{z}(t) \geq \boldsymbol{w}(t) = \boldsymbol{v} \geq \boldsymbol{0}$$

as required.

Appendix B.3. Lemma B.3

Lemma B.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 $i \in \{1, ..., n\}$,

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

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

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

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

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

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

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

for some matrix M, where, from the previous results

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

Thus, from Lemma B.2,

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

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.

Appendix B.4. Lemma B.4

Lemma B.4. *For each i,*

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

Proof: Firstly, suppose

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

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 B.3. Thus, $Q = 0$ and so there exists some sequence t_n such that

$$\lim_{n\to\infty}(t_n) = \infty \quad \text{and} \quad \lim_{n\to\infty}(I(t_n)) = 0. \tag{B.1}$$

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 $t \to \infty$ must exist and be finite, so in particular

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

must exist. Thus, by (B.1), the only possible limit is 0 so

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

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

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

Appendix B.5. Lemma B.5

Lemma B.5. *Suppose that $I_i(t) > 0$ for some $t \geq 0$ and some $i \in \{1, ..., n\}$. 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 μ_i^2 instead of a μ_i^1).

Appendix B.6. Lemma B.6

Lemma B.6. *Define*

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

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 Π^0 in M as follows. The index $i \in \{1, ..., n\}$ belongs to C if any only if there is some sequence $a_1, ..., a_k$ such that

$$a_j \in \{1, ..., n\} \quad \forall j \in \{1, ..., k\},$$

$$M_{a_1,a_2} M_{a_2,a_3} ..., M_{a_{k-1}a_k} > 0$$

and

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

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 Π^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 Π^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 $i \in P^{k+1}/P^k$. Then, there exists some j such that

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

Thus, by continuity, for sufficiently small τ,

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

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

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

Now, choose any $\epsilon \in [0, \tau]$. By Boundedness Theorem, $I_i(t)$ achieves is bounded and achieves its maximum, θ_ϵ 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 B.5,

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

Thus, as ϵ 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 B.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

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

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.

Appendix B.7. Lemma B.7

Lemma B.7. *Suppose that $f : \Re^n \to \Re$ is differentiable with bounded derivatives. Define C to be any closed bounded subset of $\Re^n$. Then, f is Lipschitz continuous on C - that is, there exists some $L > 0$ such that*

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

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. Then,

$$\begin{aligned}
|f(\boldsymbol{x}) - f(\boldsymbol{y})| &= \left| \sum_{i=1}^{n} \int_{s_i=x_i}^{s_i=y_i} \frac{\partial f}{\partial x_i}(\boldsymbol{s}) ds \right| \\
&\leq \sum_{i=1}^{n} \left| \int_{s_i=x_i}^{s_i=y_i} \frac{\partial f}{\partial x_i}(\boldsymbol{s}) ds \right| \\
&\leq \sum_{i=1}^{n} \sup_{\boldsymbol{s} \in C} \left| \frac{\partial f}{\partial x_i}(\boldsymbol{s}) \right| |y_i - x_i| \\
&\leq M \sum_{i=1}^{n} |y_i - x_i|
\end{aligned}$$

and hence, the required Lipschitz continuity holds with $M = L$.

Appendix B.8. Lemma B.8

Lemma B.8. *Define the set of functions*

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

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

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

initial conditions

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

and vaccination policy $\boldsymbol{U}(t;\epsilon)$. 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 $i \in \{1, ..., n\}$ and $\epsilon \geq 0$,

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

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

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

Proof: To begin, it is helpful to note that, by Lemma B.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

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

$$\begin{aligned}\frac{dS_i^O(t;\epsilon)}{dt} = &-\sum_{j=1}^n \left[\left(\beta_{ij}^1(\epsilon)I_j(t;\epsilon) + \beta_{ij}^2(\epsilon)I_j^V(t;\epsilon)\right)S_i(t;\epsilon)\right] \\ &-\sum_{j=1}^n \left[\left(\beta_{ij}^3(\epsilon)I_j(t;\epsilon) + \beta_{ij}^4(\epsilon)I_j^V(t;\epsilon)\right)\left(S_i^O(t;\epsilon) - S_i(t;\epsilon)\right)\right].\end{aligned}$$

