The Economic Dynamics of Antibiotic Efficiency under Open Access

Markus Herrmann,

Université de Montréal

(preliminary version, April 20th 2006,

do not quote without permission)

Société canadienne de science économique

Congrès annuel 2006

1 Introduction

This paper studies some economic aspects of the problem raised by the fact that antibiotic consumption tends to deplete the efficiency of many antibiotics in combating bacterial infections, as the bacteria develops resistance to the antibiotic. We analyze the exploitation of antibiotic efficiency in a market subject to open access on the part of antibiotic producers to the common pool of efficiency and compare it to the social optimum. We derive the demand function for the antibiotic under the assumptions that individuals differ with respect to their valuation of being in good health and are aware of the probability that the antibiotic treatment is effective. Antibiotic efficiency is modelled as a common pool renewable resource. The dynamics of the antibiotic efficiency is based on an epidemiological model (the SIS-model) borrowed from the biological literature, with one antibiotic treatment for one particular bacterial infection. The open access equilibrium and the social optimum are derived subject to the demand for the antibiotic and the underlying epidemiological model, which describes the dynamic interaction between the level of efficiency of the antibiotic and the level of infection in the population.

We find that the industry under open access does not completely deplete the stock of antibiotic efficiency. Rather, the equilibrium levels of efficiency and infection tend to a steady state where the efficiency renews itself to maintain the steady state. This market outcome is however not socially optimal. The optimal solution may combine extreme controls (treat all of the infected population or treat none) and singular controls (treat a fraction which changes over time), or it may involve only an extreme control, depending on the parameters

of the model. The socially optimal solution also leads to a steady state, which will in general be different from the open access steady state. The steady state is reached faster in the social optimum than in the open access equilibrium. As the social planner also values individuals that are in good health, he has an interest to decrease infection faster, as this will reduce the spread of the infection in the future. This positive intertemporal externality drives the working of the optimal solution and leads to a lower steady-state level of efficiency than in open access. The steady-state levels of infection are in general the same as in open access.

Analyzing the problem of bacterial resistance due to antibiotic treatment in the context of economic models is relatively new. Some contributions, beyond others, that have been made are from Laxminarayan and Brown (2001), Rowthorn and Brown (2003) as well as Wilen and Msangi (2003). These contributions all use a version of a SIS-model and differ with respect to either the renewability of the antibiotic efficiency or the number of antibiotic treatments available. We review these contributions here shortly.

With the help of analytical case studies, Laxminarayan and Brown (2001) show in the context of a single bacterial infection and two antibiotic treatments having non-renewable efficiency, that the more effective treatment should be applied first if both antibiotics have the same treatment cost, and that the less costly should go first if both antibiotics have initially the same efficiency. The second antibiotic should only come into use if either it has the same efficiency as the one used first or if the implicit value associated to the initially more effective antibiotic has sufficiently decreased to compensate for the higher treatment cost of the second antibiotic. The proportion of the antibiotic in simultaneous use is inversely related to the rapidity with which efficiency decreases and the cost differential.

Rowthorn and Brown (2003) study the case of two antibiotics used to treat two infections, each antibiotic being only effective with respect to one infection. Society must decide on the allocation of its resources between the two treatments ; it is supposed to be capacity constraint as there cannot be double treatments within infected individuals. The allocation depends on the treatment costs and on the social cost associated to each of the infections. The authors present an arbitrage condition between the less expensive antibiotic and the infection implying a higher social cost due to its predominance.

Wilen and Msangi (2003) define an optimal control problem with constraints from the SIS-model referring to only one antibiotic treatment available for a single infection. Efficiency of the antibiotic is supposed

renewable. The authors show that after a period of an initially extreme control, a singular control applies : the fraction of the population treated tends to the level where efficiency is renewable and can remain constant. This singular control is determined analytically as a feedback solution with respect to the current state of efficiency and of infection with respect to their steady state values.

Our paper is structured as follows. In section 2 we present the SIS-model from which we define antibiotic efficiency and derive the constraints for the optimal control problem of the social planner. In section 3 we derive the demand for the antibiotic on the basis of a model of vertical differentiation. The market equilibrium under open access is derived in section 4. In section 5, we characterize the socially optimal outcome. In section 6, we compare market outcome and social optimum. We conclude in section 7.

2 The Epidemiological Constraints

Before deriving the market demand for the antibiotic, we have to consider the epidemiological constraints. We suppose that total population at time t is given by $N(t) = S(t) + I(t) = S(t) + I_w(t) + I_r(t)$, where S refers to the population being in good health and I is the population infected. I can be further partitioned into those individuals infected with a drug-susceptible strain (I_w) and those infected with a drug-resistant strain (I_r). An individual being infected ignores whether he is infected with a drug sensitive strain or a drug resistant strain.[1] Population dynamics leading to the SIS-model have been derived first by Ross (1911) and Kermack and McKendrick (1927) in order to model the spread of malaria. The evolution of the partition into healthy and infected population is captured by the following equations (the time indices are omitted) :

$$\frac{dS}{dt} = E - nS - \beta S(I_w + I_r) + r_w I_w + r_r I_r + f I_w r_f \quad (1)$$

$$\frac{dI_w}{dt} = (\beta S - m - r_w - f r_f) I_w \quad (2)$$

$$\frac{dI_r}{dt} = (\beta S - m - r_r) I_r \quad (3)$$

[1] It is of course possible to verify the sensitivity of the infectious strain with respect to the antibiotic, but we suppose that this is not generally done because of costs for analyzing the strain in a laboratory, costs of delay, and the potential difficulty to take a sample from the herd of infection.

E refers to the number of entries into the system (births), n and m are death rates of the healthy and infected population respectively. The parameter β represents the transmission rate of the infection between healthy and infected population. r_r and r_w refer to the natural recovery rates - *i.e.* without any treatment - from the drug-susceptible and the drug-resistant strain respectively, while r_f is the augmented recovery treatment for effective antibiotic treatment. In the equations stated above f refers to the fraction of the infected population that takes treatment. For total population to be constant over time, we impose $E = n = m = 0$.

It is clear that because of the constant population, equation (1) will be redundant and we can restrict our analysis on equations (2) and (3). We want to derive the dynamics for the state variables of interest $w(t)$ and $I(t)$, where $w(t)$ is defined as the proportion of the infected population that is suffering from a drug-susceptible strain and can be interpreted as the efficiency (rate) of the antibiotic treatment. We have

$$\begin{aligned} \frac{dw}{dt} &= \frac{d\left(\frac{I_w}{I}\right)}{dt} = \frac{\dot{I}_w I - I_w \dot{I}}{I^2} \text{ and} \\ \frac{dI}{dt} &= \frac{dI_w}{dt} + \frac{dI_r}{dt}. \end{aligned}$$

Replacing $S(t) = N - I(t)$, and using equations (2) and (3), latter equations can be rewritten to obtain :

$$\frac{dw}{dt} = w(1-w)(\Delta r - r_f\ f) \tag{4}$$

$$\frac{dI}{dt} = (\beta(N-I) - r_r)I + wI(\Delta r - r_f\ f) \tag{5}$$

where $\Delta r = r_r - r_w$ refers to the fitness cost of infection.[2] When the fitness cost equals zero, then efficiency of the antibiotic drug can be considered as a nonrenewable resource as equation (4) shows. For positive values of f, $\frac{dw}{dt}$ is negative. However, when the fitness cost is positive, the efficiency of the antibiotic becomes a renewable resource. The fitness cost is a crucial element in the analysis of antibiotic resistance. Our formulation allows to treat all cases such that $\Delta r = 0$.

For later analysis we characterize the steady states, w^{SS} and I^{SS}, at which $\frac{dw}{dt} = \frac{dI}{dt} = 0$. As can be seen

[2] The term "fitness cost" has been before in the economic literature on antibiotic resistance and was borrowed from the epidemiological vocabulary. It refers to the fact that resistant bacterial strains naturally clear more frequently (at rate r_r) as do their antibiotic susceptible twins (at rate r_w). The feature of resistance thus comes at a cost of survival for the antibiotic resistant strains.

from equation (4), three different steady states for efficiency are possible : (1) $w^{SS} = 0$, (2) $w^{SS} = 1$ and (3) any value for efficiency as long as the fraction of the infected population receiving the antibiotic is $f^{SS} = \frac{\Delta r}{r_f}$. Combining the information on the three steady states for efficiency with equation (5) and setting the latter equal to zero, we obtain the steady state values for infection. The steady states are summarized in the following table :

Steady State	w^{SS}	I^{SS}
(1)	0	$\frac{\beta N - r_r}{\beta}$
(2)	1	$\frac{\beta N - r_w - r_f \ f^{SS}}{\beta}$
(3)	n.d. with $f^{SS} = \frac{\Delta r}{r_f}$	$\frac{\beta N - r_r}{\beta}$

It turns out that the system of differential equations is nonautonomous. In our context, this means that it is not possible to draw a unique phase diagram for changing values of $f \in [0, 1]$, which, we remind is the fraction of the infected population being treated with the antibiotic. This would actually imply the construction of a three-dimensional phase diagram, which is a rather cumbersome job.

In order to construct the phase diagrams in the plane (w, I) with respect to f, we concentrate on certain values that f may take and characterize the isoclines for w and I given by $\frac{dw}{dt} = 0$ and $\frac{dI}{dt} = 0$. It is directly seen, that $\frac{dw}{dt} = 0$ implies $w = 0$, $w = 1$ or $f = \frac{\Delta r}{r_f}$ (refer to the steady states before). The isocline for I is obtained after setting equation (5) equal to zero and solving for, say w. This gives for $r_f f \neq \Delta r$:

$$w = \frac{r_r - \beta N}{\Delta r - r_f \ f} + \frac{\beta}{\Delta r - r_f \ f} I \tag{6}$$

It is important to recognize that equation (6) is not defined for $f = \frac{\Delta r}{r_f}$ and that it represents a line with negative or positive steepness depending on whether $f > \frac{\Delta r}{r_f}$ or $f < \frac{\Delta r}{r_f}$. In the case of $f = \frac{\Delta r}{r_f}$, the isocline for I will be a vertical line passing through the steady state $I^{SS} = \frac{\beta N - r_r}{\beta}$.

We can now draw the phase diagrams for $f \in [0, \frac{\Delta r}{r_f})$, $f \in (\frac{\Delta r}{r_f}, 1]$ or $f = \frac{\Delta r}{r_f}$. Consider Figures 1.a, 2 and 3 respectively.[3]

[3]Figures1.a and 2 have already been presented in Msangi and Wilen (2003) for the limit values for f, *i.e.* $f = 0$ and $f = 1$. We want to point out that it is possible to contruct a respective phase diagram for all values of f belonging to a particular

The forces driving the state variables when not on the isoclines are indicated by arrows. Latter are obtained for w, by noting that when $f > \frac{\Delta r}{r_f}$, then w is decreasing (respectively, $f < \frac{\Delta r}{r_f}$ implies w to be increasing), and $f = \frac{\Delta r}{r_f}$ implies w to remain at any initial value. We can thus trace the vertical arrow in the phase diagram (consider Figure 1.a, b and 2). In Figure 3, a vertical movement is not possible as $\frac{dw}{dt} = 0$ everywhere.

In order to derive the forces that are driving I in the phase diagram, we can calculate the derivative $\frac{d(\dot{I}/I)}{dw}$. It is easily verified that $\frac{d(\dot{I}/I)}{dw} = \Delta r - r_f f$. The derivative is thus positive (negative) in the Figure 1.a (in Figure 2) and we can check the direction of the horizontal arrow. In the case of $f = \frac{\Delta r}{r_f}$ (Figure 3), equation (5) simplifies to

$$\frac{\dot{I}}{I} = \beta(N - I) - r_r,$$

which implies $\frac{d(\dot{I}/I)}{dI} = -\beta < 0$. The arrows thus point towards the vertical isocline which represents also all steady states for w and I.

From this analysis, we can conclude that the system will converge for a given value of f to steady state (1), at which $w^{SS} = 0$, if $f \in (\frac{\Delta r}{r_f}, 1]$ (see Figure 2), and that the system will converge to steady state (2), at which $w^{SS} = 1$, if $f \in [0, \frac{\Delta r}{r_f})$ (see Figure 1.a). Figure 3 shows the location of steady state (3) for $f = \frac{\Delta r}{r_f}$.

We have mentioned before that the system is nonautonomous. Graphically, this characteristic of the dynamic system is captured by the fact that the I-isocline pivots for changing values of f.

As an example, consider two values of f, with $f_i \in [0, \frac{\Delta r}{r_f}), i = 1, 2.$ and $f_1 < f_2 < \frac{\Delta r}{r_f}$. f_i remains constant over time. Then, for the two values of f, the phase diagram has the form that has been specified in Figure 1.a. In Figure 1.b we characterize the pivoting I-isocline : it has a steeper slope for values of f that are closer to $\frac{\Delta r}{r_f}$. We have also indicated the forces driving the system when out of the isoclines (dotted lines are for $f = f_2$). The system will converge to steady state (2), given by $w = 1$ and $I^{SS}_{(2)}(f_i) = \frac{\beta N - r_w - r_f f_i}{\beta}$, with $I^{SS}_{(2)}(f_2) < I^{SS}_{(2)}(f_1)$ respectively. Indeed, when a relatively higher fraction of people is treated with the antibiotic (here f_2), the overall recovery rate from infection is higher and thus less individuals will be infected at the steady state. This implies that for any value of efficiency different from its steady state value, say $\bar{w}$,

intervall.

the level of infection at which $\dot{I} = 0$ and indicated by $\bar{I}\mid_{f_i}$ should be lower, when a higher fraction of the infected population is treated. Thus, the I-isocline should have a lower slope when a lower fraction of the infected population is treated.

Dynamics become even more complicated when f changes over time, say from lower values (like f_1) to higher values (like f_2). The I-isocline will pivot and become steeper as f rises. It is however impossible to draw a characteristic path leading to the steady state (2), as has been done in Figure 1.a, although we know that the system still reaches this steady state, and that, when f approaches $\frac{\Delta r}{r_{\mathrm{f}}}$, the I-isocline becomes the vertical shown in Figure 3.

We have described the nonautonomous system for values $f_i \in [0, \frac{\Delta r}{r_{\mathrm{f}}})$. The same reasoning applies also to f_i $\in (\frac{\Delta r}{r_{\mathrm{f}}}, 1]$.

To which one of the three steady states the system will tend, depends on the intertemporal evolution of f. Before characterizing the evolution of f in the case of open access and the social optimum, we have to develop the demand for the antibiotic.

