PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2020 – 21

Optimal Prevention and Eradication of Epidemics

Hippolyte d'Albis
Emmanuelle Augeraud-Véron

JEL Codes: I18, C61
Keywords: Epidemics, Optimal Control.

Funded by a French government subsidy managed by the ANR under the framework of the Investissements d'avenir programme reference ANR-17-EURE-001

Optimal Prevention and Eradication of Epidemics[1]

Hippolyte d'ALBIS[2]
Paris School of Economics, CNRS

Emmanuelle AUGERAUD-VÉRON[3]
GREThA, University of Bordeaux

May 1st, 2020

[1]We are grateful to Jean-Pierre Amigues, Eric Benoît, Raouf Bouccekine, Martial Dupaigne, Albert Marcet and participants in seminars at Universities of Exeter, Paris and Toulouse, at European University Institute and ETH Zurich for their valuable comments and suggestions. The usual disclamer applies.

[2]Correspondence: 48 boulevard Jourdan, 75014 Paris, France. E-mail: hdal-bis@psemail.eu

[3]Correspondence: 6 Avenue Léon Duguit, 33600 Pessac, France. E-mail: emmanuelle.augeraud@u-bordeaux.fr

Abstract

This article studies the optimal intertemporal allocation of resources devoted to the prevention of deterministic epidemics that admit an endemic steady-state. In a stylized economy, the dynamics of the optimal prevention depends on the interplay between the epidemiological characteristics of the disease, the labour productivity and the level of intergenerational equity. A minimal level of labour productivity is shown to be necessary to reduce in the long run the prevalence rate of the epidemic. If this threshold is not reached, the prevention is then at best temporary, simply slowing down the spread of the epidemic disease. However, it may not optimal to undertake temporary prevention. Conversely, if labour productivity is sufficiently high, permanent allocation of resources to prevention is feasible but not necessarily optimal. If it is the case, the prevention monotonically increases with time for low initial prevalence rate, while it is decreasing or hump-shaped otherwise. The global analysis of the dynamics permits to uncover paths that yield to the eradication of the epidemic disease. It is shown that such paths are optimal if the pure discount rate is sufficiently low. Along those paths, the prevention is initially large and reaches zero in finite time.
JEL Classification: I18, C61

Keywords: Epidemics, Optimal Control.

1 Introduction

Epidemic diseases constitute major health issues for which there is large consensus on the legitimacy of governments interventions. Yet barely no analysis has been undertaken to determine the socially optimal allocation of resources to control the evolution of epidemics. In this article, we derive a welfare criterion from individual preferences and give precise foundations on a generally held belief, according to which intervention should begin as soon as possible and involve large expenses. We focus our analysis on expenditure that reduce the number of contaminative contacts per unit of time and, consequently, reduce the spread of the epidemic. They include prevention campaigns that modify individuals' behaviors, diffusion of masks or condoms that reduce the probability of a contact to be infectious, and any measures that reduce physical contacts such as lockdowns of populations. All those policies will be named prevention measures below, and our objective will be to determine the optimal intensity over time of the prevention in a economy that faces an epidemics.

Epidemics' prevention is a legitimate topic for economics since, as argued by Bloom and Canning [9], there is little doubt that resource constraints play an important role in the spread of epidemics. Moreover diseases importantly affect labor and capital markets and thus growth. However, and despite the fact that past epidemics were recognized as economic tragedies, Gersovitz and Hammer [23] pointed out that it is only recently that economists have entered the field.

Most articles adopt a positive approach focusing on private behaviors, such as individual choices on self-exposure to the risk (as in Geoffard and Philipson [25] and Kremer [32]), health expenditure (Momota et al. [36]), human capital accumulation (Bell and Gersbach [6], Corrigan et al. [17] and Boucekkine et al. [11]), fertility (Young [44]), income distribution ([12]) or biodiversity (Bosi and Desmarchelier [14]). Some papers analyzes the effect of public policies on private behavior (e.g. Geoffard and Philipson [26]), while others consider optimal policy correcting for the obvious externalities the epidemics generate (e.g. Gersovitz and Hammer [24], Francis [21] or Bethune et al. [8]).

Even though only few papers in economics adopt a normative perspective, there exists a rich mathematical epidemiology literature, that dates back from Bernouilli [7]. Following Sethi [40] and Wickwire [43], it is common to use optimal control techniques to define the desirable timing of vaccination, screening or health promotion campaigns. In most cases these studies use as a criterion a convex combination of the dynamic costs of the control and of the number of infected individuals. Moreover, the time horizon is usually finite and, in analytical models, the problem is linear with respect to the control. Based on this approach Behncke [5] finds that the optimal solution is, in general, such that the prevention effort is maximal on some initial time interval and then set to zero (see also Morris et al.). When the case of disease eradication is considered, the problem is more complex since terminal conditions are free (Barrett and Hoel [4]). Some economic studies (in particular Gersovitz and Hammer [23], Francis

[21], Feichtinger et al. [20] and Alvares and al. [1]) modify the criterion and use the present discounted value of total income net of the costs of the disease and of the control. To conclude, we can safely state that most of the epidemiological literature relies on economic calculus and ignore approaches based on welfare.

We propose an optimal control model in the tradition of Ramsey [39], Cass [15] and Koopmans [33] in which the whole population is affected by an epidemic disease, whose dynamics is rather general while admitting an endemic steady-state. The social welfare function is the present discounted value of the product of individual utility and the size of the population. We notably show how this criteria relies on preferences. The introduction of population in the objective is key to avoid the problem stressed by the optimal population literature (see notably Dasgupta [18]) about maximizing the welfare of alive individuals only. Our work extends key contributions of Delfino and Simmons [19], Boucekkine et al. [10] and Goenka and Liu [27], Goenka et al. [28], who consider specific epidemic dynamic processes and study local dynamics, and Gersovitz and Hammer [24], who proceed by simulations. A second important assumption we make is about the production structure which is of the 'yeoman farmer' type and allow us to completely solve the model despite a general formulation for the epidemic dynamics. The dynamics of the optimal prevention then depends on the interplay between the epidemiological characteristics of the disease, labour productivity and intergenerational equity. We believe it's important to characterize the global dynamics of the optimal system as it permits to visualize both

the objective that should be reached in the short run, and some trajectories that could be hidden while performing local long-run analysis.

We find that it may be optimal to reduce the prevalence rate of the epidemic in the long run only if labour productivity is above some minimal level. If this threshold is not reached, prevention is then at best temporary, simply slowing down the spread of the epidemic disease. However, it may not be optimal to undertake temporary prevention. When instead labour productivity is sufficiently high, permanent allocation of resources to prevention is feasible though not necessarily optimal. If permanent prevention is socially optimal, the prevention effort monotonically increases with time for low initial prevalence rate, and is hump-shaped or decreasing otherwise. Hence, our paper establishes that under a welfare criterion for social intertemporal optimization a "the-sooner-the-better" strategy may not be the optimal one, in contrast to Behncke [5]. This statement is however reversed when we consider paths that may yield to the eradication of the epidemic disease. We first show that a simple modification of traditional epidemic models can be made to allow for eradication. Trajectories that drive to eradication are characterized by an increasing prevention for a finite interval of time and, once the epidemics is eradicated, the prevention is zero. We show that upon existence, such paths are optimal if the pure discount rate is sufficiently small. In that case, it consequently is socially desirable that prevention should begin as soon as possible. Moreover, the effort initially devoted to prevention is larger for eradication than for any other dynamics. This

is a theoretical foundation for an "whatever it costs" response to an epidemic disease.

We begin by presenting the dynamics of the population affected by an epidemic in section 2. The epidemiological assumptions are put forward and discussed using standard examples of HIV epidemics and Flu epidemics. In section 3 we set up the social planner's problem, then prove the existence of a solution and characterize it. The dynamics of the optimal prevention is analyzed in section 4. The question on whether it is socially optimal to eradicate the epidemic disease is studied section 5, and illustrated with a hand solved example. Section 6 concludes.

2 Epidemic dynamics

This section presents the general characterization of the epidemic dynamics that will be used in the paper. The evolution of the disease is constrained by some assumptions that are satisfied, for instance, in classical compartmental epidemic models (Kermack and Mac Kendrick [31]). We also propose to introduce an arbitrarily small threshold above which the dynamics is valid and below which the epidemics is considered as eradicated. This modelling feature will be key in the subsequent analysis.

2.1 A two-classes framework

Time is continuous and indexed by $t \in \mathbb{R}^+$. The population at time t, whose size is denoted P_t, is affected by an epidemic disease and is thus decomposed in two

classes of individuals: the susceptible, who are healthy, and the infected, who may transmit the disease. The number of individual of each class is respectively denoted $S_t > 0$ and $I_t \geq 0$, and satisfy: $P_t = S_t + I_t$. It will be convenient to define the relative share of infected individuals with respect to share of susceptible ones as follows: $a_t = I_t/S_t$. This ratio is a monotonically increasing transformation of the prevalence rate of the epidemic, given by: $I_t/P_t = a_t/(1 + a_t)$, and will be named as the prevalence index throughout the remaining of the paper. Moreover, we consider the following general law of motion for a_t:

$$\dot{a}_t = g(h_t, a_t)\, a_t, \tag{1}$$

where the dot indicates the first derivative with respect to time and where h_t stands for the per capita expenditure devoted to the epidemic's control. These expenditure can be interpreted as prevention campaigns and any preventive measures that reduce the number of contaminative contacts per unit of time and, consequently, reduce the spread of the epidemic (see, for instance, Castilho [16]). Note that these expenditure can not stand for vaccination campaigns or screening and isolation strategies, as they produce a third class of individuals within the population; a case that is not considered here. In the rest of the paper, h_t will be referred as the level of prevention at time t.

The growth rate of the prevalence index, characterized by function g, is supposed to satisfy:

H1. $g : \mathbb{R}^2_+ \to \mathbb{R}$, is $\mathcal{C}^2$, $g'_1(h_t, a_t) < 0$, $g'_2(h_t, a_t) < 0$ and $g''_{11}(h_t, a_t) > 0$. There

exist $a^* > 0$ and $h^* > 0$ such that $g(0, a^*) = g(h^*, 0) = 0$.

Assumption H1 fully describes the kind of epidemic we are dealing with. The growth rate of the prevalence index is supposed to decrease with a_t which, of course, does not imply a monotonic relationship between $\dot{a}_t$ and a_t. If there is no prevention, the epidemic lasts forever but stabilizes within the population: the prevalence index converges to the endemic steady-state a^*. The other steady-state, given for $a = 0$, and named in the epidemiological literature the disease-free equilibrium, can easily be check as being unstable as Assumption H1 implies $g(0, 0) > 0$. The reproduction number[1], i.e. the number of secondary cases that can appears when one infected individual is added in a population of fully susceptible individuals for the duration of her illness, denoted $\mathcal{R}_0$, is here larger than 1 (van den Driessche [42]). Figure 1a represents this dynamics.

