Optimal Treatment of an SIS Disease with Two Strains

Selma Telalagic*

August 16, 2011

Abstract

This paper explores optimal treatment of an SIS (Susceptible-Infected-Susceptible) disease that has two strains with different infectivities. When we assume that neither eradication nor full infection are possible, it is shown that there are two categories of equilibria. First, there are two continua of interior equilibria characterised by a fixed, positive total level of infection, where both strands of the disease prevail. Second, there are two sets of equilibria where one strand of the disease is eradicated asymptotically. The feasibility of equilibria depends on parameter assumptions; a combination of low natural rate of recovery and large difference between infectivities renders the interior equilibria and one of the asymptotic equilibria infeasible. Optimal policy under different parameter assumptions is analysed by means of simulations.

Keywords: Epidemiological modelling, Optimal control theory, Simulations

JEL Classification: I19, C61, C63

*Faculty of Economics, University of Cambridge. The author is grateful to Prof Robert Rowthorn and Flavio Toxvaerd for their valuable guidance.

1 Introduction

Epidemiology, as it is studied today, originated in the early 20th century and has since developed into a multi-faceted field that combines the skills of mathematicians, biologists and, most recently, economists. One predominant area of epidemiology focuses on transmission system models. These models are built on differential equations that describe the evolution of disease prevalence over time as a function of parameters. Often a point of criticism, these models assume homogeneous mixing within populations, identical agents and no behavioural adaptation. Although this produces the benefit of parsimony, allowing significant predictive power and the ability to work with data, there is a strand of the literature that argues this simplicity comes at the price of applicability (Epstein 2009). Nevertheless, few advances have been made in other approaches to epidemiology that have received the backing that these types of models have.

In the standard Susceptible-Infected-Susceptible (SIS) model and its numerous permutations, the probability of an individual catching a disease when he encounters an infected person depends on an exogenous infectivity or transmission parameter. This parameter is predominantly assumed to be homogeneous, a simplification that does not allow for policy differentiation if there exist several strands of the infection. Infections in reality are frequently present in more than one form. To motivate an infection stratified by transmission parameter, consider the case of influenza as an example. Influenza has several strands and some strands are more infectious than others; typically, strains that originate in humans are more infectious than those that originate in other animals, such as birds. How does a policymaker deal with the presence of several variants of the infection in the population? Supposing she can differentiate policy by infection type, does she treat the more infectious or less infectious first? What is the prevalence of the different infection strands in equilibrium?

The goal of this paper is to explore an SIS model with two infection strains and answer these important policy questions. This paper shows that this model has two categories of steady state. First, there are two continuum of steady states where both infection strains prevail. Second, there are asymptotic equilibria where one of the strains is eliminated asymptotically, while the other is endemic. Under certain parameter assumptions, it is optimal to asymptotically eliminate the less infectious strand while allowing the high infectivity strand to prevail. This interesting case is explored by way of simulations, where optimality under fixed policy and variable policy is explored. The role of cost of treatment in governing optimal policy is explored in detail. The rest of the paper is structured as follows. Section 2 reviews the relevant literature. Section 3 introduces the basic model and Section 4 develops the model to encompass two infection types. Section 5 provides examples of simulations. Section 6 concludes.

2 Relevant literature

The literature on epidemiological modelling is primarily found in the mathematical and biological fields. Variations on standard epidemiological models are common in the mathematical literature in particular, where researchers detail dynamics and equilibria but do not look at optimality and intervention. In direct relevance to this paper, Castillo-Chavez, Huang and Li (1999) develop an SIS model with a two-strand disease where individuals are genetically predisposed to a specific strand. They derive stability conditions on the various equilibria of the model, which include boundary (one or both strands eradicated) and coexistence (both strands prevail) equilibria. These are equilibria the system tends towards when left on its own. Hyman and Li (1997) analyse an SIS STD model with multiple groups where interaction between groups is behaviourally variable and depends on prevalence levels in the different groups. The development of the infection is complex and depends on how these interactions take place; individuals may reduce their contacts with individuals in high prevalence groups, which may reduce overall prevalence. However, this is not guaranteed. The biological literature primarily applies existing models to data and attempts to model the evolution of specific diseases. In relation to the present study, the mathematical epidemiology literature is most relevant. This paper develops a mathematical model akin to those studied by mathematical epidemiologists, while simultaneously introducing the economic consideration of optimal intervention.

Economic research into epidemiology is scant but fast-growing. Research has focused on three main areas: behavioural response, spatial analyses and empirical work. Modelling behaviour is tackled by some of the literature. Agents make decisions about risky and preventative behaviours as a function of perceived risks of contracting infections. Philipson and Posner (1993) introduce an expected-utility model of STDs where agents evaluate whether or not to use protection in sexual activity based on perceived probabilities of being infected and the partner being infected. They examine the idea that epidemiological models predict incidence will rise without bound until all individuals are infected, while economic models predict that incidence first rises then falls, once risky activities become too costly. There may even be a threshold prevalence level after which risky activities are again increasingly demanded due to "fatalistic" beliefs.

Geoffard and Philipson (1996) use data from the San Francisco Men's Health Study to support the idea that behaviour responds to prevalence levels. Kremer (1996) looks at behavioural choices made by individuals about how many partners to have relations with depending on observed infection prevalence rates. The model shows that increased prevalence could lead to even worse prevalence rates due to unfavourable behavioural responses. Building on these findings, Auld (2006) examines changes in risky sexual behaviour in response to changes in local HIV prevalence using the San Francisco study of the 1980s. His results are consistent with Kremer's theoretical predictions. Auld estimates a 5% reduction in the rate of partner change in response to a 10% increase in preva-

lence. Another approach is to examine the role of agents' beliefs in partner and relationship choice. Greenwood et al (2010) examine individuals' decisions on whether to engage in different types of relationships, using a beliefs-based model that is parameterised to data on Malawi. A significant amount of literature addresses the role of behavioural response in epidemiology, with broadly consistent findings that behaviour is prevalence-dependent.

Several extensions can be made to basic epidemiology models that look at aspects other than behaviour. Spatial considerations are a primary factor. Rowthorn, Laxminarayan and Gilligan (2009) focus on the spatial dynamics of disease. They answer the question of optimal control of infections via treatment in the case of metapopulations, defined as subpopulations within a population that mix at a lower rate than individuals within each subpopulation. Although intuition may suggest that equalising infection rates across subpopulations leads to the highest level of welfare, this turns out to be the worst possible solution. Another policy-relevant aspect is the role of budget constraints. Rowthorn (2004) examines this in the context of optimal control of a disease using treatment. Funds should never be retained as long as there are people that can be treated.

Significant empirical work has been carried out on infectious disease. Several studies have been carried out aiming to verify the responsiveness of sexual behaviour to changes in perceived risk of contracting infections such as HIV/AIDS. St. Lawrence et al (1991) look at differences in risky behaviour across two cities with different prevalence rates. They find startling differences with risky behaviour being as much as three times more common in the low-prevalence city as compared to the high-prevalence city. Similarly, Dupas (2005) looks at whether a public health information program that teaches teenagers about relative risks of contracting HIV/AIDS depending on partner age group has an effect on their behaviour. Dupas finds that the information campaign reduces childbearing by 1.7% in the treatment group, representing a 31% decrease in childbearing. In terms of age group, there is a reduction in cross-generational pregnancies of 65%.

Oster (2005) provides a detailed simulation-based analysis on the effects of changes in transmission rates and partner choice on national HIV/AIDS prevalence levels. Using actual transmission rates and sexual behaviour parameters, the paper predicts an HIV/AIDS infection prevalence of 0.23% in the United States and 12.7% in Africa, close to actual prevalence rates of 0.15% and 11.9%. Estimates are then carried out using US sexual behaviour parameters but Sub-Saharan African transmission rates. This results in an estimated prevalence rate of over 11% for the United States, suggesting that it is the transmission rate that is driving the higher HIV/AIDS prevalence rates observed in Africa when compared to the United States. This shows that the transmission rate is an important predictor of prevalence levels and should be measured as accurately as possible.

The economic literature has focused on behaviour, spatial factors and empirical studies. Although these are not the focus of the present study, existing studies are useful to bear in mind as they have important intuition and ideas

that may become relevant here.

3 The basic SIS model

3.1 Overview

It has become standard in the economic epidemiology literature to assume random mixing between individuals, despite the large literature on the importance of behavioural response. The basic SIS model with treatment is derived under the assumption of random mixing in Section 3.1, following Rowthorn (2004) and Goldman and Lightwood (2002). This basic model also assumes a homogeneous transmission parameter. This assumption does not provide an accurate representation of the way disease spreads when it exists in different forms. Indeed, awareness of the improved predictions resulting from accurate transmission parameters has been raised by Oster (2005). These observations provide the impetus for an extension to the basic SIS model, which is presented in Section 4.

We begin by examining the standard SIS model. The model is in continuous time. There are two possible states: individuals are susceptible (proportion S of the total population) or infected (proportion I). They can move between the two states an unlimited number of times. Agents are homogeneous and the population is closed. Perfect or homogeneous mixing is assumed between agents, with a uniform transmission probability (β). A proportion f of infected individuals is treated, with the success rate of treatment (which can be interpreted as a rate of recovery) given by the parameter α. There is also the possibility of spontaneous or natural recovery, at rate τ. The evolution of the two populations, susceptible and infected, is described by the following differential equations:

$$\dot{I}(t) = I(t)S(t)\beta - I(t)(f(t)\alpha + \tau), \tag{1}$$

$$\dot{S}(t) = I(t)(f(t)\alpha + \tau) - I(t)S(t)\beta. \tag{2}$$

This type of model has been studied for decades by mathematical and biological epidemiologists. The economic approach to epidemiology has focused on two different adjustments to this model: adding an objective function, and introducing behavioural response. Behavioural response is not considered here. Instead, we focus on optimal policy via the introduction of an objective function. Objective functions can take many forms, from a social planner's welfare maximisation function, to an individual's utility maximisation function. There is also the possibility of cost minimisation, prevalence minimisation, and so on

and so forth. One natural objective function to add to this model is social welfare, determined by the proportion of infected and susceptible individuals and the expenditure on treatment:

$$W = \int_0^\infty e^{-\delta t}[pN(1 - I(t)) - cf(t)I(t)]dt. \tag{3}$$

In this simple case, infected individuals have a value of zero while susceptible individuals have a value of p; treatment has a constant marginal cost of c per instant per individual. The problem is solved as a Hamiltonian optimal control problem, normalising population to 1: $S(t) + I(t) = N = 1$ for all t. This allows (1) and (2) to collapse to one constraint.

