Infectious Diseases, Optimal Health Expenditures and Growth *

Aditya Goenka[†] (National University of Singapore)
Lin Liu[‡] (University of Rochester)
Manh-Hung Nguyen[§] (Toulouse School of Economics)

October 1, 2010

Preliminary Version

Abstract: This paper develops a general framework to study the economic impact of infectious diseases by integrating epidemiological dynamics into a continuous time neo-classical growth model. There is a two way interaction between the economy and the disease: the incidence of the disease affects labor supply and investment in health capital can affect the incidence and recuperation from the disease. Thus, both the disease incidence and the income levels are endogenous. The dynamics of the disease make the control problem non-convex and thus, a new existence theorem is given. We fully characterize the local dynamics of the model. A disease free steady state always exists, but it can become unstable and there can be multiplicity of steady states. If the disease is endemic, the optimal health expenditure can be positive or zero depending on the parameters of the model. We show there can be an endogenous positive relationship between output and health expenditures.

Keywords: Epidemiology; Infectious Diseases; Health Expenditures; Economic Growth; Bifurcation; Existence of equilibrium.

JEL Classification: C61, D51, E13, O41, E32.

1 Introduction

This paper intends to provide a canonical theoretical framework modeling the joint determination of both income and disease prevalence by integrating epidemiological dynamics into a continuous time neo-classical growth model. It allows us to address the issue of what is the optimal investment in health from a social planner's point of view when there is a two way interaction between the disease transmission and the economy: the disease transmission affects the labor force and thus, economic outcomes, while economic choices on investment in health expenditures affect the disease transmission. It sheds light on explaining two important empirical facts. One is the correlation between economic variables and disease incidence. The literature tries to quantify the impact of infectious diseases on the economy, which mainly focus on solving the endogeneity issue of disease prevalence (see Acemoglu and Johnson (2007), Ashraf, *et al* (2009), Bell, *et al* (2003), Bleakley (2007), Bloom, *et al* (2009), Young (2005)), but the results are rather mixed. We show the reduced form estimation by assuming linear relationship is not well justified as non-linearity is an important characteristic of models associated with the disease transmission, and

*We would like to thank Murali Agastya, Michele Boldrin, Russell Cooper, Atsushi Kajii, Takashi Kamihigashi, Cuong Le Van, Francois Salanie, Karl Shell and seminar participants at the 2007 Asian General Equilibrium Theory Workshop, Singapore; 2008 European General Equilibrium Workshop, Paestum; FEMES 2008; NUS Macro Brown Bag Workshop; University of Paris I; University of Cagliari; City University London; SWIM Auckland 2010; and Kyoto Institute of Economic Research for helpful comments comments and suggestions. The usual disclaimer applies.

†Correspondence to A. Goenka, Department of Economics, National University of Singapore, AS2, Level 6, 1 Arts Link, Singapore 117570, Email: goenka@nus.edu.sg

‡Department of Economics, Harkness Hall, University of Rochester, Rochester, NY 14627, USA. Email: linliu@rochester.edu

§LERNA-INRA, Toulouse School of Economics, Manufacture des Tabacs, 21 Allée de Brienne, 31000 Toulouse, France. EMail: mhnguyen@toulouse.inra.fr

nonlinearity in disease transmission can become a source of non-linearities in economic outcomes. The other empirical fact is the rising health expenditure share along with income growth in U.S. and other OECD countries. We show the rising health expenditure is economically justified in a very standard economic model without resorting to non-standard preferences (life extension in utility, see Hall and Jones (2007)), complicated institutional or insurance structures, technological progress, etc. We find that variations in the discount rate (which could be interpreted as change in longevities) and birth rate can lead to both rising health expenditure and income growth. Observationally it may appear that expenditures on health are a luxury good, but the mechanism is through changes in marginal productivities and not preferences.

This paper is related to some of the theoretical literature on the optimal control of diseases which develops models to evaluate welfare gains of disease control and eradication (e.g. Barrett and Hoel (2004), d'Albis and Augeraud-Véron (2008), Geoffard and Philipson (1997), Gersovitz and Hammer (2004), Goenka and Liu (2010)). The difference between this paper and other literature are: first, most of other papers address optimal private health expenditure and under-investment problem due to externality inherent in disease controlling problem. In this paper we would like to know what is the best that society can do in controlling the disease transmission by taking into the externality. Thus, we look at the social planning problem (see Hall and Jones (2007) which takes a similar approach for non-infectious diseases). We show a steady state with disease prevalence and zero health expenditure could even be optimal as it depends on the relative magnitude of marginal product of physical capital investment and health expenditure. Second, these papers model either disease dynamics or the accumulation of capital, but not both. In modeling the interaction between infectious diseases and the macroeconomy, we expect savings behavior to change in response to changes in disease incidence. Thus, it is important to incorporate this into the dynamic model to be able to correctly assess the impact of diseases on capital accumulation and hence, growth and income. As the prevalence of diseases is affected by health expenditure, which is an additional decision to the investment and consumption decision, this has to modeled as well. Without modeling both physical and health capital accumulation and the evolution of diseases at the same time, it is difficult to understand the optimal response to disease incidence[1]. As the literature does not model both disease dynamics and capital accumulation explicitly, the existing models are like a black-box: the very details of disease transmissions and the capital accumulation process that are going to be crucial in understanding their effects and for the formulation of public policy, are obscured. We find that even when the strong assumption of log-linear preferences is made (which is usually invoked to justify fixed savings behavior) there can be non-linear and non-monotonic changes in steady state outcomes.

In order to model the disease transmission explicitly we integrate the epidemiology literature (see Anderson and May (1991), Hethcote (2000), Hethcote (2009)) into dynamic economic analysis. In this paper we examine the effect of the canonical epidemiological structure for recurring diseases - SIS dynamics - in a continuous time growth model. SIS dynamics characterize diseases where upon recovery from the disease there is no subsequent immunity to the disease. This covers many major infectious diseases such as flu, tuberculosis, malaria, dengue, schistosomiasis, trypanosomiasis (human sleeping sickness), typhoid, meningitis, pneumonia, diarrhoea, acute haemorrhagic conjunctivitis, strep throat and sexually transmitted diseases (STD) such as gonorrhea, syphilis, etc (see Anderson and May (1991)). As mentioned above, in our model we endogenize the epidemiological parameters by making them dependent on health capital: increases in health capital reduce the infectivity rate and increases the recovery rate from the disease.

In analyzing optimal behavior there are two sources of difficulties. First, the disease dynamics are non-convex reflecting the externalities inherent in disease transmission. This implies that Arrow-Mangasarian sufficiency conditions in optimal control problems may not hold.[2] In this paper, we address the issue directly. We show that a solution to the optimal control problem does indeed exist. The conditions we use are weaker than those in the literature (Chichilinksy (1981), d'Albis et al. (2008), Romer (1986)). Second, the system dynamics is of high dimension. Thus, we can only examine the local stability properties of the system. We show that there is a trans-critical bifurcation of the disease free steady state: As the net birth rate falls the disease free steady state ceases to be locally stable. A steady state where disease is endemic emerges and becomes locally stable. In Goenka and Liu (2010) there is a one way interaction,

[1]The model in Delfino and Simmons (2000) is an exception but it also uses fixed savings behavior and thus does not permit welfare comparisons. It does not include health capital.

[2]Gersovitz and Hammer (2004) rely on simulations to argue that the first order conditions are in fact sufficient, while d'Albis and Augeraud-Véron (2008) assume that the disease dynamics are convex so that the problem does not arise in the first place.

the disease affects the labor force participation, but not vice versa. The dynamics are two dimensional which allowed analysis of the global dynamics.

We find that there are multiple steady states: a disease free steady state always exists. It is unique when the net birth rate is high. The basic intuition is that individuals enter the economy at a faster rate than they contract the disease so that eventually it dies out. As the net birth rate decreases (holding the discount rate constant), there can be a steady state where the disease is endemic but there is no expenditure on health. Here due to the relatively high birth rate, the marginal returns to investing physical capital always dominate that of health capital: The high birth rates imply that there is low per capital physical capital on the one hand and the cost of an additional worker falling ill is low. As the net birth rate decreases further the rate of return dominance ceases to hold and in the endemic steady state there are positive health expenditures. Further decreases in the net birth rate increase health expenditures. The intuition is that it becomes increasing costly for society if an additional worker falls ill, and thus, social health expenditures increase. The negative relationship between birth rates and income is well known (see for example Brander and Dowrick (1994)). We also characterize the optimal solutions for combinations of the discount rate (which indexes longetivity) and the net birth rate, and thus are able to study how the optimal health expenditures change as either is varied. We show that in an endemic steady state it is socially optimal not to invest in health capital if the birth rate and the discount rate is sufficiently high, while there are positive health expenditures if these are low.

In this paper we abstract away from disease related mortality. This is a significant assumption as it shuts down the demographic interaction. This assumption is made for two reasons. First, several SIS diseases have low mortality so there is no significant loss by making this assumption. These include several strains of influenza, meningitis, STDs (syphilis, gonorrhea), dengue, conjunctivitis, strep throat, etc. Secondly, from an economic modeling point of view we can use the standard discounted utility framework with an exogenous discount rate if mortality is exogenous. In the paper we also consider the effect of changes in the discount rate on the variables of interest. As has been noted in the literature, increase in longetivity reduces discounting, and thus captures some effects of change in mortality.

The paper is organized as follows: Section 2 describes the model and in Section 3 we establish existence of an optimal solution. Section 4 studies the steady state equilibria, and Section 5 contains the stability and bifurcation analysis of how the nature of the equilibria change as parameters are varied. Section 6 does comparative statics of steady states while varying discount and birth rates, and the last section concludes.