3 The Demand

The demand function for the antibiotic at time t is derived from a classical model of vertical differentiation. Suppose a consumer has a valuation θ of being in good health. θ is distributed over population N with distribution function $F(\theta)$.

Suppose further that a consumer can decide whether to buy an antibiotic at price p or not when he is infected. The utility of the assumed type θ can be represented as :

$$u(\theta) = \left\{ \begin{array}{c} \theta, \text{ if in good health} \\ \\ EU(\theta, w, p), \text{ if infected} \end{array} \right\}$$

with $EU(\theta, w, p)$ being the utility resulting from the expected utility maximization problem of the infected individual, where the decision variable is whether to take or not to take the antibiotic if infected.[4] It is given

[4] Actually, it is the physician who prescribes the antibiotic to the patient. The patient then decides whether to purchase the antibiotic or not. This decision probably also depends on the insurance coverage of the patient. In the model that we present, we abstract from this fact. It could be easily introduced into the model, when the physician, following a "rule of thumb", prescribes the antibiotic to a constant fraction of the infected population. He may however have a strategy, such as to minimize

by :

$$EU(\theta, w, p) = \max\{\pi(w)\theta, [\pi(w) + r_f w]\theta - p\}$$

where $\pi(w) = wr_w + (1-w)r_r$ represents the probability of recovering without antibiotic treatment when infected and $[\pi(w) + r_f w]$ is the probability of recovery when infected and taking the antibiotic. We obtain $\pi(w)$ in the following way : an individual that is infected has a risk of $\frac{I_r}{I} = 1 - w$ of being infected with a drug-resistant strain. He thus recovers with "probability" r_r whether he takes the antibiotic or not. He has a risk of being infected with the drug-susceptible strain given by : $\frac{I_w}{I} = w$ and a recovery rate of r_w in that case. The mean recovery rate when not buying the antibiotic is thus given by the sum $\pi(w)$. In the case where the individual buys the antibiotic, he increases his chances of recovery only if the drug is effective, *i.e.* the bacterial strain he is suffering from is susceptible to treatment. Thus, we have to add the term $r_f w$ for the recovery probability when buying the antibiotic.

Denote by $\tilde{\theta}$ the consumer who is indifferent between buying the antibiotic or not. $\tilde{\theta}$ is given by :

$$\begin{aligned} \pi(w)\tilde{\theta} &= [\pi(w) + r_f w]\tilde{\theta} - p \Leftrightarrow \\ \tilde{\theta} &= \frac{p}{r_f w}. \end{aligned} \tag{7}$$

Individuals with $\theta > \tilde{\theta}$ will thus buy the antibiotic, and those with $\theta < \tilde{\theta}$ will not. If the whole population N were infected, the proportion of individuals willing to buy the antibiotic would be $[1 - F(\tilde{\theta})]$. But this is not the case : uninfected individuals won't buy the antibiotic. If we assume that the infection spreads equally over the population N (being infected and having a certain valuation θ are independent events[5]), then the fraction of the infected population willing to buy the antibiotic is given by : $\frac{I}{N}\left[1 - F\left(\tilde{\theta}\right)\right]$. Since consumers

the infection level at time t, without questioning the effects on future infection levels. Or he may make the fraction depend on the levels of efficiency and infection, for instance. This point merits more attention and is left for future research. It points out the question of whether an "intelligent" strategy for the physician exists allowing the society to attain the socially optimal outcome.

[5] Define the joint probability of an individual i to be infected ($i = I$) and heaving a valuation higher than $\tilde{\theta}$ as $\Pr(i = I, \theta_i \geq \tilde{\theta})$. Then, by independence, we have $\Pr(i = I, \theta_i \geq \tilde{\theta}) = \Pr(i = I)\Pr(\theta_i \geq \tilde{\theta}) = \frac{I}{N}\left[1 - F\left(\frac{p}{r_f w}\right)\right]$.

have a unitary demand, total demand is given by :

$$\begin{aligned} Q &= N\frac{I}{N}\left[1-F\left(\tilde{\theta}\right)\right] \\ &= I\left[1-F\left(\frac{p}{r_f w}\right)\right]. \end{aligned}$$

Therefore the inverse demand function is :

$$P\left(\frac{Q}{I},w\right) = r_f w F^{-1}\left(1-\frac{Q}{I}\right) \tag{8}$$

If we assume a unitary distribution for θ, the inverse demand function becomes :

$$P\left(\frac{Q}{I},w\right) = r_f w\left(1-\frac{Q}{I}\right) \tag{9}$$

Because of unitary demand, $\frac{Q}{I}$ represents the fraction of the infected population treated and is thus equal to the parameter f in the dynamic constraints (4) and (5). The inverse demand function depends on the fraction of the infected population being treated ($\frac{Q(t)}{I(t)}$) as well as the efficiency of the antibiotic drug ($w(t)$). Our formulation of the antibiotic demand supposes that individuals are not perfectly informed about the kind of infection they suffer from, but they know the state $w(t)$ of the system, as does the producer of the antibiotic.

For later analysis, we calculate the marshallian consumer surplus ($CS(Q)$) associated to the inverse demand function (9). It is given by :

$$\begin{aligned} CS(Q) &= \frac{1}{2}\left[r_f w - r_f w\left(1-\frac{Q}{I}\right)\right]Q \\ &= \frac{1}{2} r_f w \frac{Q^2}{I}. \end{aligned} \tag{10}$$

4 Open Access Equilibrium

Suppose that the antibiotic industry inherits a stock of efficiency and infection at time 0 of (w_0, I_0). Under open access, producers will enter until, at equilibrium, price equals average production costs. Let Q^{oa} be the

total amount of the antibiotic sold in the market under open access, and q_i^{oa} be the amount sold by producer i. Assuming the firms to be identical, we must have, in equilibrium,

$$q_i^{oa} = \frac{Q^{oa}}{n} \equiv q^{oa}$$

such that

$$P(\frac{nq^{oa}}{I}, w) = \frac{C(q^{oa})}{q^{oa}}. \tag{11}$$

If n need not be an integer, this condition implies zero producer surplus for each firm at any time.[6]

Suppose as a simple case that the cost function is given by $C(q_i) = cq_i$, where $c > 0$.[7] Then the open access condition (11) becomes

$$P(\frac{Q^{oa}}{I}, w) = c \tag{12}$$

and we can easily derive Q^{oa} to be :

$$Q^{oa}(t) = I(t)\left(1 - \frac{c}{r_f w(t)}\right) \tag{13}$$

Thus, the fraction of the infected population treated with the antibiotic under open access is given by :

$$f^{oa}(t) = \frac{Q^{oa}(t)}{I(t)} = \left(1 - \frac{c}{r_f w(t)}\right), \tag{14}$$

which is strictly positive, as long as $r_f w(t) > c$. Consequently, when this last condition is verified for a given value of $w(t)$, the antibiotic is economically viable.[8]

We first analyze the regime of open access at the different steady states. Second, we analyze whether the

[6] If we want n to be an integer only, then n is the biggest number such that a firm turns over, while $n+1$ would imply negative profits for each firm. We may then have positive profits for each incumbent firm under open access as long as the antibiotic's efficiency is economically viable.

[7] This allows us not to consider the number of firms n in the market.

[8] As the maximum value for w is one, we must have as a necessary condition for a non-trivial problem : $r_f > c$.

system converges to such a steady state.

4.1 Steady State under Open Access

At steady state (1) with $w_{(1)}^{SS} = 0$ and $I_{(1)}^{SS} = \frac{\beta N - r_r}{\beta}$ no firm would like to enter the market as it would incur negative profits. In this case demand has vanished and therefore, in equilibrium, $Q_{(1)}^{oa} = 0$.

At steady state (2), $w^{SS} = 1$ and $I^{SS} = \frac{\beta N - r_w - r_f \ f^{ss}}{\beta}$. Replacing the value for f^{oa} from condition (13) gives :

$$w_{(2)}^{SS} = 1 \tag{15}$$

$$I_{(2)}^{SS} = \frac{\beta N - r_w - r_f + c}{\beta} \tag{16}$$

$$Q_{(2)}^{oa} = \frac{\beta N - r_w - r_f + c}{\beta}\left(1 - \frac{c}{r_f}\right) \tag{17}$$

At steady state (3), the value for w hasn't been determined yet. The open access condition (13) in combination with the steady state condition for efficiency, $f^{oa} = \frac{Q^{oa}}{I} = \frac{\Delta r}{r_f}$ give : $\frac{\Delta r}{r_f} = 1 - \frac{c}{r_f w}$. This implies

$$w_{(3)}^{SS} = \frac{c}{r_f - \Delta r} \tag{18}$$

$$\text{and } I_{(3)}^{SS} = \frac{\beta N - r_r}{\beta}. \tag{19}$$

The sustainable amount of antibiotics produced is given by equation (13) after replacing (18) and (19) as :

$$Q_{(3)}^{oa} = \frac{\beta N - r_r}{\beta}\left(1 - \frac{r_f - \Delta r}{r_f}\right), \tag{20}$$

which is independent of the marginal cost of production c, but one should remind that the fraction served by the antibiotic produces f^{oa} depends on c. This is so because the inverse demand function, *i.e.* the valuation of the antibiotic, which must equal c in every point of time, can be expressed in terms of the fraction f, and thus depends only on the fraction and the level of efficiency. At steady state terms cancel out such that the quantity does not depend on the marginal cost of production.

4.2 Evolution Towards Equilibrium

The firms that are in the market under open access start at time $t = 0$ with a stock of efficiency, $w(0) = w_0$, and a stock of infected population, $I(0) = I_0$. As long as $r_f w > c$, the antibiotic remains economically viable and firms produce a positive amount. In what follows, we will show that the state variables will converge asymptotically to steady state 3 from any couple of initial interior values (w_0, I_0).[9] Figure 4 shows four representative paths of the state variables w, I leading to steady state 3 from any initial state (w_0, I_0) lying in regions I, II, III and IV. We distinguish between these regions as they allow us to specify, whether the fraction f^{oa} is bigger or smaller than its critical efficiency-renewable value $\frac{\Delta r}{r_f}$. Clearly, in regions I and II, $w_0 > w^{SS}_{(3)}$, while in regions III and IV, $w_0 < w^{SS}_{(3)}$. The respective limit between regions I and II on the one hand, and III and IV on the other hand, will be determined by the pivoting I-isocline. In Figure 4, we show the partition of regions with respect to the I-isocline when it is a vertical line through steady state value $I^{SS}_{(3)}$.

Evaluate $f^{oa}(t)$ from equation (14) at (w_0, I_0). In regions I and II, initial values (w_0, I_0) are such that :

$$f^{oa}(0) = \frac{Q^{oa}(0)}{I_0} = 1 - \frac{c}{r_f w_0} > \frac{\Delta r}{r_f},$$

which implies that the I-isocline has a negative slope (recall equation 6) and that the forces driving the dynamic system initially are those that have already been shown in Figure 2. As $f^{oa} > \frac{\Delta r}{r_f}$, efficiency will be decreasing, *i.e.* $\dot{w} < 0$ and will lead to a decreasing level of the fraction f^{oa} served. While efficiency will monotonously decrease down to its steady state value $w^{SS}_{(3)} = \frac{c}{r_f - \Delta r}$, where f^{oa} attains $\frac{\Delta r}{r_f}$, the evolution of the level of infection depends on whether one starts to the left or to the right of the I-isocline. In any case, the I-isocline is pivoting to the north-east as the system is non-autonomous and f^{oa} changes. When I_0 lies to the left of the I-isocline, infection will monotonously increase up to its steady state value. When I_0 lies to the right of the I-isocline, the level of infection will perform what one may call "undershooting", which means that it temporarily falls below its steady state value but attains it finally. These cases are depicted in

[9] By "interior" we mean that the initial state of efficiency is not on its defintion limits, *i.e.* $w_0 \neq 0$, $w_0 \neq 1$. Also, for the level of infection, we must have $I_0 \neq 0$. Differently, the whole population remains in good health for all time (consider equation 5) and there is no market for the antibiotic.

Figures 5.a and 5.b.

In Figure 5.a, (w_0, I_0) lies in region I and forces drive the system in the south-eastern direction, which implies a depletion of efficiency and an increase in the infected population. However, the state variables always remain to the left of the pivoting I-isocline. Efficiency and infection attain the steady state asymptotically.

For the efficiency variable, this can be shown by linearizing the law of motion equation for w evaluated at f^{oa} around steady state 3. Combining equations (4) and (14) gives the linearized equation :

$$\begin{aligned} \dot{w}(t) &= \left((r_f - \Delta r)(w_{(3)}^{SS} - 1)\right)\left[w(t) - w_{(3)}^{SS}\right] \\ &\equiv \varkappa_1 \left[w(t) - w_{(3)}^{SS}\right]. \end{aligned} \tag{21}$$

The change in w at time t is the product of the distance $\left[w(t) - w_{(3)}^{SS}\right]$ and the term $\varkappa_1 < 0$, with $|\varkappa_1| < 1$. For equation (21), this implies that $\dot{w} < 0$ if $w(t) > w_{(3)}^{SS}$, that $\dot{w} > 0$ if $w(t) < w_{(3)}^{SS}$ and that efficiency approaches the steady state level only asymptotically.

The level of infection cannot neither attain its steady state value in finite time. Suppose infection were to attain its steady state in finite time T and that, in term, $\dot{I}(T) = 0$. Then, from Figure 5.a, the efficiency level would necessarily be greater than its steady state level, *i.e.* $w(T) = \frac{c}{r_f - \Delta r - \epsilon} > w_{(3)}^{SS}$, with $\epsilon > 0$. From equation (14), the fraction served would be

$$f^{oa}(T) = \frac{\Delta r}{r_f} + \frac{\epsilon}{r_f}.$$

Replacing latter expression in equation (5) for $\dot{I}(t)$, we clearly have $\dot{I}(T) \neq 0$, which is in contradiction to our hypothesis that infection has already attained its steady state level. Thus, neither efficiency, nor infection can attain their steady state in finite time. This logic of this proof applies for all the regions of the initial state (w_0, I_0).