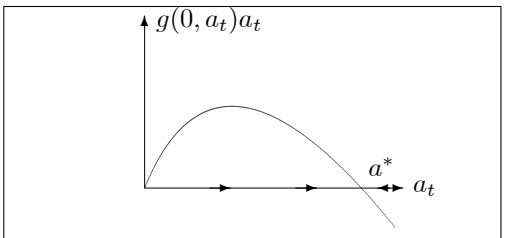


Figure 1a: The dynamics of a_t when $h_t = 0$

Prevention may modify the dynamics of the epidemics. To fix ideas, let assume that h_t is an exogenous constant. If this constant belongs to $(0, h^*)$, the prevalence index converges to another stable steady-state characterized by

[1] Formally, we can write: $ag(h, a) = F(h, a) + V(a)$ where F include the terms related to the transmission of the disease and V those related to other causes (death, etc.). The derivative at $a = 0$ of the pervious equation is given by: $g(h, 0) = F'(h, 0) + V'(0)$. Then the (controled) reproduction number is defined as $\mathcal{R}_c = -F'(h, 0)/V'(0)$, while the (classical) reproduction number is $\mathcal{R}_0 = -F'(0, 0)/V'(0)$.

a prevalence rate that is lower than a^*; the reproduction number with positive control $\mathcal{R}_c$ remains larger than 1. However, if the constant is larger than h^*, the epidemic dynamics are dramatically modified as the prevalence index converges to zero; the control is sufficient to eliminate the disease in the long run and we now have $\mathcal{R}_c < 1$. Figure 1b illustrates such dynamics with two different exogenous h_t: $h_1 \in (0, h^*)$ and $h_2 > h^*$. In the remaining of the paper, h_t will be the control variable optimally chosen by a benevolent social planner.

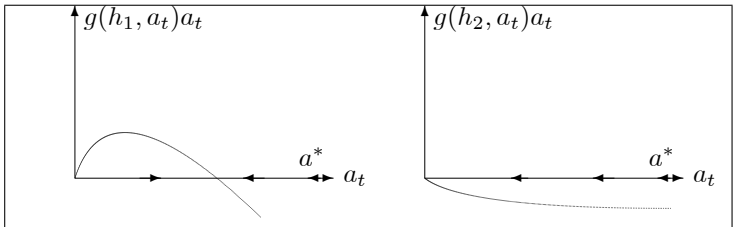


Figure 1b: The dynamics of a_t when $h_t > 0$

Let us remark that it is easy to extend our approach to other kinds of epidemic dynamics. An alternative dynamics without prevention could be obtained by assuming: $a^* < 0$. Such an epidemic would not exhibit an endemic steady-state. The unique stable steady-state being the one with a zero prevalence index. Prevention may then be used to accelerate the convergence process. Moreover, assuming that $a^* \to +\infty$, allow to consider an epidemic whose prevalence indefinitely increases if there is no control of it.

By definition, the population growth rate writes as follows:

$$\frac{\dot{P}_t}{P_t} = \frac{1}{1+a_t}\left(\frac{\dot{S}_t}{S_t} + \frac{\dot{I}_t}{I_t}a_t\right). \tag{2}$$

For computational reasons, we assume that the growth rate is characterized by

a function denoted $n(.)$ that satisfies:

H2. $\dot{P}_t/P_t = n(a_t)$. $n : \mathbb{R}_+ \to [\underline{n}, n(0)]$, is C^2, $n'(a_t) \leq 0$.

Assuming that the population growth rate does not depend on (S_t, I_t, h_t) is rather strong but is widely used in the epidemiological literature, notably in the examples presented below. Moreover, the assumption concerning non positivity of $n'(a_t)$ is not only more realistic -epidemics increase the death rates- but also, as it will be discussed throughout the paper, more meaningful.

2.2 A minimal threshold for the epidemics

We now consider a simple generalization of equation (1), which is given by:

$$\dot{a}_t = \begin{cases} g(h_t, a_t)\, a_t & \text{if } a_t > a_{\min} \\ 0 & \text{if } a_t \leq a_{\min}, \end{cases} \tag{3}$$

where $0 \leq a_{\min} \ll 1$. Moreover, Assumption H1 is left unchanged. Equation (3) introduces a threshold below which the epidemic is considered as eliminated (less than one individual is infected). Below it, the prevalence index is negligible and the epidemic can not survive and spread within the population. Above this threshold, the dynamics has the same qualitative property as the one we have considered before. Technically, the main difference between the case $a_{\min} > 0$ and the case $a_{\min} = 0$ that is usually considered in the epidemiology literature, is that steady-state $a_{\min}$ can be reached in finite time, which will appear as a key property below. We will then assume that when $a_{\min}$ is reached, the disease is eradicated.

The dynamics of a_t is represented in Figure 2, in the case of no prevention.

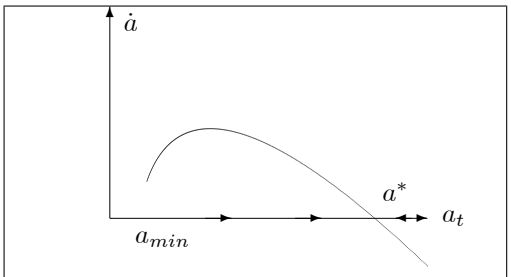


Figure 2: The dynamics of a_t with $a_{\min} > 0$

Our framework generalizes most works in economics that have analytically studied the optimal dynamics of an epidemic: Delfino and Simmons [19], Gersovitz and Hammer [23] and Barrett and Hoel [4] study a dynamics similar to (1), but for a function g which is specified, while Boucekkine et al. [10] consider a shock on the initial stock of the population describing an epidemic with instantaneous effects and no endemic steady-state.

2.3 Examples

Our assumptions on functions g and n are now confronted to two examples: a SI model, which can describe the HIV epidemic and a SIS model, which can describe flu epidemics. In SIS models, infected individuals may recover for the disease but are not immunized and may again be infected. This is relevent for flu if one considers a mutating virus. However, most current models of the SARS-CoV-2 assume this is not the case and use three-compartments models of the SIR type (Avery et al. [3]). Such models can be covered by our analysis provided that the size of the population is kept constant.

Example 1. SI model (HIV epidemic)

Let us first suppose that $a_{\min} = 0$ and ignore the possibility to control the epidemics. The natural growth rate of the susceptible population is given by $\beta - \mu$, where $\beta > 0$ and $\mu > 0$ respectively stand for the birth and the death rates, while the growth rate of the infected population is $\beta - \mu - \gamma$; parameter $\gamma \geq 0$ measures the over-mortality yield by the disease. Both vertical and horizontal contamination are considered: first a proportion $\pi \in [0, 1]$ of the children of infected people are born healthy, while the others are infected. Moreover, as in May and Anderson [34], it is assumed that the incidence of the epidemic follow a law in frequency: contamination is proportional, up to a parameter $\sigma > 0$, to the density of infected individuals in the total population. The dynamics of each subpopulation is therefore given by:

$$\dot{S}_t = (\beta - \mu) S_t + \beta\pi I_t - \sigma S_t \frac{I_t}{S_t + I_t}, \tag{4}$$

$$\dot{I}_t = [\beta (1 - \pi) - \mu - \gamma] I_t + \sigma S_t \frac{I_t}{S_t + I_t}. \tag{5}$$

It can be easily shown that this system does not generically admit a steady-state except: $S_t = I_t = 0$. Hence, the dynamics of the epidemic is better understood using per capita variables or, equivalently, by using our prevalence index a_t (defined as $a_t = I_t/S_t$). Using (4) and (5), the dynamics of a_t solves:

$$\dot{a}_t = [\sigma - \beta\pi - \gamma - \beta\pi a_t] a_t, \tag{6}$$

which is a logistic equation. Therefore, if $\pi \in (0,1]$, the dynamics of a_t writes:

$$a_t = \frac{\frac{(\sigma-\beta\pi-\gamma)}{\beta\pi} a_0}{a_0 + \left(\frac{(\sigma-\beta\pi-\gamma)}{\beta\pi} - a_0\right) e^{-(\sigma-\beta\pi-\gamma)t}}. \tag{7}$$

Equation (6) admits (i) two steady-states if $\sigma > \beta\pi + \gamma$: namely, $\hat{a} = 0$, which is unstable and $a^* = (\sigma - \beta\pi - \gamma)/\beta\pi$, which is stable, (ii) one steady-state if $\sigma \leq \beta\pi + \gamma$: namely, $\hat{a} = 0$, which is stable. Consequently, if the contamination coefficient σ is sufficiently low, the epidemic ultimately disappears. Conversely, if σ is high, the epidemic survives as the prevalence index stabilizes.[2] If there is full vertical transmission (i.e. if $\pi = 0$), the only possible steady state is $\hat{a} = 0$, whose stability is given by the sign of $\sigma - \gamma$.

Using (4) and (5), the population growth rate $n(a_t)$ solves:

$$n(a_t) = \beta - \mu - \frac{\gamma a_t}{1 + a_t}, \tag{8}$$

which decreases in a_t; moreover, $\beta - \mu - \gamma \leq n(a_t) < \beta - \mu$. The higher bound for $n(.)$ being the population growth rate without epidemic.

Suppose now that prevention may affect the epidemic dynamics through the contamination parameter by reducing the number of effective contacts by time unit. Let h_t be the per capita expenditure devoted to prevention and the contamination coefficient at time t be a function that writes $\sigma(h_t)$ and satisfies

[2] Alternatively, one may present the stability results using the reproduction number which is here given by $\mathcal{R}_0 = (\sigma - \beta\pi)/\gamma$. The dynamics of a can be rewritten as:

$$\dot{a}_t = \left[\mathcal{R}_0 - 1 - \frac{\beta\pi}{\gamma} a_t\right] a_t \gamma,$$

which implies that $a^* > 0$ if and only if $\mathcal{R}_0 > 1$.

$\sigma'(h_t) < 0$. The dynamics of a_t is now given by:

$$\dot{a}_t = [\sigma(h_t) - \beta\pi - \gamma - \beta\pi a_t] a_t. \tag{9}$$

It immediate to check that Assumption H1 is satisfied only if $\pi > 0$ and if $\sigma(0) > \beta\pi + \gamma$ and that Assumption H2 is always satisfied. Hence, full vertical transmission is excluded here.

Remarkably, our example easily extend to the case $a_{\min} > 0$, for which equation (7) holds for $a_t \geq a_{\min}$ only. The stability analysis is modified as follows: if $\sigma > \beta\pi + \gamma$, a^* remains a stable steady-state while if $\sigma < \beta\pi + \gamma$, $a_{\min}$ is stable and can be reached in finite time.

Example 2. SIS model (Flu epidemic)

Following Hethcote and Yorke [30], let us now consider the following dynamics for the susceptible and the infected:

$$\dot{S}_t = \beta(S_t + I_t) - \mu S_t + \delta I_t - \sigma S_t \frac{I_t}{S_t + I_t}, \tag{10}$$

$$\dot{I}_t = -(\delta + \mu + \gamma) I_t + \sigma S_t \frac{I_t}{S_t + I_t}. \tag{11}$$

The main difference are that there is no vertical transmission of the disease, and that infected individuals recover at rate $\delta > 0$. Then, the dynamics of the prevalence index satisfies:

$$\dot{a}_t = [\sigma - \delta - \gamma - \beta - a_t(\delta + \beta)] a_t, \tag{12}$$

while the population growth rate is the same as (8). The stability analysis is

similar to that of the previous example.[3] Again, we can assume that σ is a decreasing function of the level of prevention h.

3 The optimal control problem

This section establishes the optimal control problem we are going to study. It is an infinite horizon framework with an economic structure and the population dynamics described in section 2. We first present and discusses the social welfare function and then prove the existence of an optimal solution.