It is assumed that

$$\begin{aligned} f &\in [0, 1] \\ I(0) &= I^0 > 0 \text{ given} \end{aligned}$$

The current value Hamiltonian function is

$$H = p(1 - I) - cfI + \gamma I((1 - I)\beta - f\alpha - \tau) \tag{4}$$

where γ is the shadow price of infection. Differentiating the Hamiltonian with respect to the control variable gives us the solution, which is of "bang-bang" form as the control enters the problem in linear fashion:

$$f^* \begin{Bmatrix} = 0 \\ \in (0, 1) \\ = 1 \end{Bmatrix} if\ \gamma^* \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} - \frac{c}{\alpha}. \tag{5}$$

Policy can either be at an interior level $f^* \in (0, 1)$, or at a boundary level, $f^* = 0$ or 1. The interpretation is as follows. The multiplier is the shadow price of another infected individual. The higher is this shadow price in absolute terms, the more costly it is to social welfare to have an additional infected person. On the other hand, $\frac{c}{\alpha}$ is the relative price of treating an infected individual - it is the ratio of cost to treatment effectiveness. Bearing in mind that the Hamiltonian condition compares the shadow price to the *negative* of the price of treatment, it is clear that if the price of infection exceeds the price of treatment, everyone is treated. Similarly, when the price of treatment is higher than the price of infection, no one is treated. When they are equal, any interior level of treatment is optimal subject to parameters. The equation of motion for the multiplier is

$$\dot{\gamma} = \delta\gamma - \frac{\partial H}{\partial I} = p + cf - \gamma((1-2I)\beta - f\alpha - \tau - \delta). \tag{6}$$

Let us examine the cases of interior and boundary policies more closely.

3.2 Policy is interior

For an interior policy to be optimal, the Hamiltonian conditions require that $\gamma = -\frac{c}{\alpha}$. Differentiating this gives us $\dot{\gamma}(t) = 0$. Further, it must be that $\dot{I}(t) = 0$ if we are in steady state. These three conditions give us steady state solutions $I = I^*, \gamma = \gamma^*$ and $f = f^*$:

$$I^* = \frac{\alpha p + c(\beta - \delta - \tau)}{2c\beta}, \tag{7}$$

$$f^* = \frac{c(\beta + \delta - \tau) - \alpha p}{2c\alpha}, \tag{8}$$

$$\gamma^* = -\frac{c}{\alpha}. \tag{9}$$

Thus, a path with interior policy has $I = I^*, \gamma = \gamma^*$ and $f = f^*$. Note that f^* may lie outside the range $(0, 1)$, in which case no feasible interior policy exists.

3.3 Policy is at a boundary

There are two feasible boundary policies that can be optimal in steady state: $f = 0$ or $f = 1$. Consider the case where $\gamma > -\frac{c}{\alpha}$. Under this policy it must be that $f = f^{**} = 0$. Solving $\dot{I}(t) = 0$ yields

$$I^{**} = 1 - \frac{\tau}{\beta}. \tag{10}$$

The disease is endemic as long as $\tau < \beta$. It is eradicated if $\tau \geqslant \beta$. Setting $\dot{\gamma}(t) = 0$ yields

$$\gamma^{**} = \frac{p}{\tau - \beta - \delta}. \tag{11}$$

Another possibility is that $\gamma > -\frac{c}{\alpha}$. In this case, $f = f^{***} = 1$. Solving $\dot{I}(t) = 0$ and $\dot{\gamma}(t) = 0$ yields

$$I^{***} = 1 - \frac{\alpha + \tau}{\beta},$$

$$\gamma^{***} = \frac{p + c}{\alpha + \tau - \beta - \delta}.$$

The disease is endemic as long as $\alpha + \tau < \beta$. It is eradicated if $\alpha + \tau \geqslant \beta$.

3.4 Optimal policy

Policy can be either at one of the boundaries or at an interior level, depending on the value of the shadow price. Rowthorn (2004) and Goldman and Lightwood (2002) show that optimal policy will take on one of the two boundary values. It is never optimal to treat partially. This is because the shadow price is a single-valued function of the state variable, so optimal policy can have at most one switch point. The interior steady state can only be reached by a path that zig-zags back on itself. In contrast, each of the boundary steady states can be reached by a path with at most one switch point, with the precise path depending on the initial infection level. Which policy of the two boundaries is optimal will depend on the value of parameters.

4 The SIS model with two strains of infection

4.1 Set-up

In the previous section, the transmission rate β was uniform and there was one policy instrument. Suppose there are two variants of infection, one more infectious than the other. The more infectious variant H has transmission rate β_H while the less infectious variant L is characterised by transmission rate β_L. The policymaker has two policy instruments at her disposal (f_H and f_L), each targeting one of the infection strands. There is an implicit assumption that the policymaker can distinguish the two strains and therefore target therapy perfectly. Individuals at the outset can catch either infection strand, and when infected they transmit the strand that they themselves are infected with. The two strands are mutually exclusive, in the sense that individuals cannot become infected with both at the same time. Similar to the previous section, there is a

possibility of exogenous recovery. If individuals recover, they are again susceptible to either infection strand. The proportion of the total population infected with H is I_H. The proportion infected with L is I_L. The total population is normalised to size 1. The policymaker maximises the social welfare function

$$V(I_H^0, I_L^0) = \int_0^\infty e^{-\delta t}(p(1-I_H(t)-I_L(t))-c(f_H(t)I_H(t)+f_L(t)I_L(t)))dt \quad (12)$$

subject to the equations of motion for the two infection types:

$$\begin{aligned} \dot{I}_H &= \beta_H I_H(t)(1-I_H(t)-I_L(t)) - \tau I_H(t) - \alpha f_H(t) I_H(t), & (13)\\ \dot{I}_L &= \beta_L I_L(t)(1-I_H(t)-I_L(t)) - \tau I_L(t) - \alpha f_L(t) I_L(t). & (14) \end{aligned}$$

All parameters are strictly positive. Further,

$$\begin{aligned} f_H, f_L &\in [0,1]\\ I_H(0) &= I_H^0 > 0 \text{ given}\\ I_L(0) &= I_L^0 > 0 \text{ given}\\ I_H^0 + I_L^0 &< 1 \end{aligned}$$

In addition,

$$\beta_H > \beta_L > \tau + \alpha \quad (15)$$

The above inequalities ensure that neither variant of the disease can be eliminated even asymptotically by treating all infected people. Thus, at any fixed point $I_H, I_L > 0$. They also ensure that $I_H(t) + I_L(t) < 1$ for all t.

The current value Hamiltonian is

$$\begin{aligned} H &= p(1-I_H-I_L) - c(f_H I_H + f_L I_L)\\ &\quad +\lambda_H(\beta_H I_H(1-I_H-I_L) - \tau I_H - f_H \alpha I_H)\\ &\quad +\lambda_L(\beta_L I_L(1-I_H-I_L) - \tau I_L - f_L \alpha I_L) \quad (16) \end{aligned}$$

The first order-conditions yield the following solution:

$$f_H^* \left\{ \begin{array}{c} =0 \\ \in (0,1) \\ =1 \end{array} \right\} if\ \lambda_H^* \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} - \frac{c}{\alpha}, \tag{17}$$

$$f_L^* \left\{ \begin{array}{c} =0 \\ \in (0,1) \\ =1 \end{array} \right\} if\ \lambda_L^* \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} - \frac{c}{\alpha}. \tag{18}$$

The equations of motion for the two costate variables are

$$\begin{aligned} \dot{\lambda}_H &= \delta\lambda_H - \frac{\partial H}{\partial I_H} \\ &= p + cf_H - \lambda_H \left(-\delta + \beta_H(1 - I_H - I_L) - \tau - \alpha f_H\right) \\ &\quad +(\lambda_H \beta_H I_H + \lambda_L \beta_L I_L), \end{aligned} \tag{19}$$

$$\begin{aligned} \dot{\lambda}_L &= \delta\lambda_L - \frac{\partial H}{\partial I_L} \\ &= p + cf_L - \lambda_L \left(-\delta + \beta_L(1 - I_H - I_L) - \tau - \alpha f_L\right) \\ &\quad +(\lambda_H \beta_H I_H + \lambda_L \beta_L I_L). \end{aligned} \tag{20}$$

4.2 Fixed points

4.2.1 The set of feasible fixed points

Definition 1 *A fixed point is a solution* $(f_H^*, f_L^*, I_H^*, I_L^*, \lambda_H^*, \lambda_L^*)$ *satisfying equations (13), (14), (17), (18), (19) and (20) as well as* $\dot{I}_H = \dot{I}_L = \dot{\lambda}_H = \dot{\lambda}_L = \dot{f}_H = \dot{f}_L = 0$.

There are nine potential fixed points, listed below. The notation A_{ab} denotes the fixed point with policy $f_H^* = a, f_L^* = b$ for $a, b = 0$ or 1. The notation $a, b = 2$ denotes an interior policy.

$$
\begin{aligned}
A_{00} &: \quad f_H = 0, f_L = 0 \\
A_{01} &: \quad f_H = 0, f_L = 1 \\
A_{02} &: \quad f_H = 0, f_L \in (0,1) \\
A_{10} &: \quad f_H = 1, f_L = 0 \\
A_{11} &: \quad f_H = 1, f_L = 1 \\
A_{12} &: \quad f_H = 1, f_L \in (0,1) \\
A_{20} &: \quad f_H \in (0,1), f_L = 0 \\
A_{21} &: \quad f_H \in (0,1), f_L = 1 \\
A_{22} &: \quad f_H \in (0,1), f_L \in (0,1)
\end{aligned}
$$

Definition 2 *An asymptotic fixed point (AFP) is a solution* $(f_H^*, f_L^*, I_H^*, I_L^*, \lambda_H^*, \lambda_L^*)$ *where at least one component in each of the pairs* $(I_H, I_L), (\lambda_H, \lambda_L)$ *comes arbitrarily close to its solution but only converges to it in the limit. At least one equality in each of the following pairs does not hold:* $\{\dot{I}_H = 0, \dot{I}_L = 0\}, \{\dot{\lambda}_H = 0, \dot{\lambda}_L = 0\}$. *The condition* $\dot{f}_H = \dot{f}_L = 0$ *holds.*

There are two potential types of AFPs, each encompassing several potentially optimal policies.

$$
\begin{aligned}
A_{13} &: \quad I_H \to 0, I_L = I_L^*, f_H^* = 1, f_L = f_L^* \in [0,1] \\
A_{31} &: \quad I_H = I_H^*, I_L \to 0, f_H = f_H^* \in [0,1], f_L^* = 1
\end{aligned}
$$

Lemma 3 *All fixed points are of type* $A_{10}, A_{12},$ *or* A_{20}.

Proof. Consider A_{22}. Suppose $f_H \in (0,1)$ and $f_L \in (0,1)$ during a finite interval of time. Then $\lambda_H = \lambda_L = -\frac{c}{\alpha}$ and thus $\dot{\lambda}_H = \dot{\lambda}_L = 0$ within this interval. Subtracting (20) from (19) yields:

$$(c/\alpha)(\beta_H - \beta_L)(1 - I_L - I_H) = 0$$

This is not possible since the left hand side is strictly positive. This demonstrates that A_{22} does not satisfy the Hamiltonian conditions and is not feasible. Thus, the Hamiltonian conditions imply that at least one of the control variables at a steady state is on the boundary.