Then, one can define

$$\boldsymbol{v} := (\boldsymbol{S}^O, \boldsymbol{I}, \boldsymbol{I}^V, \boldsymbol{R}, \boldsymbol{R}^V)^T$$

and $\boldsymbol{p}(\epsilon)$ to be a vector of the elements of $\mathcal{P}$ at some $\epsilon \geq 0$. Then, (substituting for $\boldsymbol{S}$), the model equations can be written in the form

$$\frac{d\boldsymbol{v}(t;\epsilon)}{dt} = \boldsymbol{\Phi}(\boldsymbol{v}(t;\epsilon), \boldsymbol{W}(t;\epsilon), \boldsymbol{p}(\epsilon))$$

where Φ is a smooth function. Thus, from Lemma B.7, there exists some constant L such that, for $\boldsymbol{v}$, $\boldsymbol{W}$ and $\boldsymbol{p}$ within the closed bounded feasible set of values and any $j \in \{1, ..., 5n\}$,

$$|\boldsymbol{\Phi}(\boldsymbol{v}, \boldsymbol{W}, \boldsymbol{p})_j - \boldsymbol{\Phi}(\boldsymbol{v}^*, \boldsymbol{W}^*, \boldsymbol{p}^*)_j| \leq L \left(\sum_{i=1}^{5n} |v_i - v_i^*| + \sum_{i=1}^{n} |W_i - W_i^*| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right).$$

Thus, in particular, this means that

$$\begin{aligned}
\frac{d}{dt}\Big(|v_j(t;\epsilon) - v_j(t;0)| \Big) &\leq \left| \frac{d}{dt}\Big(v_j(t;\epsilon) - v_j(t;0) \Big) \right| \\
&\leq \Big| \boldsymbol{\Phi}(\boldsymbol{v}(t;\epsilon), \boldsymbol{W}(t;\epsilon), \boldsymbol{p}(\epsilon))_i - \boldsymbol{\Phi}(\boldsymbol{v}(t;0), \boldsymbol{W}(t;0), \boldsymbol{p}(\epsilon))_i \Big| \\
&\leq L \left(\sum_{i=1}^{5n} |v_i(t;\epsilon) - v_i(t;0)| + \sum_{i=1}^{n} |W_i(t;\epsilon) - W_i(t;0)| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right).
\end{aligned}$$

Now, adding these $5n$ inequalities together, one seems that

$$\begin{aligned}
&\frac{d}{dt}\Big(\sum_{i=1}^{5n} |v_i(t;\epsilon) - v_i(t;0)| \Big) \\
&\leq 5nL \left(\sum_{i=1}^{5n} |v_i(t;\epsilon) - v_i(t;0)| + \sum_{i=1}^{n} |W_i(t;\epsilon) - W_i(t;0)| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right)
\end{aligned}$$

and hence

$$\begin{aligned}
\sum_{i=1}^{5n} \left[\frac{d}{dt} \left(e^{-5nLt} |v_i(t;\epsilon) - v_i(t;0)| \right) \right] &\leq 5nLe^{-5nLt} \left(\sum_{i=1}^{n} |W_i(t;\epsilon) - W_i(t;0)| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| \right) \\
&\leq (15n^2 + 20n^3) L\epsilon e^{-5nLt}.
\end{aligned}$$

Thus, integrating (and using the fact that the initial conditions differ by at most ϵ)

$$\begin{aligned}
e^{-5nLt} \sum_{i=1}^{5n} |v_i(t;\epsilon) - v_i(t;0)| &\leq \sum_{i=1}^{5n} |v_i(0;\epsilon) - v_i(0;0)| + (3n + 4n^2)\epsilon(1 - e^{-5nLt}) \\
&\leq 5n\epsilon + (3n + 4n^2)\epsilon(1 - e^{-5nLt})
\end{aligned}$$

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 $i \in \{1, ..., 5n\}$

$$|v_i(t;\epsilon) - v_i(t;0)| \leq 5n\epsilon e^{5nLt} + (3n + 4n^2)\epsilon(e^{5nLt} - 1).$$

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 $\boldsymbol{I}$, $\boldsymbol{I}^V$, $\boldsymbol{R}$ and $\boldsymbol{R}^V$ for all $s \leq t$. Now, note also that $S_i(t;\epsilon)$ is a smooth function of $W_i(t;\epsilon)$, $\boldsymbol{v}(\epsilon)$, $S_i(0;\epsilon)$ and $\boldsymbol{p}$ so that there exists an L' such that

$$\begin{aligned}
|S_i(t;\epsilon) - S_i(0;\epsilon)| &< L' \left(\sum_{i=1}^{5n} |v_i - v_i^*| + \sum_{i=1}^{n} |W_i - W_i^*| + \sum_{i=1}^{4n^2+2n} |p_i - p_i^*| + |S_i(0;\epsilon) - S_i(0;0)| \right) \\
&< L'\epsilon \left[5ne^{5nLt} + (3n + 4n^2)(e^{5nLt} - 1) + (3n + 4n^2) + 1 \right] \\
&:= \chi(t)\epsilon
\end{aligned}$$

and so, as $\chi(t)$ is non-decreasing in t, taking

$$\epsilon < \frac{\delta}{\chi(t)}$$

gives the required inequalities for $\boldsymbol{S}$ 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 $\boldsymbol{S}^V$ for all times $s \leq t$ and hence completes the proof.