2 The Model

In this paper we study the canonical deterministic SIS model which divides the population into two classes: susceptible (S) and infective (I) (see Figure 1). Individuals are born healthy but susceptible can contract the disease - becoming infected and capable of transmitting the disease to other, i.e. infective. Upon recovery, individuals do not have any disease conferred immunity, and move back to the class of susceptible individuals. Thus, there is horizontal incidence of the disease. This model is applicable to infectious diseases which are absent of immunity or which mutate rapidly so that people will be susceptible to the newly mutated strains of the disease even if they have immunity to the old ones. As there is no disease conferred immunity, there typically do not exist robust vaccines for diseases with SIS dynamics. There is homogeneous mixing so that the likelihood of any individual contracting the disease is the same, irrespective of age. Let $S(t)$ be the number of susceptibles at time t, $I(t)$ be the number of infectives and $N(t)$ be the total population size. The fractions of individuals in the susceptible and infected class are $s(t) = S(t)/N(t)$ and $i(t) = I(t)/N(t)$, respectively. Let α be the average number of adequate contacts of a person to catch the disease per unit time or the contact rate. Then, the number of new cases per unit of time is $(\alpha I/N)S$. This is the standard model (also known as frequency dependent) used in the epidemiology literature (Hethcote (2009)). The basic idea is that the pattern of human interaction is relatively stable and what is important is the *fraction of infected people* rather than the total number: If the population increases the pattern of interaction is going to be invariant. Thus, only the proportion of infectives and not the total size is relevant for the spread of the disease. The parameter α is the key parameter and reflects two different aspects of disease transmission: the biological infectivity of the disease and the pattern of social interaction. Changes in either will change α. The recovery of individuals is governed by the parameter γ and the total number of individuals who recover from the disease at time t is γI.

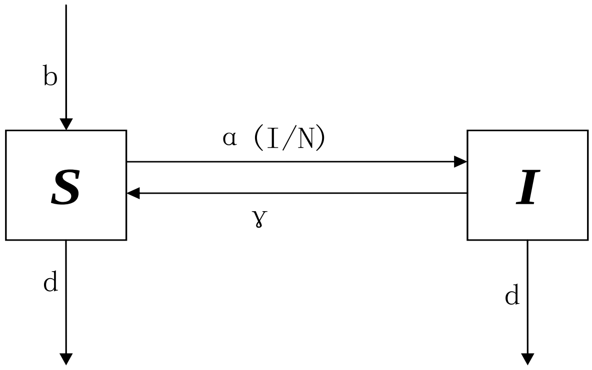

Figure 1: The transfer diagram for the SIS epidemiology model

Many epidemiology models assume total population size to be constant when the period of interest is short, i.e. less than a year, or when natural births and deaths and immigration and emigration balance each other. As we are interested in long run effects, we assume that there is a constant birth rate b, and a constant (natural) death rate d.

Assumption 1 *The birth rate b and death rate d are positive constant scalars with $b \geq d$.*

Thus, the SIS epidemiology we have described so far is the same in the epidemiology literature and given by the following system of differential equations (Hethcote, 2009):

$$\begin{aligned} &dS/dt = bN - dS - \alpha SI/N + \gamma I \\ &dI/dt = \alpha SI/N - (\gamma + d)I \\ &dN/dt = (b - d)N \\ &S, I, N \geq 0 \, \forall t; S_0, I_0, N_0 > 0 \, \text{given with } N_0 = S_0 + I_0. \end{aligned}$$

Since $N(t) = S(t) + I(t)$, we can simplify the model in terms of the susceptible fraction s_t:

$$\dot{s_t} = (1 - s_t)(b + \gamma - \alpha s_t) \tag{1}$$

with the total population growing at the rate $b - d$. In this pure epidemiology model, there are two steady state equilibria ($\dot{s_t} = 0$) given by: $s_1^* = 1$ and $s_2^* = \frac{b+\gamma}{\alpha}$. We notice s_1^* (the disease-free steady state) exists for all parameter values while s_2^* (the endemic steady state) exists only when $\frac{b+\gamma}{\alpha} < 1$. Linearizing the one-dimensional system around its equilibria and the Jacobians are $Ds|_{s_1^*} = \alpha - \gamma - b$ and $Ds|_{s_2^*} = \gamma + b - \alpha$. Thus, if $b > \alpha - \gamma$ the system only has one disease-free steady state, which is stable, and if $b < \alpha - \gamma$ the system has one stable endemic steady state and one unstable disease-free steady state (refer to Figure 2). Hence, there is a bifurcation point, i.e. $b = \alpha - \gamma$, where the existence and stability of the equilibria changes. Equation (1) can be solved analytically [3] and these dynamics are global.

In this paper, we endogenize the parameters α and γ in a two sector growth model. The key idea is that the epidemiology parameters, α, γ, are not immutable constants but are affected by (public) health expenditures. As there is an externality in the transmission of infectious diseases, there may be underspending on private health expenditures, and in any case due to the contagion effects, private expenditures may not be sufficient to control incidence of the disease[4]. We want to look at the best possible outcome which will increase social welfare. Thus, we study the social planner's problem and in this paper concentrate on public health expenditures (see the discussion in Hall and Jones (2007) who also concentrate on the planning problem). In this way, the externalities associated with the transmission of the infectious diseases can be taken into account in the optimal allocation of health expenditures.

[3] Since $\dot{s_t} = (1 - s_t)(b + \gamma - \alpha s_t)$, with initial value $s_0 < 1$, is a Bernoulli differential equation, we can solve it and get an explicit unique solution: $s_t = 1 - \frac{e^{[\alpha-(\gamma+b)]t}}{\frac{\alpha}{\alpha-(\gamma+b)} e^{[\alpha-(\gamma+b)]t} + \frac{1}{1-s_0} - \frac{\alpha}{\alpha-\gamma-b}}$ (for $b \neq \alpha - \gamma$) and $s_t = 1 - \frac{1}{\alpha t + \frac{1}{1-s_0}}$ (for $b = \alpha - \gamma$).

[4] The literature on rational epidemics as in Geoffard and Philipson (1996), Kremer (1996), Philipson (2000) looks at changes in epidemiology parameters due to changes in individual choices. Individual choice is more applicable to disease which transmit by one-to-one contact, such as STDs.

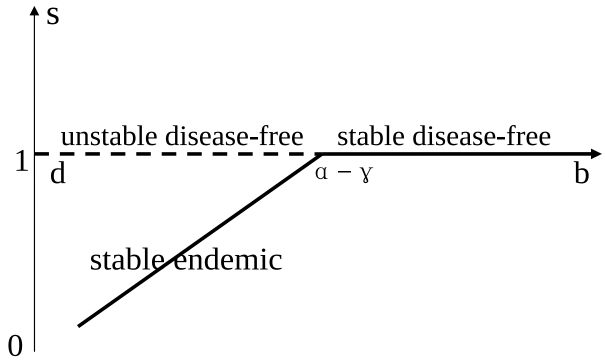

Figure 2: The bifurcation diagram for SIS model

There is a population of size $N(t)$ growing over time at the rate of $b - d$. Each individual's labor is indivisible: We assume infected people cannot work and labor force consists only of healthy people with labor supplied inelastically.[5] Thus, in time period t the labor supply is $L(t) = N(t) - I(t) = S(t)$ and hence, $L(t)$ inherits the dynamics of $S(t)$, that is,

$$\dot{l_t} = (1 - l_t)(b + \gamma - \alpha l_t),$$

in terms of the fraction of effective labor $l_t = L_t/N_t$. We allow for health capital to affect the epidemiology parameters, hence, allowing for a two-way interaction between the economy and the infectious diseases. We endogenize them by treating the contact rate and recovery rate as functions of health capital per capita h_t. This takes into account intervention to control the transmission of infectious diseases through their preventive or therapeutic actions. When health capital is higher people are less likely to get infected and more likely to recover from the diseases. We assume that the marginal effect diminishes as health capital increases. We further assume that the marginal effect is finite as health capital approaches zero: a small public health expenditure will not have a discontinuous effect on disease transmission.

Assumption 2 *The epidemiological parameter functions $\alpha(h_t)$ and $\gamma(h_t)$: $\Re_+ \to \Re_+$ satisfy:*

1. *$\alpha(h_t)$ is a C^∞ function with $\alpha'(h_t) \leq 0$, $\alpha''(h_t) \geq 0$, $\lim_{h_t \to 0} |\alpha'(h_t)| < \infty$, $\lim_{h_t \to \infty} \alpha'(h_t) \to 0$, $\alpha(h_t) \to \overline{\alpha}$ as $h_t \to 0$ and $\alpha(h_t) \to \underline{\alpha}$ as $h_t \to +\infty$;*
2. *$\gamma(h_t)$ is a C^∞ function with $\gamma'(h_t) \geq 0$, $\gamma''(h_t) \leq 0$, $\lim_{h_t \to 0} \gamma'(h_t) < \infty$, $\lim_{h_t \to \infty} \gamma'(h_t) \to 0$, $\gamma(h_t) \to \underline{\gamma}$ as $h_t \to 0$ and $\gamma(h_t) \to \overline{\gamma}$ as $h_t \to +\infty$.* [6]

We assume physical goods and health are generated by different production functions. The output is produced using capital and labor, and is either consumed, invested into physical capital or spent in health expenditure. The health capital is produced only by health expenditure.[7] For simplicity, we assume the depreciation rates of two capitals are the same and $\delta \in (0, 1)$. Thus, the physical capital k_t and health capital h_t are accumulated as follows.

$$\dot{k_t} = f(k_t, l_t) - c_t - m_t - \delta k_t - k_t(b - d)$$
$$\dot{h_t} = g(m_t) - \delta h_t - h_t(b - d).$$

The physical goods production function $f(k_t, l_t)$ and health capital production function $g(m_t)$ are the usual neo-classical technologies. The health capital production function is increasing in health expenditure but the marginal product is decreasing. The marginal product is finite as health expenditure approaches zero as discussed above.