Since $I_H, I_L > 0$ due to our parameter assumptions, we can rewrite the equations of motion as follows:

$$\frac{\dot{I}_H}{I_H} = \beta_H(1 - I_H - I_L) - \tau - f_H\alpha \quad (21)$$

$$\frac{\dot{I}_L}{I_L} = \beta_L(1 - I_H - I_L) - \tau - f_L\alpha \quad (22)$$

At a fixed point the right hand sides of the above equations must be zero. This implies that

$$f_H = \frac{\beta_H(1 - I_H - I_L) - \tau}{\alpha}, \quad (23)$$

$$f_L = \frac{\beta_L(1 - I_H - I_L) - \tau}{\alpha}. \quad (24)$$

Subtracting (24) from (23) yields

$$f_H - f_L = \frac{(\beta_H - \beta_L)(1 - I_H - I_L)}{\alpha} > 0. \quad (25)$$

This is not satisfied by fixed points A_{00}, A_{01}, A_{11}, A_{02} and A_{21}. This reduces the set of feasible fixed points to $F = \{A_{10}, A_{12}, A_{20}\}$. ■

4.2.2 Analysis of fixed points A_{10} and A_{12}

Fixed points A_{10} and A_{12} are the case when $f_H^* = 1$ and $f_L^* = 1$ in the former while $f_L^* \in (0, 1)$ in the latter. Setting $\dot{I}_H = 0$, $\dot{I}_L = 0$ and $f_H^* = 1$, we obtain that the fixed points A_{10} and A_{12} are characterised by the following treatment levels:

$$f_H^* = 1 \quad (26)$$

$$f_L^* = 1 - \frac{(\beta_H - \beta_L)}{\beta_H}\frac{\tau + \alpha}{\alpha}, \quad (27)$$

the latter only being feasible if $1 \geq \frac{(\beta_H - \beta_L)}{\beta_H}\frac{\tau+\alpha}{\alpha}$. Except in the special case of strict equality, $1 > \frac{(\beta_H - \beta_L)}{\beta_H}\frac{\tau+\alpha}{\alpha}$ and the fixed point is of the form A_{12}. The fixed point A_{10} is a boundary fixed point and will be addressed in Section 4.2.5. At fixed point A_{12}, we can characterise the total level of infection:

$$I_H^* + I_L^* = 1 - \frac{\tau + \alpha}{\beta_H}. \quad (28)$$

Further, the equation of motion for λ_H is given by

$$\begin{aligned}\dot{\lambda}_H &= 0 \\ &= p + cf_H^* + \delta\lambda_H^* - \lambda_H^* \left(\frac{\dot{I}_H}{I_H^*}\right) + (\lambda_H^*\beta_H I_H^* + \lambda_L^*\beta_L I_L^*) \\ &= p + cf_H^* + \delta\lambda_H^* + (\lambda_H^*\beta_H I_H^* + \lambda_L^*\beta_L I_L^*).\end{aligned} \tag{29}$$

where the second equality follows from the fact that $\dot{I}_H = 0$ at a steady state. Similarly,

$$\begin{aligned}\dot{\lambda}_L &= 0 \\ &= p + cf_L^* + \delta\lambda_L^* - \lambda_L^* \left(\frac{\dot{I}_L}{I_L^*}\right) + (\lambda_H^*\beta_H I_H^* + \lambda_L^*\beta_L I_L^*) \\ &= p + cf_L^* + \delta\lambda_L^* + (\lambda_H^*\beta_H I_H^* + \lambda_L^*\beta_L I_L^*).\end{aligned} \tag{30}$$

By subtraction,

$$c(f_H^* - f_L^*) + \delta(\lambda_H^* - \lambda_L^*) = 0.$$

Since f_L^* is interior, it must be that $\lambda_L^* = -\frac{c}{\alpha}$. Thus,

$$\lambda_H^* = -\frac{c}{\alpha}\left[1 + \frac{(\beta_H - \beta_L)}{\beta_H}\frac{\tau + \alpha}{\delta}\right].$$

Since $\beta_H - \beta_L > 0$, it follows that $\lambda_H^* < -\frac{c}{\alpha}$, as required by the Hamiltonian conditions. Thus, there is a line of fixed points in (I_H, I_L) space of type A_{12} that satisfies the Hamiltonian conditions with the following properties:

$$\begin{aligned}I_H^* + I_L^* &= 1 - \frac{\tau + \alpha}{\beta_H} \\ \lambda_H^* &= -\frac{c}{\alpha}\left[1 + \frac{(\beta_H - \beta_L)}{\beta_H}\frac{\tau + \alpha}{\delta}\right] < -\frac{c}{\alpha} \\ \lambda_L^* &= -\frac{c}{\alpha} \\ f_H^* &= 1 \\ f_L^* &= 1 - \frac{(\beta_H - \beta_L)}{\beta_H}\frac{\tau + \alpha}{\alpha}\end{aligned}$$

4.2.3 Analysis of fixed point A_{20}

Fixed point A_{20} is the case when $f_H^* \in (0,1)$ and $f_L^* = 0$. At fixed point A_{20}, setting $f_L^* = 0$, $\dot{I}_H = 0$ and $\dot{I}_L = 0$ yields the following treatment levels:

$$\begin{aligned} f_H^{**} &= \frac{\beta_H - \beta_L}{\beta_L}\frac{\tau}{\alpha} \\ f_L^{**} &= 0 \end{aligned} \tag{31}$$

the latter only being feasible if $1 \geq \frac{(\beta_H - \beta_L)}{\beta_L}\frac{\tau}{\alpha}$. Except in the special case of strict equality, $1 > \frac{(\beta_H - \beta_L)}{\beta_L}\frac{\tau}{\alpha}$. Note that in the case of strict equality, this fixed point becomes A_{10}.

Rearranging $\dot{I}_H = 0$ and $\dot{I}_L = 0$ gives us the total level of infection,

$$I_H^{**} + I_L^{**} = 1 - \frac{\tau}{\beta_L}. \tag{32}$$

The equations of motion for the costate variables are,

$$\begin{aligned} \dot{\lambda}_H &= 0 \\ &= p + cf_H^{**} + \delta\lambda_H^{**} - \lambda_H^{**}\left(\frac{\dot{I}_H}{I_H^{**}}\right) + (\lambda_H^{**}\beta_H I_H^{**} + \lambda_L^{**}\beta_L I_L^{**}) \\ &= p + cf_H^{**} + \delta\lambda_H^{**} + (\lambda_H^{**}\beta_H I_H^{**} + \lambda_L^{**}\beta_L I_L^{**}), \end{aligned}$$

$$\begin{aligned} \dot{\lambda}_L &= 0 \\ &= p + cf_L^{**} + \delta\lambda_L^{**} - \lambda_L^{**}\left(\frac{\dot{I}_L}{I_L^{**}}\right) + (\lambda_H^{**}\beta_H I_H^{**} + \lambda_L^{**}\beta_L I_L^{**}) \\ &= p + cf_L^{**} + \delta\lambda_L^{**} + (\lambda_H^{**}\beta_H I_H^{**} + \lambda_L^{**}\beta_L I_L^{**}). \end{aligned}$$

Employing the same method as in the previous section, we subtract to yield

$$c(f_H^{**} - f_L^{**}) + \delta(\lambda_H^{**} - \lambda_L^{**}) = 0.$$

Since f_H^{**} is interior it must be that $\lambda_H^{**} = -c/\alpha$. Thus,

$$\lambda_L^{**} = -\frac{c}{\alpha}\left[1 - \frac{(\beta_H - \beta_L)}{\beta_L}\frac{\tau}{\delta}\right]. \tag{33}$$

Since $\beta_H - \beta_L > 0$ it follows that $\lambda_L^* > -\frac{c}{\alpha}$ always holds, as required by the Hamiltonian conditions. Thus, there is a line of fixed points of type A_{20} that satisfy the Hamiltonian conditions with the following properties:

$$\begin{aligned} I_H^{**} + I_L^{**} &= 1 - \frac{\tau}{\beta_L} \\ \lambda_H^{**} &= -\frac{c}{\alpha} \\ f_H^{**} &= \frac{(\beta_H - \beta_L)}{\beta_L}\frac{\tau}{\alpha} \\ \lambda_L^{**} &= -\frac{c}{\alpha}\left[1 - \frac{(\beta_H - \beta_L)}{\beta_L}\frac{\tau}{\delta}\right] \\ f_L^{**} &= 0 \end{aligned}$$

Let us label fixed points A_{12} and A_{20} as Interior Fixed Points (IFPs) for ease of exposition, as they are fixed points that induce one policy instrument to be at an interior level.

4.2.4 Analysis of asymptotic fixed points

For the analysis of asymptotic fixed points we need to define the concept of a Most Rapid Approach Path (MRAP).

Definition 4 *An MRAP is a path with a policy that ensures convergence to the fixed point in less time than any other policy.*

First, consider A_{13}. I_H tends asymptotically towards zero and I_L converges to some equilibrium level:

$$A_{13} : I_H \longrightarrow 0, I_L = I_L^*$$

For I_H to asymptotically tend to zero, we require $\frac{\dot{I}_H}{I_H} < 0$ at all points in time, for which the MRAP is $f_H^* = 1$. Combining these features gives $\frac{\dot{I}_H}{I_H} =$

$\beta_H(1 - I_H - I_L^*) - \alpha - \tau \approx \beta_H(1 - I_L^*) - \alpha - \tau$, for I_H sufficiently close to zero. This needs to be negative, so the condition required for this to be a feasible AFP is

$$1 - \frac{\alpha + \tau}{\beta_H} < I_L^*. \tag{34}$$

Similarly, I_L converges to I_L^*, which requires $\frac{\dot{I}_L}{I_L} = 0$. Using this we can solve for I_L^*:

$$I_L^* = 1 - \frac{\tau + \alpha f_L^*}{\beta_L}. \tag{35}$$

Thus, (34) simplifies to

$$\frac{\tau + \alpha f_L^*}{\beta_L} < \frac{\alpha + \tau}{\beta_H}. \tag{36}$$

Any f_L^* chosen to satisfy this will cause I_H to converge asymptotically to zero and I_L to I_L^* as defined above. Thus there exists a fixed point of type A_{13}, which involves asymptotic convergence of I_H to zero:

$$\begin{aligned} I_H &\longrightarrow 0 \\ I_L^* &= 1 - \frac{\tau + \alpha f_L^*}{\beta_L} \\ f_H^* &= 1 \end{aligned}$$

Similarly, consider fixed point A_{31}, where I_L tends towards zero asymptotically and I_H converges to I_H^*:

$$A_{31} : I_H = I_H^*, I_L \longrightarrow 0$$

For I_L to tend asymptotically to zero, we require $\frac{\dot{I}_L}{I_L} < 0$ at all points in time, for which the MRAP is $f_L^* = 1$. Combining these features gives $\frac{\dot{I}_L}{I_L} = \beta_L(1 - I_H^* - I_L) - \alpha - \tau \approx \beta_L(1 - I_H^*) - \alpha - \tau$, for I_L sufficiently close to zero. This needs to be negative, which requires

$$1 - \frac{\alpha + \tau}{\beta_L} < I_H^*. \tag{37}$$

Similarly, I_H converges to I_H^*, which requires $\frac{\dot{I}_H}{I_H} = 0$. Using this we can solve for I_H^*:

$$I_H^* = 1 - \frac{\tau + \alpha f_H^*}{\beta_H}. \tag{38}$$

Thus, (37) simplifies to

$$\frac{\tau + \alpha f_H^*}{\beta_H} < \frac{\alpha + \tau}{\beta_L}. \tag{39}$$

Any f_H^* chosen to satisfy this will cause I_L to converge asymptotically to zero and I_H to I_H^* as defined above. Thus, there exists a fixed point of type A_{31}, which involves asymptotic convergence of I_L to zero:

$$\begin{aligned} I_L &\longrightarrow 0 \\ I_H^* &= 1 - \frac{\tau + \alpha f_H^*}{\beta_H} \\ f_L^* &= 1 \end{aligned}$$

Note that both asymptotic fixed points can be feasible at the same time. Rearranging (36) and (39) gives:

$$\begin{aligned} \tau(\frac{1}{\beta_L} - \frac{1}{\beta_H}) &< \frac{\alpha}{\beta_H} - \frac{\alpha f_L^*}{\beta_L}, \\ \tau(\frac{1}{\beta_L} - \frac{1}{\beta_H}) &> \frac{\alpha f_H^*}{\beta_H} - \frac{\alpha}{\beta_L}. \end{aligned}$$

Both conditions can be satisfied as long as $f_H^*, f_L^* < 1$. The case of $f_H = f_L = 1$ deserves further attention and is examined more fully in Section 4.2.6.