3.1 The social welfare function

The social welfare function we introduce is derived from the aggregation of individual's preferences. Each individual is supposed to belong to a dynasty of altruistic individuals. Without epidemics, the growth rate of the dynasty is the constant $n(0)$. However, at each point of time, the epidemic disease may kill the dynasty. Denote by λ_t the probability as of time $t = 0$ that the dynasty is still alive at time t. If alive at time t, the utility of a dynasty member depends on consumption c_t and is independent of the health status. The utility function is then $u(c_t)$. If not alive at time t, the utility is supposed to be the constant $u(0) \gg -\infty$. The expected utility of the dynasty at time $t = 0$ is therefore:

$$\int_0^{+\infty} e^{-(\rho - n(0))t} \left[\lambda_t u(c_t) + (1 - \lambda_t) u(0)\right] dt, \tag{13}$$

[3] The reproduction number is now given by $\mathcal{R}_0 = \sigma / (\delta + \beta + \gamma)$ and the dynamics of the prevalence index can be rewritten as:

$$\dot{a}_t = (\delta + \gamma + \mu) \left[\mathcal{R}_0 - 1 - a_t \frac{\delta + \beta}{\delta + \gamma + \mu}\right] a_t.$$

where ρ is the pure discount rate which satisfies the following restriction:

H3. $\rho > n(0)$.

Moreover, function u satisfies the following assumption:

H4. $u : \mathbb{R}_+ \to \mathbb{R}_+$, $u \in C^3$, $u' > 0$, $u'' < 0$ and $\lim_{c \to 0} u'(c) = +\infty$.

To obtain the social welfare function, we assume that the population is composed by a continuum of identical dynasties, whose total size at time $t = 0$ is P_0. By the law of large numbers, the probability λ_t is, at the aggregate level, the ratio between the size of the population and the size that would prevail without epidemic. Thus: $\lambda_t = P_t / P_0 e^{n(0)t} = e^{\int_0^t [n(a_s) - n(0)] ds}$. The social welfare function at time $t = 0$ is therefore simply obtained by multiplying the function (13) by the initial size of the population P_0, and rearranging to obtain:

$$P_0 \int_0^{+\infty} e^{-\int_0^t [\rho - n(a_s)] ds} \left[u(c_t) - u(0) \right] dt + \frac{P_0 u(0)}{\rho - n(0)}. \tag{14}$$

Maximizing this latter function is, in fact, equivalent to maximizing:

$$\int_0^{\infty} e^{-\rho t} P_t u(c_t) \, dt. \tag{15}$$

The social planner function is thus the discounted value of the product of the size of the population, P_t, and of the instantaneous utility of each individual. For a given path of consumption, a larger population hence increases the social welfare. Consequently, the assumption $n'(a_t) < 0$, implies that reducing the

number of infected individuals increases welfare, everything being equal. Observe finally that the limit case $n'(a_t) = 0$, which features epidemics that have no impact of the population growth, implies $\lambda_t = 1$ and a standard objective function.

3.2 The social planner's program

The social planner faces the resource constraint of the economy. There is one material good produced using labor and it is assumed that the productivity of an infected individual is lower than the one of an susceptible individual. Production per capita writes: $\alpha f(a_t)$ where f is C^1 a non increasing function and $\alpha > 0$ is a measure of the productivity of labor. The larger α, the wealthier the economy. For instance, one may consider that the production function is linear with respect to labor, with the productivity being equal to α for susceptible individuals and to $\eta\alpha$ for infected individual (with $0 \leq \eta \leq 1$). Then, in this example: $f(a_t) = (1 + \eta a_t) / (1 + a_t)$.

The produced good can be used for consumption or for the expenditure devoted to the control of the epidemics. The resource constraint written in per capita units is therefore:

$$c_t + h_t = \alpha f(a_t). \tag{16}$$

Moreover, consumption, prevention and the prevalence index should be non negative. The program of the social planner is to maximize (15) subject to (1)

and (16). It writes:

$$\max_{h_t} \int_0^{\infty} e^{-\int_0^t \theta(a_s)ds} u\left(\alpha f\left(a_t\right) - h_t\right) dt,$$

$$s.t. \left| \begin{array}{l} \dot{a}_t = \left| \begin{array}{ll} g\left(h_t, a_t\right) a_t & \text{if } a_t > a_{\min} \\ 0 & \text{if } a_t \leq a_{\min}, \end{array} \right. \\ \\ 0 \leq h_t \leq \alpha f\left(a_t\right) \text{ and } a_0 \in \left(a_{\min}, a^*\right) \text{ given.} \end{array} \right. \tag{17}$$

where $\theta\left(a_t\right) := \rho - n\left(a_t\right)$. We note that since n is a decreasing function and given Assumption H3, one has $\theta\left(.\right) > 0$ and $\theta'\left(.\right) > 0$.

The problem is formally similar to an optimal growth model with endogenous discounting (see also Boucekkine et al. [13]). To reduce the length of the proofs and to focus on the meaningful cases, we solved the problem for $a_0 \in \left(a_{\min}, a^*\right)$ but the analysis can be generalized to $a_0 > a^*$. Let us notice that the problem is trivial if the initial prevalence of the epidemic is below the minimal threshold (i.e. $a_0 \leq a_{\min}$): the optimal consumption is equal to the production $\alpha f\left(a_{\min}\right)$ and the prevention is equal to zero.

The intertemporal trade-off is the following: an increase in h_t yields a reduction of both the immediate per-capita consumption and the prevalence of the epidemic. The latter implies first an increase in future per-capita production and therefore the expenditure devoted to prevention can be understood as an investment. Moreover, reducing prevalence leads to a modification of the spread between the discount rate and the population growth rate. As $\theta'\left(a_t\right) > 0$, an increase in h_t implies a reduction of the spread, meaning a more equal treatment between individuals of different generations as it increases the weight associated

to the utility of future generations.

3.3 An existence result

The program (17) is a non autonomous problem with endogenous discounting but can be equivalently analyzed as an exogenous discounting problem using the virtual time method described by Uzawa [41]. The following results are then derived.

Lemma 1 *There exists an optimal solution to program (17). The optimal solution satisfies* $h_t < \alpha f(a_t)$.

Proof. See Appendix.

We now turn to the characterization of the solution by analyzing first the case such that $a_{\min} = 0$, which is standard in mathematical epidemiology. The main goal is to characterized the dynamics of the optimal prevention strategy. Then, we turn to the case $a_{\min} > 0$ to analyze the issue of eradication.

4 Is prevention optimal?

This section studies the system of equations (18) that characterizes the dynamics of prevalence and prevention when there is no minimal threshold for the epidemics (i.e. for $a_{\min} = 0$). Local dynamics around steady-state are first studied and a geometrical analysis using phase diagrams is then provided.

4.1 The optimal dynamics for $a_{\min} = 0$

Let us first give the equations describing the optimal trajectories.

Lemma 2 *An optimal path is necessarily a solution of the following system:*

$$\begin{cases} \dot{a}_t = \frac{g(h_t,a_t)a_t}{\theta(a_t)} & \\ \dot{h}_t = \frac{\Phi(h_t,a_t,c_t)}{-\theta(a_t)\left[\frac{u''(\alpha f(a_t)-h_t)}{u'(\alpha f(a_t)-h_t)}+\frac{g''_{11}(h_t,a_t)}{g'_1(h_t,a_t)}\right]} & \text{if } h_t > 0, \\ \dot{a}_t = \frac{g(0,a_t)a_t}{\theta(a_t)} & \text{if } h_t = 0, \end{cases} \tag{18}$$

with $c_t = \alpha f(a_t) - h_t$ *and:*

$$\begin{aligned} \Phi(h_t,a_t,c_t) &= \left[-\alpha f'(a_t)\frac{u''(c_t)}{u'(c_t)} + \frac{g''_{12}(h_t,a_t)}{g'_1(h_t,a_t)} + \frac{\theta'(a_t)}{\theta(a_t)}\right] g(h_t,a_t)\, a_t \\ &\quad + \left[-\alpha f'(a_t) + \frac{u(c_t)}{u'(c_t)}\frac{\theta'(a_t)}{\theta(a_t)}\right] g'_1(h_t,a_t)\, a_t \\ &\quad - g'_2(h_t,a_t)\, a_t + \theta(a_t). \end{aligned} \tag{19}$$

Proof. See the Appendix.

From Lemma 2, it is straightforward to derive the optimal dynamics of consumption, which is always positive (see Lemma 1).

4.2 Prevention in the long run

Let us first consider the local dynamics in the neighborhood of the steady-states of (18). We define a 'corner steady-state' as a steady-state for which the optimal prevention is zero and an 'interior steady-state' as a steady-state for which the optimal prevention is positive.

Given Assumption H1, the pairs $(a^*, 0)$ and $(0, 0)$ are the two corner steady-states of our system. The first one satisfies the following properties:

Lemma 3 *Provided that there exists an optimal path that converges to the corner steady-state* $(a^*, 0)$*, it satisfies* $h_t = 0$ *for* t *large enough.*

Proof. See the Appendix.

Whatever its initial dynamics, the optimal prevention is thus equal to zero after a finite date if the prevalence index converges to an endemic steady-state without prevention. The intuition is that $(a^*, 0)$ is not a steady-state for the interior dynamics of system (18). Optimal prevention may hence not converge to zero but only reach zero in a finite time. Since it is never optimal to reach a^* in a finite time, we conclude that $h_t = 0$ in the neighborhood of $(a^*, 0)$. Note that this argument rules out local indeterminacy.

Concerning the second corner steady-state, we obtained the following:

Lemma 4 *There is no optimal path that converges to* $(0, 0)$.

Proof. See the Appendix.

The eradication of the epidemics is not a possible output of the model with a traditional representation of the epidemics. The intuition, which is illustrated in the phase diagrams below is the following. A trajectory that would converges to the desease-free-equilibrium would necessary imply $\dot{h}_t > 0$ forever, which is not possible because of the budget constraint.

Let us now study interior steady-states, upon which prevention is positive. Using (18), such a steady-state is a pair (a, h) that satisfies $a \in (0, a^*)$, $h \in$

$(0, \alpha f(a))$ and solves:

$$g(h,a) = 0, \tag{20}$$

$$-\left[g_2'(h,a)\,a - \theta(a)\right] = -\left[-\alpha f'(a) + \frac{u(c)}{u'(c)}\frac{\theta'(a)}{\theta(a)}\right] g_1'(h,a)\,a, \tag{21}$$

where $c = \alpha f(a) - h$. Then, a necessary condition for existence of an interior steady-state is the positivity of equation (21)'s right hand side, which rewrites as follows:

$$\frac{d}{da}\left(\frac{u(\alpha f(a) - h)}{\theta(a)}\right) < 0. \tag{22}$$

Condition (22) means that the discounted welfare of a generation in the long run should be increased by a reduction of the epidemic. The increase in utility implied by a marginal decreases of a (measured by $-\alpha f'(a)\,u'(c)/\theta(a)$) should be larger than the negative impact on the endogenous discount (given by $\theta'(a)\,u(c)/(\theta(a))^2$). Necessary and sufficient conditions (20) and (21) may be rewritten in the following way:

$$\frac{u'(u(\alpha f(a) - \eta(a)))}{\theta(a)} = \frac{d\dot{a}}{dh} \times \frac{d}{da}\left(\frac{u(\alpha f(a) - \eta(a))}{\theta(a)}\right), \tag{23}$$

where $\eta(a)$ is the implicit relation between h and a derived from Assumption H1. The contemporaneous desutility induced by a marginal increase in prevention should equal the benefits of the reduction of the epidemic.

The following lemma studies the existence of interior steady-states.