In Figure 5.b, the initial state (w_0, I_0) lies in region II, *i.e.* to the right of the I-isocline ($\dot{I} = 0|_{t=0}$). Because of the relatively high value of initial efficiency, the fraction f^{oa} is also high. This leads to a decrease in efficiency, implying a change in f over time and a pivoting I-isocline. The trajectory crosses the I-isocline ($\dot{I} = 0|_{t=t_1}$) at time $t_1$and remains to the left of the pivoting isocline thereafter (example : $\dot{I} = 0|_{t=t_2}$). The dynamics

lead the state variables to their respective steady state values simultaneously but asymptotically (see Figure 5.b). While efficiency decreases monotonously, the path of infection is characterized by an "undershooting" pattern. This means, that the level of efficiency decreases initially below its steady state value (in Figure 5.b, the minimum is attained at t_1), before it increases up to its steady state $I^{SS}_{(3)}$. Our many numeric simulations have been characterized by this undershooting pattern when starting from region II. The intuition of this undershooting pattern is that due to the relative high value of efficiency, and the dynamic forces which tend to decrease infection in region II, the level of infection can be brought temporarily below its steady state level (up to t_1). As efficiency continues to decrease, the level of infection starts rising from t_1 on to obtain its steady state asymptotically.

We turn now to initial values (w_0, I_0) lying in regions III and IV. We thus have :

$$f^{oa}(0) = \frac{Q^{oa}(0)}{I_0} = 1 - \frac{c}{r_f w_0} < \frac{\Delta r}{r_f},$$

which implies that the I-isocline has a positive slope and that the forces driving the system are those already shown in Figure 1.a.

The level of efficiency w_0 may be below the level of economic viability $\frac{c}{r_f}$. In such a case, $f^{oa} = 0$ initially, which, if fitness cost is positive, allows efficiency to increase to the level where it becomes economically viable. Once economic viability is reached, potential trajectories for the evolution of state variables are shown in Figures 6.a and 6.b.

In Figure 6.a, initial state (w_0, I_0) is located in region IV and the dynamics drive the system towards the north-east. The increasing value of efficiency allows the fraction f^{oa} to increase over time and to approach $\frac{\Delta r}{r_f}$. The changing value for f^{oa} induces the isocline to pivot to the north-west (e.g. $\dot{I} = 0|_{t=t_1}$), hence becoming steeper until it reaches the vertical line. This will occur when $f^{oa} = \frac{\Delta r}{r_f}$. While the evolution for efficiency is monotonously increasing, there will be "overshooting" for the level of infection. The level of infection increases above its steady state value up to time t_1, where the system crosses the $\dot{I} = 0|_{t=t_1}$ isocline. From there on, the system will remain to the right of the I-isoclines (e.g. $\dot{I} = 0|_{t=t_2}$) and infection decreases down to its steady state level.

In Figure 6.b we show a case, in which the initial state (w_0, I_0) is to the right of the initial I-isocline (in

region III). As $f^{oa} < \frac{\Delta r}{r_f}$, efficiency can increase and thus f^{oa} also increases. This induces the I-isocline to pivot (consider $\dot{I} = 0|_{t=t_1}$). The trajectory of efficiency and infection doesn't cross an I-isocline and will remain in region III up to convergence. The level of efficiency increases monotonously as does the level of infection.

Consider finally the case where the initial state (w_0, I_0) is located on the limit between the regions I, II and III, IV, which implies $w_0 = w_{(3)}^{SS}$. Thus $f^{oa} = \frac{\Delta r}{r_f}$ and the forces driving the dynamics are those described in figure 3. Infection I converges to steady state asymptotically, while efficiency remains at its steady state level.

An important conclusion is that, under open access, the dynamic system will converge to steady state (3), $w_{(3)}^{SS} = \frac{c}{r_f - \Delta r}$, $I_{(3)}^{SS} = \frac{\beta N - r_f}{\beta}$, no matter the initial interior state of the system. As long as the fitness cost is positive, efficiency at equilibrium will be greater than its economically viable level ($w(t) = \frac{c}{r_f}$). If the fitness cost is zero ($\Delta r = 0$), then efficiency will be depleted down to economic viability ($w(t) = \frac{c}{r_f}$). This case is illustrated in Figure 7 for two different initial values (w_0^i, I_0^i) with $i = A, B$.

5 The Social Planner : First Best Outcome

The social planner maximizes social welfare, given by the sum of individual utilities minus the cost of producing the antibiotic, by choosing the sequence of quantities $\{Q(t)\}$ optimally over time. Given $I(t)$ at time t, we can change the control variable to $f(t) = \frac{Q(t)}{I(t)}$. Omitting the time indices for convenience, the maximization problem of the social planner is given by :

$$\max_{\{f(t)\}} \int_0^\infty \{N \int_{\theta_{min}}^{\theta_{max}} u(\theta) dF(\theta) - C(fI)\} e^{-\rho t} dt \tag{22}$$

subject to

$$\frac{dw}{dt} = w(1-w)(\Delta r - r_f\ f) \tag{23}$$

$$\frac{dI}{dt} = I(\beta(N-I) - r_r + w(\Delta r - r_f\ f)) \tag{24}$$

$$0 \leq f \leq 1 \tag{25}$$

$$0 \leq w \leq 1 \tag{26}$$

$$0 \leq I \leq N \tag{27}$$

and the initial conditions $w(0) = w_0$ and $I(0) = I_0$. In what follows, we assume a uniform distribution of θ with supports $\theta_{\max} = 1$ and $\theta_{\min} = 0$. We thus have $dF(\theta) = f(\theta)d\theta = d\theta$.

When choosing $f(t) < 1$, it is optimal for the social planner to serve those individuals with the highest valuation of being in good health, as this will maximize the objective function. Using our earlier derivation of the indifferent consumer type (see (7)), we have $\tilde{\theta}(p) = \frac{P(\frac{Q}{I},w)}{r_f w} = \frac{r_f w(1-\frac{Q}{I})}{r_f w} = (1-\frac{Q}{I}) = (1-f)$ and the sum of individual utilities can be written as :

$$\begin{aligned}
& N\int_0^1 u(\theta)d\theta \\
= & (N-I)\int_0^1 \theta d\theta + I\int_0^{\tilde{\theta}(p)} \pi(w)\theta\ d\theta + I\int_{\tilde{\theta}(p)}^1 \{[\pi(w)+r_f w]\theta - P(f,w)\}\ d\theta \\
= & \frac{1}{2}(N-I) + (1-f)I[\frac{1}{2}\pi(w)(1-f)] + fI\left\{\pi(w)\left(1-\frac{1}{2}f\right) + \frac{1}{2}r_f w f\right\} \\
= & \frac{1}{2}(N-I) + I\left[\frac{1}{2}\pi(w)\right] + \frac{1}{2}r_f w I f^2
\end{aligned} \tag{28}$$

The surplus accruing to the population is thus the sum of three terms. The first term represents the mean valuation, $\theta = \frac{1}{2}$, of the number of healthy individuals given by $S = N - I$. The second term refers to the mean valuation for being in good health by those who are infected. Infected individuals recover naturally with probability $\pi(w)$. Finally, the third term captures the surplus of being in good health due to antibiotic treatment. As can be easily verified, it is equal to the Marshallian consumer surplus, measured under the inverse demand curve and calculated in (10).

Substituting $C(fI) = cfI$, the current value Hamiltonian associated to (22) is as follows :

$$\begin{aligned} H(.) &= \frac{1}{2}(N-I) + I(\frac{1}{2}\pi(w)) + \frac{1}{2}r_f wIf^2 - cfI \\ &\quad +\mu w(1-w)(\Delta r - r_f\ f) + \lambda I[(\beta(N-I) - r_r + w(\Delta r - r_f\ f)]. \end{aligned} \tag{29}$$

Notice that the Hamiltonian is convex in f, as $\frac{\partial^2 H}{\partial f^2} = r_f wI \geq 0$, such that corner solutions may be optimal. Define the switching function $\Omega(t)$ as the difference in the Hamiltonians evaluated at the corners of the control variable, *i.e.* at $f = 0$ and $f = 1$:

$$\Omega(t) = H(w, I, f, \mu, \lambda)|_{f=1} - H(w, I, f, \mu, \lambda)|_{f=0} \tag{30}$$

It is optimal to choose

$$f^*(t) = \left\{ \begin{array}{c} 0,\ \text{if } \Omega(t) < 0 \\ \hat{f}(t), \text{if } \Omega(t) = 0 \\ 1,\ \text{if } \Omega(t) > 0 \end{array} \right\} \tag{31}$$

where $\hat{f}(t)$ represents a singular solution in the particular case when $H(w, I, f, \mu, \lambda)|_{f=1} = H(w, I, f, \mu, \lambda)|_{f=0}$. We will turn to the characterization of $\hat{f}(t)$ below.[10]

Further necessary conditions for an optimal solution in addition to (23) and (24) are :

$$\begin{aligned} \dot{\mu} - \rho\mu &= -\frac{\partial H}{\partial w} \\ &= \frac{1}{2}\Delta rI - \frac{1}{2}r_f If^2 + (\Delta r - r_f\ f)[\mu(2w-1) - \lambda I] \end{aligned} \tag{32}$$

$$\begin{aligned} \dot{\lambda} - \rho\lambda &= -\frac{\partial H}{\partial I} \\ &= \frac{1}{2}(1-\pi(w)) - \frac{1}{2}r_f wf^2 + cf + \lambda[r_r - w(\Delta r - r_f\ f) - \beta(N-2I)] \end{aligned} \tag{33}$$

[10] An "interior" solution $\hat{f}(t) \in (0,1)$ cannot be optimal, as the Hamiltonian is convex in the control variable f. Only extreme controls can be optimal. Whenever $\Omega(t) = 0$, the singular control $\hat{f}(t)$ to which we refer here should be reinterpreted as a probability measure(or a frequency) with which $f = 1$ is optimal, while $(1 - \hat{f}(t))$ represents the probability with which $f = 0$ is optimal. Such a control is generally called a chattering control or optimal sliding control. See Lee and Markus (1967) and Borisov and Zelikin (1994) for further reference.

which represent the evolution of the shadow values along the path of $f(t) = 0$ or $f(t) = 1$.

For a singular solution to be optimal, we must have over a certain period of time $\Omega(t) = 0.$, *i.e.* the state and costate variables have to adjust in a way to maintain the switching function equal to zero. The switching function $\Omega(t)$ is given by :

$$\begin{aligned}\Omega(t) &= \frac{1}{2}r_f wI - cI \\ &+\mu|_{f=1} w(1-w)(\Delta r - r_f) + \lambda|_{f=1}\ I\ [(\beta(N-I) - r_r) + w(\Delta r - r_f)] \\ &-\mu|_{f=0} w(1-w)\Delta r - \lambda|_{f=0}\ I\ [(\beta(N-I) - r_r) + w\Delta r]\end{aligned} \tag{34}$$

At a point of time t, it depends on the current state (w, t) and the value of the costate variables evaluated at the two extreme controls. If the switching function is equal to 0 over a certain period of time, then its derivative with respect to time must also equal zero, *i.e.*

$$\dot{\Omega}(t) = \frac{\partial \Omega}{\partial w}\dot{w} + \frac{\partial \Omega}{\partial I}\dot{I} + \frac{\partial \Omega}{\partial \mu_0}\dot{\mu}_0 + \frac{\partial \Omega}{\partial \mu_I}\dot{\mu}_I + \frac{\partial \Omega}{\partial \lambda_0}\dot{\lambda}_0 + \frac{\partial \Omega}{\partial \lambda_I}\dot{\lambda}_I = 0 \tag{35}$$

where we write $\mu|_{f=1} = \mu_I$, $\mu|_{f=0} = \mu_0$ and respectively $\lambda|_{f=1} = \lambda_I$ and $\lambda|_{f=0} = \lambda_0$ for convenience.

The last equation can be written as

$$\begin{aligned}&\left\{\frac{1}{2}r_f I + (1-2w)\Delta r(\mu_I - \mu_0) - r_f(1-2w)\mu_I + I\Delta r(\lambda_I - \lambda_0) - r_f I\lambda_I\right\}\dot{w}(\hat{f}) \\ &+\left\{\frac{1}{2}r_f w - c + [\beta(N-2I) - r_r + w\Delta r]\,(\lambda_I - \lambda_0) - r_f w\lambda_I\right\}\dot{I}(\hat{f}) \\ &+w(1-w)\Delta r(\dot{\mu}_I - \dot{\mu}_0) - w(1-w)r_f\dot{\mu}_I \\ &+I\,[\beta(N-2I) - r_r + w\Delta r]\,(\dot{\lambda}_I - \dot{\lambda}_0) - r_f wI\dot{\lambda}_I \\ \\ =\ &0\end{aligned}$$

The singular solution $\hat{f}(t)$ is characterized by the former equation as it appears in the terms $\dot{w}(.)$ and $\dot{I}(.)$.

By replacing for the laws of motion of state variables from equations (23) and (24), we obtain :

$$\hat{f} = \frac{\frac{\partial \Omega}{\partial w} w(1-w)\Delta r + \frac{\partial \Omega}{\partial I} I(\beta(N-I) - r_r + w\Delta r) + \frac{\partial \Omega}{\partial \mu_0}\dot{\mu}_0 + \frac{\partial \Omega}{\partial \mu_1}\dot{\mu}_I + \frac{\partial \Omega}{\partial \lambda_0}\dot{\lambda}_0 + \frac{\partial \Omega}{\partial \lambda_1}\dot{\lambda}_I}{r_f w \left(\frac{\partial \Omega}{\partial w}(1-w) + \frac{\partial \Omega}{\partial I} I\right)} \tag{36}$$

We would like to characterize the optimal solution further, but for doing this, we must evaluate $\dot{\mu}_0$, $\dot{\mu}_I$, $\dot{\lambda}_0$, and $\dot{\lambda}_I$. Despite considerable effort to find a tractable analytical solution, we did not achieve this goal. We thus made use of numerical simulations using the method of reverse shooting, which is explained in detail in the appendix, in order to characterize the switching function $\Omega(t)$ in the (w, I)-space. The main idea of the reverse shooting approach is that we solve the dynamic system backward for either $f(t) = 1$ or $f(t) = 0$ $\forall t$ for a fine grid of initial states (w_0, I_0) and then calculate for which states the switching function is equal to zero.

Our simulations were run for a big variety of parameter values, notably the economic parameters which are unitary treatment cost c, and the discount rate ρ. The simulation evidence shows for a wide range of parameter values, that the (w, I)-space can be partitioned into two regions, in the first of which it is optimal to treat the whole infected population, *i.e.* $f(t) = 1$, and in the second of which nobody gets treatment, *i.e.* $f(t) = 0$. The limit between these regions is determined by the level line along which the switching function is zero, $\Omega(t) = 0$. Consider as an example Figure 8.

We have run numerous simulations, notably with respect to the economic parameters, which are treatment cost c and discount rate ρ, in order to qualify the location of the zero level line of the switching function.