4.2.5 Further analysis of feasibility: fixed points

Define the constant K as

$$K = \frac{\beta_H - \beta_L}{\beta_L} \frac{\tau}{\alpha}$$

The parameter space can now be divided into three regimes in terms of feasibility of the various fixed points.

Proposition 5 *If $K < 1$, there exist a line of fixed points of type A_{12} and A_{20}. If $K = 1$, there exists a fixed point of type A_{10}. If $K > 1$, there are no fixed points.*

Proof. The conditions for the two kinds of interior fixed points to exist are as follows:

$$\begin{aligned} A_{12} &: \quad f_H = 1, f_L \in (0,1) \text{ needs } 1 > \frac{(\beta_H - \beta_L)}{\beta_H} \frac{\tau + \alpha}{\alpha} \\ A_{20} &: \quad f_H \in (0,1), f_L = 0 \text{ needs } 1 > \frac{(\beta_H - \beta_L)}{\beta_L} \frac{\tau}{\alpha} \end{aligned}$$

These two conditions are, in fact, identical. To see this, consider the following rearrangement of the condition for A_{12}:

$$\begin{aligned} 1 \quad &> \quad \frac{(\beta_H - \beta_L)}{\beta_H} \frac{\tau + \alpha}{\alpha} \\ &\Leftrightarrow \quad \beta_H \alpha > (\beta_H - \beta_L)(\tau + \alpha) \\ &\Leftrightarrow \quad \beta_L(\tau + \alpha) > \beta_H \tau \\ &\Leftrightarrow \quad \beta_L \alpha > (\beta_H - \beta_L)\tau \\ &\Leftrightarrow \quad 1 > \frac{(\beta_H - \beta_L)}{\beta_L} \frac{\tau}{\alpha} \end{aligned}$$

This demonstrates that both conditions are equivalent to $\beta_L(\tau + \alpha) > \beta_H \tau$. The condition $1 > \frac{(\beta_H - \beta_L)}{\beta_L} \frac{\tau}{\alpha}$ is identical to $K < 1$.

If $K = 1$, this implies that $\frac{(\beta_H - \beta_L)}{\beta_L} \frac{\tau}{\alpha} = \frac{(\beta_H - \beta_L)}{\beta_H} \frac{\tau + \alpha}{\alpha} = 1$, and both A_{12} and A_{20} become the fixed point of type A_{10}. Total infection is characterised by the equation:

$$I_H + I_L = 1 - \frac{\tau + \alpha}{\beta_H} = 1 - \frac{\tau}{\beta_L}.$$

If $K > 1$, none of the conditions for A_{12}, A_{20} nor A_{10} are satisfied. Therefore there are no fixed points. ■

The proof of this Proposition shows that if $K < 1$, there are two lines of fixed points with total infection levels:

$$\begin{aligned} I_H^* + I_L^* &= 1 - \frac{\tau + \alpha}{\beta_H}, \\ I_H^{**} + I_L^{**} &= 1 - \frac{\tau}{\beta_L}. \end{aligned}$$

Subtracting,

$$(I_H^{**} + I_L^{**}) - (I_H^* + I_L^*) = \frac{\alpha}{\beta_H}\left(1 - \frac{(\beta_H - \beta_L)}{\beta_L}\frac{\tau}{\alpha}\right) > 0$$

This shows that A_{20} always has higher total infection than A_{12}. This is obvious as in the latter, both treatment levels are higher.

4.2.6 Further analysis of feasibility: AFPs

We examine further the role of K in the feasibility of the AFPs. Let us denote A_{13}^0 as the AFP A_{13} when $f_H^* = 1$ and $f_L^* = 0$. Further denote the AFP A_{13} when $f_H^* = 1$ and $f_L^* = 1$ as A_{13}^1. Last, A_{13}^i is the AFP A_{13} when $f_H^* = 1$ and $f_L^* \in (0, 1)$. Symmetrically, we can define A_{31}^0, A_{31}^i and A_{31}^1 as the AFP A_{31} when $f_H^* = 0, f_H^* \in (0, 1)$ and $f_H^* = 1$ respectively and $f_L^* = 1$ in all cases.

Proposition 6 *If $K < 1$, there exist AFPs of type A_{31}^0, A_{31}^i, A_{31}^1, A_{13}^0 and A_{13}^i. If $K \geq 1$, there exist AFPs of type A_{31}^0, A_{31}^i, and A_{31}^1.*

Proof. First, note that the necessary condition for the feasibility of A_{31} ($\frac{\tau + \alpha f_H^*}{\beta_H} < \frac{\alpha + \tau}{\beta_L}$) is satisfied independently of the value of K.

Next, consider $K < 1$ and A_{13}. The necessary condition for this AFP to be feasible is $\frac{\tau + \alpha f_L^*}{\beta_L} < \frac{\alpha + \tau}{\beta_H}$. This is never satisfied for $f_L^* = 1$. However, it may be satisfied for small enough f_L^*. In particular, when $K < 1$, it is satisfied when $f_L^* = 0$. Thus, the AFPs that are feasible when $K < 1$ are $A_{31}^0, A_{31}^i, A_{31}^1, A_{13}^0$ and A_{13}^i. Here, A_{13}^i is defined such that f_L^* is small enough to satisfy the feasibility condition for this AFP.

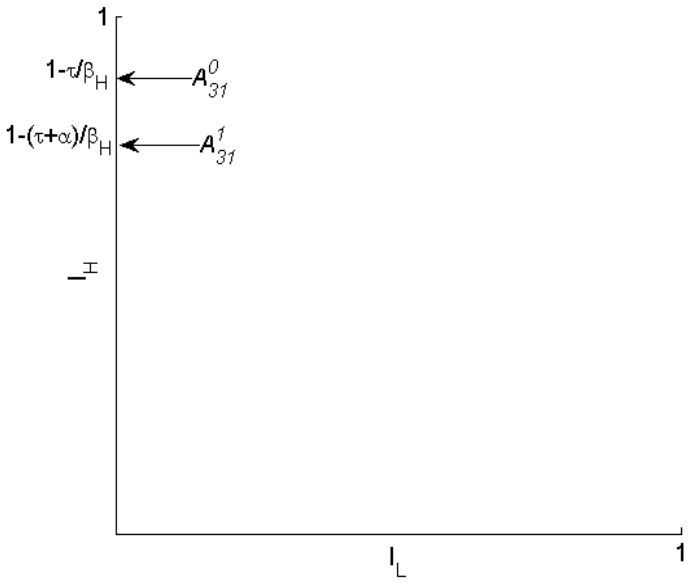

Figure 1: The set of feasible fixed points and AFPs when $K > 1$.

Next, consider $K \geq 1$. For A_{13} to be a feasible equilibrium, we require $\frac{\tau+\alpha}{\beta_H} > \frac{\tau+\alpha f_L^*}{\beta_L}$, which is violated for all values of f_L^* when $K \geq 1$. Thus, the set of AFPs that are feasible when $K \geq 1$ is A_{31}^0, A_{31}^i and A_{31}^1. ■

The set of feasible fixed points and AFPs when $K > 1$ is depicted in Figure 1. The feasible set when $K = 1$ is shown in Figure 2. Figure 3. shows the set of feasible fixed points and AFPs when $K < 1$. The AFPs with interior policies are not depicted in these graphs as they are only pinned down once the treatment levels are known.

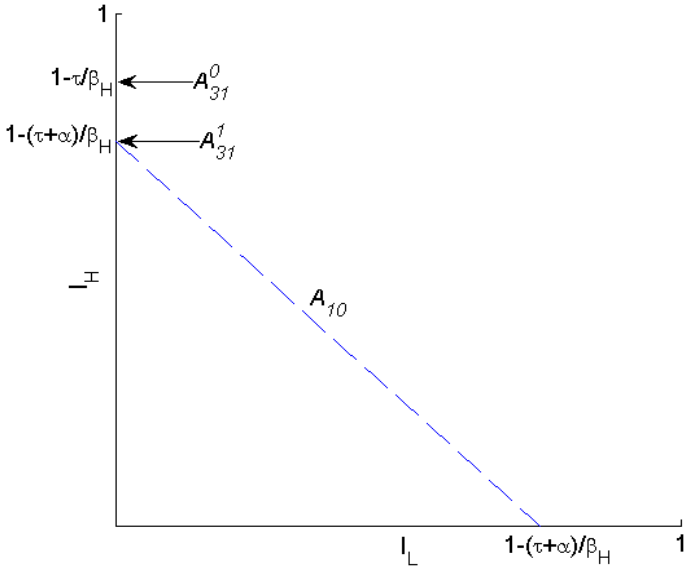

Figure 2: The set of feasible fixed points and AFPs when $K = 1$.

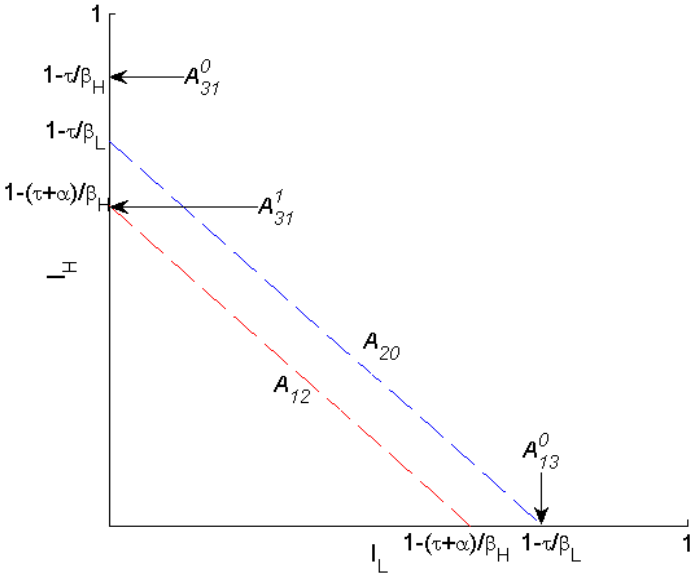

Figure 3: The set of feasible fixed points and AFPs when $K < 1$.

4.2.7 Feasible policy along the path

It is necessary to consider the path towards each of the steady states, and in particular which policies are feasible under which conditions. Policies along the path will always be boundary policies, as these are Most Rapid Approach Paths (MRAPs). We take each of the boundary policies in turn and examine the feasibility conditions required for I_H and I_L to converge to their steady state values. Let P_{ab} denote the policy $f_H = a$, $f_L = b$.