Lemma 5 *i) There exists $\bar{\alpha} > 0$, such that there is no interior steady-states if $\alpha \leq \bar{\alpha}$. ii) There exists $\hat{\alpha} > \bar{\alpha}$, such that there exist interior steady-states if*

$\alpha \geq \hat{\alpha}$. iii) Upon existence, interior steady-states are locally unstable.

Proof. See the Appendix.

Lemma 5 shows the importance of labor productivity, or equivalently, of the level of wealth per capita, on the prevalence index in steady-state: there are thresholds below which there is an high prevalence and no prevention, and above which there is lower prevalence with some prevention. The intuition of this result hinges on the concavity and on the Inada condition imposed on the utility function: when the average production is low, resources are exclusively devoted to consumption since a marginal decrease of it produces a large desutility and since the marginal impact of prevention is independent of the level of productivity. Consequently, an interior steady-state is more likely to exist if the labor productivity is increased: the immediate marginal desutility of prevention (i.e. the LHS of (23)) is then lowered while its impact on future generations' discounted utility (i.e. the RHS of (23)) is increased.

Importantly, interior steady-states are locally unstable. As it will be done below, it is possible to characterize, under further conditions, saddle-path steady-states. Hence, from Lemmas 3 and 5, we conclude there are at most two kinds of paths that converge to a steady-state. First, we have a family of paths converging to corner steady-states. Second, upon existence of a saddle-path steady-state, we have the stable arm converging to one interior steady-state. All paths are candidates for optimality.

Suppose now there exists at least one saddle-path steady-state. Let us compare the unique path that converges to this steady-state to the family of those which converge to the corner steady-state.

Lemma 6 *Suppose there exists a saddle-path interior steady-state, denoted by* $\left(\bar{a},\bar{h}\right)$. *The stable arm that converges to* $\left(\bar{a},\bar{h}\right)$ *may not be optimal.*

Proof. See the Appendix.

Lemma 6 shows that the existence of an interior steady-state does not implies that the stable arm is necessarily optimal. In the proof, we show that using a particular case where the long run cost of prevention is higher that the benefit in terms of production of having a lower share of infected individuals. Then, there exist sets of initial condition such that the intertemporal utility yield by the stable arm is lower than the one yield by the path converging to the corner steady-state. Conversely, there exists initial conditions, upon which the saddle-path is always preferred to those converging to the corner steady-state.

4.3 The global dynamics of prevention

We now propose a geometrical representation of the results that have been previously established by drawing the phase diagram associated to the system of equations (18). To reduce the length of the proofs, an additional set of restrictions is assumed:

H5. $g''_{12}(h,a)=0$, and $-u'''(c)/u''(c)\geq -u''(c)/u'(c)$.

Assumption H5 is sufficient to ensure that isocline $\dot{h} = 0$, derived from function Φ given in equation (19) is well defined for $g > 0$. The epidemic dynamics considered in Assumption H1 is now constrained by further assumptions on the impact of prevention on $\dot{a}_t$: it is now convex and proportional to a_t. Moreover, the utility function restricts to a representative individual with absolute risk aversion lower than absolute prudence, a property satisfied by standard utility functions including those with harmonic absolute risk aversion. We established the following properties.

Lemma 7 *i) The* $\dot{a} = 0$ *locus is downward slopping in the plane* (a, h) *and is such that* $\dot{a} > 0$ *below the locus. ii) For* $g > 0$*, the* $\dot{h} = 0$ *locus is defined by a function* $h = \chi(a)$ *that satisfies* $\lim_{a \to 0} \chi(a) = -\infty$ *and is such that* $\dot{h} > 0$ *above the locus. iii) As* α *increases, the* $\dot{h} = 0$ *moves upward in* $\left\{(a, h) \in \mathbb{R}^2,\ g(h, a) > 0\right\}$.

Proof. See the Appendix.

Lemma 7 gives enough information to draw the phase diagram associated to system (18). Depending on the existence of interior steady-states, optimal paths can hence be represented. Possible phase diagrams are given by figure 3a if there

is no interior steady-states and by figure 3b if there are two interior steady-states.

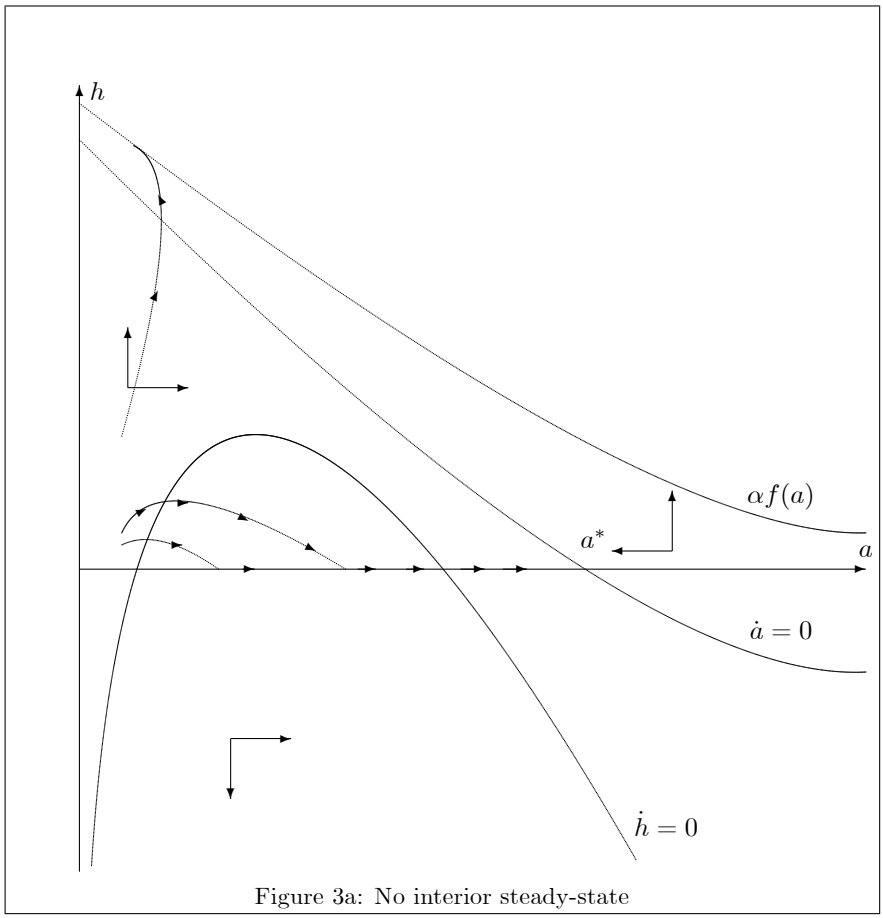


Figure 3a: No interior steady-state

Note that we have represented the upper limit for h, given by function $\alpha f(a)$, above the $\dot{a} = 0$ locus, but it as well can be below for some value of a. For an initial condition $a_0 \in (0, a^*)$, there is hence a family of feasible paths: they converge toward the corner steady-state $(0, a^*)$ and provided that the $\dot{h} = 0$ locus is above the horizontal axis, a positive level of prevention is possible for a finite interval of time. The dynamics of the epidemic may be slowed down for a while but, ultimately, the prevalence index reaches the long-run level with no intervention. Another family of paths is drawn in Figure 2a: they move to

the vertical axis with a high level of prevalence. These paths are however not feasible since the prevention monotonously increases with time and reaches in finite time the upper limit given by $\alpha f(a)$. The consumption is there equal to zero and, consequently, the path is not optimal.

Labor productivity, which has been shown in the previous section to be crucial for the existence of interior steady-states, has an impact on the dynamics. Geometrically, as α increases, the $\dot{h} = 0$ locus and the constraint $\alpha f(a)$ move up. Since the $\dot{a} = 0$ locus is left unchanged, interior steady-states are more likely to appear, as stated in lemma 5 and drawn in Figure 3b.

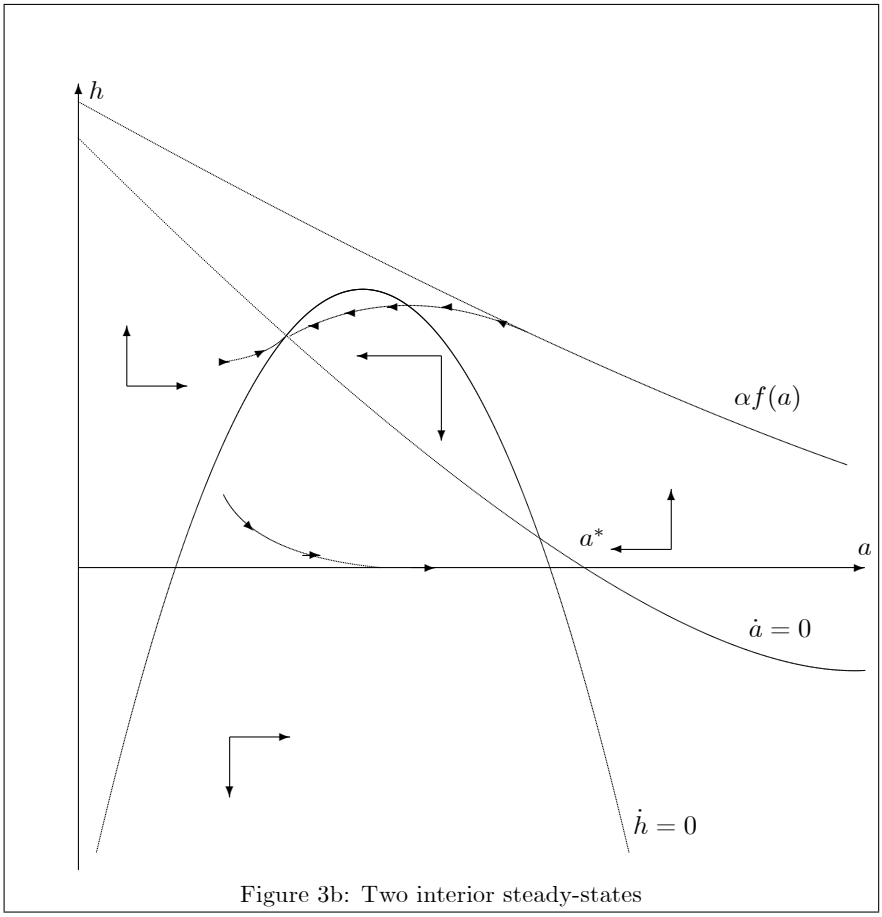


Figure 3b: Two interior steady-states

Two interior steady-states are represented in the phase diagram of Figure 3: a first one with a higher level of prevention and a lower prevalence, which is saddle-path and a second one which is a repulsive cycle. Hence, in addition to the families of paths that have been considered in the case without interior steady-states, there is a unique path that converges to the saddle-path steady-state. Remark that if there exist more than two interior steady-states, the phase diagram would exhibit alternatively saddle-paths and cycles. If the stable arm converging to the steady-state $\left(\bar{h}, \bar{a}\right)$ is optimal, the prevention monotonically increase with time if $a_0 < \bar{a}$ and can be an hump shaped function of time if $a_0 > \bar{a}$.

In this phase diagram analysis, we have insisted on the fact that on the neighborhood of the vertical axis, the growth rate of prevention is always positive (i.e. $\dot{h} > 0$). This behavior is induced, using (18), by the assumption of a "not too concave" relationship between the epidemic growth rate and the prevention. More precisely, it is true if:

$$\lim_{h \to 0} \frac{g_{11}''(h,0)}{g_1'(h,0)} < -\frac{u''(\alpha f(0))}{u'(\alpha f(0))}, \tag{24}$$

which is satisfied in Assumption H5.