On the one hand, lowering the treatment cost c or the discount rate ρ tends to push the zero level line downward, such that in the limit the extreme control $f = 1$ is optimal in the whole (w, I)-space. In this case, the system will converge to steady state (1) with $w^{SS}_{(1)} = 0$ and $I^{SS}_{(1)} = \frac{\beta N - r_r}{\beta}$. Clearly, a decreasing treatment cost makes the treatment cheaper and thus more infected individuals should be treated in the social optimum. A decreasing discount rate represents a higher valuation for the future by society. As a smaller infected population today will infect less individuals tomorrow, and thus lead to a more healthier population tomorrow - which represents a positive externality to society while the individual is not influenced in its choice of buying the antibiotic - more people should be treated today.

On the other hand, increasing the treatment cost c tends to push the zero level line of the switching function

upward as does an increase in the discount rate ρ. In the limit, an increase in the treatment cost results in the extreme control $f = 0$ to be optimal in the whole (w, I)-space, as treatment becomes too expensive with respect to the social benefits of reduced infection. The system would converge to steady state (2) with $w_{(2)}^{SS} = 1$ and $I_{(2)}^{SS} = \frac{\beta N - r_w}{\beta}$. Increasing the discount rate up to infinity results in the static optimization problem. It is easily verified that the static optimization problem, which is still convex, will admit the corner solution

$$\begin{aligned} f^* &= 1, \text{ if } w > 2\frac{c}{r_f} \\ f^* &= 0, \text{ else.} \end{aligned}$$

Depending on the model parameters c and r_f, the critical value $2\frac{c}{r_f}$ is bigger (or equal) or smaller as one, implying either a unique extreme control of $f = 0$ in the whole state space, or a separation of the state space into two regions of extreme controls. For $f = 0$ to be optimal in the whole state space, c must be relatively high compared to the additional recovery rate r_f that procures antibiotic treatment. As future generations of healthy individuals are not valuable in such a case to the society living today, it does not care about the positive externality effect, that less infected individuals today imply also less infection tomorrow. In term, the antibiotic should not be applied at all in the social optimum.

In what follows we concentrate on the case in which the state space can be partitionned into two regions of either extreme controls, our analysis being done with the help of phase diagrams. We treat two interesting cases, in which the initial state $A = (w_0^A, I_0^A)$ falls into the region in which $f(t) = 1$ and the initial state $B = (w_0^B, I_0^B)$ falls into the region in which $f(t) = 0$.

Consider Figure 9. In this case, the initial level of efficiency w_0^A is relatively high, as is the initial level of infection I_0^A. The phase diagram shows two different regimes characterized by the two isoclines for infection, $\dot{I}|_{f=0}$ and $\dot{I}|_{f=1}$. At initial state A, we have $f = 1$. From there on, the state variables evolute in the south east direction, crossing the isocline $\dot{I}|_{f=1}$and try to achieve steady state (1) at which efficiency would be completely depleted, *i.e.* $w_{(1)}^{SS} = 0$. By doing so, the optimal path must necessarily reach at some point of time, say t_1, the 0-level line of the switching function, such that $\Omega(t_1) = 0$. At t_1, dynamic forces of both

regimes of the extreme controls, *i.e.* forces out of the 0-level line, prevent the state variables to leave the 0-level line again. Thus, from t_1 on, an interior solution is optimal. Numerical simulations have shown us that the singular solution tends monotonously to $\frac{\Delta r}{r_f}$ from above. The system thus reaches steady state (3), at which $\hat{f}^{SS} = \frac{\Delta r}{r_f}$, $I^{SS}_{(3)} = \frac{\beta N - r_r}{r_f}$, and the level of efficiency $w^{SS}_{(3)}$ being determined by the intersection of the I-isocline and the 0-level line of the switching function.[11]

We present another optimal control path if efficiency is relatively low initially, such that $f = 0$. Consider steady state $B = (w^B_0, I^B_0)$ in Figure 10. Dynamic forces drive the system first in the north west direction, efficiency increases, while infection decreases. The path crosses now the isocline $\dot{I}|_{f=0}$ from which point of time infection starts to rise again. The path tends to reach steady state (2), at which we would have $w^{SS}_{(2)} = 1$, but the path necessarily reaches the 0-level line of the switching function first, say at some point of time t_1. As before, a singular solution $\hat{f}$ applies from there on and the system converges to steady state (3), at which $\hat{f}^{SS} = \frac{\Delta r}{r_f}$, $I^{SS}_{(3)} = \frac{\beta N - r_r}{r_f}$, and $w^{SS}_{(3)}$ being determined by the intersection of $\dot{I}|_{f=\Delta r/r_f}$ with the 0-level line of the switching function. The only difference is that the singular solution now approaches $\frac{\Delta r}{r_f}$ monotonously from below.

Figure 10+ shows numerical simulations of the optimal control paths for the two former cases and two more.

6 Comparison Between Social Optimum and Open Access

The regime of open access is not socially optimal. As we have seen, an extreme control is optimal from a social point of view, *at least initially* before converging to the singular solution $\hat{f}$, exception being made if the initial state (w_0, I_0) is located on the 0-level line of the switching function $\Omega(t)$ and a singular solution applies from the beginning on. The system converges to steady state (3) if the state space (w, I) can be partitioned into two non-empty regions of optimal extreme control. We then have at the steady state :

$$\begin{aligned} f^{SS}_{sp} &= \frac{\Delta r}{r_f} \\ I^{SS}_{sp} &= \frac{\beta N - r_r}{\beta} \end{aligned}$$

[11] We remind that the singular solution $\hat{f}(t)$ to which we refer here should be interpreted as a probability measure for the chattering control, *i.e.* the probability of interchanging extreme controls. This probability determines the frequency with which the control path switches around the zero-level-line of the switching function. We are presently working on the formulation of an auxiliary problem allowing to analyze the chattering control in more detail.

and the level of efficiency w_{sp}^{SS} is determined by the intersection of the 0-level line of the switching function and the vertical line passing through I_{sp}^{SS}.

In contrast to this, an extreme control cannot be the result of open access, as this would either make the price decline down to zero in the case of $f = 1$ and thus imply zero profits for the antibiotic producers, nor, in the case of $f = 0$, as it would then be profitable for any producer to sell at a price above marginal production cost and to serve the whole market. The fraction of the infected population served by the antibiotic industry under open access is given by

$$f^{oa}(t) = 1 - \frac{c}{r_f w(t)},$$

if the antibiotic is economically viable, *i.e.* $r_f w(t) > c$. We have shown that the system tends to steady state (3) at which

$$\begin{aligned} w_{oa}^{SS} &= \frac{c}{r_f - \Delta r} \left\{ \begin{array}{l} <1 \text{ if } c < r_f - \Delta r \\ =1 \text{ else} \end{array} \right\} \\ I_{oa}^{SS} &= \frac{\beta N - r_r}{\beta} \\ f_{oa}^{SS} &= \frac{\Delta r}{r_f}. \end{aligned}$$

Thus, what is decisive for the open access outcome are the production cost c, the level of efficiency $w(t)$ and the additional recovery rate r_f, which determine the valuation of the antibiotic by the infected population, as well as the fitness cost Δr. The industry under open access is notably not concerned with the discount rate ρ, a parameter influencing the socially optimal control path, nor is it concerned with the valuation of being in good health by the uninfected population, the mean value which is equal to $\frac{1}{2}(N - I)$ (see equation (28)).

What the open access and the social optimum have in common is that for a large class of parameter values, the state variables tend to steady state (3), at which efficiency is renewable and the level of infection is equal to $\frac{\beta N - r_r}{\beta}$. However, the level of efficiency is different in general and even higher in the open access as in the social optimum. This seems counterintuitive at first sight, as one would think that the open access industry

overexploits the efficiency of the antibiotic. But as both regimes lead to the same level of infection, we must also question the speed with which that level is achieved. As the social planner considers the valuation of being in good heath by those who are uninfected today and in the future, he is interested in reducing the number of infected individuals more rapidly.

By assessing further the similarities and differences between the open access and the social optimum, we distinguish between three possible outcomes for which we found evidence. They are discussed in the form of case studies below.

1) The antibiotic is applied under open access and in the social optimum, however the steady states reached under both regimes are different.

2) The antibiotic is applied under open access and in the social optimum and the steady states under each regime are the same, they both tend to steady state (3) and we have $w_{oa}^{SS} = w_{sp}^{SS}$.

3) The antibiotic is not economically viable, but is applied in the social optimum.

Case Study (1). Consider Figure 11 as an example. For this simulation, we have $c = 0.18$ and $\rho = 0.03$. Epidemiological parameters are $\beta = 0.6$, $r_r = 0.3$, $r_w = 0.15$, $\Delta r = 0.15$ and $r_f = 0.4$. Economic viability of the antibiotic is at a level of $\frac{c}{r_\mathrm{f}} = 0.45$, while $w_{oa}^{SS} = \frac{c}{r_\mathrm{f} - \Delta r} = 0.72$. The level of efficiency reached in the social optimum is below 0.2. We have shown analytically before that the steady state under open access is only achieved asymptotically. In our simulations, the steady state was generally reached with an accuracy of about 10^{-5} or more after 10000 periods, while the steady state in the social optimum was reached with a comparable accuracy in only 350 periods, which indicates that the socially optimal steady state is probably to be achieved in finite time.[12] It is clear that the state path must evolute faster in the social optimum in this case compared to the open access outcome, as it is initially socially optimal to treat every infected individual, while the antibiotic is not economically viable under open access. More individuals thus benefit from the antibiotic and thus infection decreases faster over time. This continues to be true, even when the singular solution applies in the social optimum, as $\hat{f}$ converges in this particular case to $\frac{\Delta r}{r_\mathrm{f}}$ from above (as simulations have shown us), while it converges to $\frac{\Delta r}{r_\mathrm{f}}$ from below once it is economically viable under open access. The steady state is thus more rapidly reached in the social optimum as in the open access : infected

[12] For the reader who is not familiar with numeric simulations, notably the reverse shooting approach, we want to point out that the model must be formulated in discrete time in order to be applied numerically. For more details see the appendix.

individuals get treatment faster in the social optimum reducing infection in the future.

Case Study (2). Consider Figure 12. Here we have $w_{oa}^{SS} = w_{sp}^{SS} = 0.72$, in which case the steady state under open access is socially optimal, but the path leading to it is not. The parameter values that were used in the simulation are $c = 0.18$, $\rho = 2.9$, the remaining parameters remaining the same as in the previous case study. As the discount rate increases, the zero level line of the switching function also increases, up to the point where the steady states in the social optimum and under open access coincide. This is the case at a discount rate of 290%, which corresponds to a society more impatient, but not as much as the open access industry.

Case Study (3). Consider Figure 13. Here we have $c = r_f = 0.4$, implying that the viability of the antibiotic is equal to $\frac{c}{r_f} = 1$. The discount rate and all other model parameters are again equal to those specified in case study (1). In this case, the antibiotic is not economically viable under open access, and isn't neither applied initially in the social optimum. Open access and socially optimal paths coincide, up to the point of time where the zero level line of the switching function is reached. From here on, the socially optimal path follows the zero level line up to convergence to steady state (3). The singular solution that applies converges to $\frac{\Delta r}{r_f}$ from below. The open access path leads to steady state (2), which is characterized by a higher level of infection, but also by an efficiency level equal to one.

7 Future research

We have analyzed the market outcome under open access, and we have shown that it is not optimal from a societal point of view. We have to further assess the similarities in the steady states under open access and social optimum, as well as the differences between the two regimes. A more realistic model of antibiotic exploitation takes into account that, before an antibiotic is sold by a competitive industry, it is protected by a patent in general and thus generates monopoly power to the exploiting firm. We are presently analyzing the optimal control problem for the monopolist. Further more, while the monopolist faces an exogenously given patent duration, the patent protection agency will have to determine the optimal patent duration. Our intuition is that in the case of antibiotic efficiency, the optimal patent duration may be positive even without taking into account R&D costs that would have to be recovered by the developing firm. A monopolist may

better manage the exploitation of the efficiency of the antibiotic as a competitive industry does. This point clearly merits more attention

8 Appendix

8.1 The Reverse Shooting Approach

The approach of reverse shooting allows us to keep the formulation of the optimal control problem stated so far, exception being made for the continuous time aspect. The general idea of this numerical approach is that we start at the "end" of the problem (*i.e.* where the steady state has been reached and where we know the values taken by the state and costate variables in either case with $f(t) = 0$ or $f(t) = 1$, $\forall t$) and solve the problem backward.[13]

For the sake of completeness of this section we restate the current value Hamiltonian in terms of the fraction treated, $f(t)$.

$$\begin{aligned} H(.) &= \frac{1}{2}(N-I) + I(\frac{1}{2}\pi(w)) + \frac{1}{2}r_f w I f^2 - cIf \\ &\quad + \mu w(1-w)(\Delta r - r_f\ f) + \lambda I[(\beta(N-I) - r_r) + w(\Delta r - r_f\ f)]. \end{aligned} \tag{37}$$

The laws of motion for the state and costate variables are

$$\dot{w} = w(1-w)(\Delta r - r_f\ f) \tag{38}$$

$$\dot{I} = I[\beta(N-I) - r_r + w(\Delta r - r_f\ f)] \tag{39}$$

$$\dot{\mu} - \rho\mu = \frac{1}{2}\Delta r I - \frac{1}{2}r_f I f^2 + (\Delta r - r_f\ f)[\mu(2w-1) - \lambda I] \tag{40}$$

$$\dot{\lambda} - \rho\lambda = \frac{1}{2}(1-\pi(w)) - \frac{1}{2}r_f w f^2 + cf + \lambda[r_r - w(\Delta r - r_f\ f) - \beta(N-2I)] \tag{41}$$

A singular solution is characterized by

$$\hat{f} = \frac{\frac{\partial\Omega}{\partial w}w(1-w)\Delta r + \frac{\partial\Omega}{\partial I}I(\beta(N-I) - r_r + w\Delta r) + \frac{\partial\Omega}{\partial\mu_0}\dot{\mu}_0 + \frac{\partial\Omega}{\partial\mu_1}\dot{\mu}_I + \frac{\partial\Omega}{\partial\lambda_0}\dot{\lambda}_0 + \frac{\partial\Omega}{\partial\lambda_1}\dot{\lambda}_I}{r_f w\left(\frac{\partial\Omega}{\partial w}(1-w) + I\frac{\partial\Omega}{\partial I}\right)} \tag{42}$$

[13] Judd (1998) contains an instructing analysis on this approach and many more.