First, consider P_{00}. This implies

$$\begin{aligned} \frac{\dot{I}_H}{I_H} &= \beta_H(1 - I_H - I_L) - \tau, \\ \frac{\dot{I}_L}{I_L} &= \beta_L(1 - I_H - I_L) - \tau. \end{aligned} \tag{40}$$

There are two ways of approaching a fixed point with this policy. First, we can have $\dot{I}_H, \dot{I}_L > 0$ (i.e. I_H and I_L are increasing towards I_H^* and I_L^*). The required conditions for this are

$$\begin{aligned} 1 - \frac{\tau}{\beta_L} &> I_H + I_L \\ 1 - \frac{\tau}{\beta_H} &> I_H + I_L \end{aligned}$$

which collapse to

$$1 - \frac{\tau}{\beta_L} > I_H + I_L. \tag{41}$$

Similarly, we can have $\dot{I}_H, \dot{I}_L < 0$ (i.e. I_H and I_L are decreasing towards I_H^* and I_L^*). The required conditions for this are

$$\begin{aligned} 1 - \frac{\tau}{\beta_L} &< I_H + I_L \\ 1 - \frac{\tau}{\beta_H} &< I_H + I_L \end{aligned}$$

which collapse to

$$1 - \frac{\tau}{\beta_H} < I_H + I_L. \tag{42}$$

Next, consider P_{10}. For $\dot{I}_H, \dot{I}_L > 0$, the required conditions are

$$\begin{aligned} 1 - \frac{\tau}{\beta_L} &> I_H + I_L \\ 1 - \frac{\tau + \alpha}{\beta_H} &> I_H + I_L \end{aligned}$$

where the overriding condition is

$$1 - \frac{\tau}{\beta_L} > I_H + I_L. \tag{43}$$

For $\dot{I}_H, \dot{I}_L < 0$, we require

$$\begin{aligned} 1 - \frac{\tau}{\beta_L} &< I_H + I_L \\ 1 - \frac{\tau + \alpha}{\beta_H} &< I_H + I_L \end{aligned}$$

both of which are satisfied when

$$1 - \frac{\tau + \alpha}{\beta_H} < I_H + I_L. \tag{44}$$

Third, take P_{11}. For $\dot{I}_H, \dot{I}_L > 0$, we need to satisfy

$$\begin{aligned} 1 - \frac{\tau + \alpha}{\beta_H} &> I_H + I_L \\ 1 - \frac{\tau + \alpha}{\beta_L} &> I_H + I_L \end{aligned}$$

where the overriding condition is

$$1 - \frac{\tau + \alpha}{\beta_L} > I_H + I_L. \tag{45}$$

For $\dot{I}_H, \dot{I}_L < 0$, we require

$$\begin{aligned} 1 - \frac{\tau + \alpha}{\beta_H} &< I_H + I_L \\ 1 - \frac{\tau + \alpha}{\beta_L} &< I_H + I_L \end{aligned}$$

both of which are satisfied when

$$1 - \frac{\tau + \alpha}{\beta_H} < I_H + I_L. \tag{46}$$

Last, consider P_{01}. For $\dot{I}_H, \dot{I}_L > 0$, we need to satisfy

$$\begin{aligned} 1 - \frac{\tau}{\beta_H} &> I_H + I_L \\ 1 - \frac{\tau + \alpha}{\beta_L} &> I_H + I_L \end{aligned}$$

where the overriding condition is

$$1 - \frac{\tau + \alpha}{\beta_L} > I_H + I_L. \tag{47}$$

For $\dot{I}_H, \dot{I}_L < 0$, we require

$$\begin{aligned} 1 - \frac{\tau}{\beta_H} &< I_H + I_L \\ 1 - \frac{\tau + \alpha}{\beta_L} &< I_H + I_L \end{aligned}$$

both of which are satisfied when

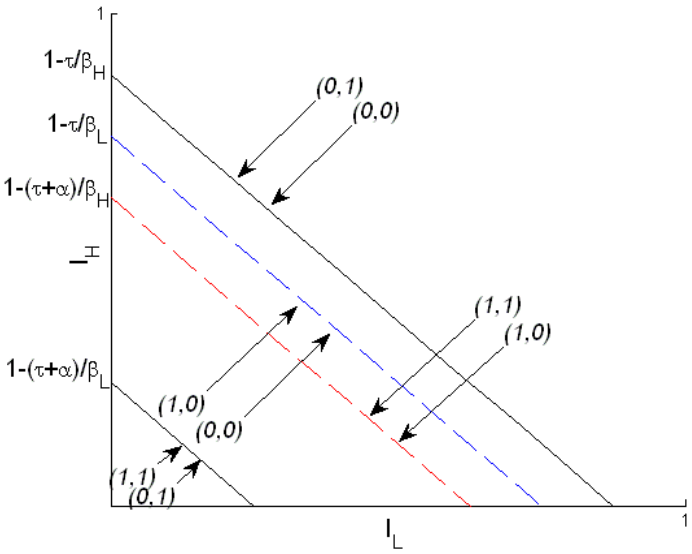

Figure 4: Feasible policies depicted for the case where $K < 1$.

$$1 - \frac{\tau}{\beta_H} < I_H + I_L.$$

These facts are summarised in the table below:

Table 1 (Feasible policies along the path)

	$\dot{I}_H, \dot{I}_L > 0$	$\dot{I}_H, \dot{I}_L < 0$
P_{00}	$1 - \frac{\tau}{\beta_L} > I_H + I_L$	$1 - \frac{\tau}{\beta_H} < I_H + I_L$
P_{10}	$1 - \frac{\tau}{\beta_L} > I_H + I_L$	$1 - \frac{\tau+\alpha}{\beta_H} < I_H + I_L$
P_{11}	$1 - \frac{\tau+\alpha}{\beta_L} > I_H + I_L$	$1 - \frac{\tau+\alpha}{\beta_H} < I_H + I_L$
P_{01}	$1 - \frac{\tau+\alpha}{\beta_L} > I_H + I_L$	$1 - \frac{\tau}{\beta_H} < I_H + I_L$

Figure 4 shows which policies are feasible in different regions of initial infection levels.

4.3 Optimal policy

4.3.1 Optimal policy in the neighbourhood of the IFPs

Having derived feasibility conditions for the various policies, the obvious question is which policies are optimal. We explore the behaviour of the path in approaching each of the interior fixed points. Specifically, what is the policy along the path near to the fixed point? We know that policies along the path will be at a boundary, as these are MRAPs. Therefore, we examine those policies that are at an interior level at the steady state, as they are likely to have a switch point along the path. The approach is to perturb the fixed point slightly and derive the policy in the neighbourhood of the fixed point.

First, take A_{12}. Let us perturb the solution by changing f_L from f_L^* to $f_L^* + \Delta f_L$ whilst leaving f_H unchanged at its steady state value. Immediately following this change, $\dot{I}_H = 0, \dot{I}_L = -\alpha I_L^* \Delta f_L \neq 0, \dot{\lambda}_H = 0$ and $\dot{\lambda}_L = 0$. Differentiating (20) yields

$$\begin{aligned}
\ddot{\lambda}_L &= (c + \alpha\lambda_L^*)\dot{f}_L - \dot{\lambda}_L\left(-\delta + \beta_L(1 - I_L^* - I_H^*) - \tau - \alpha f_L^*\right) + \lambda_L^*\beta_L(\dot{I}_H + \dot{I}_L) \\
&\quad + (\dot{\lambda}_H\beta_H I_H^* + \dot{\lambda}_L\beta_L I_L^*) + (\lambda_H^*\beta_H\dot{I}_H + \lambda_L^*\beta_L\dot{I}_L) \\
&= 2\lambda_L^*\beta_L\dot{I}_L \\
&= -2\frac{c}{\alpha}\beta_L(-\alpha I_L^*)\Delta f_L \\
&= 2c\beta_L I_L^*\Delta f_L \neq 0
\end{aligned}$$

Thus, there is a policy switch. To see this, consider the following. If $\Delta f_L > 0$ then $\ddot{\lambda}_L > 0$ and $\dot{I}_L < 0$. Since we require $\lambda_L = -\frac{c}{\alpha}$ at the fixed point, this implies that $\lambda_L < -\frac{c}{\alpha}$ when approaching the fixed point from above. The Hamiltonian conditions imply that $f_L = 1$ along this segment of the path. Likewise, if $\Delta f_L < 0$ then $\ddot{\lambda}_L < 0$ and $\dot{I}_L > 0$. This implies that $\lambda_L > -\frac{c}{\alpha}$ when approaching the fixed point from below, and hence from the Hamiltonian conditions it must be that $f_L = 0$. Since $\dot{\lambda}_H = 0$ and $\lambda_H < -\frac{c}{\alpha}$ at the fixed point, it must be that $\lambda_H < -\frac{c}{\alpha}$ holds on either side of the fixed point, by continuity. This demonstrates that there is a Hamiltonian path to A_{12} which involves boundary values of f_H and f_L until it reaches the fixed point, when it switches to an interior value of f_L. There is no change in f_H.

Next, take A_{20}. Perturb the solution by altering f_H from f_H^{**} to $f_H^{**} + \Delta f_H$, leaving $f_L^* = 0$. Immediately following this change, $\dot{I}_H = -\alpha I_H^{**}\Delta f_H \neq 0$, $\dot{I}_L = 0$, $\dot{\lambda}_H = 0$ and $\dot{\lambda}_L = 0$. Differentiating $\dot{\lambda}_H$ we see that in the proximity to this fixed point,

$$\begin{aligned}
\ddot{\lambda}_H &= (c+\alpha\lambda_H^{**})\dot{f}_H - \dot{\lambda}_H\left(-\delta+\beta_H(1-I_H^{**}-I_L^{**})-\tau-\alpha f_H^{**}\right)+\lambda_H^{**}\beta_H(\dot{I}_H+\dot{I}_L) \\
&\quad +(\dot{\lambda}_H\beta_H I_H^{**}+\dot{\lambda}_L\beta_L I_L^{**})+(\lambda_H^{**}\beta_H\dot{I}_H+\lambda_L^{**}\beta_L\dot{I}_L) \\
&= 2\lambda_H^{**}\beta_H\dot{I}_H \\
&= -2\frac{c}{\alpha}\beta_H(-\alpha I_H^{**})\Delta f_H \\
&= 2c\beta_H I_H^{**}\Delta f_H \neq 0
\end{aligned}$$

Again, there will be a policy switch. If $\Delta f_H > 0$ then $\ddot{\lambda}_H > 0$ and $\dot{I}_H < 0$. This implies that $\lambda_H < -\frac{c}{\alpha}$ when approaching the fixed point from above, and hence from the Hamiltonian conditions it must be that $f_H = 1$. Likewise, if $\Delta f_H < 0$ then $\ddot{\lambda}_H < 0$ and $\dot{I}_H > 0$. This implies that $\lambda_H > -\frac{c}{\alpha}$ when approaching the fixed point from below, and hence it must be the case that $f_H = 0$. By continuity, $\lambda_L > -\frac{c}{\alpha}$ and $f_L = 0$ on both sides of the fixed point. Thus, there is a Hamiltonian path to this fixed point which involves boundary values of f_H and f_L and a switch to an interior value of f_H on reaching the fixed point, while retaining the boundary value for f_L.