Note finally that any path that may converge to the vertical axis, and thus to the eradication of the epidemic, necessarily reaches in finite time the resource constraint and is therefore not optimal. Assuming the opposite inequality than that of (24) would not change the statement about the impossibility to eradicate,

but simply the phase diagram. In the next section, we assume that $a_{\min} > 0$, which allows to consider the eradication of the epidemic.

5 The eradication of the epidemic

Lemma 4 has shown that eradication was not possible in the case a $a_{\min} = 0$. We now consider the case with a positive threshold below which the epidemic remains constant at a negligible prevalence rate. Below, we first show that a trajectory that converges to $a_{\min} > 0$ is feasible and then give the conditions for its optimality.

5.1 The trajectory toward eradication

We start with the following result.

Lemma 8 *An optimal path that converges to the corner steady-state* $(a_{\min}, 0)$ *may exist. It satisfies* $h_t = 0$ *for* t *large enough.*

Proof. See the Appendix.

If $a_{\min} > 0$ is allowed, the eradication is feasible in our model. Since $a_{\min}$ can be reached in finite time, h_t also reaches zero in finite time, as prevention has no impact when $a_t = a_{\min}$.

5.2 Global dynamics with eradication

Let us now see how our phase diagrams are modified: a new feasible path appears, which yields to the eradication of the epidemic in finite time. Figure

4a represents a possible phase diagram when there is no interior steady-state.

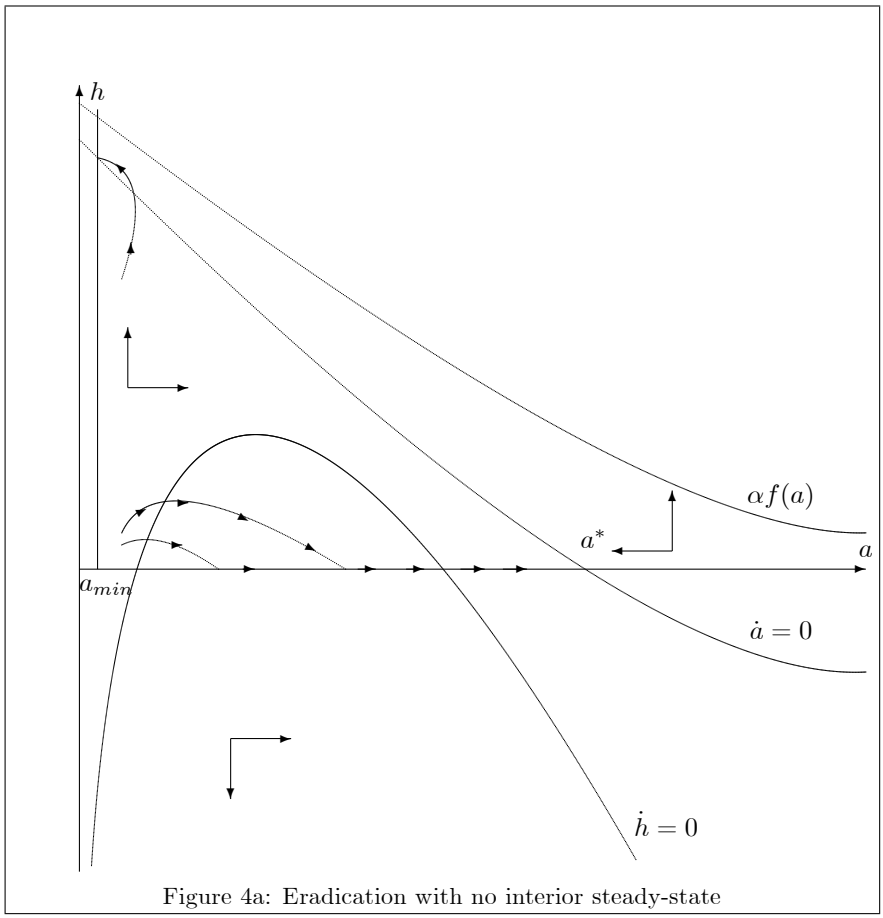


Figure 4a: Eradication with no interior steady-state

The family of paths that converges to the corner steady-state $(0, a^*)$ is still represented in Figure 4a. Moreover, there is a path that reaches the vertical axis in finite time, along which prevention monotonically increases. Moreover, the path is unique as it necessarily goes through the intersection between the vertical line at $a_{\min}$ and the $\dot{a} = 0$ locus. A necessary condition for the existence of such a path is simply that those coordinates are feasible, meaning that the $\breve{h}$ such that $g\left(\breve{h}, a_{\min}\right) = 0$ satisfies the following condition: $\breve{h} < f\left(a_{\min}\right)$. A path leading to eradication also appears when there are interior steady-states,

as it is shown by Figure 4b.

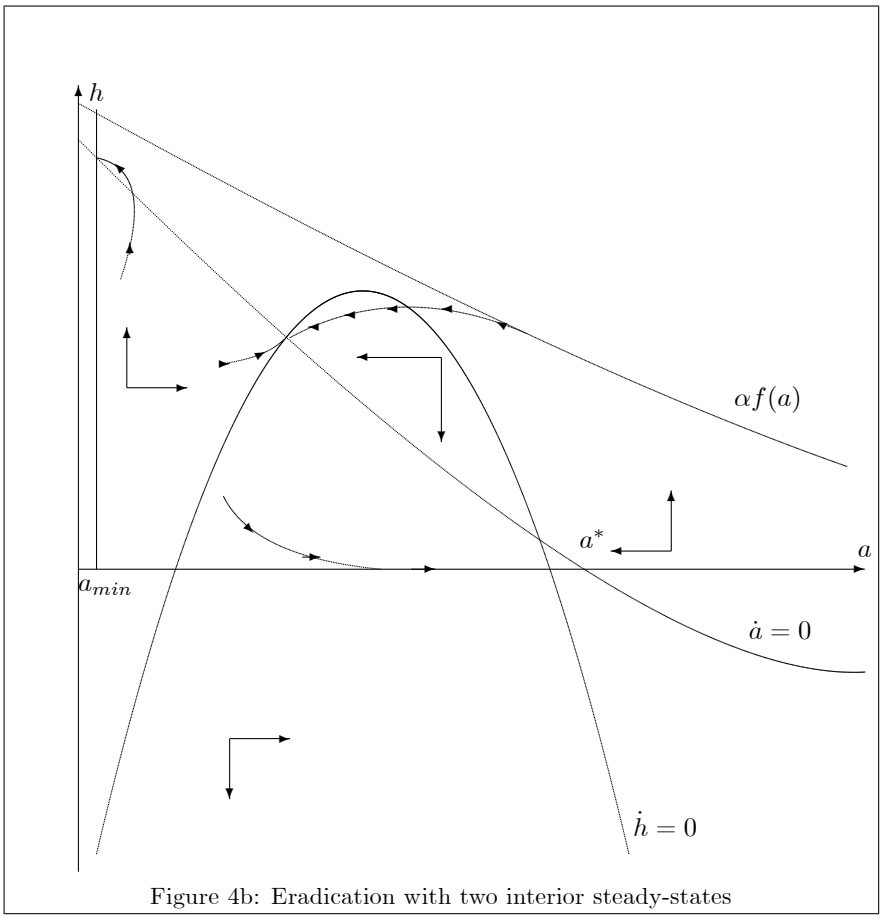


Figure 4b: Eradication with two interior steady-states

In the following lemma, we give a condition for the optimality of the eradication of the epidemic.

Lemma 9 *Suppose there exists a path that leads to the eradication of the epidemic. There is a threshold for the pure discount rate, denoted* $\bar{\rho}$*, such that for* $\rho < \bar{\rho}$*, this path is optimal.*

Proof. See the Appendix.

When eradication is achieved, the discounted utility of generation t, (i.e. $u(c_t)/\theta(a_t)$) is the highest possible since there is no prevention and the prevalence index is zero. This however implies that generations that have lived before the eradication have supported large reduction of their welfare due to the necessarily high levels of prevention devoted to the control of the epidemics. The path that yields to eradication is thus optimal if the pure discount rate is sufficiently small.

5.3 A hand solved example

Let us illustrate our dynamics with a simple example that encompass the two cases described in Section 2. Concerning the epidemics, we first assume that the over-mortality rate of infected persons is zero ($\gamma = 0$), which considerably simplifies the problem as $n'(a) = 0$. Moreover, to reduce the number of parameters, we assume $n(0) = 0$, which simply means that the death rate equals the birth rate. Both restrictions are consistent with H1 and H3. We also suppose that the impact of prevention on the infectious contacts is linear by writing: $\sigma(h) = \sigma_0 - h$, and assuming that $h \in [0, \sigma_0]$. The dynamics of the prevalence index is thus given by: $\dot{a}_t = a_t[\sigma_0 - h_t - \zeta(1 + a_t)]$ where $\zeta = \beta\pi$ for the model describing the HIV epidemic and $\zeta = \delta + \beta$ for the model describing an epidemics of flu. The endemic steady-state without control is thus $a^* = \sigma_0/\zeta - 1$, which is positive provided that $\sigma_0 > \zeta$. Concerning the economy, we assume $u(c) = \ln c$ and $f(a) = 1 - a$, which imposes $a \leq 1$.

The optimal control problem is thus:

$$\begin{aligned}
&\int_0^{\infty} e^{-\rho t} \ln c_t dt \\
&\qquad \dot{a}_t = \begin{cases} a_t \left[(\sigma_0 - \alpha - \zeta) + c_t - (\zeta - \alpha) a_t \right] & \text{if } a_t > a_{\min} \\ 0 & \text{if } a_t \leq a_{\min} \end{cases} \\
&s.t. \\
&\qquad \max\left\{0, \alpha - \alpha a_t - \sigma_0\right\} \leq c_t \leq \alpha - \alpha a_t \\
&\qquad a_0 \in \left(a_{\min}, \min\left\{a^*, 1\right\}\right), \text{ given.}
\end{aligned} \tag{25}$$

For $a_t > a_{\min}$, the optimal dynamics is given by the following system:

$$\begin{cases} \dot{c}_t = -c_t \left[\rho + (\zeta - \alpha) a_t \right] \\ \dot{a}_t = a_t \left[(\sigma_0 - \alpha - \zeta) + c_t - (\zeta - \alpha) a_t \right] \end{cases} \tag{26}$$

while the optimal prevention is given by $h_t = \alpha (1 - a_t) - c_t$.

Two different cases shall be considered $\alpha \leq \zeta$ and $\alpha > \zeta$. For $\alpha \leq \zeta$, the phase diagram is presented in Figure 5a. The diagram represents two lines: the isocline $\dot{a}_t = 0$ and the line given by $c_t = \alpha (1 - a_t)$, which represents the upper limit for c_t and the situation such that $h_t = 0$; the isocline $\dot{c}_t = 0$ does not appear here as it is in the negative quadrant. In this diagram, there is thus no interior steady-state, which is consistent with Lemma 7 as the level of development given by α is low. For an initial prevalence a_0, two trajectories are feasible, which converge to the corner steady-states, $a_{\min}$ and a^*. The first one features a temporary decline in consumption till the prevalence reaches $a_{\min}$; then, consumption jumps to its highest possible level. In this trajectory, prevention first increases with time and fall to zero when the epidemics is eradicated (Lemma 8). We see that such a trajectory is not likely to be optimal if $a_{\min} = 0$ as consumption would converge to zero (Lemma 4). The second

trajectory features a infinite decline in consumption as the epidemic converges to the endemic steady-state; the prevention is always zero in that case (Lemma 3). Among those two trajectories, one is optimal and Figure 5a is a nice illustration of Lemma 9. Compared to the second one, the trajectory that lead to eradication indeed imposes initially a lower consumption that is compensated later by a higher one. Such a strategy can be optimal only if the discount rate is low.