[5] See Goenka and Liu (2010) for a model with an endogenous labor supply. This paper shows the dynamics are invariant to introduction of endogenous labor supply choice under certain conditions.

[6] For analysis of the equilibria C^2 is required and for local stability and bifurcation analysis at least C^5 is required. Thus, for simplicity we assume all functions to be smooth functions.

[7] This health capital production function could depend on physical capital as well. If this is the case, there will be an additional first order condition equating marginal product of physical capital in the two sectors and qualitative result of the paper still hold. We assume that the production function of health capital does not depend on labor or in effect that its production is more capital intensive than the production of the consumption good to avoid problems associated with factor intensity reversals.

Assumption 3 *The production function* $f(k_t, l_t) : \Re^2_+ \to \Re_+$*:*

1. $f(\cdot,\cdot)$ *is* C^∞ *and homogenous of degree one;*
2. $f_1 > 0, f_{11} < 0,\ f_2 > 0, f_{22} < 0,\ f_{12} = f_{21} > 0$ *and* $f_{11}f_{22} - f_{12}f_{21} > 0$*;*
3. $\lim_{k_t \to 0^+} f_1 = \infty$, $\lim_{k_t \to \infty} f_1 = 0$ *and* $f(0, l_t) = f(k_t, 0) = 0$*.*

Assumption 4 *The production function* $g(m_t) : \Re_+ \to \Re_+$ *is* C^∞ *with* $g' > 0$*,* $g'' < 0$*,* $\lim_{m_t \to 0} g' < \infty$ *and* $g(0) = 0$*.*

We further assume that all individuals are identical. Utility function depends only on current consumption, c_t, is additively separable, and is discounted at the rate $\theta > 0$.

Assumption 5: *The instantaneous utility function* $u(c_t) : \Re_+ \to \Re_+$ is C^∞ with $u' > 0$, $u'' < 0$ and $\lim_{c_t \to 0^+} u' = \infty$.

As discussed above, we look at the optimal solution where the social planner maximizes the discounted utility of the representative consumer. Given concavity of the period utility function, any efficient allocation will involve full insurance.[8] Thus, the consumption of each individual is the same irrespective of health status and we do not need to keep track of individual health histories. The social planner's problem is

$$\max_{c,m} \int_0^\infty u(c) e^{-\theta t} dt$$

subject to

$$\begin{aligned}
\dot{k} &= f(k,l) - c - m - \delta k - k(b-d) && (2)\\
\dot{h} &= g(m) - \delta h - h(b-d) && (3)\\
\dot{l} &= (1-l)(b + \gamma(h) - \alpha(h)l) && (4)\\
k &\geq 0, m \geq 0, h \geq 0, 0 \leq l \leq 1 && (5)\\
& k_0 > 0, h_0 \geq 0, l_0 > 0 \text{ given.} && (6)
\end{aligned}$$

It is worthwhile noting here that we have irreversible health expenditure as it is unlikely that the resource spent on public health can be recovered. For simplicity, we drop time subscript t when it is self-evident.

3 Existence of an optimal solution

In the problem we study, the law of motion of the labor force is not concave reflecting the increasing returns of infections. This can be seen from the Hessian:

$$\begin{pmatrix} 2\alpha(h) & -(\gamma'(h) - \alpha'(h)l) - \alpha'(1-l) \\ -(\gamma'(h) - \alpha'(h)l) - \alpha'(1-l) & (1-l)(\gamma''(h) - \alpha''(h)l) \end{pmatrix}$$

In addition the maximized Hamiltonian, H^*, may not be concave as it is possible that $\frac{\partial^2 H^*}{\partial^2 l} > 0$.[9] Thus, the Arrow sufficiency conditions do not apply. Hence, we directly show the existence of a solution with less stringent conditions in the literature, which is appropriate for the problem at hand. The argument for existence of solutions relies on compactness of the feasible set and some form of continuity of objective function. We first prove the uniform boundedness of the feasible set (which are assumptions

[8] Alternatively instead of maximizing the representative agent's welfare we could maximize the total welfare by using $\int_0^\infty e^{-\theta t} e^{(b-d)t} N_0 u(c_t) dt$ (see the discussion in Arrow and Kurz (1970)). It is equivalent to having a lower discount factor. The qualitative results of this paper still remain although the optimal allocation may vary slightly.

[9] See Gersovitz and Hammer (2004) for more on sufficiency conditions in *SIS* dynamics models.

in Romer (1986) and in d'Albis et al (2008)) that deduces the Lebesgue uniformly integrability. Let us denote by $L^1(e^{-\theta t})$ the set of functions f such that $\int_0^\infty |f(t)|\, e^{-\theta t}dt < \infty$. Recall that $f_i \in L^1(e^{-\theta t})$ weakly converges to $f \in L^1(e^{-\theta t})$ for the topology $\sigma(L^1(e^{-\theta t}), L^\infty)$ (written as $f_i \rightharpoonup f$) if and only if for every $q \in L^\infty$, $\int_0^\infty f_i q e^{-\theta t}dt$ converges to $\int_0^\infty f q e^{-\theta t}dt$ as $i \to \infty$. (written as $\int_0^\infty f_i q e^{-\theta t}dt \longrightarrow \int_0^\infty f q e^{-\theta t}dt$). When writing $f_i \longrightarrow f^*$, we mean that for every $t \in [0,\infty)$, $\lim_{i\to\infty} f_i(t) = f^*(t)$.

We make the following assumption:

Assumption 5 *There exists $\kappa \geq 0, \kappa \neq \infty$ such that $-\kappa \leq \dot{k}/k$.*

This reasonable assumption implies that it is not possible that the growth rate of physical capital converges to $-\infty$ rapidly and is weaker than those used in the literature (see, e.g, Chichilnisky (1981), LeVan and Vailakis (2003), d'Albis et al (2008)). LeVan and Vailakis (2003) use this assumption in a discrete-time optimal growth model with irreversible investment: $0 \leq (1-\delta)k_t \leq k_{t+1}$ or $-\delta \leq (k_{t+1} - k_t)/k_t$ ($\delta > 0$ is the physical depreciation rate in their model, and thus is equivalent to κ). Let us define the net investment : $I = \dot{k} + (\delta + b - d)k = f(k,l) - c - m$. A.6 then implies there exist $\kappa \geq 0, \kappa \neq \infty$ such that $I + [\kappa - (\delta + b - d)]k \geq 0$.

If the standard assumption 2 (v) in Chichilnisky (1981) holds (non-negative investment, $I \geq 0$) then A.6 holds with $\kappa = \delta + b - d$. Therefore, assuming non-negative investment is stronger than A.6 in the sense that κ can take any value except for infinity. We divide the proof into two lemmas. The first lemma proves the relatively weak compactness of the feasible set. For this we show that the relevant variables are uniformly bounded and hence, are uniformly integrable. Using the Dunford-Pettis Theorem we then have relatively weak compactness of the feasible set.

Lemma 1 *Let us denote by $\mathcal{K} = \{(c,k,h,l,m,\dot{k},\dot{h},\dot{l})\}$ the feasible set satisfying (2)-(6). Then $\mathcal{K}$ is relatively weak compact in $L^1(e^{-\theta t})$.*

Proof. See Appendix A for the proof. ■

Since $\mathcal{K}$ is relatively compact in the weak topology $\sigma(L^1(e^{-\theta t}), L^\infty)$, a sequence $\{c_i, k_i, h_i, l_i, m_i, \dot{k}_i, \dot{h}_i, \dot{l}_i\}$ in $\mathcal{K}$ has convergent subsequences (denoted by $\{c_i, k_i, h_i, l_i, m_i, \dot{k}_i, \dot{h}_i, \dot{l}_i\}$ for simplicity of notation) which weakly converge to limit points in $L^1(e^{-\theta t})$.

The following Lemma shows that the control variables and derivatives of state variables weakly converge in the weak topology $\sigma(L^1(e^{-\theta t}), L^\infty)$, while the state variables converge pointwise.

Lemma 2 *i) Let $k_i, h_i, l_i, \dot{k}_i, \dot{h}_i, \dot{l}_i$ in $\mathcal{K}$ and suppose that $k_i \rightharpoonup k^*, h_i \rightharpoonup h^*, l_i \rightharpoonup l^*$. Then $k_i \longrightarrow k^*, h_i \longrightarrow h^*, l_i \longrightarrow l^*$ as $i \to \infty$. Moreover, $\dot{k}_i \rightharpoonup \dot{k}^*, \dot{h}_i \rightharpoonup \dot{h}^*, \dot{l}_i \rightharpoonup \dot{l}^*$ for the the topology $\sigma(L^1(e^{-\theta t}), L^\infty)$.*

ii) Let $\mathbf{x}_i = (c_i, m_i, \dot{k}_i, \dot{h}_i, \dot{l}_i)$ and suppose that $\mathbf{x}_i \rightharpoonup \mathbf{x}^*$ in $\sigma(L^1(e^{-\theta t}), L^\infty)$. Then there exists a function $\mathcal{N} : N \to N$ and a sequence of sets of real numbers $\{\omega_{i(n)} \mid i = n, ..., \mathcal{N}(n)\}$ such that $\omega_{i(n)} \geq 0$ and $\sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} = 1$ such that the sequence v_n defined by $v_n = \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} \mathbf{x}_i$ converges pointwise to $\mathbf{x}^*$ as $n \to \infty$.

Proof. i) For any $x_i \in \mathcal{K}$ and $x_i \rightharpoonup x^*$. We first claim that, for $t \in [0,\infty)$, $\int_0^t x_i dt \to \int_0^t x^* dt$. Note that $x_i \rightharpoonup x^*$ for the topology $\sigma(L^1(e^{-\theta t}), L^\infty)$ if and only if for every $q \in L^\infty$, $\int_0^\infty x_i q e^{-\theta t}dt \to \int_0^\infty x^* q e^{-\theta t}dt$.