Comparing these optimal policies to the feasibility conditions of the previous section, we find that all of the optimal policies are feasible [is this trivial?]. From the conditions derived for policy along the path, it is clear that, for example, the upper line $I_H^{**} + I_L^{**} = 1 - \frac{\tau}{\beta_L}$ is attainable from the top using both P_{11} and P_{10}. It is also clear that P_{11} is the MRAP. However, we find that P_{10} is the optimal policy. The intuition for this is as follows. This is because if $f_L^* = 0$ at the fixed point, then $\lambda_L^* > -\frac{c}{\alpha}$. Since λ_L is continuous it must be that $\lambda_L^* > -\frac{c}{\alpha}$ in the vicinity of the fixed point. Hence $f_L^* = 0$ in the vicinity of the fixed point and it cannot be optimal to reach this fixed point with P_{11}. Similar intuition applies for the optimal policy for A_{12}.

4.3.2 Skiba Hypothesis

A "Skiba curve" is a curve of indifference along which policy is indifferent between the available options. In our case, it is a curve along which conditions prescribe indifference between selecting the path towards A_{20} versus A_{12}. Note that these paths are the optimal paths derived in the previous section. Following on from these results, we hypothesise that there is a Skiba curve lying between the two lines of fixed points. If the initial point $\left(I_H^0, I_L^0\right)$ lies between the origin and the Skiba curve, then optimal policy is

$$
\begin{aligned}
f_H &= 1, f_L = 0 \text{ for } I_H^0 + I_L^0 < 1 - \frac{\tau + \alpha}{\beta_H} \\
f_H &= 1, f_L = 1 \text{ for } I_H^0 + I_L^0 > 1 - \frac{\tau + \alpha}{\beta_H} \\
f_H &= 1, f_L = 1 - \frac{(\beta_H - \beta_L)}{\beta_H} \frac{\tau + \alpha}{\alpha} \text{ for } I_H^0 + I_L^0 = 1 - \frac{\tau + \alpha}{\beta_H}
\end{aligned}
$$

If the initial point $\left(I_H^0, I_L^0\right)$ lies on the opposite side of the Skiba curve from the origin, then optimal policy is

$$
\begin{aligned}
f_H &= 0, f_L = 0 \text{ for } I_H^0 + I_L^0 < 1 - \frac{\tau}{\beta_L} \\
f_H &= 1, f_L = 0 \text{ for } I_H^0 + I_L^0 > 1 - \frac{\tau}{\beta_L} \\
f_H &= \frac{(\beta_H - \beta_L)}{\beta_L} \frac{\tau}{\alpha}, f_L = 0 \text{ for } I_H^0 + I_L^0 = 1 - \frac{\tau}{\beta_L}
\end{aligned}
$$

These sets of optimal policies are depicted in Figure 5. Skiba curves cannot be derived analytically. Their presence can only be detected by means of simulations.

4.3.3 Optimal policy in the neighbourhood of the AFPs

We have already shown that optimal policy for A_{31} will involve $f_L^* = 1$, as this is the MRAP. Similarly, optimal policy for A_{13} will involve $f_H^* = 1$. The question is which policy is optimal of the range available to f_H in A_{31} and f_L in A_{13}. In order to draw conclusions on this we observe that the asymptotic fixed points always involve one strand of the infection that is asymptotically eradicated. As a result, the behaviour of the system in the neighbourhood of the fixed point can be approximated by the behaviour of a one-infection system. This is because the behaviour of the system for small I_L is very similar to the behaviour when $I_L = 0$. Naturally this holds for I_H close to zero as well.

The behaviour of a one-infection system has been analysed in Section 3. As was discussed, Rowthorn (2004) shows that only extreme values for policy are optimal. This is because a one-infection system has a costate variable that is single-valued in the infection level along the optimal path, which implies that the optimum path cannot be a spiral. In our case the costate variable λ_H is defined as

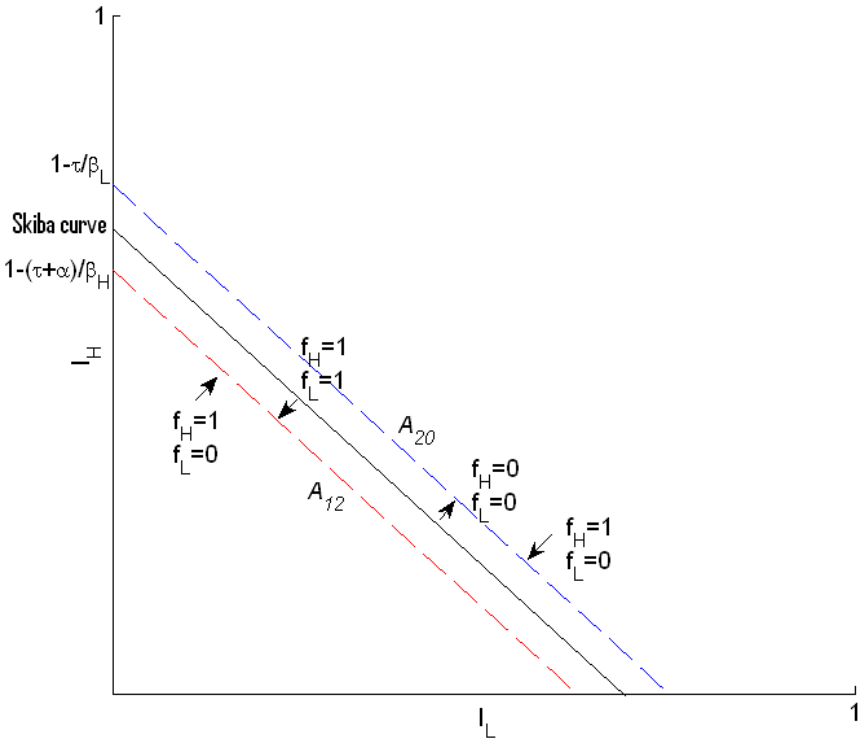

Figure 5: Optimal policies and Skiba curve.

$$\lambda_H = \frac{\partial V(I_H, I_L)}{\partial I_H},$$

which for small I_L is single-valued along the optimal path. Similarly,

$$\lambda_L = \frac{\partial V(I_H, I_L)}{\partial I_L},$$

which for small I_H is single-valued along the optimal path. Interior policies involve spirals. This implies that optimal policy for the AFPs will only ever involve boundary values, which allows us to eliminate A^i_{31} and A^i_{13} as steady states that are never optimal. Therefore, when $K > 1$, the set of feasible fixed points and AFPs is $F = \{A^0_{31}, A^1_{31}\}$, one of which will be optimal. When $K < 1$, the set of feasible fixed points and AFPs is $F = \{A_{12}, A_{20}, A^0_{31}, A^1_{31}, A^0_{13}\}$, one of which will be optimal. Similarly, when $K = 1$, the feasible set is $F = \{A_{10}, A^0_{31}, A^1_{31}\}$. We cannot make any further conclusions on the optimality of these remaining feasible fixed points. Optimality will depend on parameter values. This will be explored by way of simulations in Section 5.

4.4 Extensions

4.4.1 Asymptotic eradication

Proposition 7 *Both variants of the disease cannot be simultaneously eradicated even asymptotically in equilibrium, i.e. we cannot have both $I^*_H \to 0$ and $I^*_L \to 0$, if we assume that*

$$\begin{aligned} \frac{\tau + \alpha}{\beta_H} &< 1, \\ \frac{\tau}{\beta_L} &< 1, \\ \frac{\tau + \alpha}{\beta_L} &< 1. \end{aligned}$$

Proof. To see this, first consider the case of the interior fixed points. Here, it is trivial. In the case of A_{12}, $I^*_H + I^*_L = 1 - \frac{\alpha+\tau}{\beta_H}$. We cannot have $I^*_H + I^*_L = 0$. Similarly for A_{20} where $I^*_H + I^*_L = 1 - \frac{\tau}{\beta_L}$, it is not possible to have $I^*_H + I^*_L = 0$. Next, consider the asymptotic fixed points. In the case of A_{31}, we know that $I^*_L \to 0$ so the question is what happens to I^*_H. For I_L to tend towards zero asymptotically, the necessary condition is $I_H > 1 - \frac{\alpha+\tau}{\beta_L}$. Clearly $I_H \neq 0$ is

necessary for this to be satisfied. Similarly, fixed point A_{13} implies that $I_H^* \to 0$. The necessary condition for this is $I_L > 1 - \frac{\alpha+\tau}{\beta_H}$, which can only be satisfied if $I_L \neq 0$. Thus, both variants of the disease cannot be eradicated in equilibrium, even asymptotically. ■

5 Simulations

The purpose of simuations is to enable the identification of optimal policy under different parameters. We provide examples of optimal policy in the case of various parameter assumptions. Simulations are carried out using the fourth-order Runge-Kutta method. Recalling that we defined $K = \frac{\beta_H - \beta_L}{\beta_L} \frac{\tau}{\alpha}$, there are three cases that can be evaluated: $K < 1$, $K = 1$ and $K > 1$. We focus on the case when $K > 1$ and there are only two feasible fixed points: A_{31}^0 and A_{31}^1. This case is interesting because it suggests that with an appropriate set of parameters, it may be optimal to only eradicate the low infectivity strain, while allowing the high infectivity strain to be endemic, with full or maybe even no treatment. Further, this case will allow the clearest policy recommendations as the number of possible optimal policies is small. The following parameter assumptions ensure that $K > 1$:

Table 2 (Parameter values)

Parameter	Value
β_H	0.95
β_L	0.4
τ	0.15
α	0.2

In addition, we assume that $p = 1$ and $\delta = 0.111$.

5.1 Paths with fixed policy

The goal is to evaluate whether, under different scenarios, it is better to move towards A_{31}^0 or A_{31}^1. In Section 4.2.4 it was shown that in the neighbourhood of A_{31}^0, optimal policy is $(f_H^*, f_L^*) = (0, 1)$. In the neighbourhood of A_{31}^1, optimal policy is $(f_H^*, f_L^*) = (1, 1)$. These policies may not be optimal along the entire path towards these fixed points. However, we begin with a simple thought experiment where we assume that the policymaker can only choose one policy and cannot change it. This may happen in reality, for example, if the policymaker commits to a certain treatment level and purchases the requisite amount of material. Organising additional treatment may take time. Further, there

may be political factors as agencies responsible for treatment may not be able to secure additional funds from governements in the short run. We carry out simulations where we assume that this is the case. In the next section, we allow for flexibility of treatment across time.

Optimal policy is evaluated based on the value of the integral, V, under each policy. We fix policy at the beginning and allow the system to converge to steady state. In order to analyse policy under various scenarios, we focus on the cost paremeter c, which we vary. We find that there are three regions of values for c, each of which involve a different optimal policy. These are shown in the table below:

Table 3 (Regions of optimal policy as c varies)

Region	c	f_H^*	f_L^*
I	$c < 0.2875$	1	1
II	$0.2875 \leq c \leq 0.3006$	0 or 1, depending on I_H^0	1
III	$c > 0.3006$	0	1

Let us look at examples from each region and compare the value of V when starting at different initial points I_H^0 and setting $f_H = 0$ or $f_H = 1$. Note that when we are at A_{31}^1, $I_H^* = 1 - \frac{\tau+\alpha}{\beta_H} = 0.6316$. When we are at A_{31}^0, $I_H^* = 1 - \frac{\tau}{\beta_H} = 0.8421$. We take five initial infection levels for the H type, distributed evenly across the interval $I_H^0 \in [0.6316, 0.8421]$. The initial value for the L infection is constant across all simulations and is set at a value close to zero: $I_L^0 = 0.1$.