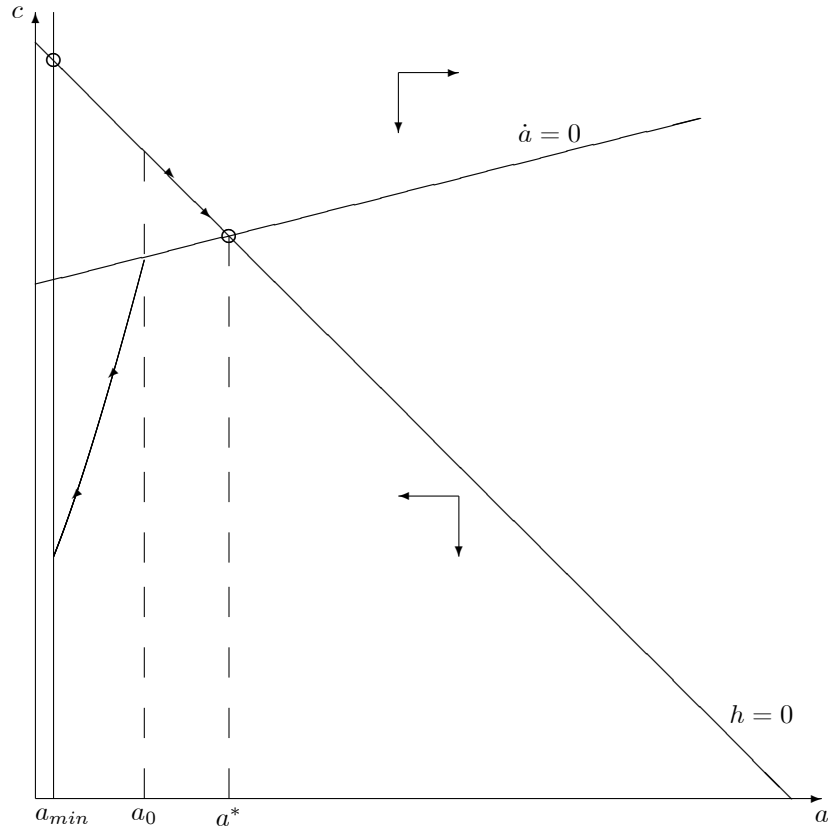


Figure 5a: Possible trajectories an initial prevalence index (case $\alpha < \zeta$)

For $\alpha > \zeta$ (i.e. if the level of development is sufficiently large), the phase

diagram analysis (Figure 5b[4]) now represents the two isoclines $\dot{a}_t = 0$ and $\dot{c}_t = 0$ and the line $c_t = \alpha(1 - a_t)$. We see there exists an interior steady-state, which is given by $\bar{a} = \rho/(\alpha - \zeta)$, $\bar{c} = \alpha + \zeta - \rho - \sigma_0$, and $\bar{h} = \sigma_0 - \zeta[1 + \rho/(\alpha - \zeta)]$, and is saddle-path (as established in Lemma 5). We represented the trajectories for two initial prevalence index a_0^l and a_0^h. The trajectory that converges to the interior steady-state features a consumption that decreases over time for the low initial prevalence and increase for the large one. There are, moreover, two additional trajectories that converge to $a_{\min}$ or a^*. The second corner steady-state can be reached only if the initial prevalence is larger than $\bar{a}$. As in the case with a low α, the choice between those trajectories depends on the pure discount

[4] The diagram represents three lines: the isoclines $\dot{c}_t = 0$ and $\dot{a}_t = 0$ and the $c_t = \alpha - \alpha a_t$, which represents the upper limit for c_t and the situation such that $h_t = 0$.

rate: the more the initial prevention, the higher the long run consumption.

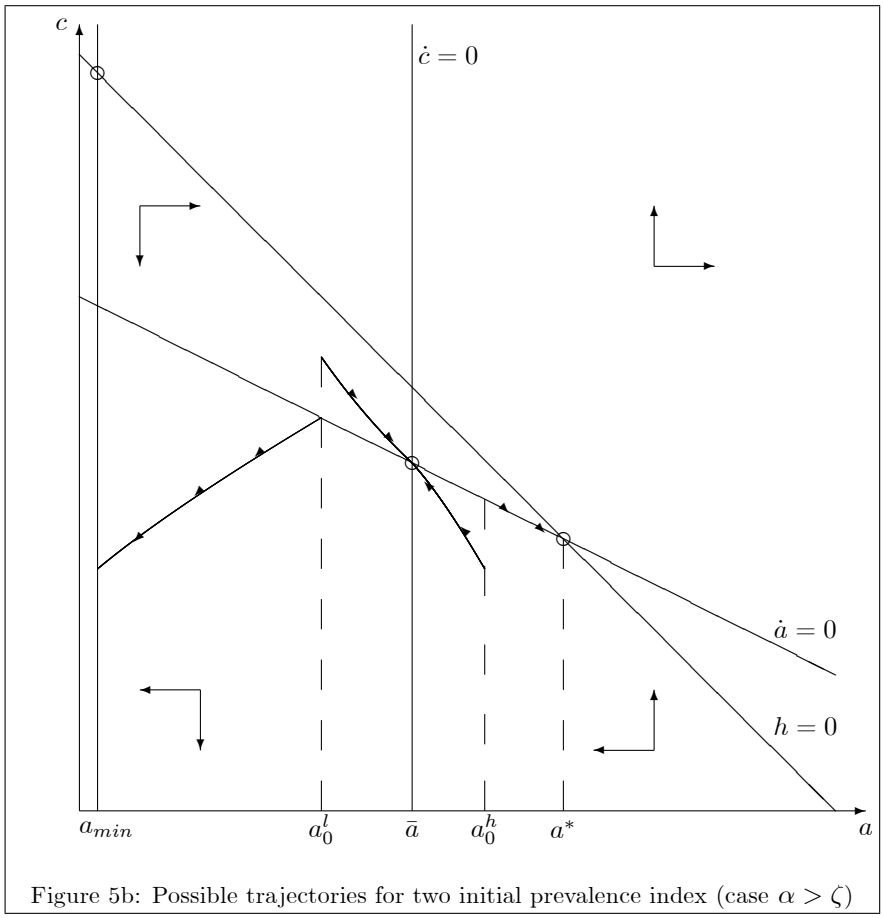


Figure 5b: Possible trajectories for two initial prevalence index (case $\alpha > \zeta$)

6 Conclusion

In this article, we have exhibited the relative role of resource constraints and individual preferences on the dynamic of the optimal prevention. Resources constraints are crucial for defining which paths are feasible, while preferences, and notably the discount rate, are used to characterize optimality. In the limit case of no pure discounting, as in Ramsey [39], "the-sooner-the-better" strategy is always optimal, provided there is an minimal threshold below which the epidemics is considered as eradicated. This strategy is initially very costly in term

of prevention but permits to set future expenses to zero.

Possible extensions of the present work may include the traditional decomposition of the population in three classes, a production economy and/or some endogenous growth factors. The size of the dynamic system would then increase making the analytical resolution of the model unlikely. More importantly, future researches should incorporate delay and age structure effects as they are crucial for most epidemic diseases. The next step is also to incorporate individual behaviors and decentralize the optimum.

Appendix

Proof of Lemma 1. The program (17) is a non autonomous problem with endogenous discounting. Let us rewrite it as an exogenous discounting problem using the virtual time method[5] notably used by Uzawa [41]. This method is based on a change of time scale, which can be applied as the state dynamics is autonomous. First define: $\delta_t = \int_0^t \theta(a_s)\, ds$; since δ is invertible from $\mathbb{R}_+$ to $\mathbb{R}_+$, it is possible to characterize $t = \tau(\delta)$; moreover, $d\delta = \theta(a_t)\, dt$. According to Assumption H3, $\theta(a) > 0$. The social planner's program (17) is then equivalent to:

$$\begin{aligned} &\max_{\tilde{h}_\delta} \int_0^\infty e^{-\delta} \frac{u\left(\alpha f(\tilde{a}_\delta) - \tilde{h}_\delta\right)}{\theta(\tilde{a}_\delta)} d\delta, \\ &s.t. \left| \begin{array}{l} \frac{d\tilde{a}_\delta}{d\delta} = \left| \begin{array}{ll} \frac{g\left(\tilde{h}_\delta, \tilde{a}_\delta\right)\tilde{a}_\delta}{\theta(\tilde{a}_\delta)} & \text{if } \tilde{a}_\delta > a_{\min} \\ 0 & \text{if } \tilde{a}_\delta \leq a_{\min}, \end{array} \right. \\ 0 \leq \tilde{h}_\delta \leq \alpha f(\tilde{a}_\delta) \text{ and } \tilde{a}_0 \in (a_{\min}, a^*) \text{ given.} \end{array} \right. \end{aligned} \tag{27}$$

where $\left(\tilde{a}_\delta, \tilde{h}_\delta\right) \equiv \left(a_{\tau(\delta)}, h_{\tau(\delta)}\right)$. Since there is no ambiguity, we will nevertheless keep the usual notations (a_t, h_t). Let us denote $a(., t_0, a_0, h(.))$ the unique solution of the state dynamics with initial condition a_0 at time t_0. Let

$$K = \left\{h(.) \text{ piecewise continuous such that } 0 \leq h(t) \leq \alpha f\left(a(., t_0, a_0, h(.))\right)\right\}. \tag{28}$$

The existence problem is then standard (Aniţa et al. [2]). Let us denote

$$J(h) = \int_0^\infty e^{-t} \frac{u\left(\alpha f\left(a^h\right) - h\right)}{\theta\left(a^h\right)} dt \tag{29}$$

[5] Francis and Kompas [22] propose a nice presentation of the method and of its conditions of application.

where a^h is the solution of

$$\dot{a} = \begin{cases} \frac{g(h,a)a}{\theta(a)} & \text{if } a > a_{\min} \\ 0 & \text{if } a \leq a_{\min}, \end{cases} \tag{30}$$

and $a(0) = a_0 \in (a_{\min}, a^*)$. Observe first that the application $a \longmapsto g(h,a)\, a/\theta(a)$ is locally lipschitz for $a \in (a_{\min}, a^*)$ and that, given the assumptions on g, for any admissible control,

$$\dot{a} \leq \frac{g(0, a_{\min})\, a_{\min}}{\theta(a_{\min})}. \tag{31}$$

We conclude that: $a_{\min} \leq a^h \leq a_0 e^{\frac{g(0,a_{\min})a_{\min}}{\theta(a_{\min})}} < a^*$. Moreover, according to the properties of u, f and θ, one has:

$$0 < \frac{u\left(\alpha f\left(a_t^h\right) - h_t\right)}{\theta\left(a_t^h\right)} < \frac{u\left(\alpha f\left(a_t^h\right)\right)}{\theta(a_{\min})} < \frac{u(\alpha f(a_{\min}))}{\theta(a_{\min})}, \tag{32}$$

from which we deduce that $\max_{h \in K} J(h)$ is positive and finite. We conclude the existence of an optimal control h^* and its relative optimal state a^{h^*} by using the fact that K is a closed and convex set.