Pick any t in $[0,\infty)$ and let

$$q(s) = \begin{cases} \frac{1}{e^{-\theta t}} \text{ if } s \in [0,t] \\ 0 \text{ if } s > t. \end{cases}$$

Therefore, $q \in L^\infty$ and we get $\int_0^t x_i ds = \int_0^\infty x_i q e^{-\theta s} ds \to \int_0^\infty x^* q e^{-\theta s} ds = \int_0^t x^* ds$.

Now, given that $k_i \rightharpoonup k^*$ and $\dot{k}_i \rightharpoonup y^*$ weakly in $L^1(e^{-\theta t})$. By the claim, for all $t \in [0,\infty)$ we have $\int_0^t \dot{k}_i ds \to \int_0^t y^* ds$. This implies, for a fixed t, $k_i \to \int_0^t y^* ds + k_0$. Thus $\int_0^t y^* ds + k_0 = k^*$. Therefore, $\dot{k}^* = y^*$ or $\dot{k}_i \rightharpoonup \dot{k}^*$. The same reasoning applies for h and l to get the conclusion.

ii) A direct application of Mazur's Lemma. ■

We are now in a position to prove the existence of solution to the to the social planner's problem.

Theorem 6 *Under Assumptions A.1-A.6, there exists a solution to the social planner's problem.*

Proof. Since u is concave, for any $\bar{c} > 0$, $u(c) - u(\bar{c}) \le u'(\bar{c})(c_i - \bar{c})$. Thus, if $c \in L^1(e^{-\theta t})$ then $\int_0^\infty u(c)e^{-\theta t}dt$ is well defined because

$$\int_0^\infty u(c)e^{-\theta t}dt \le \int_0^\infty [u(\bar{c}) - u'(\bar{c})\bar{c}]e^{-\theta t}dt + u'(\bar{c})\int_0^\infty ce^{-\theta t}dt < +\infty.$$

Let us define $S :\overset{def}{=} \sup_{c\in\mathcal{K}} \int_0^\infty u(c)e^{-\theta t}dt$. Assume that $S > -\infty$ (otherwise the proof is trivial). Let $c_i \in \mathcal{K}$ be the maximizing sequence of $\int_0^\infty u(c)e^{-\theta t}dt$ so $\lim_{i\to\infty}\int_0^\infty u(c_i)e^{-\theta t}dt = S$.

Since $\mathcal{K}$ is relatively weak compact, suppose that $c_i \rightharpoonup c^*$ for some c^* in $L^1(e^{-\theta t})$. By Mazur's Lemma, there is a sequence of convex combination

$$x_n = \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}c_{i(n)} \to c^*, \omega_{i(n)} \ge 0, \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} = 1.$$

Because u is concave, we have

$$\begin{aligned}
\limsup_{n\to\infty} u(x_n) &= \limsup_{n\to\infty} u(\sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}c_{i(n)}) \\
&\le \limsup_{n\to\infty} [u(c^*) + u'(c^*)(\sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}c_{i(n)} - c^*)] = u(c^*).
\end{aligned}$$

Since this holds for almost t, integrate w.r.t $e^{-\theta t}dt$ to get

$$\int_0^\infty \limsup_{n\to\infty} u(x_n)e^{-\theta t}dt \le \int_0^\infty u(c^*)e^{-\theta t}dt.$$

Using Fatou's lemma we yield

$$\limsup_{n\to\infty}\int_0^\infty u(x_n)e^{-\theta t}dt \le \int_0^\infty \limsup_{n\to\infty} u(x_n)e^{-\theta t}dt \le \int_0^\infty u(c^*)e^{-\theta t}dt. \tag{7}$$

Moreover, by Jensen's inequality we get

$$\limsup_{n\to\infty}\int_0^\infty u(x_n)e^{-\theta t}dt \ge \limsup_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} \int_0^\infty u(c_{i(n)})e^{-\theta t}dt. \tag{8}$$

But since $\int_0^\infty u(c_{i(n)})e^{-\theta t}dt \to S$, (7) and (8) imply $\int_0^\infty u(c^*)e^{-\theta t}dt \ge S$.

So it remains to show that c^* is feasible (because $\mathcal{K}$ is only relatively weak compact, it is not straightforward that $c^* \in \mathcal{K}$).

The task is now to show that there exists some (k^*, l^*, h^*, m^*) in $\mathcal{K}$ such that $(c^*, k^*, l^*, h^*, m^*)$ satisfies (2)-(6).

Consider a feasible sequence $(k_{i(n)}, l_{i(n)}, h_{i(n)}, m_{i(n)})$ in $\mathcal{K}$ associated with $c_{i(n)}$ we have

$$\begin{aligned}
c^* &= \lim_{n\to\infty} x_n = \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}c_{i(n)} \\
&\le \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}[f(k_{i(n)}, l_{i(n)}) - m_{i(n)} - k_{i(n)}(\delta + b - d) - \dot{k}_{i(n)}] \\
&= \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}[f(\lim_{n\to\infty} k_{i(n)}, \lim_{n\to\infty} l_{i(n)}) - (\delta + b - d)\lim_{n\to\infty} k_{i(n)}] \\
&\quad - \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}\dot{k}_{i(n)} - \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)}m_{i(n)}.
\end{aligned}$$

According to Lemma 2, there exists k^*, l^* such that $\lim_{n\to\infty} k_{i(n)} = k^*, \lim_{n\to\infty} l_{i(n)} = l^*$.

By Lemma 2, $\dot{k}_{i(n)} \rightharpoonup \dot{k}^*$ and since $m_{i(n)})$ in $\mathcal{K}$, there exists m^* such that $m_{i(n)} \rightharpoonup m^*$. Thus it follows from Mazur's Lemma that

$$\lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} \dot{k}_{i(n)} \to \dot{k}^*, \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} m_{i(n)} \to m^*.$$

Therefore,

$$c^* \leq f(k^*, l^*) - \dot{k}^* - m^* - \delta k^* - k^*(b - d).$$

Since $\dot{l}_i \rightharpoonup \dot{l}^*$, by Mazur's Lemma, there exists $v_n = \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} \dot{l}_{i(n)} \to \dot{l}^*$ as $n \to \infty$. Thus,

$$\dot{l}^* = \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} \dot{l}_{i(n)} \leq \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} [(1 - l_{i(n)})(b + \gamma(h_{i(n)}) - \alpha(h_{i(n)}) l_{i(n)})].$$

In view of Lemma 2 , $h_{_{i(n)}} \longrightarrow h^*, l_{_{i(n)}} \longrightarrow l^*$ as $n \to \infty$ and $\gamma(h_{_{i(n)}}), \alpha(h_{_{i(n)}})$ are continuous, we get

$$\begin{aligned} \dot{l}^* &\leq \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} [(1 - l^*)(b + \gamma(h^*) - \alpha(h^*) l^*)] \\ &= (1 - l^*)(b + \gamma(h^*) - \alpha(h^*) l^*). \end{aligned}$$

Applying a similar argument and using Jensen's inequality yields

$$\begin{aligned} \dot{h}^* &= \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} \dot{h}_{i(n)} \leq \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} [g(m_{i(n)}) - \delta h_{i(n)} - h_{i(n)}(b - d)] \\ &\leq g(\lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} m_{i(n)}) - \lim_{n\to\infty} \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} (\delta + b - d) h^*_{i(n)} \\ &= g(m^*) - \delta h^* - h^*(b - d). \end{aligned}$$

The proof is done. ■

We have proven that the control variables c, m and derivatives of state variables weakly converge in the weak topology $\sigma(L^1(e^{-\theta t}), L^\infty)$, while the state variables converge pointwise (Lemma 2). The problem is that even if we have a weakly convergent sequence, the limit point may not be feasible. For pointwise convergent sequences, the continuity is all that is necessary to prove the feasibility. Therefore, concavity is not needed for state variables. Theorem 1 shows that the limit point is indeed optimal in the original problem. For weakly convergent sequence, Mazur's Lemma is used to change into pointwise convergence. Jensen's inequality is used to eliminate the convex-combination-coefficients to prove the feasibility. Thus, concavity with respect to control variables is crucial. Our proof is adapted from work of Chichilnisky (1981), Romer (1986) and d'Albis et al (2008) to SIS dynamic model with less stringent assumptions and a nonconvex technology. Chichilnisky (1981) used the theory of Sobolev weighted space and imposed a Caratheodory condition on utility function, Romer (1986) made assumptions that utility function has an integrable upper bound, satisfies a growth condition and d'Albis et al (2008) assumed feasible paths are uniformly bounded and the technology is convex with respect to the control variables.

4 Characterization of Steady State Equilibria

To analyze the equilibria, we look at first order conditions to the optimal solution. This is valid as we know that these conditions are necessary and a solution exists, and thus a solution must satisfy these conditions. Note that we allow for corner solutions. As we will see for some parameters there is a unique (steady state) solution to the first order conditions. For others, there are multiple steady state solutions.