First, consider Region I. Let $c = 0.1$. The table below gives the prevalence of each infection type when steady state is reached, and the value of the integral of moving to that steady state. The policy with the higher value of V - the optimal policy - is emphasised in bold.

Table 4 ($c = 0.1$)

	$f_H = 1$ (path towards A_{31}^1)			$f_H = 0$ (path towards A_{31}^0)		
I_H^0	I_L^*	I_H^*	V	I_L^*	I_H^*	V
0.6667	0.0000110	0.6316	**2.6126**	0.00000027	0.8421	1.5425
0.7018	0.0000109	0.6316	**2.5632**	0.00000026	0.8421	1.4968
0.7369	0.0000108	0.6316	**2.5161**	0.00000026	0.8421	1.4531
0.7719	0.0000107	0.6316	**2.4711**	0.00000026	0.8421	1.4114
0.8070	0.0000106	0.6316	**2.4278**	0.00000026	0.8421	1.3714

In this scenario, policy is independent of the initial value. It is always optimal to set $f_H^* = 1$ and treat everyone. As costs rise, we enter Region II. As an example of policy evaluation for costs in this region, we set $c = 0.295$. The table below shows details of the value of the integral and the infection levels for this parameter combination:

Table 5 ($c = 0.295$)

	$f_H = 1$ (path towards A^1_{31})			$f_H = 0$ (path towards A^0_{31})		
I^0_H	I^*_L	I^*_H	V	I^*_L	I^*_H	V
0.6667	0.0000110	0.6316	**1.4806**	0.00000027	0.8421	1.4636
0.7018	0.0000109	0.6316	**1.4223**	0.00000026	0.8421	1.4184
0.7369	0.0000108	0.6316	1.3668	0.00000026	0.8421	**1.3753**
0.7719	0.0000107	0.6316	1.3138	0.00000026	0.8421	**1.3342**
0.8070	0.0000106	0.6316	1.2628	0.00000026	0.8421	**1.2947**

From the simulations it is clear that for $I^0_H \leq 0.7018$, the policy $f^*_H = 1$ is optimal. For $I^0_H \geq 0.7369$, the policy $f^*_H = 0$ is optimal. However, we can be more specific than this. In the region $I^0_H \in (0.7018, 0.7369)$, there is a point of indifference where the initial value is such that policy is indifferent between setting $f^*_H = 0$ and $f^*_H = 1$. Simulations show that this value is $\tilde{I}^0_H = 0.7125$, where $V = 1.4051$ for both policies. Optimal policy when $c = 0.295$ is summarised in the table below:

Table 6 (Optimal policy when $c = 0.295$)

I^0_H	f^*_H	f^*_L
$I^0_H < 0.7125$	1	1
$I^0_H = 0.7125$	0 or 1	1
$I^0_H > 0.7125$	0	1

The remaining region to be considered is Region III, where $c > 0.3006$ and optimal policy is $f^*_H = 0$. Let us take $c = 0.5$ as an example. The table below details the values of the relevant variables from the simulations:

Table 7 ($c = 0.5$)

	$f_H = 1$ (path towards A^1_{31})			$f_H = 0$ (path towards A^0_{31})		
I^0_H	I^*_L	I^*_H	V	I^*_L	I^*_H	V
0.6667	0.0000110	0.6316	0.2900	0.00000027	0.8421	**1.3806**
0.7018	0.0000109	0.6316	0.2228	0.00000026	0.8421	**1.3361**
0.7369	0.0000108	0.6316	0.1585	0.00000026	0.8421	**1.2936**
0.7719	0.0000107	0.6316	0.0971	0.00000026	0.8421	**1.2530**
0.8070	0.0000106	0.6316	0.0381	0.00000026	0.8421	**1.2140**

When $c = 0.5$, the optimal policy is $f^*_H = 0$. This is the optimal policy for any c in Region III. Note that all of the above simulations show the same qualitative results for smaller values of I^0_L, namely $I^0_L = 0.01$, $I^0_L = 0.001$ and $I^0_L = 0.0001$.

These simulations show a fairly intuitive result, namely that as costs rise, optimal policy is more likely to be treating no one infected with the H strand. They also demonstrate an interesting finding, whereby there is a small range of costs for which optimal policy is dependent on initial prevalence of infection.

5.2 Hamiltonian paths with variable policy

In this section we allow policy to vary. We also look at strictly Hamiltonian paths i.e. those that satisfy the Hamiltonian conditions for optimality. In order to check that paths are Hamiltonian paths, costate variables are required. We examine each initial point studied above and solve for values of the costate variables at these points using the facts that

$$\begin{aligned} \lambda_H &= \frac{\partial V(I_H, I_L)}{\partial I_H}, \\ \lambda_L &= \frac{\partial V(I_H, I_L)}{\partial I_L}. \end{aligned}$$

These partial derivatives can be approximated by perturbing the infection levels slightly. Thus, for initial infection levels I_H^0 and I_L^0,

$$\lambda_H^0 \approx \frac{V(I_H^0 + \Delta, I_L^0) - V(I_H^0, I_L^0)}{\Delta}, \tag{48}$$

$$\lambda_L^0 \approx \frac{V(I_H^0, I_L^0 + \Delta) - V(I_H^0, I_L^0)}{\Delta}, \tag{49}$$

for small Δ. In these simulations we set $\Delta = 0.001$. Table 8 depicts the Hamiltonian conditions required for our two potential policies to be optimal:

Table 8 (Hamiltonian optimality conditions)

Policy	Condition
$f_H^* = 1, f_L^* = 1$	$\lambda_H^* < -\frac{c}{\alpha}, \lambda_L^* < -\frac{c}{\alpha}$
$f_H^* = 0, f_L^* = 1$	$\lambda_H^* > -\frac{c}{\alpha}, \lambda_L^* < -\frac{c}{\alpha}$

In order to test whether our paths are Hamiltonian paths, for each initial infection level we find the initial costate variables, λ_H^0 and λ_L^0, using (48) and (49). We then test whether either of the two candidate policies satisfies the Hamiltonian conditions. If one does, we simulate the path from this initial point, using values for our costate variables to test for optimal policy at each time increment. This allows policy to vary optimally. We then plot graphs of the evolution of the policy variables over time along with the state variables I_H and I_L. This will allow us to see whether there are any switch points (i.e. changes) in policy, and at what levels of I_H and I_L they occur.

Let us begin with the lowest costs, $c = 0.1$. Solving for the costate variables and checking the Hamiltonian conditions shows that there is a Hamiltonian

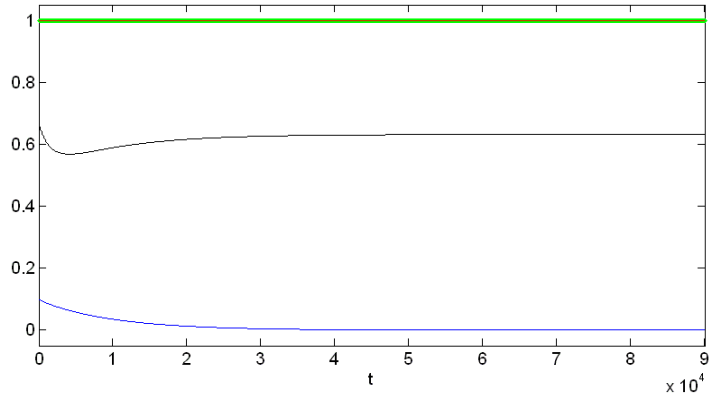

Figure 6: Evolution of system towards fixed point. $I_H^0 = 0.6667, c = 0.1$. Legend: Green=f_H^*, Red=f_L^*, Black=I_H, Blue=I_L

path from each initial point, with optimal policy $(1,1)$ along the entire path. There are no switch points. One example of such a path is depicted in Figure 6. This path is simple. The control variables do not change over time and the system gradually moves towards the fixed point, with I_H dipping slightly before converging to the steady state level. Fixing policy is optimal. This is in line with our findings from the previous section, where moving to A_{31}^1 was optimal when $c = 0.1$.

Further simulations are carried out on paths when $c = 0.295$. Each of our five initial points has a Hamiltonian path. The path with initial value closest to the fixed point has no switch points, while the remaining four paths have one switch point. Simulations show that all paths converge to fixed point A_{31}^1, somewhat unexpectedly as, for some, initial optimal policy is $(0,1)$. These details are shown in Table 9. It is interesting to compare this to the results of the previous section. For the initial point $I_H^0 = 0.6667$, the results are the same. Policy $(1,1)$ is optimal throughout. For the other fixed points, we find that $(0,1)$ is initially optimal, but early on there is an optimal switch to the policy $(1,1)$. The system never optimally converges to A_{31}^0 in contrast to the case of fixed policy. Policy is dependent on initial values, but not in the same way that we observed in the previous section.

Table 9 (Switch points for Hamiltonian paths when $c = 0.295$)

I_H^0	Initial (f_H^*, f_L^*)	Δf_H^*	Δf_L^*	I_H at switch	I_L at switch	I_H^*
0.6667	$(1,1)$	None	None	N/A	N/A	0.6316
0.7018	$(0,1)$	$0 \to 1$	None	0.7408	0.0728	0.6316
0.7362	$(0,1)$	$0 \to 1$	None	0.7558	0.0722	0.6316
0.7719	$(0,1)$	$0 \to 1$	None	0.7626	0.0717	0.6316
0.8070	$(0,1)$	$0 \to 1$	None	0.7692	0.0713	0.6316

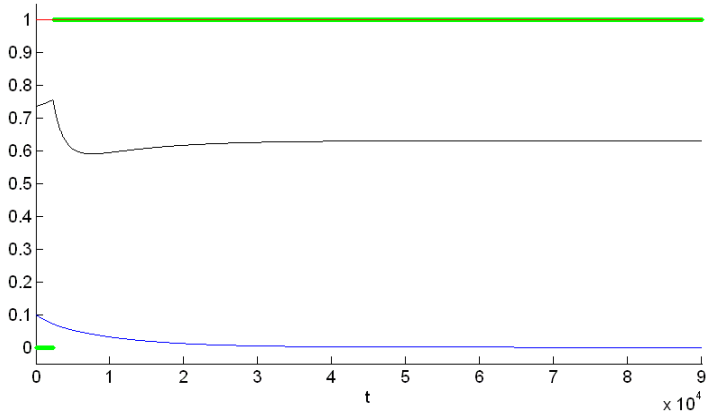

Figure 7: Evolution of system towards fixed point. $I_H^0 = 0.7369, c = 0.295$. Legend: Green=$f_H^*$, Red=$f_L^*$, Black=$I_H$, Blue=$I_L$

An example of one of these paths is depicted in Figure 7. The behaviour is different to what we observed in Figure 6. There is a switch point early on, after which optimal policy remains at $(1, 1)$. Prior to the switch point, prevalence of the H strand rises. It then undershoots, growing slightly to converge to the low prevalence steady state. The intuition behind the switch point is that initially, prevalence of the H strand is not high enough to justify full treatment - the marginal cost of an additional infected person is lower than the relative cost of treatment. As I_H rises, there comes a point when this marginal cost exceeds the relative cost of treatment. At this point, policy switches to treating everyone.