If the antibiotic is given to the whole infected population, *i.e.* $f(t) = 1$, $\forall t$, the dynamic system will reach steady state (1), at which $w^{SS} = 0$, $I^{SS} = \frac{\beta N - r_r}{\beta}$. Replacing the steady state values for efficiency and infection in equation (40), with $\dot{\mu}|_{f=1} = 0$, and in equation (41), with $\dot{\lambda}|_{f=1}$, we obtain the respective steady state values for the costate variables. They are

$$\begin{aligned} \mu^{SS}|_{f=1} &= I^{SS}(\Delta r - r_f)\frac{\left[\lambda^{SS}|_{f=1} - \frac{1}{2}\right]}{\rho - (\Delta r - r_f)}, \\ \lambda^{SS}|_{f=1} &= \frac{\frac{1}{2}(r_r - 1) - c}{\beta N - r_r + \rho}. \end{aligned}$$

If the antibiotic is not given to the infected population over time, *i.e.* $f(t) = 0$, $\forall t$, the dynamic system will reach steady state (2), at which $w^{SS} = 1$, $I^{SS} = \frac{\beta N - r_w}{\beta}$. The respective steady state values for the costate variables are

$$\begin{aligned} \mu^{SS}|_{f=0} &= I^{SS}\Delta r\frac{\left[\lambda^{SS}|_{f=0} - \frac{1}{2}\right]}{\rho + \Delta r} \\ \lambda^{SS}|_{f=0} &= \frac{\frac{1}{2}(r_w - 1)}{\beta N - r_w + \rho}. \end{aligned}$$

We have now completely specified the steady states if the antibiotic were to be applied with $f(t) = 0$ or $f(t) = 1$, $\forall t$. Having the laws of motion for the state variables and costate variables at hand (equations (38) to (41)) we can transform the system into its discrete time approximative system. For any variable $x(t)$ we have

$$\frac{dx(t)}{dt} = \lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t}.$$

Applied to our system, we obtain as an approximating system :

$$w(t + \Delta t) = w(t)\left[\Delta t(1 - w(t))\left(\Delta r - r_f\ f(t)\right) + 1\right] \tag{43}$$

$$I(t + \Delta t) = I(t)\left[\Delta t\left(\beta(N - I(t)) - r_r + w(t)\left(\Delta r - r_f\ f(t)\right)\right) + 1\right] \tag{44}$$

$$\begin{pmatrix} \mu(t + \Delta t) \\ \lambda(t + \Delta t) \end{pmatrix} = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix}\begin{pmatrix} \mu(t) \\ \lambda(t) \end{pmatrix} + \begin{pmatrix} b_\mu(t) \\ b_\lambda(t) \end{pmatrix}, \tag{45}$$

with

$$
\begin{aligned}
a_{11}(t) &= 1 + \Delta t \left(\rho + \left(\Delta r - r_f\ f(t)\right)\left(2w(t) - 1\right)\right) \\
a_{12}(t) &= -\Delta t I(t) \left(\Delta r - r_f\ f(t)\right) \\
a_{21}(t) &= 0 \\
a_{22}(t) &= 1 + \Delta t \left(\rho + 2\beta I(t) - \beta N + r_r - w(t)\left(\Delta r - r_f\ f(t)\right)\right) \\
b_\mu(t) &= \frac{1}{2}\Delta t \left(\Delta r I(t) - r_f\ I f(t)^2\right) \\
b_\lambda(t) &= \Delta t \left(\frac{1}{2}(1 - \pi(t)) - \frac{1}{2} r_f w(t) f(t)^2 + c f(t)\right)
\end{aligned}
$$

In the reverse shooting approach, we calculate the evolution of efficiency and infection from some initial condition (w_0, I_0) for a given extreme control f from equations (43) and (44). Knowing the evolution of w and I, and the steady states for the costate variables if $f(t) = 0$ and $f(t) = 1$, $\forall t$, we can use equation (45) to calculate backward the evolution of the costate variables. The equation admits a unique solution in the unknown variables $\mu(t)$, $\lambda(t)$ and particularly $\mu(0), \lambda(0)$.

Figures A1 and A2 show the evolution of the state and costate variables from initial state (w_0, I_0) with $w_0 = 0.5$ and $I_0 = 0.9$ for $f(t) = 0$ and $f(t) = 1$, $\forall t$ respectively. The parameters are grouped in the following table :

$\beta = 0.6$	$\rho = 0.05$
$r_r = 0.3$	$c = 0.3$
$r_w = 0.15$	
$\Delta r = 0.15$	$N = 1$
$r_f = 0.4$	$\Delta t = 1$

In Figure A1, steady state (1) is reached, at which $w^{SS} = 0$ and $I^{SS} = 0.5$. As efficiency decreases, its shadow value μ increases. Infection decreases initially below its steady state level, its shadow value λ mirrors that evolution. In Figure A2, steady state (2) is reached, at which $w^{SS} = 1$ and $I^{SS} = 0.75$. No use is made of the antibiotic and thus the steady state value of infection is higher in steady state (2) compared to steady

state (1). The shadow value of efficiency is negative although efficiency represents a good to society. An interpretation for this is that as the antibiotic is never used in this case to fight the infection, and as there is no other use for the antibiotic in our model, its efficiency has a negative intrinsic value : it would actually cost nothing to decrease the resource.

We can now turn to the algorithm of reverse shooting itself. It can be stated as follows :

Algorithm 1 *Reverse Shooting*

1) From the initial condition (w_0, I_0)*, calculate the evolution of efficiency and infection for the two cases with* $f(t) = 0$ *and* $f(t) = 1$*,* $\forall t$ *up to convergence to the respective steady states (1) and (2).*

2) Use the laws of motion for the costate variables μ *and* λ*, their respective final values* μ^{SS} *and* λ^{SS}*, as well as the series of* w *and* I *from step (1) in order to calculate the evolution of the costate variables backward.*

3) At (w_0, I_0)*, evaluate the switching function* $\Omega(t)$ *and determine the optimal control* f^**. If* $\Omega(t)$ *is zero, we calculate the singular solution* $\hat{f}(t)$ *from a time discrete version of equation (42).*

4) Calculate $w(1), I(1)$ *using* f^* *from step (3) and start from step (1).*

5) Repeat steps (1) to (4) in order to generate a time series of necessary length to reach steady state (3).

Références

[1] Borisev, V.F. and M.I. Zelikin (1994). *Theory of Chattering Control*. Birkhäuser.

[2] Judd, Kenneth L. 1998. *Numercial Methods in Economics*. Cambridge, MA : MIT Press.

[3] Kermack, W.O. and A.G. McKendrick (1927). A contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character,* 115 (772), 700-721.

[4] Laxminarayan, R. and G.M. Brown (2001). Economics of Antibiotic Resistance : A Theory of Optimal Use. *Journal of environmental Economics and Management*, 42, pp. 183-206.

[5] Lee E.B. and L. Markus (1967). *Foundations of Optimal Control Theory*. New York, Wiley.

[6] Ross, R. (1911). *The Prevention of Malaria*, Second Edition. London : Murray.

[7] Rowthorn, R. and G. M. Brown (2003). Using Antibiotics When Resistance Is Renewable. *Battling Resistance to Antibiotics and Pesticides : An Economic Approach.* R. Laxminarayan, editor. Chapter 2. Resources for the Future, Washington, DC.

[8] Wilen, J.E. and S. Msangi (2003). Dynamics of Antibiotic Use : Ecological versus Interventionist Strategies to Manage Resistance to Antibiotics. *Battling Resistance to Antibiotics and Pesticides : An Economic Approach.* R. Laxminarayan, editor. Chapter 1. Resources for the Future, Washington, DC.

Figure 1.a : $f \in \left[0, \frac{\Delta r}{r_f}\right)$

($\Delta r > 0$)

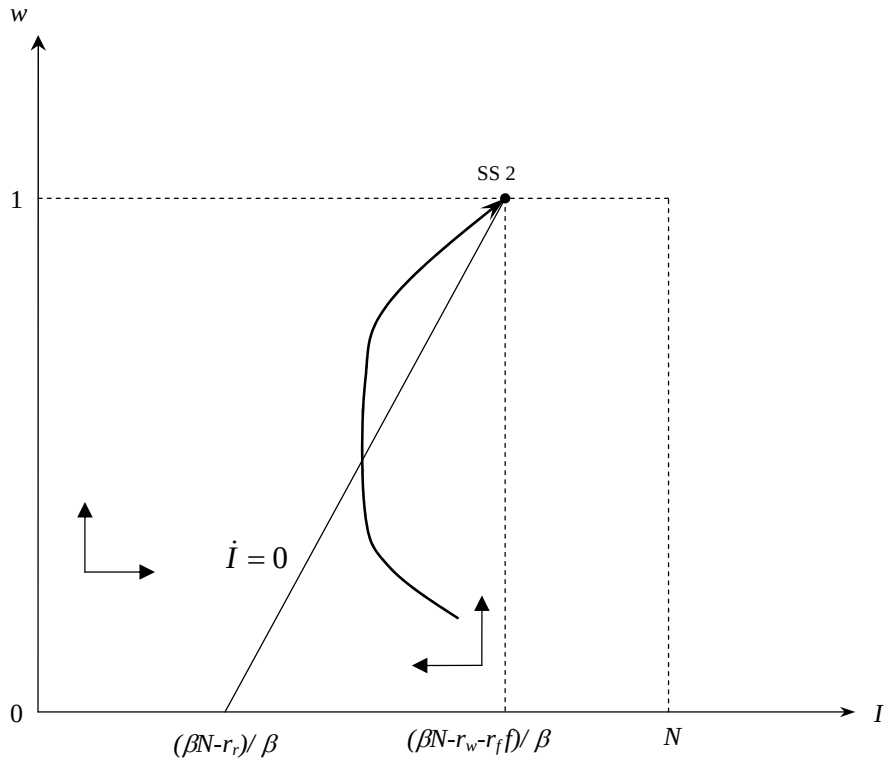


Figure 1.b: $f_1 < f_2 < \dfrac{\Delta r}{r_f}$

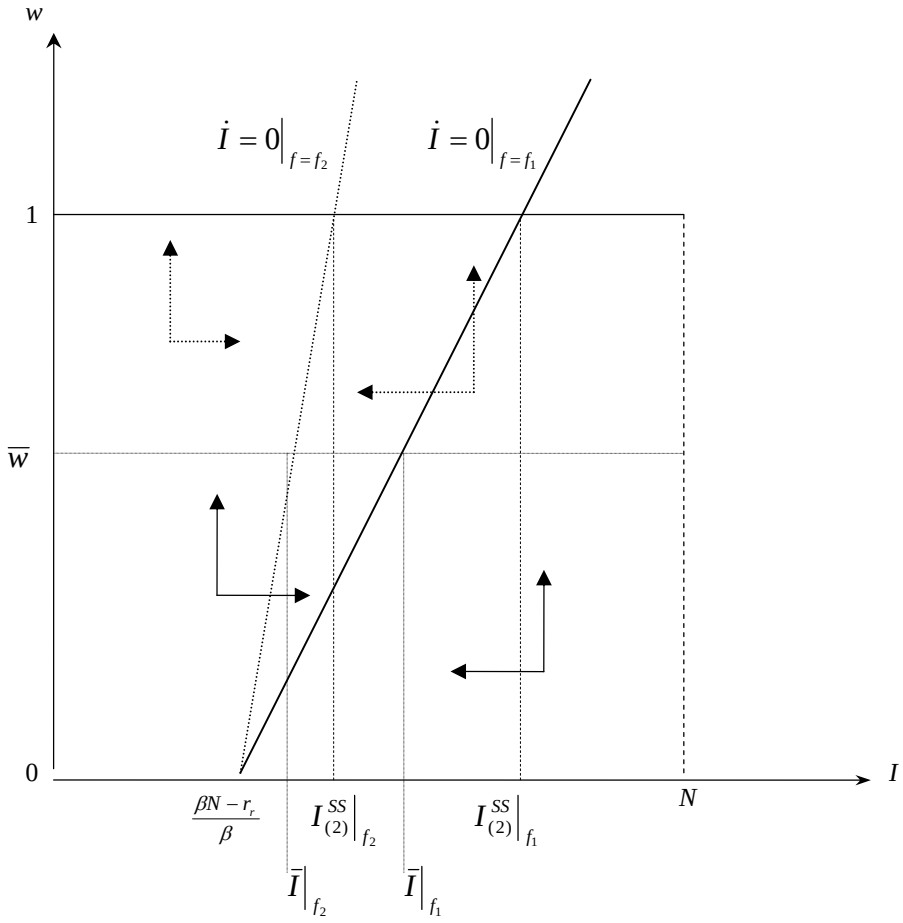


Figure 2: $f \in \left(\frac{\Delta r}{r_f}, 1 \right]$

($\Delta r > 0$)

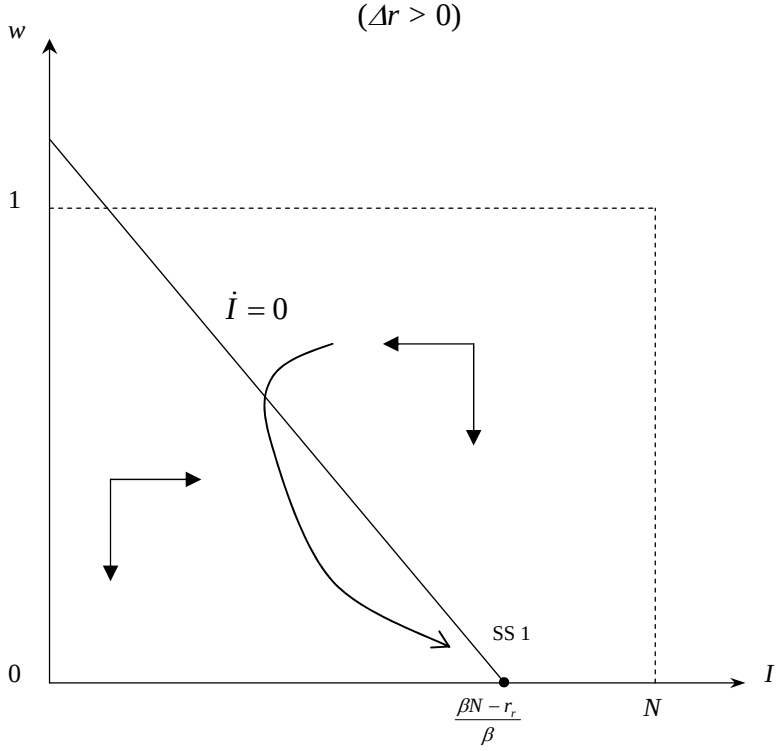


Figure 3: $f = \frac{\Delta r}{r_f}$

($\Delta r > 0$)