From the Inada conditions we can rule out $k = 0$, and the constraint $l \geq 0$ is not binding since $\dot{l} = b + \gamma > 0$ whenever $l = 0$. The constraint $h \geq 0$ can be inferred from $m \geq 0$, and hence, can be ignored. Now consider the central planner's maximization problem with irreversible health expenditure $m \geq 0$ and the inequality constraint $l \leq 1$. The current value Lagrangian for the optimization problem above is:

$$\begin{aligned}\mathcal{L} &= u(c) + \lambda_1[f(k,l) - c - m - \delta k - k(b-d)] + \lambda_2[g(m) - \\ &\quad - \delta h - h(b-d)] + \lambda_3(1-l)(b + \gamma(h) - \alpha(h)l) + \mu_1(1-l) + \mu_2 m\end{aligned}$$

where λ_1, λ_2, λ_3 are costate variables, and μ_1, μ_2 are the Lagrange multipliers. The Kuhn-Tucker conditions and transversality conditions are given by

$$c: \quad u'(c) = \lambda_1, \tag{9}$$

$$m: \quad m(\lambda_1 - \lambda_2 g') = 0 \qquad m \geq 0 \qquad \lambda_1 - \lambda_2 g' \geq 0 \tag{10}$$

$$k: \quad \dot{\lambda}_1 = -\lambda_1(f_1 - \delta - \theta - (b-d)) \tag{11}$$

$$h: \quad \dot{\lambda}_2 = \lambda_2(\delta + \theta + b - d) - \lambda_3(1-l)(\gamma' - \alpha' l) \tag{12}$$

$$l: \quad \dot{\lambda}_3 = -\lambda_1 f_2 + \lambda_3(\theta + b + \gamma + \alpha - 2\alpha l) + \mu_1 \tag{13}$$

$$\mu_1 \geq 0 \qquad 1 - l \geq 0 \qquad \mu_1(1-l) = 0 \tag{14}$$

$$\lim_{t\to\infty} e^{-\theta t}\lambda_1 k = 0 \qquad \lim_{t\to\infty} e^{-\theta t}\lambda_2 h = 0 \qquad \lim_{t\to\infty} e^{-\theta t}\lambda_3 l = 0. \tag{15}$$

The system dynamics are given by equations (2)-(6) and (9)-(15). If x is a variable, we use x^* to denote its steady state value. In the epidemiology literature and Goenka and Liu (2010), α is the key parameter which is varied. In this paper α is endogenous. Thus, we characterize steady state equilibria in terms of the pair of exogenous parameters $(b, \theta) \in [d, \infty) \times (0, \infty)$.

Define $\underline{l} := \min\{\frac{b+\underline{\gamma}}{\overline{\alpha}}, 1\}$, $\underline{k}$ such that $f_1(\underline{k}, \underline{l}) = \delta + b - d + \theta$ and $\overline{k}$ such that $f_1(\overline{k}, 1) = \delta + b - d + \theta$. Clearly $\overline{k} \geq \underline{k}$ for each (b, θ).

Proposition 1 *Under A.1 – A.6,*

1. *There exists a unique disease-free steady state with $l^* = 1$, $m^* = 0$, $h^* = 0$, and $k^* = \overline{k}$ for any $(b, \theta) \in [d, \infty) \times (0, \infty)$;*

2. *There exists an endemic steady state ($l^* < 1$) if and only if $b < \overline{\alpha} - \underline{\gamma}$ and there is a solution (l^*, k^*, m^*, h^*) to the following system of equations:*

$$l^*(h^*) = \frac{\gamma(h^*) + b}{\alpha(h^*)} \tag{16}$$

$$f_1(k^*, l^*) = \delta + \theta + b - d \tag{17}$$

$$g(m^*) = (\delta + b - d)h^* \tag{18}$$

$$m^*(f_1(k^*, l^*) - f_2(k^*, l^*)l'_\theta(h^*)g'(m^*)) = 0 \tag{19}$$

$$m^* \geq 0 \tag{20}$$

$$f_1(k^*, l^*) \geq f_2(k^*, l^*)l'_\theta(h^*)g'(m^*), \tag{21}$$

where we define $l'_\theta(h^*) := \frac{(1-l^*)(\gamma'(h^*) - \alpha'(h^*)l^*)}{\theta + \alpha(h^*) - b - \gamma(h^*)}$.

Proof. From $\dot{l} = 0$ we have either $l^* = 1$ (disease-free case) or $l^* = \frac{\gamma(h^*)+b}{\alpha(h^*)} < 1$ (endemic case).

Case 1: $l^* = 1$. Since $\dot{\lambda}_2 = \lambda_2(\delta + b - d + \theta) = 0$, $\lambda_2^* = 0$. As g' is finite by assumption, $\lambda_1^* - \lambda_2^* g' = u'(c^*) > 0$, which implies $m^* = 0$ by equation (10). Since $g(0) = 0$, $h^* = 0$ from equation (3). From $\dot{\lambda}_1 = 0$, $k^* = \overline{k}$. So the model degenerates to neo-classical growth model. Moreover $l^* = 1$ exists for all parameter values.

Case 2: $l^* < 1$. This steady state exists if and only if there exists $h^* \geq 0$ such that $l^* = \frac{\gamma(h^*)+b}{\alpha(h^*)} < 1$ and (l^*, k^*, m^*, h^*) is a steady state solution to the dynamical system (3)- (4), (9)- (15). For the former, by assumption $A.2$, $l(h)$ is increasing in h. So if $\frac{b+\gamma}{\overline{\alpha}} < 1$, that is, $b < \overline{\alpha} - \underline{\gamma}$, we could find $h \geq 0$ such that endemic steady state exists. For the latter, since $l^* < 1$, $\mu_1 = 0$. From $\dot{\lambda_2} = 0$ and $\dot{\lambda_3} = 0$, we have:

$$\lambda_2^* = \frac{u'(c^*) f_2(k^*, l^*)}{f_1(k^*, l^*)} \frac{(1 - l^*)(\gamma'(h^*) - \alpha'(h^*) l^*)}{\theta + \alpha(h^*) - b - \gamma(h^*)}$$

So equation (10) could be written as equations (19)-(21). Moreover by letting $\dot{h} = 0$, $\dot{\lambda_1} = 0$ and $\dot{l} = 0$ we have equations (16)-(18). ■

Therefore, the economy has a unique disease-free steady state in which the disease is completely eradicated and there is no need for any health expenditure. In this case, the model reduces to the standard neo-classical growth model. Note that the disease-free steady state always exists. Furthermore, when birth rate is smaller than $\overline{\alpha} - \underline{\gamma}$, in addition to the disease-free steady state, there exists an endemic steady state in which the disease is prevalent and there is non-negative health expenditure. The L.H.S. of equation (21) is the marginal benefit of physical capital investment while the R.H.S. is marginal benefit of health expenditure. To see this, on the R.H.S. the first term $f_2(k^*, l^*)$ is the marginal productivity of labor, the middle term $l'_\theta(h^*)$ can be interpreted as the marginal contribution of health capital on labor supply and the last term $g'(m^*)$ is the marginal productivity of health expenditure. Essentially we can think there is an intermediate production function which transforms one unit of health expenditure into labor supply through the effect on endogenous disease dynamics. Equations (19)-(21)says that if the marginal benefit of physical capital investment is higher than the marginal benefit of health expenditure, there will be no health expenditure ($m^* = 0$).

We want to characterize the endemic steady state further.

Assumption 7 *$\alpha(\alpha''(\gamma + b) - \gamma''\alpha) > 2\alpha'(\alpha'(\gamma + b) - \gamma'\alpha)$.*

By assumption 7 we can show

$$\begin{aligned} l''_\theta(h) &:= \frac{\partial l'_\theta(h)}{\partial h} \\ &= -\frac{(\alpha - \gamma - b + \theta)(\alpha - \gamma - b)[\alpha(\alpha''(\gamma + b) - \gamma''\alpha) - 2\alpha'(\alpha'(\gamma + b) - \gamma'\alpha)] + \alpha\theta(\alpha'(\gamma + b) - \gamma'\alpha)(\alpha' - \gamma')}{\alpha^3(\alpha - \gamma - b + \theta)^2} \\ &< 0 \end{aligned}$$

From equations (16)-(18), we could write (l^*, k^*, m^*) as a function of h. We have $l^*(h)$ given by equation (16) with $\frac{\partial l^*(h)}{\partial h} = \frac{\gamma'\alpha - (\gamma + b)\alpha'}{\alpha^2} > 0$. $m^*(h) > 0$ is given by equation(18) with $\frac{\partial m^*(h)}{\partial h} = \frac{\delta + b - d}{g'(m)} > 0$. $k^*(h)$ is determined by equation (17), that is, at the steady state marginal productivity of physical capital equals to the marginal cost. Since f_1 is strictly decreasing and lies in $(0, +\infty)$ for each $l^*(h)$, we can always find a unique $k^*(h)$ and $\frac{\partial k^*(h)}{\partial h} = -f_{12} \frac{\partial l^*(h)}{\partial h} / f_{11} > 0$. Since $\frac{\partial f_2(k^*(h), l^*(h))}{\partial h} = \frac{f_{11} f_{22} - f_{12} f_{21}}{f_{11}} \frac{\partial l^*(h)}{\partial h} < 0$, $l''_\theta(h) < 0$ and $\frac{\partial g'(m^*(h))}{\partial h} = g'' \frac{\partial m^*(h)}{\partial h} < 0$, the R.H.S. of equation (21) decreases as h increases. That is, we have diminishing marginal product of health capital under assumption 7, which guarantees the uniqueness of endemic steady state.

From equation (20), there are two cases: $m^* = 0$ and $m^* > 0$. The first is termed as the endemic steady state without health expenditure and the second the endemic steady state with health expenditure. For the endemic steady state without health expenditure, $\dot{h} = 0$ implies $h^* = 0$. Equation (21) reduces to

$$f_1(\underline{k}, \underline{l}) \geq f_2(\underline{k}, \underline{l}) l'_\theta(0) g'(0), \tag{22}$$

where $l'_\theta(0) := \frac{(1 - \underline{l})(\gamma'(0) - \alpha'(0)\underline{l})}{\theta + \overline{\alpha} - b - \underline{\gamma}}$. Due to diminishing marginal product of health capital mentioned above, a unique endemic steady state without health expenditure exists if and only if equation (22) is satisfied. Otherwise an endemic steady state with health expenditure exists.