Next, we turn to the example of high costs, when $c = 0.5$. Only three of our five initial points have a Hamiltonian path. Despite each path beginning with an optimal policy of $(0, 1)$, similar to our results in the previous section, all three paths switch to the policy $(1, 1)$ after a short period and converge to the fixed point with lower infection level. Thus, the paths we derived in the previous section when $c = 0.5$ were not Hamiltonian along their entirety. The details of the switch points when $c = 0.5$ are given in the table below:

Table 10 (Switch points for Hamiltonian paths when $c = 0.5$)

I_H^0	Initial (f_H^*, f_L^*)	Δf_H^*	Δf_L^*	I_H at switch	I_L at switch	I_H^*
0.6667	$(0, 1)$	$0 \to 1$	None	0.7608	0.0619	0.6316
0.7018	$(0, 1)$	$0 \to 1$	None	0.7640	0.0617	0.6316
0.7369	$(0, 1)$	$0 \to 1$	None	0.7671	0.0615	0.6316

Figure 8 depicts an example of such a path. The behaviour of the system is similar to the case when $c = 0.295$. There is a sharper rise in I_H than in the previous example, but the system still converges to A_{31}^1 after the policy switch.

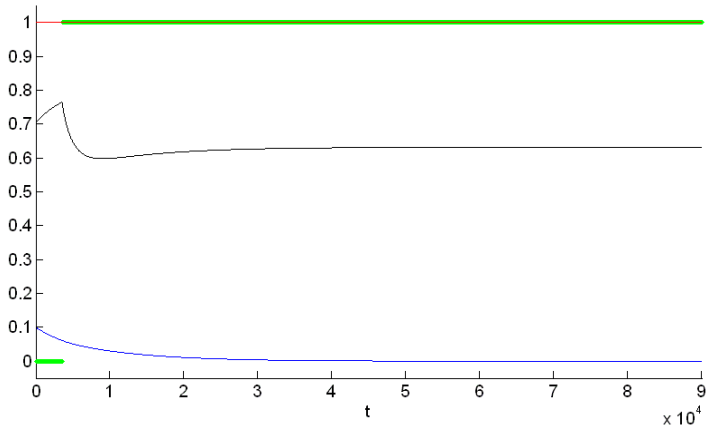

Figure 8: Evolution of system towards fixed point. $I_H^0 = 0.7018, c = 0.5$. Legend: Green=f_H^*, Red=f_L^*, Black=I_H, Blue=I_L

There are several points to take away from these simulations. First, when policy is variable, all simulations converge to the fixed point A_{31}^1. This is interesting because, despite the policy $(0,1)$ being optimal on segments of some of the paths, it is still optimal to converge to the low infection state. In constrast, a policymaker who has to fix policy in advance is more likely to choose not to treat the H strand as costs rise, allowing it to converge to a higher steady state prevalence level. Thus, the simulations suggest that it is always better to attempt to lower prevalence of the H strand as much as possible. Another observation to note is that when costs are sufficiently low, our paths are simple, with no switch points. They retain the same optimal policy that they began with, namely $(1,1)$. Thus, fixed policy is optimal at low cost levels. As costs rise, the paths become more complex. We observe switch points and the policymaker is better off if she has flexibility in her actions.

6 Conclusion

This paper has explored an SIS model with two variants of infection differentiated by transmission risk. It has been shown that there are two types of steady states. First, there is a set of fixed points with one treatment level at the boundary and one at an interior level. These fixed points form two lines in (I_H, I_L) space and are only feasible under certain parameter combinations. Only the total level of infection is pinned down here; the distribution of this total infection between the two strains will depend on initial levels. Optimal policy for these steady states is derived; along the path, optimal policy is always

at a boundary, after which it may switch to an interior level when steady state is reached. There are also asymptotic fixed points that involve asymptotic eradication of one strand, while the other strand remains endemic. Under the same parameter combinations that eliminate the interior fixed points, we are left only with those asymptotic fixed points that asymptotically eradicate the L strand, leaving the H strand to prevail. This is interesting as it suggests that sometimes it may be optimal for the policymaker to focus treatment on the less infective strand, which may seem counterintuitive. It was also shown that simultaneous asymptotic eradication of both strands is not possible.

Simulations focus on the case when only asymptotic eradication of the L strand is feasible. We consider two cases: when the policymaker must keep policy fixed throughout the period, and when policy is flexible. We vary costs and compare policy across different parameter combinations. The results are insightful. When policy is fixed, there is a clear relationship between costs and optimal policy. For low costs, optimal policy is always to treat everyone. As costs rise, there is a small range where optimal policy is dependent on initial value. The closer is initial infection to the no-treatment steady state, the more likely it is that optimally the H types are left untreated. As costs rise even further, it is optimal not to treat the H strain, regardless of initial prevalence. These are intuitive results; the higher the cost, the more likely we are to curb expenditure.

When policy is allowed to vary, we observe further intriguing results. In our examples, all paths converge to the steady state where the H strain is treated, even at high costs. At low costs, optimal policy is fixed and there is no added benefit from being able to vary policy. As costs rise, policy exhibits switch points. Optimal policy begins by not treating the H infection and switches to full treatment after a short period of time. Thus, as costs rise, there is added benefit from variable policy.

There are several points to take away from these results. First, there are many possible steady states, and feasibility will depend on paremeters. When there is a large difference between the infectivities of the two strains and low natural rate of recovery, it may be that the policymaker can only hope to asymptotically eradicate the L strand, while making a decision on whether or not to treat the H strand. Second, simulations show that if this is the case, optimal fixed policy is clearly related to cost of treatment. Optimal variable policy will always point the policymaker towards eventually treating the H strand and converging to the steady state with lower overall infection. Therefore, our policy recommendation is that treatment agencies should negotiate flexible terms with their suppliers and their governments, so that they have the option to change policy over time if it is optimal.

Further research should consider extending the simulations to look at other parameter combinations. In particular, the interior fixed points have not been simulated here. This is an essential next step in order to make a complete judgment on optimal policy under different parameter scenarios. It would also be interesting to extend this model to include protection via vaccination as another instrument available to the policymaker.

References

[1] Ahituv, A., V. J. Holz and T. Philipson. "The Responsiveness of the Demand for Condoms to the Local Prevalence of AIDS." *Journal of Human Resources* 31.4 (1996): 869-897.

[2] Anderson, R. M. and R. M. May. *Infectious Diseases of Humans: Dynamics and Control.* Oxford: Oxford University Press, 1991.

[3] Auld, M. C. "Estimating Behavioural Response to the AIDS Epidemic." *Contributions to Economic Analysis and Policy* 5.1 (2006).

[4] Barrett, S. "Global Disease Eradication." *Journal of the European Economic Association* 1.2-3 (2003): 591-600.

[5] Boulier, B. L., T. S. Datta and R. S. Goldfarb. "Vaccination Externalities". *B. E. Journal of Economic Analysis and Policy* 7.1 (2007)

[6] Brito, D.L., E. Sheshinski and M. D. Intriligator. "Externalities and Compulsory Vaccinations." *Journal of Public Economics* 45.1 (1991): 69-90.

[7] Castillo-Chavez, C., W. Huang and J. Li. "Competitive Exclusion and Coexistence of Multiple Strains in an SIS STD Model." *SIAM Journal of Applied Mathematics* 59.5 (1999): 1790-1811.

[8] Dupas, P. "Relative Risks and the Market for Sex: Teenagers, Sugar Daddies and HIV in Kenya." (2005).

[9] Epstein, J. M. "Modelling to Contain Pandemics." *Nature* 460.7256 (2009): 687.

[10] Geoffard, P.-Y. and T. Philipson. "Rational Epidemics and Their Public Control." *International Economic Review* 37.3 (1996): 603-624.

[11] Geoffard, P.-Y. and T. Philipson. "Disease Eradication: Private versus Public Vaccination." *American Economic Review* 87.1 (1997): 222-230.

[12] Gersovitz, M. and J. S. Hammer. "Infectious Diseases, Public Policy and the Marriage of Economics and Epidemiology." *World Bank Research Observer* 18.2 (2003): 129-157.

[13] Gersovitz, M. and J. S. Hammer. "The Economical Control of Infectious Diseases." *Economic Journal* 114.492 (2004): 1-27.

[14] Goldman, S.M. and J. Lightwood. "Cost Optimisation in the SIS Model of Infectious Disease with Treatment." *Topics in Economic Analysis and Policy* 2.1 (2002): 1-22.

[15] Greenwood, J., P. Kircher and M. Tertilt. "An Equilibrium Model of the African HIV/AIDS Epidemic." (2010).

[16] Hethcote, H. W. and J. W. Van Ark. "Epidemiological Models for Heterogeneous Populations: Proportionate Mixing, Parameter Estimation, and Immunization Programs." *Mathematical Biosciences* 84.1 (1987): 85-118.

[17] Hyman, J. M. and J. Li. "Behaviour Changes in SIS STD Models with Selective Mixing." SIAM Journal of Applied Mathematics 57.4 (1997): 1082-1094.

[18] Keeling, M. J. and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton, New Jersey: Princeton University Press, 2008.

[19] Kremer, M. "Integrating Behavioural Choice into Epidemiological Models of AIDS." *Quarterly Journal of Economics* 111.2 (1996): 549-573.

[20] Oster, E. "Sexually Transmitted Infections, Sexual Behaviour and the HIV/AIDS Epidemic." *Quarterly Journal of Economics* 120.2 (2005): 467-515.

[21] Oster, E. "HIV and Sexual Behaviour Change: Why not Africa?" (2009).

[22] Philipson, T. and R. A. Posner. *Private Choices and Public Health: An Economic Interpretation of the AIDS Epidemic*. Cambridge, MA: Harvard University Press, 1993.

[23] Reluga, T. C. "An SIS Epidemiology Game with Two Subpopulations." *Journal of Biological Dynamics* 3.5 (2009): 515-531.

[24] Rowthorn, R. E., R. Laxminarayan and C. A. Gilligan. "Optimal Control of Epidemics in Metapopulations." *Journal of the Royal Society Interface* 6 (2009): 1135-1144.

[25] Rowthorn, R. E. "Optimal Treatment of Disease Under a Budget Constraint." *Explorations in Environmental and Natural Resource Economics: Essays in Honor of Gardner M. Brown, Jr.* Ed. R. Halvorsen and D. F. Layton. Cheltenham, UK: Edward Elgar Pub., 2006. 20-35.

[26] St Lawrence, J. S., Kelly, J. A., Diaz, Y.E. et al. "HIV Risk Behaviour Reduction following Intervention with Key Opinion Leaders of Population: An Experimental Analysis." *American Journal of Public Health* 81.2 (1991): 168-171.

[27] Toxvaerd, F. "Recurrent Infection and Externalities in Protection." (2009).

[28] Toxvaerd, F. "Recurrent Infection and Externalities in Treatment" (2009).