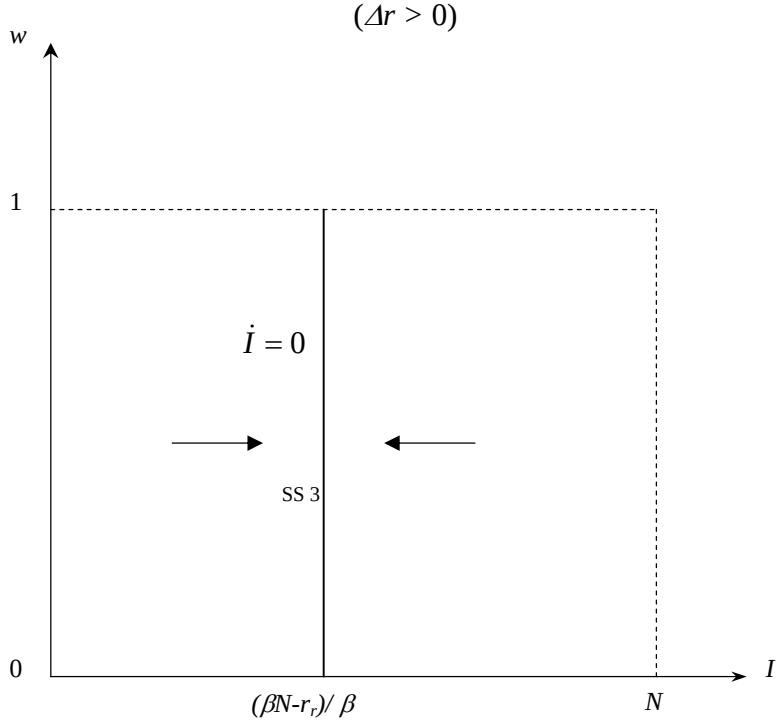


Figure 4: Convergence to Equilibrium

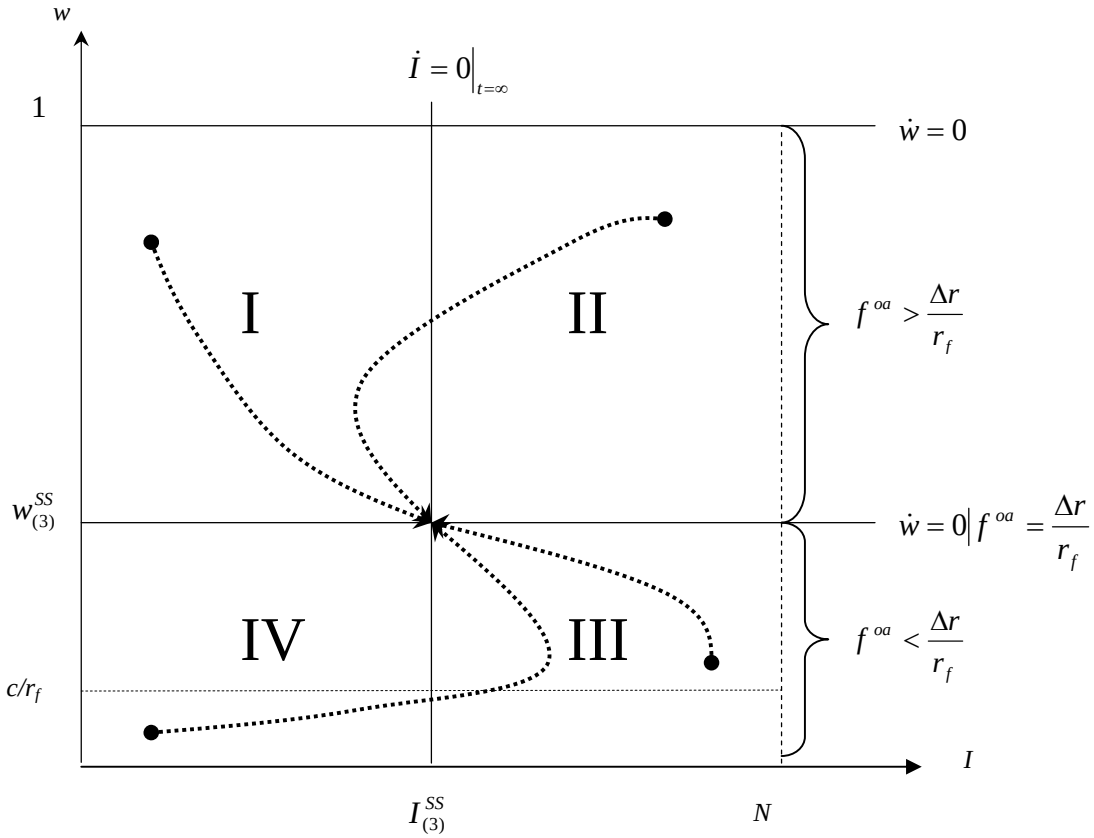


Figure 5.a: $f^{oa} > \dfrac{\Delta r}{r_f}$

$0 < t_1 < \infty$

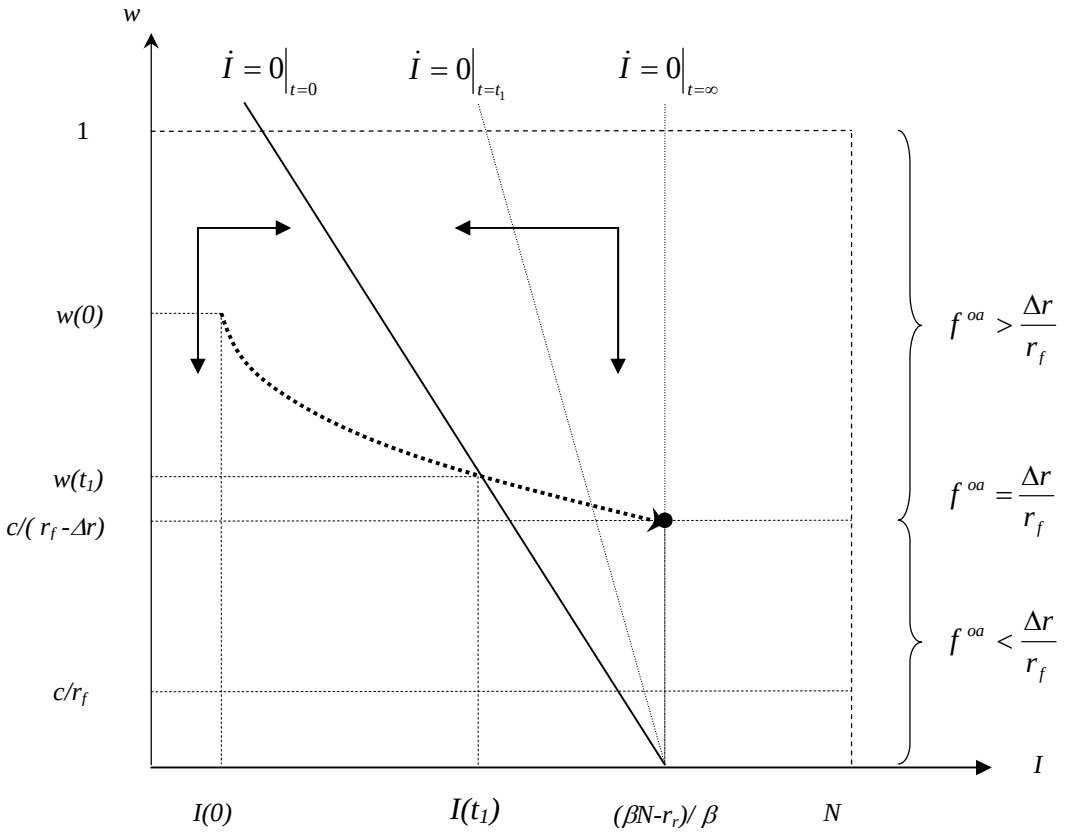


Figure 5.b: $f^{oa} > \dfrac{\Delta r}{r_f}$

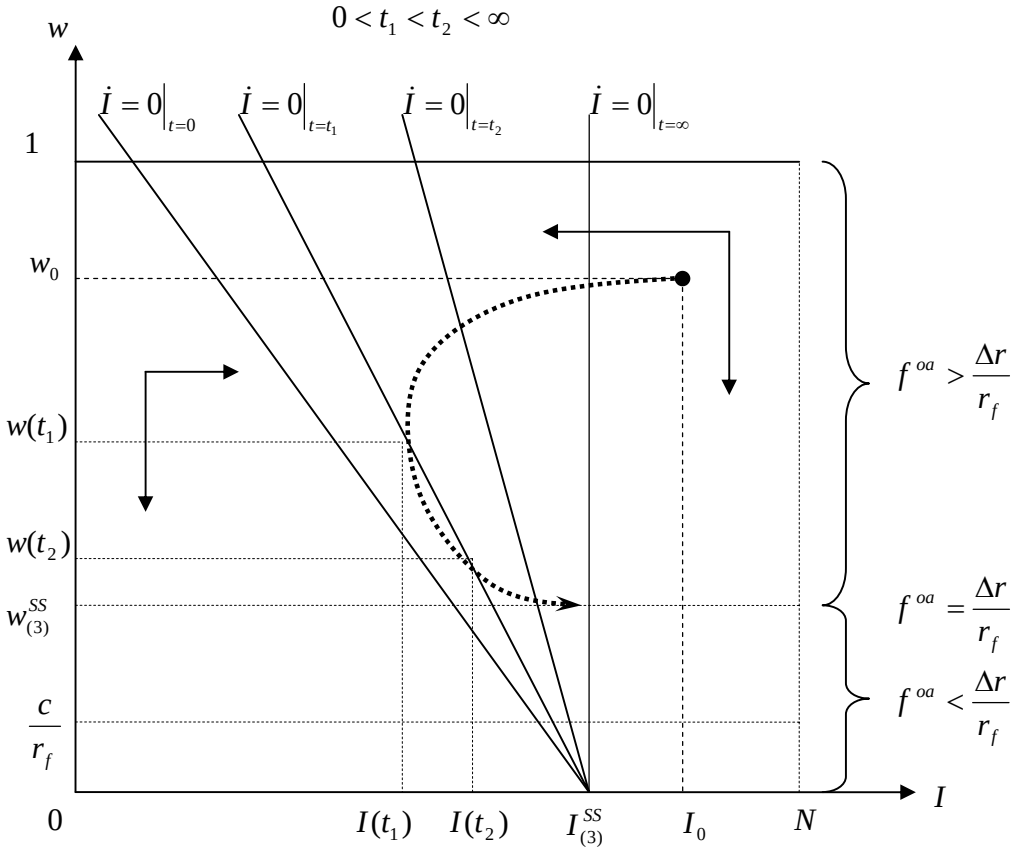


Figure 6.a: $f^{oa} < \frac{\Delta r}{r_f}$

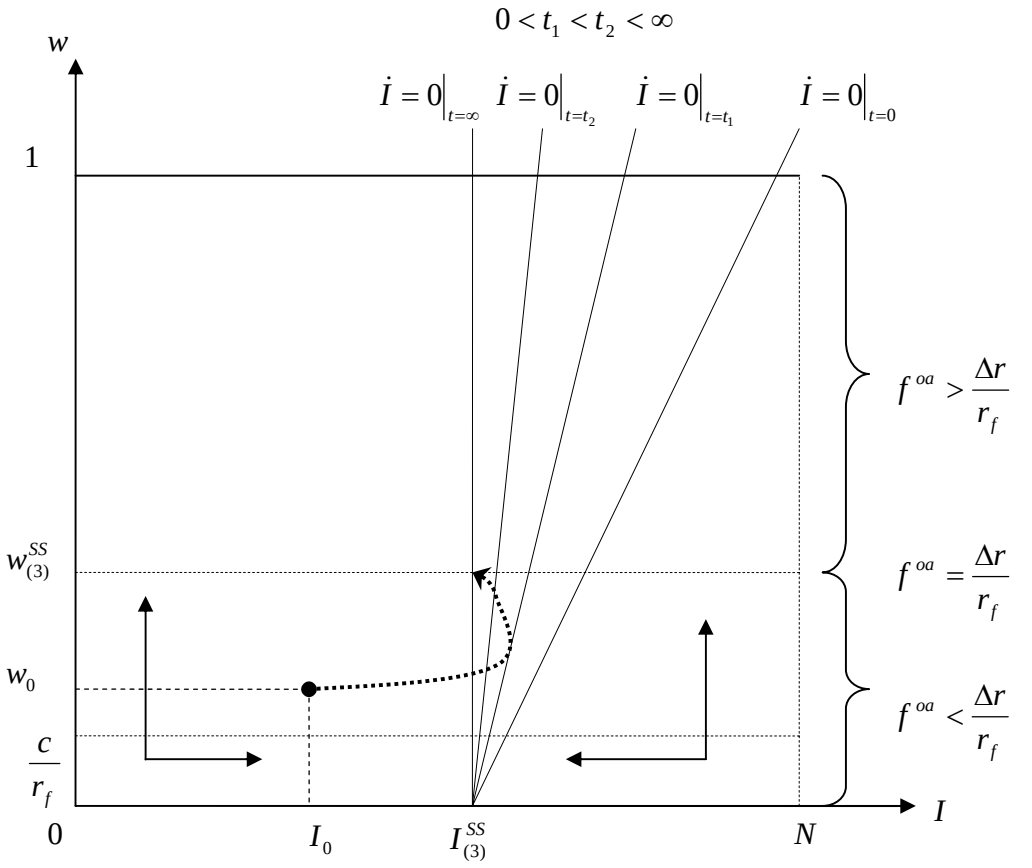


Figure 6.b: $f^{oa} < \dfrac{\Delta r}{r_f}$