Lemma 3 *For each fixed $b \in [d, \overline{\alpha} - \underline{\gamma})$, there exists a unique $\hat{\theta}(b)$, which is determined by $f_1(\underline{k}, \underline{l}) = f_2(\underline{k}, \underline{l}) l'_\theta(0) g'(0)$, such that:*

1. *If $\hat{\theta}(b) > 0$, then an endemic steady state without health expenditure exists if $\theta \geq \hat{\theta}(b)$ and an endemic steady state with health expenditure exists if $\theta < \hat{\theta}(b)$;*

2. If $\hat{\theta}(b) \leq 0$, then an endemic steady state without health expenditure exists .

Proof. An endemic steady state without health expenditure exists if and only if equation (22) is satisfied. Lets fix any $b \in [d, \overline{\alpha} - \underline{\gamma})$, L.H.S. of (22) is increasing in θ while R.H.S. of (22) is decreasing in θ. So for each b there exists a unique $\hat{\theta}(b)$ such that $f_1(\underline{k}, \underline{l}) = f_2(\underline{k}, \underline{l}) l'_\theta(0) g'(0)$. Note $\hat{\theta}(b)$ could be non-positive. Case 1: $\hat{\theta}(b)$ is positive. If $\theta \geq \hat{\theta}(b)$, equation (22) is satisfied and an endemic steady state without health expenditure exists. Otherwise an endemic steady state with health expenditure may exist. Case 2: $\hat{\theta}(b)$ is non-positive. Then equation (22) is satisfied for all discount factors and only an endemic steady state without health expenditure exists. ■

This result shows that while the disease is endemic it may be socially optimal not to spend any resources on health capital. This is because the marginal productivity of physical capital is higher than that of health expenditures. Furthermore, there is expenditure on health when the discount rate is low (people are more patient) and the birth rate is low. Next we want to study the properties of the function $\hat{\theta}(b)$ for $b \in [d, \overline{\alpha} - \underline{\gamma})$.

Assumption 8 *Elasticity of marginal contribution of health capital on labor supply with respect to birth rate is small, that is,* $\frac{\partial l'_\theta(0)/\partial b}{l'_\theta(0)/b} < b[\frac{1}{f_1} - \frac{f_{21}}{f_{11} f_2} - \frac{f_{22} f_{11} - f_{21} f_{12}}{\overline{\alpha} f_{11} f_2}]$.[10]

Lemma 4 $\hat{\theta}(b)$ *is decreasing in b. And as $b \to \overline{\alpha} - \underline{\gamma}$, $\hat{\theta}(b)$ approaches a non-positive number.*

Proof. Since $\underline{k}$ is given by $f_1(\underline{k}, \underline{l}) = \delta + b - d + \theta$, we have

$$\frac{\partial \underline{k}}{\partial \theta} = \frac{1}{f_{11}} \quad \text{and} \quad \frac{\partial \underline{k}}{\partial b} = \frac{1}{f_{11}} - \frac{f_{12}}{\overline{\alpha} f_{11}}.$$

Moreover, function $\hat{\theta}(b)$ is determined by

$$H = 1 - \frac{f_2(\underline{k}, \underline{l})}{f_1(\underline{k}, \underline{l})} l'_\theta(0) g'(0) = 0.$$

By the implicit function theorem, $\hat{\theta}(b)$ is continuous and

$$\frac{\partial H}{\partial \hat{\theta}} = -\frac{f_{21} f_1 - f_{11} f_2}{f_1^2} \frac{\partial \underline{k}}{\partial \theta} l'_\theta(0) g'(0) - \frac{f_2}{f_1} \frac{\partial l'_\theta(0)}{\partial \theta} g'(0) > 0,$$

and

$$\frac{\partial H}{\partial b} = -\left(\frac{\overline{\alpha} f_{21} + f_{22} f_{11} - f_{21} f_{12}}{\overline{\alpha} f_{11} f_1} - \frac{f_2}{f_1^2} \right) l'_\theta(0) g'(0) - \frac{f_2}{f_1} \frac{\partial l'_\theta(0)}{\partial b} g'(0) > 0$$

under $A.8$. Thus, we have $\partial \hat{\theta} / \partial b < 0$, that is $\hat{\theta}(b)$ is decreasing in b.

Let $b \to \overline{\alpha} - \underline{\gamma}$. For any $\theta > 0$, $\underline{l} \to 1$, $l'_\theta(0) \to 0$ and R.H.S. of equation (22) goes to 0. However L.H.S. of equation (22) equals to $\delta + b - d + \theta$, which is strictly positive as b approaches $\overline{\alpha} - \underline{\gamma}$. So as $b \to \overline{\alpha} - \underline{\gamma}$, equation (22) is satisfied for all $\theta > 0$, which means $\hat{\theta}(b)$ goes to some non-positive number as $b \to \overline{\alpha} - \underline{\gamma}$. ■

From the Figure 4, it is easy to see the graph $\hat{\theta}(b)$ intersects the horizontal axis at the point which lies on the left side of $b = \overline{\alpha} - \underline{\gamma}$. Let us denote $\hat{\theta}(d)$ as the intersection point of both the function $\hat{\theta}(b)$ and vertical axis $b = d$. As the function $\hat{\theta}(b)$ is a one-to-one mapping, we could write its inverse mapping as $\hat{b}(\theta)$ for $\theta \in (0, \hat{\theta}(d)]$ and define $\hat{b}(\theta) = d$ for $\theta > \hat{\theta}(d)$.

Proposition 2 *Under $A.1 - A.8$, for each $\theta > 0$ a unique endemic steady state without health expenditure exists if and only if $\hat{b}(\theta) \leq b < \overline{\alpha} - \underline{\gamma}$. The steady state is given by $l^* = \underline{l}, m^* = 0, h^* = 0$ and $k^* = \underline{k}$.*

---

[10] Under Cobb-Douglas production function $f(k, l) = Ak^a l^{1-a}$, the assumption reduces to $\frac{\partial l'_\theta(0)/\partial b}{l'_\theta(0)/b} < \frac{b}{(1-a)(\delta+b-d+\theta)}$. As $\frac{\partial l'_\theta(0)}{\partial b} = -\frac{l'_\theta(0)}{\overline{\alpha}(1-\underline{l})} - \frac{(1-\underline{l})\alpha'(0)}{\overline{\alpha}(\theta+\overline{\alpha}-b-\underline{\gamma})} + \frac{l'_\theta(0)}{\theta+\overline{\alpha}-b-\underline{\gamma}}$, the assumption is then given by $-\frac{1}{\overline{\alpha}(1-\underline{l})} - \frac{\alpha'(0)}{\overline{\alpha}(\gamma'(0)-\alpha'(0)\underline{l})} + \frac{1}{\theta+\overline{\alpha}-b-\underline{\gamma}} < \frac{1}{(1-a)(\delta+b-d+\theta)}$, which is shown to be satisfied for a wide range of parameter values.

Proof. The proof follows from Proposition 1 and Lemmas 3, 4. It is easily seen from Figure 4. ∎

Proposition 3 *Under $A.1 - A.8$, for each $\theta > 0$ a unique endemic steady state with health expenditure exists if and only if $d \leq b < \hat{b}(\theta)$. The steady state is given by $l^* = \frac{\gamma(h^*)+b}{\alpha(h^*)}$, and k^*, h^* and m^* determined by:*

$$f_1(k^*, l^*) = \delta + b - d + \theta$$
$$f_2(k^*, l^*) l'_\theta(h^*) g'(m^*) = \delta + b - d + \theta$$
$$g(m^*) = (\delta + b - d)h^*.$$

Proof. The proof follows from Proposition 1 and Lemmas 3, 4. Moreover as $m^* > 0$, equation (21) holds at equality. It implies marginal productivity of physical capital equals the marginal productivity of health capital. As l^*, k^*, m^* could be written as function of h, we only need to show there always exists a solution h^* to the following equation:

$$f_2(k^*(h), l^*(h)) l'_\theta(h) g'(m^*(h)) = \delta + b - d + \theta \tag{23}$$

Since $\lim_{h\to\infty} f_2 l'_\theta(h) g'(m) = 0$ and $\lim_{h\to 0} f_2 l'_\theta(h) g'(m) = f_2(\underline{k}, \underline{l}) l'_\theta(0) g'(0) > f_1(\underline{k}, \underline{l}) = \delta + b - d + \theta$ if $b \in [d, \hat{b}(\theta))$, equation (23) always has a solution. That is, under $A.1$-$A.8$ there exists endemic steady state with health expenditure if $b \in [d, \hat{b}(\theta))$. Moreover, since R.H.S. of equation (23) decreases as h increases, there exists a unique endemic steady state with health expenditure. ∎

Hence, an endemic steady state without health expenditure exists only when marginal productivity of physical capital is no less than marginal productivity of health capital. In other words, despite the prevalence of the disease, if marginal productivity of physical capital investment is greater than marginal productivity of health capital, there will be no investment in health. Thus, the prevalence of the disease is not sufficient (from purely an economic point of view) to require health expenditures. It is conceivable that in labor abundant economies with low physical capital this holds, and thus, we may observe no expenditure on controlling an infectious disease while in other richer economies there are public health expenditures to control it. The endemic steady state without health expenditure is the same as a neo-classical steady state but with only a smaller labor force. Thus, there is lower consumption and production in the steady state. By investing in health expenditure we are able to control infectious disease. Compared with the disease-free case the economy has lower physical capital and a smaller labor force. The production will be lower, and and there is expenditure allocated for health expenditure. Thus, clearly the consumption will be lower. It does not make too much sense to compare welfare for two endemic steady states as they do not coexist.

5 Local Stability and Bifurcation

The dynamical system is given by equations (2)- (4), (9)- (15) and there are three equilibria. In order to examine their stability we linearize the system around each of the steady states. To simplify the exposition we make the following assumption.