Let $\mathcal{H}(h_t, a_t, \mu_t)$ be the current value of the Hamitonian associated with problem (27) that is given by:

$$\mathcal{H}(h_t, a_t, \mu_t) = \frac{u(\alpha f(a_t) - h_t)}{\theta(a_t)} + \mu_t \frac{g(h_t, a_t)\, a_t}{\theta(a_t)}, \tag{33}$$

and μ_t be the associated costate variable. Denoting by p and q the Lagrange multipliers associated with the inequality constraints, the Lagrangian in the current value is written as:

$$\mathcal{L}(h_t, a_t, \mu_t) = \frac{u(\alpha f(a_t) - h_t)}{\theta(a_t)} + \mu_t \frac{g(h_t, a_t)\, a_t}{\theta(a_t)} + p_t h_t + q_t(\alpha f(a_t) - h_t). \tag{34}$$

Let us consider an optimal pair (a, h). According to the previous remark and applying the maximum principle (Proposition 3.52, Grass et al [29]), there exists a continuous function μ, which is piecewise continuously differentiable, and two piecewise continuous multipliers (q_t, p_t) that satisfy:

$$-\frac{u'\left(\alpha f\left(a_t\right)-h_t\right)}{\theta\left(a_t\right)}+\mu_t\frac{g_1'\left(h_t,a_t\right)}{\theta\left(a_t\right)}a_t+p_t-q_t=0, \tag{35}$$

and

$$\begin{aligned}\dot{\mu}_t &= -\alpha f'\left(a_t\right)\frac{u'\left(\alpha f\left(a_t\right)-h_t\right)}{\theta\left(a_t\right)}-\mu_t\left(\frac{g_2'\left(h_t,a_t\right)a_t+g\left(h_t,a_t\right)}{\theta\left(a_t\right)}-1\right)\\ &\quad -q_t\alpha f'\left(a_t\right)+\frac{\theta'\left(a_t\right)}{\theta\left(a\right)}\mathcal{H}\left(h_t,a_t,\mu_t\right)\end{aligned} \tag{36}$$

Moreover, the complementary slackness conditions are:

$$\begin{aligned}&p_t h_t=0,\ p_t\geq 0,\ h_t\geq 0,\\ &q_t\left(\alpha f\left(a_t\right)-h_t\right)=0,\ q_t\geq 0,\ \alpha f\left(a_t\right)\geq h_t\end{aligned} \tag{37}$$

while the transversality condition (see Michel [35]) is:

$$\lim_{t\to\infty}e^{-t}\mathcal{H}\left(h_t,a_t,\mu_t\right)=0. \tag{38}$$

As μ is continuous, piecewise $\mathcal{C}^1$, the condition $\lim_{c\to 0}u'\left(c\right)=\infty$ in H4 prevents the first order condition $\partial\mathcal{L}/\partial h=0$ to be satisfied at $c=0$. Consequently $q_t=0$.

□

Proof of Lemma 2. For $h_t>0$, $p_t=0$. The expression for $\dot{h}_t$ is obtained by differentiating (35) with respect to t and rearranging using (36). For $h_t=0$, the derivation of the system is immediate. □

Proof of Lemma 3. In the neighborhood V of $(a^*, 0)$ the interior solution of h_t solves:

$$\dot{h}_t\Big|_{(a^*,0)} = \frac{-\left[g_2'(0,a^*)\,a^* - \theta(a^*)\right] + \left[-\alpha f'(a^*) + \frac{u(\alpha f(a^*))}{u'(\alpha f(a^*))}\frac{\theta'(a^*)}{\theta(a^*)}\right] g_1'(0,a^*)\,a^*}{-\theta(a^*)\left[\frac{u''(\alpha f(a^*))}{u'(\alpha f(a^*))} + \frac{g_{11}''(0,a^*)}{g_1'(0,a^*)}\right]}. \tag{39}$$

Thus, $\dot{h}_t\Big|_{(a^*,0)}$ can be positive or negative, depending on the exogenous values of (α, a^*). If $\dot{h}_t\Big|_{(a^*,0)} > 0$, the constraint $h_t \geq 0$ implies $h_t = 0$. If $\dot{h}_t\Big|_{(a^*,0)} < 0$, the trajectory reaches $h = 0$ in finite time t_0 and then $h_t = 0$ for all $t \geq t_0$. □

Proof of Lemma 4. In the neighborhood V of $(0, h)$, $c > 0$ and h is finite, the interior solution of (a_t, h_t) solves:

$$\dot{h}_t\Big|_{(0,h)} = \frac{1}{-\left[\frac{u''(\alpha f(0)-h)}{u'(\alpha f(0)-h)} + \frac{g_{11}''(h,0)}{g_1'(h,0)}\right]}, \tag{40}$$

which given H1 and H4 is positive. Due to $\dot{a}_t = g(h_t, a_t)\,a_t$, $a = 0$ is not achieved in finite time. Thus the constraint $h = \alpha f(0)$ is reached in finite time. According to lemma 2, this solution is not optimal. □

Proof of Lemma 5. As a preliminary, use (20) as an implicit equation to define $h = \eta(a)$, which, given Assumption H1, satisfies $\eta'(a) < 0$ and $\eta(a^*) = 0$, and replace it in (21) to define the function $\phi(a, \alpha)$ such that:

$$\begin{aligned}\phi(a,\alpha) &= -\left[g_2'(\eta(a), a)\,a - \theta(a)\right] \\ &\quad + \left[-\alpha f'(a) + \frac{u(\alpha f(a) - \eta(a))}{u'(\alpha f(a) - \eta(a))}\frac{\theta'(a)}{\theta(a)}\right] g_1'(\eta(a), a)\,a.\end{aligned} \tag{41}$$

Function $\phi(a, \alpha) \in C^2(D(\alpha) \times \mathbb{R}^+, \mathbb{R})$ where:

$$D(\alpha) = \left\{a \in \mathbb{R}^+ / \alpha f(a) - \eta(a) > 0\right\}. \tag{42}$$

Then, an interior steady-state is a $a \in (a^*, 0)$ such that $\phi(a, \alpha) = 0$.

To prove claim *i)*, use Lemma 2 and Assumption H4 to establish the following limit:

$$\lim_{\alpha \to 0} \phi(a, \alpha) = -\left[g_2'(0, a)\, a - \theta(a)\right] > 0, \tag{43}$$

and conclude using the continuity of $\phi(.,.)$ with respect to α.

Claim *ii)*: since $\alpha \longmapsto \alpha f(a)$ is strictly increasing and since $\lim_{\alpha \to \infty} \alpha f(a) = \infty$, there exists $\acute{\alpha}$ such that $D(\acute{\alpha}) = \mathbb{R}^+$. Suppose $\alpha \geq \acute{\alpha}$. Under Assumptions H1 and H4, $\phi(a, \alpha)$ decreases with α, $\phi(0, \alpha) = \theta(0) > 0$ and:

$$\begin{aligned} \phi(a^*, \alpha) &= -\left[g_2'(0, a^*)\, a^* - \theta(a^*)\right] \\ &\quad + \left[-\alpha f'(a^*) + \frac{u(\alpha f(a^*))}{u'(\alpha f(a^*))} \frac{\theta'(a^*)}{\theta(a^*)}\right] g_1'(0, a^*)\, a^*, \end{aligned} \tag{44}$$

is negative for sufficiently large α.

Claim *iii)*: the stability property is obtained by computing the trace of the Jacobian matrix of system (18) on the neighborhood of any interior steady-state. Since:

$$\left.\frac{\partial \Phi(h_t, a_t)}{\partial h_t}\right|_{\Phi(h_t, a_t) = 0} = 1 - \frac{g_2'(h_t a_t)\, a_t}{\theta(a_t)}, \tag{45}$$

$$\left.\frac{\partial g(h_t, a_t)\, a_t}{\partial a_t}\right|_{g(h_t, a_t) = 0} = \frac{g_2'(h_t a_t)\, a_t}{\theta(a_t)}, \tag{46}$$

the trace is consequently equal to 1. The steady-states are not stable. □

Proof of Lemma 6. Let us denote by $\left(\bar{a}, \bar{h}\right)$ the interior steady-state and by $(a^*, 0)$ the corner steady-state. To prove the lemma, we compute the intertemporal utilities for each paths in a particular case.

Suppose that $a_0 = \bar{a} < a^*$. Consider two paths that are candidates for optimality: the first one is given by: $h_t = \bar{h}$ and $a_t = \bar{a}$, and the second one is given by $h_t = 0$ and $a_t = \hat{a}(\bar{a}, t)$ (where according to notations used in the proof of Lemma 1 $\hat{a}(\bar{a}, t) = a(t, 0, \bar{a}, 0)$). The intertemporal utility yield by the first path is:

$$\int_0^\infty e^{-t} \frac{u\left(\alpha f(\bar{a}) - \bar{h}\right)}{\theta(\bar{a})} dt = \frac{u\left(\alpha f(\bar{a}) - \bar{h}\right)}{\theta(\bar{a})}. \tag{47}$$

The intertemporal utility yield by the second path is denoted $U(\bar{a})$ and satisfies:

$$U(\bar{a}) = \int_0^\infty e^{-t} \frac{u(\alpha f(\hat{a}(\bar{a}, t)))}{\theta(\hat{a}(\bar{a}, t))} dt. \tag{48}$$

As $g(0, a)\, a$ is $\mathcal{C}^2$ according to Assumption H1, $\hat{a}(\bar{a}, t)$ has a derivative according to $\bar{a}$ and

$$\begin{aligned} & \frac{d}{d\bar{a}} \left(\frac{u(\alpha f(\hat{a}(\bar{a}, t)))}{\theta(\hat{a}(\bar{a}, t))} \right) \\ = \quad & \frac{d\hat{a}(\bar{a}, t)}{d\bar{a}} \left(\frac{\alpha f'(\hat{a}(\bar{a}, t))\, u'(\alpha f(\hat{a}(\bar{a}, t)))}{\theta(\hat{a}(\bar{a}, t))} - \frac{\theta'(\hat{a}(\bar{a}, t))\, u(\alpha f(\hat{a}(\bar{a}, t)))}{\theta^2(\hat{a}(\bar{a}, t))} \right) < 0 \end{aligned}$$

Since $U'(\bar{a}) < 0$, conclude that $U(\bar{a}) > U(a^*)$. Hence, the first path is not optimal if:

$$\frac{u\left(\alpha f(\bar{a}) - \bar{h}\right)}{\theta(\bar{a})} < \frac{u(\alpha f(a^*))}{\theta(a^*)}, \tag{49}$$

It is easy to check that the last inequality may not be satisfied for a pair $\left(\bar{h}, \bar{a}\right)$ that satisfies: (20), (21). For instance use the following functions:

$$g(h, a) = e^{-h} - [\beta\pi(1 + a) + \gamma + \delta], \tag{50}$$

$$f(a) = \frac{1 + \lambda a}{1 + a}, \tag{51}$$

$$u(c) = \sigma c^{\frac{1}{\sigma}}, \tag{52}$$

and the following parameters: $\alpha = 1.5$, $\mu = 0.01$, $\gamma = 0.01$, $\pi = 0.9$, $\beta = 0.5$, $\rho = 0.5$, $\lambda = 0.6$, $\sigma = 1/0.69$. $\square$

Proof of Lemma 7. We consider successively the $\dot{a} = 0$ locus and the $\dot{h} = 0$ locus.

Claim *i)*. The $\dot{a} = 0$ locus for all $a > 0$ is given by the implicit function $g(h, a) = 0$. Given Assumption H1, the locus is downward slopping in the plane (a, h), and is such that $\dot{a} > 0$ below the locus and $\dot{a} < 0$ above.