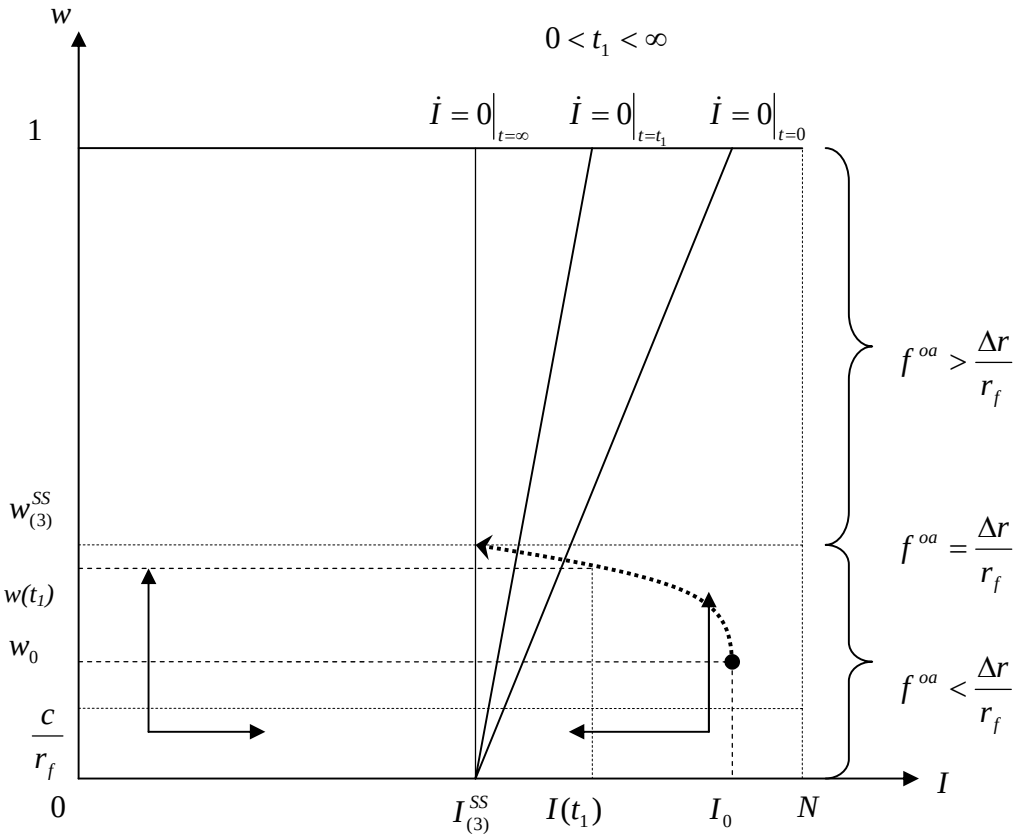


Figure 7: $\Delta r = 0$

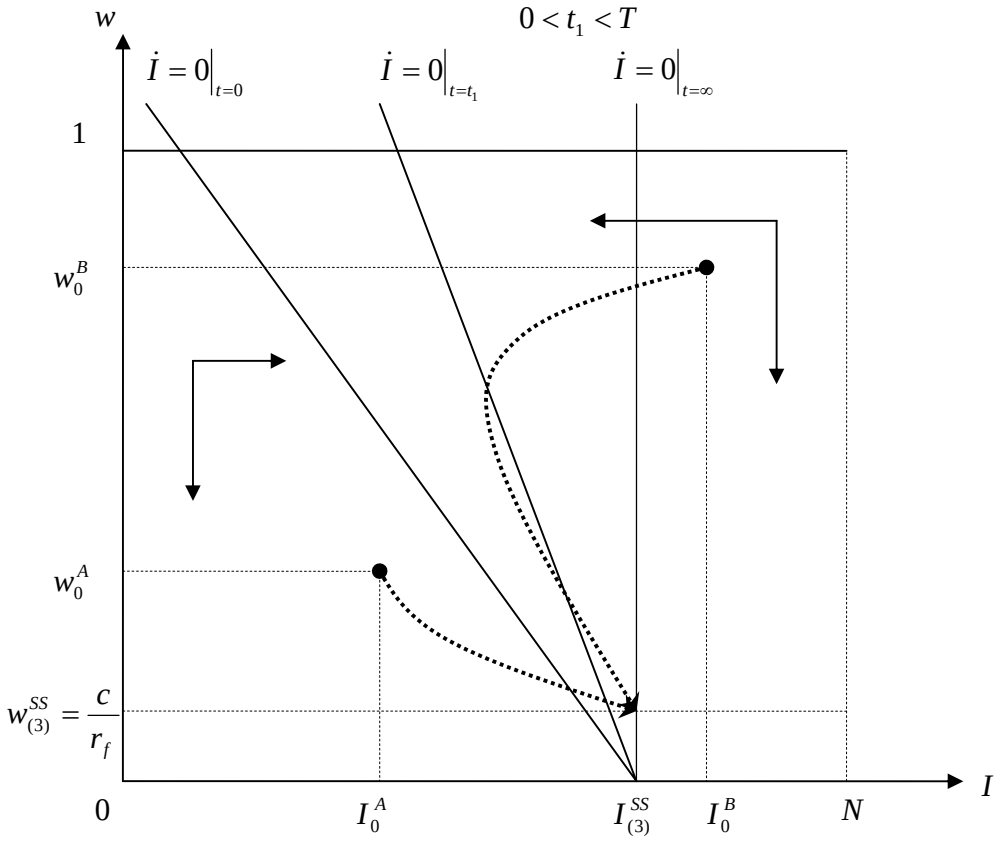


Figure 8

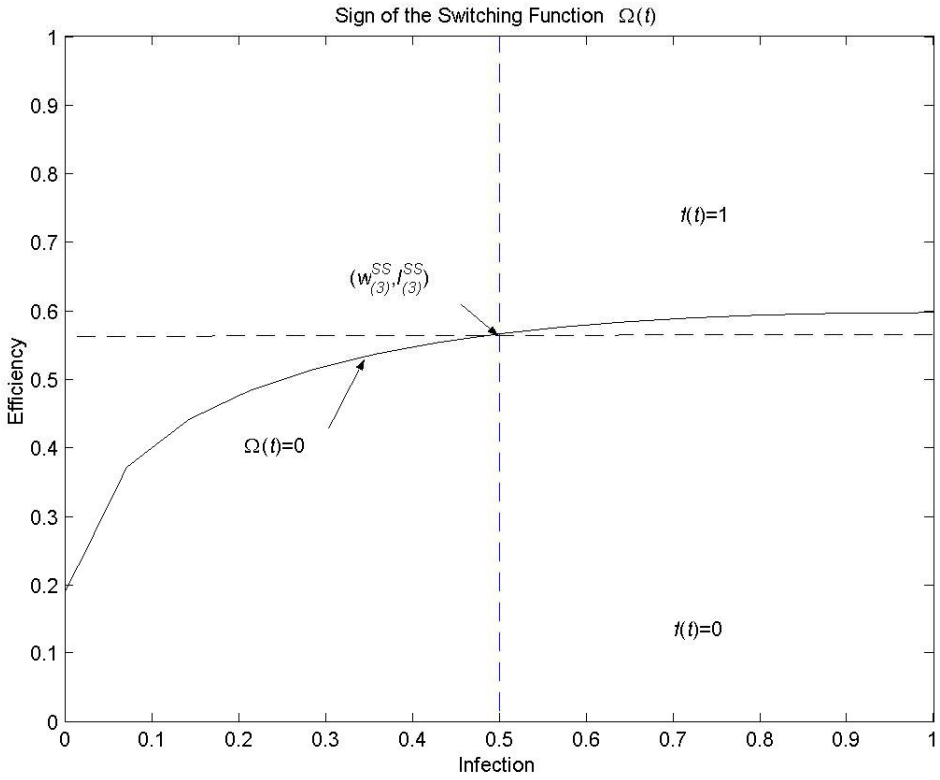

Sign of the Switching Function $\Omega(t)$
Efficiency
Infection
$I(t)=1$
$I(t)=0$
$\Omega(t)=0$
$(w^{SS}_{(3)}, I^{SS}_{(3)})$
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 9: Optimal Path with (w_0^A, I_0^A)

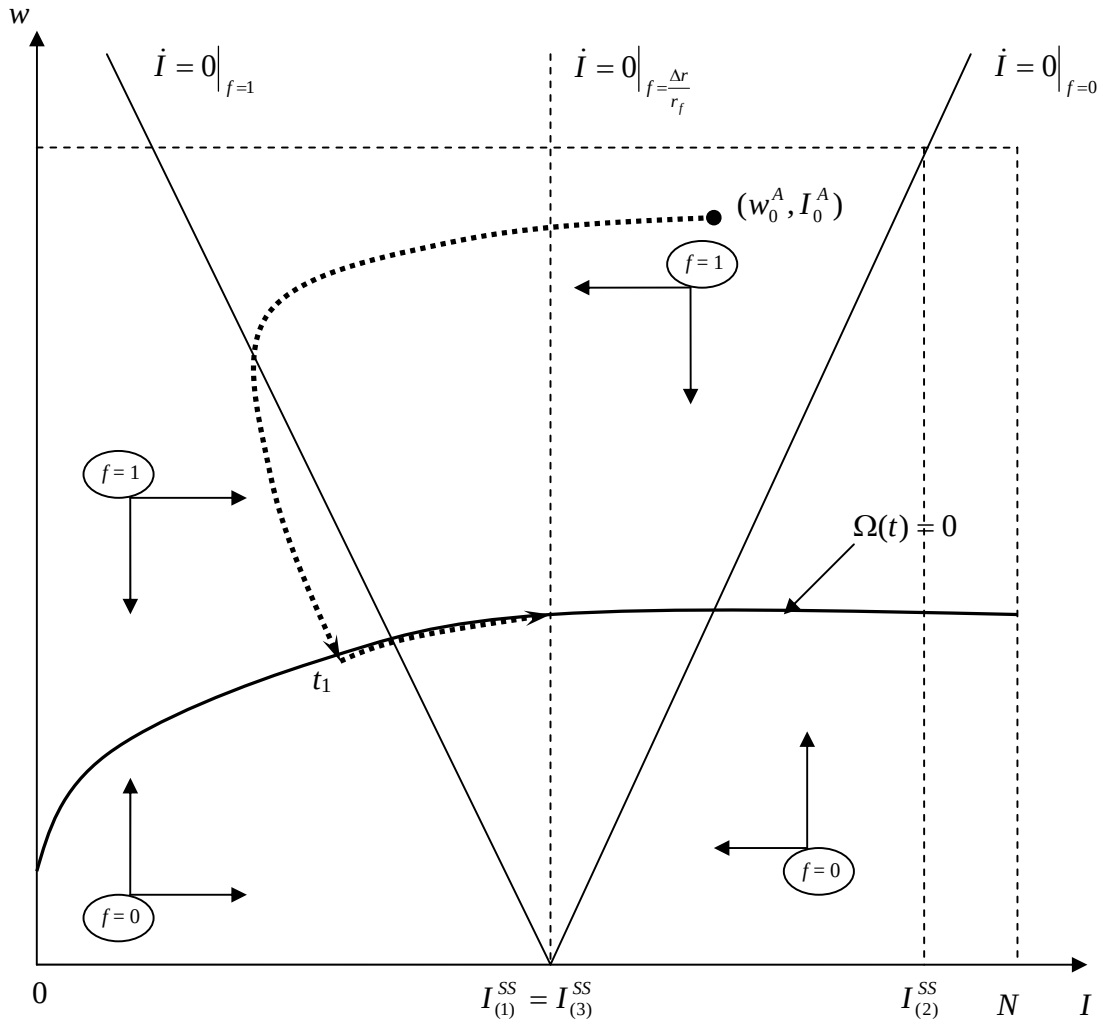


Figure 10: Optimal Path with (w_0^B, I_0^B)