Assumption 9: *The instantaneous utility function $u(c) = \log c$.*

Substituting $\lambda_1 = u'(c) = 1/c$ into equation (11), we get

$$\dot{c} = c(f_1 - \delta - \theta - (b - d)). \tag{24}$$

5.1 The Disease-Free Case

At the disease-free steady state, $\lambda_1 > \lambda_2 g'$. Since all the functions in this model are smooth functions, by continuity there exists a neighborhood of the steady state such that the above inequality still holds. Thus, $m^* = 0$ in this neighborhood. Intuitively around the steady state the net marginal benefit of health investment is negative: the disease is eradicated and health investment only serves to reduce physical capital accumulation and hence, lower levels of consumption, and thus no resources are spent on eradicating diseases. As $m^* = 0$ in the neighborhood of the steady state, we have a maximization problem with only one choice variable - consumption and the dynamic system reduces to:

$$\begin{aligned}
\dot{k} &= f(k,l) - c - \delta k - k(b-d) \\
\dot{h} &= -\delta h - h(b-d) \\
\dot{l} &= (1-l)(b - \alpha(h)l + \gamma(h)) \\
\dot{c} &= c(f_1 - \delta - \theta - (b-d)),
\end{aligned}$$

with three state variables and one choice variable. This can also be simply derived by substituting $m = 0$ into the original dynamic system. By linearizing the system around the steady state, we have:

$$\mathcal{J}_1 = \begin{pmatrix} \theta & 0 & f_2^* & -1 \\ 0 & -\delta - (b-d) & 0 & 0 \\ 0 & 0 & \overline{\alpha} - (\underline{\gamma} + b) & 0 \\ c^* f_{11}^* & 0 & c^* f_{12}^* & 0 \end{pmatrix}.$$

The eigenvalues are $\Lambda_1 = -\delta - (b-d) < 0$, $\Lambda_2 = \frac{\theta - \sqrt{\theta^2 - 4c^* f_{11}^*}}{2} < 0$, $\Lambda_3 = \frac{\theta + \sqrt{\theta^2 - 4c^* f_{11}^*}}{2} > 0$, and $\Lambda_4 = \overline{\alpha} - (\underline{\gamma} + b)$. The sign of Λ_4 depends on b. We notice if $b = \overline{\alpha} - \underline{\gamma}$, $\mathcal{J}_1$ has a single zero eigenvalue. Thus, we have a non-hyperbolic steady state and a bifurcation may arise. In other words, the disease-free steady state possesses a 2-dimensional local invariant stable manifold, a 1-dimensional local invariant unstable manifold and 1-dimensional local invariant center manifold . In general, however, the behavior of trajectories in center manifold cannot be inferred from the behavior of trajectories in the space of eigenvectors corresponding to the zero eigenvalue. Thus, we shall take a close look at the flow in the center manifold. As the zero eigenvalue comes from dynamics of l, and the dynamics of l and h are independent from the rest, we could just focus on the dynamics of l and h. By taking b as bifurcation parameter and following the procedures given by Wiggins (2002) and Kribs-Zaleta (2003), we are able to calculate the dynamics on the center manifold (See the Appendix B for details):

$$\dot{z} = \overline{\alpha} z (z - \frac{1}{\overline{\alpha}} \tilde{b}), \tag{25}$$

where $\tilde{b} = b - (\overline{\alpha} - \underline{\gamma})$.

The fixed points of (25) are given by $z = 0$ and $z = \frac{1}{\overline{\alpha}} \tilde{b}$, and plotted in figure 3. We can see the dynamics on the center manifold exhibits a transcritical bifurcation at $\tilde{b} = 0$. Hence, for $\tilde{b} < 0$, there are two fixed points; $z = 0$ is unstable and $z = \frac{1}{\overline{\alpha}} \tilde{b}$ is stable. These two fixed points coalesce at $\tilde{b} = 0$, and for $\tilde{b} > 0$, $z = 0$ is stable and $z = \frac{1}{\overline{\alpha}} \tilde{b}$ is unstable. Thus, an exchange of stability occurs at $\tilde{b} = 0$, i.e., $b = \overline{\alpha} - \underline{\gamma}$. Therefore, for the original dynamical system if $b > \overline{\alpha} - \underline{\gamma}$, there is a 3-dimensional stable manifold and a 1-dimensional unstable manifold, and if $b < \overline{\alpha} - \underline{\gamma}$, there is a 2-dimensional stable manifold and 2-dimensional unstable manifold. Moreover, while physical capital, health capital and labor force are given at any point in time, the consumption can jump. Thus, if $b > \overline{\alpha} - \underline{\gamma}$, the system is locally saddle stable and has a unique stable path; and if $b < \overline{\alpha} - \underline{\gamma}$, the system is locally unstable.

5.2 The Endemic Case Without Health Expenditures

For the endemic steady state with no health expenditures, $\lambda_1 \geq \lambda_2 g'$ and $m^* = 0$. By continuity, this will also hold in a small neighborhood of the steady state. Thus, it is similar to the disease-free case except that $l^* < 1$. Linearizing the system around the steady state:

$$\mathcal{J}_2 = \begin{pmatrix} \theta & 0 & f_2^* & -1 \\ 0 & -\delta - (b-d) & 0 & 0 \\ 0 & (1-l^*)(\gamma'^* - \alpha'^* l^*) & \overline{\alpha} - (\underline{\gamma} + b) & 0 \\ c^* f_{11}^* & 0 & c^* f_{12}^* & 0 \end{pmatrix}.$$

The eigenvalues are $\Lambda_1 = -\delta - (b-d) < 0$, $\Lambda_2 = \frac{\theta - \sqrt{\theta^2 - 4c^* f_{11}^*}}{2} < 0$, $\Lambda_3 = \frac{\theta + \sqrt{\theta^2 - 4c^* f_{11}^*}}{2} > 0$, and $\Lambda_4 = (\underline{\gamma} + b) - \overline{\alpha} < 0$. So it has 3-dimensional stable manifold and 1-dimensional unstable manifold.

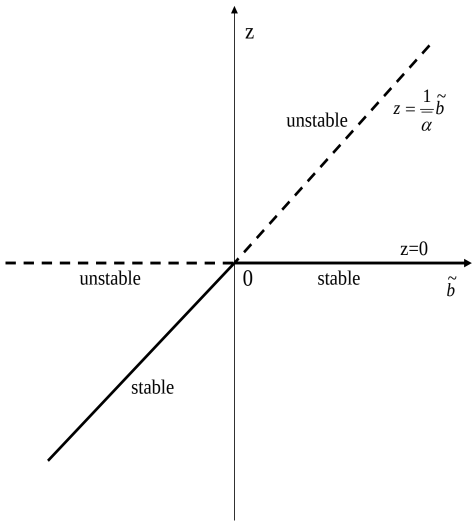

Figure 3: The transcritical bifurcation diagram

Since the system has three state variables and one choice variable, it is locally saddle stable and has a unique stable path. Moreover, this corresponds to the stable steady state $z = \frac{1}{\alpha}\tilde{b}$ when $\tilde{b} < 0$ in figure 3. This also explains why when $\tilde{b}$ decreases and crosses 0, the stable disease-free steady state undergoes a bifurcation into one unstable disease-free steady state and one stable endemic steady state without health expenditure.

5.3 The Endemic Case With Health Expenditures

For the endemic case with health expenditures, the dynamical system is given by equations (2)- (4), (9)-(15) with $\lambda_1 = \lambda_2 g'$, $m^* > 0$ and $l^* < 1$. Simplifying, the system is reduced to:

$$\begin{aligned}
\dot{k} &= f(k,l) - c - m - \delta k - k(b-d) \\
\dot{h} &= g(m) - \delta h - h(b-d) \\
\dot{l} &= (1-l)(b + \gamma(h) - \alpha(h)l) \\
\dot{c} &= c(f_1 - \delta - (b-d) - \theta) \\
\dot{m} &= (c\lambda_3 g'(m)(1-l)(\gamma' - \alpha' l) - f_1)\frac{g'(m)}{g''(m)} \\
\dot{\lambda_3} &= -\frac{1}{c}f_2 + \lambda_3\theta - \lambda_3(2\alpha(h)l - b - \gamma(h) - \alpha(h)).
\end{aligned}$$

We now have a higher dimensional system than the earlier two cases as $m > 0, h > 0$. Linearizing around the equilibrium the Jacobian is given by:

$$\mathcal{J}_3 = \begin{bmatrix}
\theta & 0 & f_2^* & -1 & -1 & 0 \\
0 & -\delta - (b-d) & 0 & 0 & g'^* & 0 \\
0 & (1-l^*)(\gamma'^* - \alpha'^* l^*) & b + \gamma^* - \alpha^* & 0 & 0 & 0 \\
c^* f_{11}^* & 0 & c^* f_{12}^* & 0 & 0 & 0 \\
-f_{11}^* \frac{g'^*}{g''^*} & \frac{f_1^*(\gamma''^* - \alpha''^* l^*)}{\gamma'^* - \alpha'^* l^*}\frac{g'^*}{g''^*} & \left(\frac{f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*)}{(1-l^*)(\gamma'^* - \alpha'^* l^*)} - f_{12}^*\right)\frac{g'^*}{g''^*} & \frac{f_1^*}{c^*}\frac{g'^*}{g''^*} & f_1^* & \frac{f_1^*}{\lambda_3^*}\frac{g'^*}{g''^*} \\
-\frac{f_{12}^*}{c^*} & -\lambda_3^*(2\alpha'^* l^* - \gamma'^* - \alpha'^*) & -\frac{f_{22}^*}{c^*} - 2\lambda_3^*\alpha^* & \frac{f_2^*}{c^{*2}} & 0 & \frac{f_2^*}{c^*\lambda_3^*}.
\end{bmatrix}.$$