Claim *ii)*. Using the definition of $\dot{h}$ in (18) and Assumption H5, the $\dot{h} = 0$ locus is given by the function $\psi(h, a, \alpha) = 0$ where:

$$\begin{aligned}\psi(h,a,\alpha) &= \left[-\alpha f'(a)\frac{u''(\alpha f(a)-h)}{u'(\alpha f(a)-h)}+\frac{\theta'(a)}{\theta(a)}\right]g(h,a)\,a \\ &\quad+\left[-\alpha f'(a)+\frac{u(\alpha f(a)-h)}{u'(\alpha f(a)-h)}\frac{\theta'(a)}{\theta(a)}\right]g_1'(h,a)\,a \\ &\quad-g_2'(h,a)\,a+\theta(a), \end{aligned} \tag{53}$$

$$\begin{aligned}\psi_1'(h,a,\alpha) &= \left[-\alpha f'(a)\frac{u^{(3)}(\alpha f(a)-h)\,u'(\alpha f(a)-h)+\left(u^{(2)}(\alpha f(a)-h)\right)^2}{\left(u'(\alpha f(a)-h)\right)^2}\right]g(h,a)\,a \\ &\quad+\frac{\theta'(a)}{\theta(a)}\left[\frac{u(\alpha f(a)-h)\,u^{(2)}(\alpha f(a)-h)}{\left(u'(\alpha f(a)-h)\right)^2}\right]g_1'(h,a)\,a \\ &\quad+\left[-\alpha f'(a)+\frac{u(\alpha f(a)-h)}{u'(\alpha f(a)-h)}\frac{\theta'(a)}{\theta(a)}\right]g_{11}'(h,a)\,a, \end{aligned} \tag{54}$$

which, using Assumptions H1, H4 and H5, satisfies $\psi_1'(h, a, \alpha) > 0$ for all (h, a) below the $\dot{a} = 0$ locus (i.e. for $g(h, a) > 0$). It is then possible to define the function χ such that $h = \chi(a, \alpha)$, which satisfies: $\lim_{a \to 0} \chi(a, \alpha) = -\infty$.

Using (18), compute, on the neighborhood of $(0, 0)$, the interior solution of

h_t to obtain:

$$\dot{h}_t\Big|_{(0,0)} = -\left[\frac{u''(\alpha f(0))}{u'(\alpha f(0))} + \frac{g''_{11}(0,0)}{g'_1(0,0)}\right]^{-1}, \tag{55}$$

which is positive given the convexity of g (Assumption H5). Conclude that $\dot{h} > 0$ above the locus and that $\dot{h} < 0$ below.

Claim *iii)*. Define $R = \left\{(a,h) \in \mathbb{R}^2,\ g(h,a) > 0\right\}$. Using Assumptions H1, H4 and H5, we have $\psi'_3(h,a,\alpha) < 0$ for all (h,a) such that $g(h,a) > 0$. As $\psi'_1(h,a,\alpha) > 0$ on R, the claim is proved. □

Proof of Lemma 8. In the neighborhood V of $(a_{\min}, h)$, $c > 0$ and h is finite, the interior solution of (a_t, h_t) solves:

$$\dot{h}_t\Big|_{(a_{\min},h)} \simeq \dot{h}_t\Big|_{(0,h)} = \frac{1}{-\left[\frac{u''(\alpha f(0)-h)}{u'(\alpha f(0)-h)} + \frac{g''_{11}(h,0)}{g'_1(h,0)}\right]} > 0, \tag{56}$$

which given H1 and H4 is positive. As the optimal trajectory (a,h) reaches $a_{\min}$ in finite time t_1, for $t > t_1$, $a(t) = a_{\min}$, which does not depends on h.Thus, for $t > t_1$, $h = 0$. □

Proof of Lemma 9. A path that leads to eradication in a finite time T can be characterized by $\{h^e_t, a^e_t\}$ for all $t < T$ and by $h_t = 0$ and $a_t = a_{\min}$ for all $t \geq 0$. This yields the following intertemporal utility:

$$\int_0^T e^{-t}\frac{u(\alpha f(a^e_t) - h^e_t)}{\theta(a^e_t)}dt + e^{-T}\frac{u(\alpha f(a_{\min}))}{\theta(a_{\min})}. \tag{57}$$

For $a_{\min} > 0$ but small, then $\theta(a_{\min})$ is close to zero, and therefore the intertemporal utility is very large while intertemporal utilities yield by any other path is smaller. We conclude by continuity. □

References

[1] F. E. Alvarez, D. Argente, and F. Lippi, A Simple Planning Problem for COVID-19 Lockdown, NBER Working Paper No. 26981, 2020

[2] S. Aniţa, V. Arnăutu and V. Capasso, An Introduction to Optimal Control Problems in Life Sciences and Economics, Birkhäuser, 2011.

[3] C. Avery, W. Bossert, A. Clark, G. Ellison, and S. Fisher Ellison, Policy implications of models of the spread of Coronavirus: Perspectives and opportunities for economists, NBER Working Paper No. 27007, 2020

[4] S. Barrett and M. Hoel, Optimal disease eradication, Environment and Development Economics 12, (2007) 627-652.

[5] H. Benhcke, Optimal control of deterministic epidemics, Optimal Control Applications and Methods 21, (2000) 269-285.

[6] C. Bell and H. Gersbach, Growth and enduring epidemic diseases, Journal of Economic Dynamics and Control 37 (2013) 2083-2103.

[7] D. Bernoulli, Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir, Memoire de mathématiques et de physique, Académie Royale des Sciences (1760) 1-45.

[8] Z. A. Bethune and A. Korinek, Covid-19 infection externalities: Trading off lives vs. livelihoods, NBER Working Paper No. 27009, 2020.

[9] D. E. Bloom and D. Canning, Epidemics and Economics, PGDA working paper No. 9, Harvard School of Public Health.

[10] R. Boucekkine, B. Diene and T. Azomahou, Growth economics of epidemics: A review of the theory, Mathematical Population Studies 15 (2008), 1-26.

[11] R. Boucekkine, B. Diene and T. Azomahou, HIV/AIDS and development: A reappraisal of the productivity and factor accumulation effects, American Economic Review 106, (2016) 472-477.

[12] R. Boucekkine and J.-P. Laffargue, On the distributional consequences of epidemics, Journal of Economic Dynamics and Control 34 (2010) 231-245.

[13] R. Boucekkine, B. Martínez and R. J. Ruiz-Tamarit, Optimal Population Growth as an Endogenous Discounting Problem: The Ramsey Case In: Lecture Notes in Economics and Mathematical Systems, pp. 321-347, Springer, Cham, 2018.

[14] S. Bosi and D. Desmarchelier, Biodiversity, infectious diseases, and the dilution effect. Environmental Modeling & Assessment (2020). https://doi.org/10.1007/s10666-020-09688-9.

[15] D. Cass, Optimum growth in an aggregative model of capital accumulation, Review of Economic Studies 32 (1965), 233-240.

[16] C. Castilho, Optimal control of an epidemic through educational campaigns, Electronic Journal of Differential Equation 125, (2006) 1-11.

[17] P. Corrigan, G. Glomm and F. Méndez, AIDS crisis and growth, Journal of Development Economics 77, (2005) 107-124.

[18] P. Dasgupta, On the concept of optimal population, Review of Economic Studies 36, (1969) 295-318.

[19] D. Delfino and P. J. Simmons, Positive and normative issues of economic growth with infectious disease, working paper, University of York, 2000.

[20] G. Feichtinger, V. M. Veliov and T. Tsachev, Maximum principle for age and duration structured systems: A tool for optimal prevention and treatment of HIV, Mathematical Population Studies 11, (2004) 3-28.

[21] P. J. Francis, Optimal tax/subsidy combinations for the flu season, Journal of Economic Dynamics and Control 28, (2004) 2037-2054.

[22] J. Francis and T. Kompas, Uzawa's transformation and optimal control problems with variable rates of time preference, working paper, 1998.

[23] M. Gersovitz and J. S. Hammer, The economical control of infectious diseases, Economic Journal 114, (2004) 1-27.

[24] M. Gersovitz and J. S. Hammer, Tax/subsidy policies toward vector-borne infectious diseases, Journal of Public Economics 89, (2005) 647-674.

[25] P.-Y. Geoffard and T. Philipson, Rational epidemics and their public control, International Economic Review 37, (1996) 603-624.

[26] P.-Y. Geoffard and T. Philipson, Disease eradication: Private versus public vaccination, American Economic Review 87, (1997) 222-230.

[27] A. Goenka and L. Lin, Infectious diseases and endogenous fluctuations, Economic Theory 50, (2012) 125-149.

[28] A. Goenka, L. Liu and M. Nguyen, Infectious diseases and economic growth. Journal of Mathematical Economics 50, (2014) 34-53.

[29] D. Grass, J. P. Caulkins, G. Feichtinger, G. Tragler and D. A. Behrens, Optimal Control of Nonlinear Processes With Applications in Drugs, Corruption and Terror, 2008 Springer-Verlag Berlin Heidelberg.

[30] H. W. Hethcote and J. A. Yorke, Gonorrhea Transmission Dynamics and Control, Lecture Notes in Biomathematics 56, 1984. Springer-Verlag Berlin.

[31] W. O. Kermarck and A. G. Mac Kendrick, A contribution to the mathematical theory of epidemics, Proc. Royal Soc. London 115, (1927) 700-721.

[32] M. Kremer, Integrating behavioral choice into epidemiological models of AIDS, Quarterly Journal of Economics 111, (1996) 549-573.

[33] T. C. Koopmans, On the concept of optimal economic growth, in "The Econometric Approach to Developpement Planning", North Holland, Amsterdam, 1965.

[34] R. M. May and R. M. Anderson, The transmission dynamics of Human Immunodeficiency Virus, in S. A. Levin, T. G. Hallam and L. J. Gross, Eds., Applied Mathematical Ecology, Springer-Verlag, Berlin, 1989.

[35] P. Michel, Some clarifications on the transversality condition, Econometrica 58, (1990) 705-723

[36] A. Momota, K. Tabata and K. Futagami, Infectious disease and preventive behavior in an overlapping generations model, Journal of Economic Dynamics and Control 29, (2005) 1673-1700.

[37] D. Morris, F., Rossine, J., Plotkin and S. Levin, Optimal, near-optimal, and robust epidemic control. Working paper. 2020. https://osf.io/rq5ct/

[38] R. Naresh, A. Tripathi and S. Omar, Modelling the spread of AIDS epidemic with vertical transmission, Applied Mathematics and Computation 178, (2006) 262-272.

[39] F. P. Ramsey, A mathematical theory of savings, Economic Journal 38, (1928) 543-559.

[40] S. P. Sethi, Optimal control of some simple deterministic epidemic models, The Journal of the Operational Research Society 29, (1978) 129-136.

[41] H. Uzawa, Time preference, the consumption function, and optimal asset holding, in J. N. Wolfe (ed.), Capital and Growth Papers in Honor of Sir John Hicks, The University of Edimburg Press, 485-504, 1968.

[42] P. van den Driessche, Reproduction numbers of infectious disease models. Infectious Disease Modelling 2(3), (2017) 288-303.

[43] K. H. Wickwire, Mathematical models for the control of pests and infectious diseases, A survey. Theoretical Population Biology 11, (1977) 182-238.

[44] A. Young, The gift of the dying: the tragedy of AIDS and the welfare of future African generations, Quarterly Journal of Economics 120, (2005) 423-466.