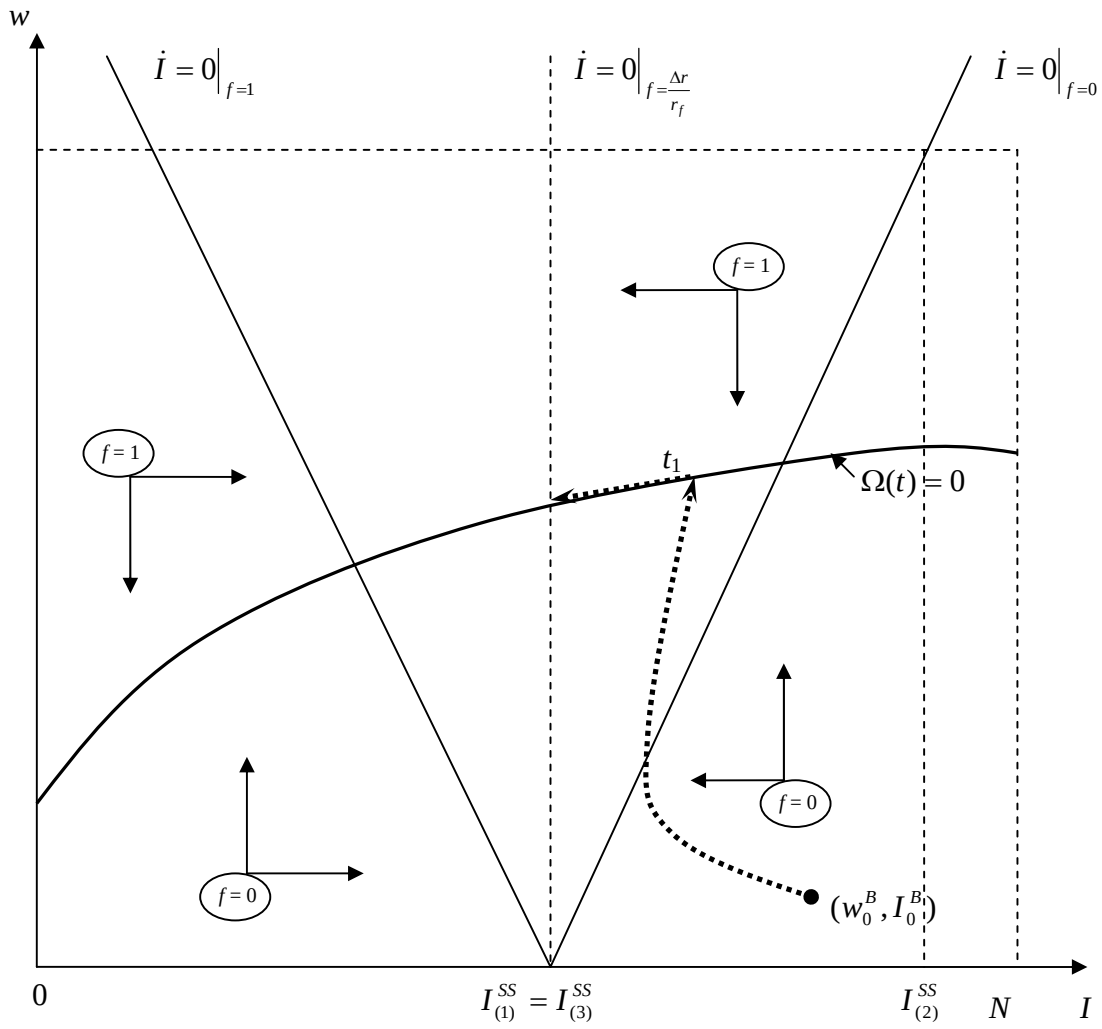


Figure 11

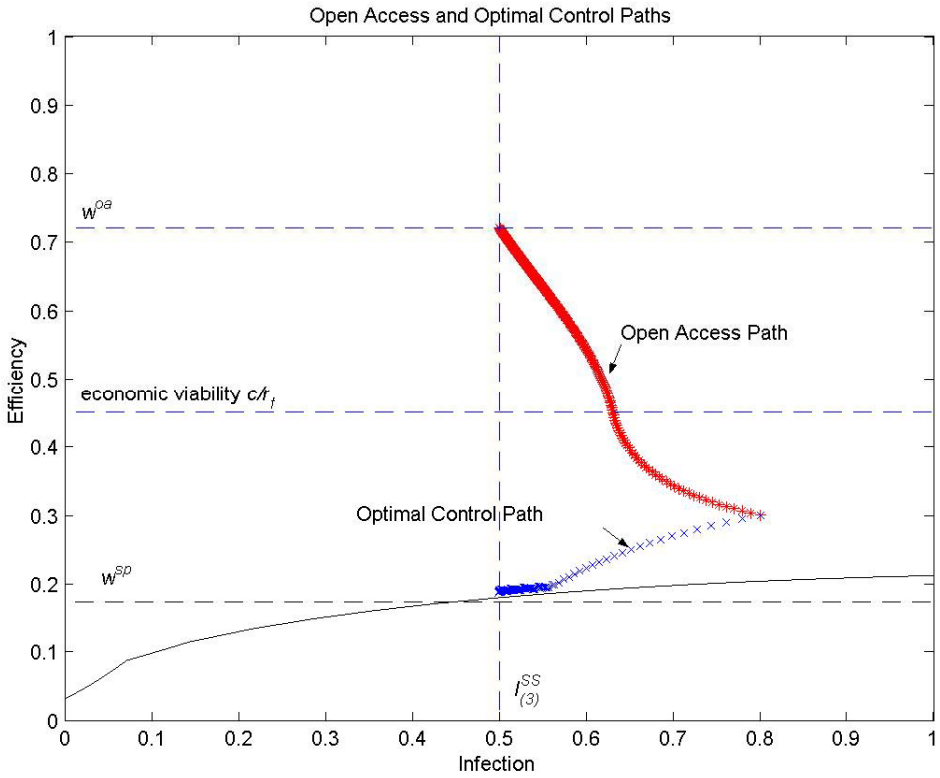

Open Access and Optimal Control Paths
Efficiency
Infection
w^{Oa}
economic viability c/f_1
Open Access Path
Optimal Control Path
w^{sp}
$I^{SS}_{(3)}$
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 12

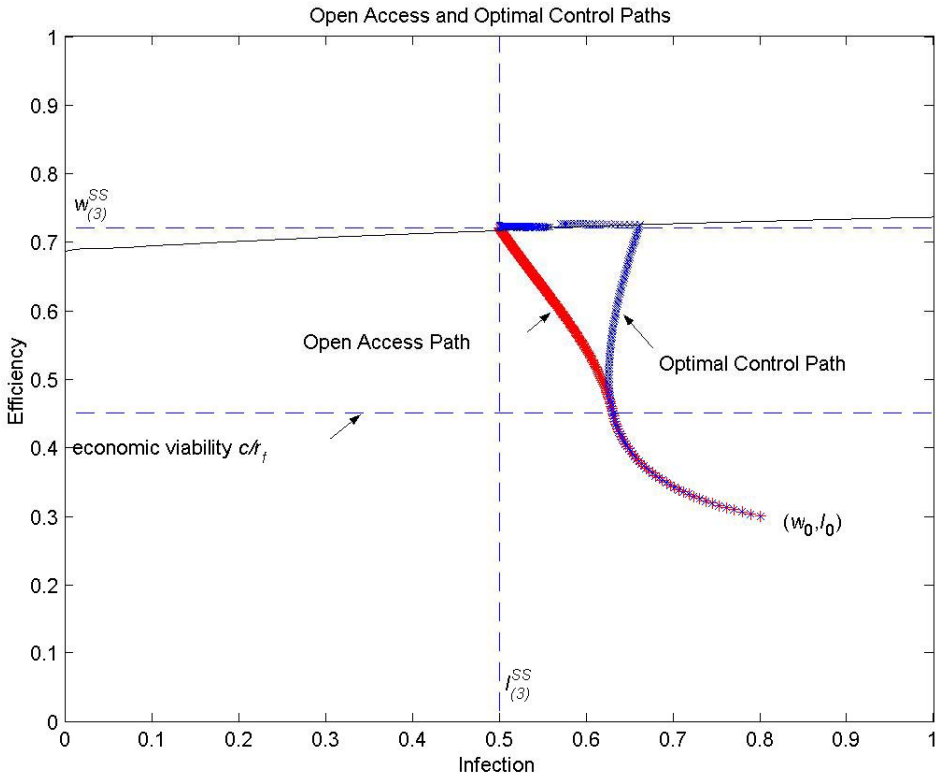

Open Access and Optimal Control Paths
Efficiency
Infection
$w^{SS}_{(3)}$
$I^{SS}_{(3)}$
Open Access Path
Optimal Control Path
economic viability c/f_1
(w_0, I_0)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 13

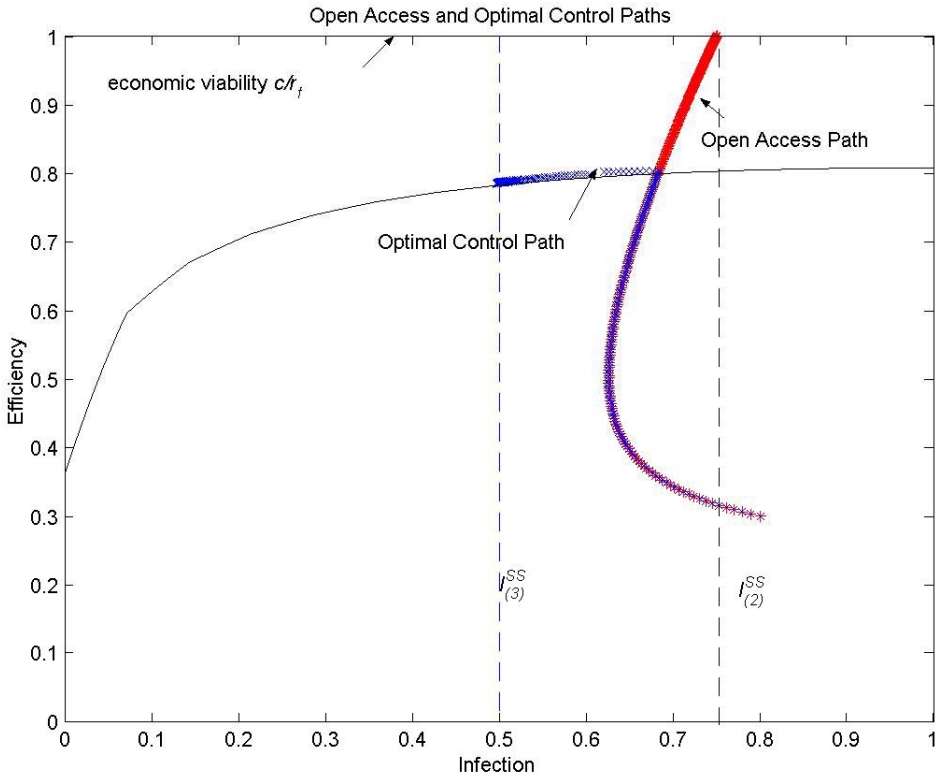

Open Access and Optimal Control Paths
economic viability c/f_I
Open Access Path
Optimal Control Path
$I^{SS}_{(3)}$
$I^{SS}_{(2)}$
Efficiency
Infection
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure A1

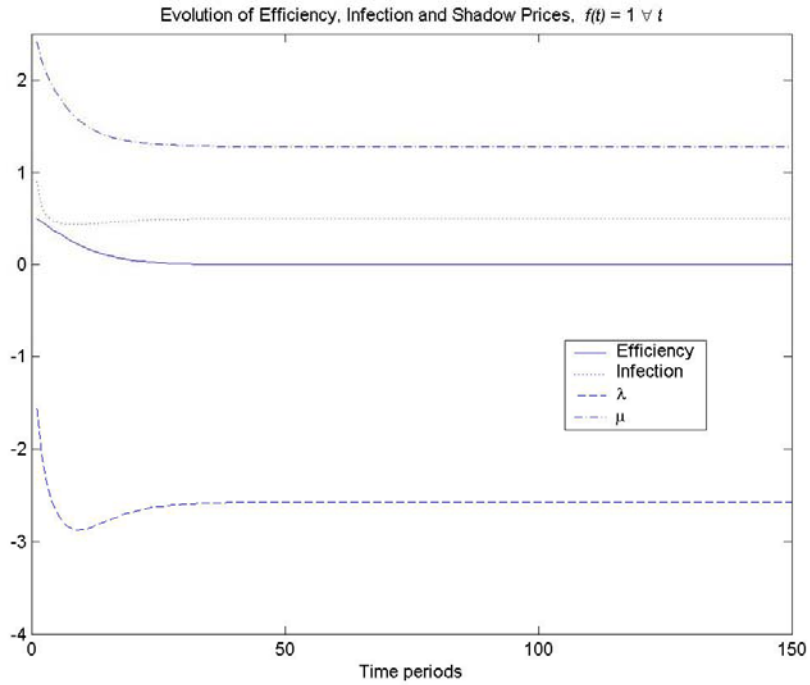


Figure A2

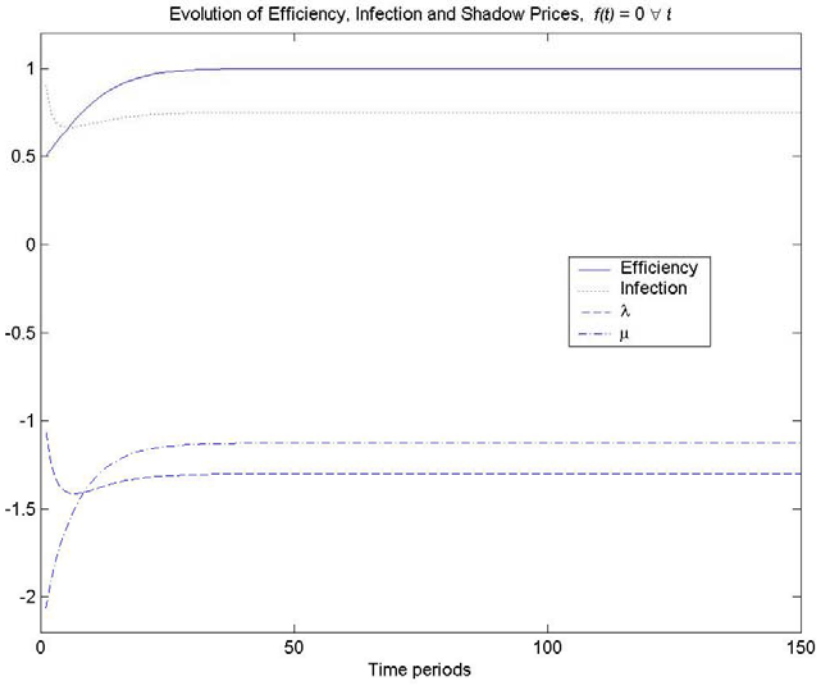
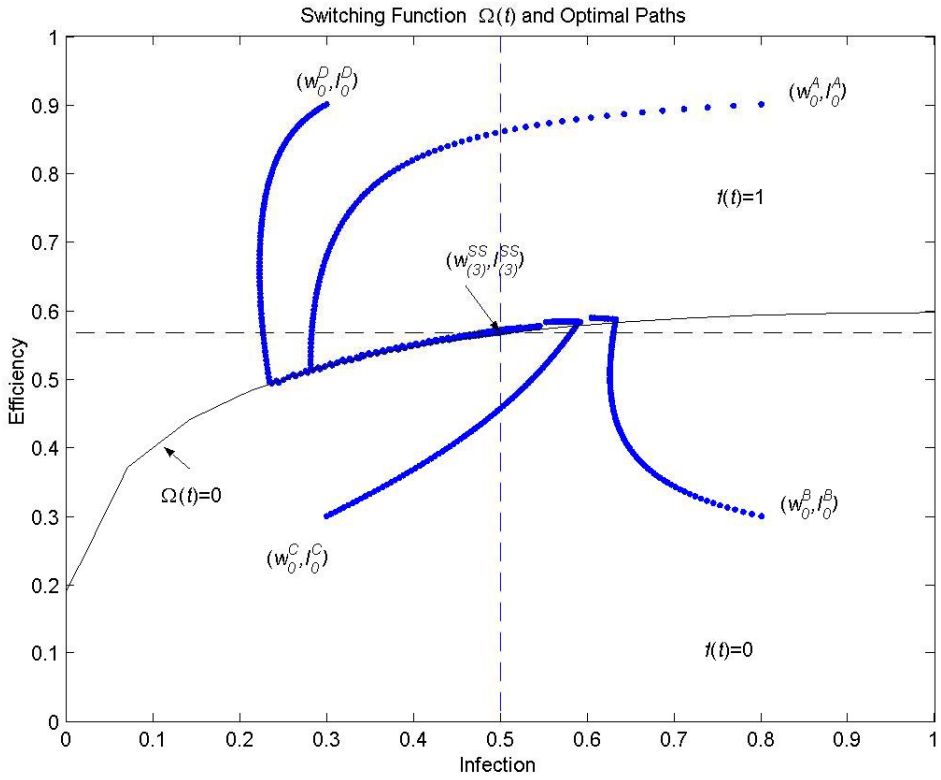


Figure 10+


Switching Function Ω(t) and Optimal Paths
Efficiency
Infection
(w_0^D, I_0^D)
(w_0^A, I_0^A)
$(w_{(3)}^{SS}, I_{(3)}^{SS})$
Ω(t)=0
(w_0^C, I_0^C)
(w_0^B, I_0^B)
t(t)=1
t(t)=0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1