Let us denote $\mathcal{J}_3$ as a matrix $(a_{ij})_{6\times 6}$ with the signs of a_{ij} given as follows:

$$\begin{bmatrix} a_{11}(+) & 0 & a_{13}(+) & -1 & -1 & 0 \\ 0 & a_{22}(-) & 0 & 0 & a_{25}(+) & 0 \\ 0 & a_{32}(+) & a_{33}(-) & 0 & 0 & 0 \\ a_{41}(-) & 0 & a_{43}(+) & 0 & 0 & 0 \\ a_{51}(-) & a_{52}(+) & a_{53} & a_{54}(-) & a_{55}(+) & a_{56}(-) \\ a_{61}(-) & a_{62} & a_{63} & a_{64}(+) & 0 & a_{66}(+) \end{bmatrix}$$

Note that as $l^* = \frac{\gamma^*+b}{\alpha^*} < 1$, at the steady state $a_{33} = b + \gamma^* - \alpha^* < 0$ and $\dot{\lambda}_3 = 0$ so we get

$$\lambda_3^* = \frac{f_2^*}{c^*(\theta - 2\alpha^* l^* + b + \gamma^* + \alpha^*)} = \frac{f_2^*}{c^*(\theta + \alpha^* - b - \gamma^*)} > 0.$$

Thus, the terms a_{53}, a_{62}, a_{63} remain to be signed. The characteristic equation, $|\Lambda I - \mathcal{J}_3| = 0$, can be expanded and written as a polynomial of λ as

$$P(\Lambda) = \Lambda^6 - D_1\Lambda^5 + D_2\Lambda^4 - D_3\Lambda^3 + D_4\Lambda^2 - D_5\Lambda + D_6 = 0$$

where the D_i are the sum of the i-th order minors about the principal diagonal of $\mathcal{J}_3$ which are explicitly defined (See Appendix).

Thus, for D_1 we have

$$\begin{aligned} D_1 &= a_{11} + a_{22} + a_{33} + a_{44} + a_{55} + a_{66} \\ &= \theta - \delta + \gamma^* + d - \alpha^* + f_1^* + \frac{f_2^*}{c^*\lambda_3^*} = 3\theta. \end{aligned}$$

which are first order minors about the diagonal.

Let denote Λ_i $(i = 1..6)$ the solutions of the characteristic equation, by Vietae's formula we have

$$\Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4 + \Lambda_5 + \Lambda_6 = D_1 = 3\theta > 0$$

which implies there exists at least one root $\Lambda_i > 0$.

We now prove that, under the following assumption, the system is saddle-point stable.

Assumption 10: *The parameters of the model satisfied*

$$i) \qquad (\alpha^* - b - \gamma^*)(\gamma'^* - \alpha'^*) < (\theta - b - \gamma^* + \alpha^*)(\gamma'^* + \alpha'^* - \frac{2\alpha'^*(b+\gamma^*)}{\alpha^*})$$

$$ii) \qquad \theta < \frac{X + Y + \sqrt{(X+Y)^2 + 32(X^2+Y^2)}}{16}$$

where $X = \delta + b - d, Y = \alpha - \gamma - b$.

Note that *A.10* holds in the leading example given below, *A.10 (i)* is satisfied when $\alpha(h)$ is constant, and *A.10 (ii)* holds when $\alpha(h)$ is large relative to θ.

It follows from *A.10* that

$$-(2\alpha'^* l^* - \alpha'^* - \gamma'^*) = \gamma'^* + \alpha'^* - \frac{2\alpha'^*(b+\gamma^*)}{\alpha^*} > 0.$$

Hence, $a_{53} = \left(\frac{f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*)}{(1-l^*)(\gamma'^* - \alpha'^* l^*)} - f_{12}^*\right)\frac{g'^*}{g''^*} \geq 0, a_{62} = -\lambda_3^*(2\alpha'^* l^* - \gamma'^* - \alpha'^*) \geq 0$. With this assumption, every sign of a_{ij} is defined except for $a_{63} = -\frac{f_{22}^*}{c^*} - 2\lambda_3^*\alpha^*$. The proof of the following proposition will be given in Appendix.

Proposition 4 *Under A.1 - A.10 (i),* $\det \mathcal{J}_3 < 0$ *and there exists at least one negative characteristic root.*

The discussion so far shows that we may have *one*, *three* or *five* number of negative roots and have at least *one* positive root. We are interested in the case of at least *three* negative characteristic roots with the case of *three* negative roots giving saddle-point stability. Note that the coefficients of characteristic equation D_i (sum of i-dimension principal minors of Jacobian matrix) $(i = 1, 2, 3, 4, 5, 6)$ are well defined.

Lemma 5 *Under A.1 - A.10 we have* $D_1D_2 - D_3 < 0$.

Proof. See Appendix ■

Proposition 5 *Under A.1-A.10.*

i) If $3\theta D_4 - D_5 > 0$ *then there exist three or* $five$ *negative characteristic roots. As a special case, if* $D_4 > 0, D_5 < 0$ *then there are exactly three negative roots.*

ii) If $3\theta D_4 - D_5 < 0$ *and* $(3\theta D_2 - D_3)D_5 < 9\theta^2 D_6$ *then there are exactly three negative roots.*

Proof. The number of negative roots of $P(\Lambda)$ is exactly the number of positive roots of

$$P(-\Lambda) = \Lambda^6 + D_1\Lambda^5 + D_2\Lambda^4 + D_3\Lambda^3 + D_4\Lambda^2 + D_5\Lambda + D_6 = 0. \tag{26}$$

We will use the Routh's stability criterion which states that the number of positive roots of equation (26) is equal to the number of changes in sign of the coefficients in the first column of the Routh's table as shown below:

$$\begin{bmatrix} 1 & D_2 & D_4 & D_6 & 0 \\ D_1 & D_3 & D_5 & 0 & 0 \\ a_1 & a_2 & D_6 & 0 & 0 \\ b_1 & b_2 & 0 & 0 & 0 \\ c_1 & D_6 & 0 & 0 & 0 \\ d_1 & 0 & 0 & 0 & 0 \\ D_6 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where

$$\begin{aligned} a_1 &= \frac{D_1D_2 - D_3}{D_1}, a_2 = \frac{D_1D_4 - D_5}{D_1} \\ b_1 &= \frac{a_1D_3 - a_2D_1}{a_1}, b_2 = \frac{a_1D_5 - D_6D_1}{a_1} \\ c_1 &= \frac{b_1a_2 - a_1b_2}{b_1}, d_1 = \frac{c_1b_2 - b_1D_6}{c_1}. \end{aligned}$$

Recall that we have $D_1 = 3\theta > 0, D_2 < 0, D_3 < 0, D_6 < 0$ and $a_1 = \frac{D_1D_2 - D_3}{D_1} < 0$.

Let us see the sign of the first column in the Routh's table.

1	D_1	a_1	b_1	c_1	d_1	D_6
+	+	−	±	±	±	−

If any of b_1, c_1, d_1 are positive, there are at least *three* changes of sign (Only *three* or $five$ of changes in sign is possible).

i) Obviously $a_2 > 0$ in this case.

Suppose that all b_1, c_1, d_1 are negative. That is $b_1 < 0$ and

$$c_1 = \frac{b_1a_2 - a_1b_2}{b_1} < 0, d_1 = \frac{c_1b_2 - b_1D_6}{c_1} < 0.$$

This implies

$$\begin{aligned} b_1a_2 &> a_1b_2 \Rightarrow b_2 > 0 \\ c_1b_2 &> b_1D_6 \Rightarrow b_2 < 0. \end{aligned}$$

A contradiction.

Thus, at least one of b_1, c_1, d_1 is positive, i.e. there are *three* or $five$ negative characteristic roots. As a special case, if $D_4 > 0, D_5 < 0$ then obviously $3\theta D_4 - D_5 > 0$. So we have at least *three* negative characteristic roots. On the other hand, we have *three* changes of sign of the coefficients of $P(-\Lambda)$ as shown below

1	D_1	D_2	D_3	D_4	D_5	D_6
+	+	−	+	+	−	−

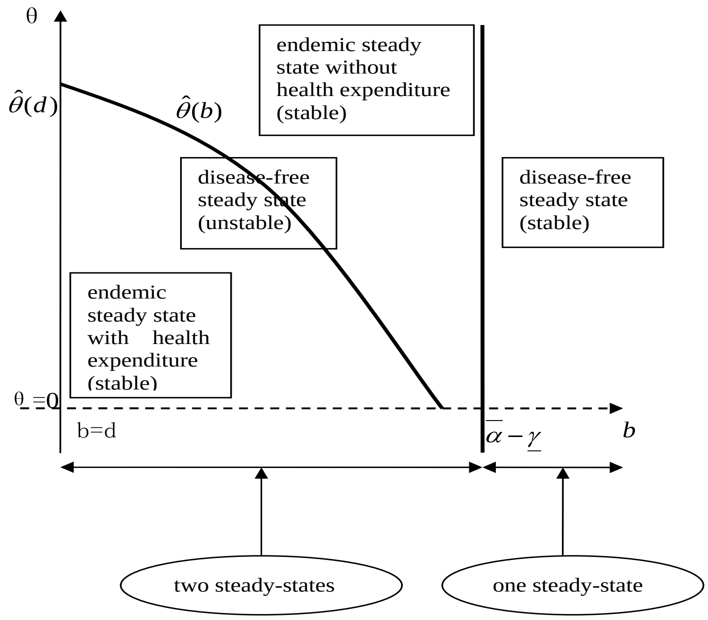

Figure 4: The local stability and bifurcation diagram

According to Descartes' rule as there are *three* changes in sign of the coefficients of $P(-\Lambda)$, $P(-\Lambda)$ has at most *three* positive roots. That means $P(\Lambda)$ has at most *three* negative roots. Therefore, in this case, there are exact *three* negative roots.

ii) In this case, clearly $a_2 < 0$, $b_1 = \frac{a_1 D_3 - a_2 D_1}{a_1} < 0$, and

$$a_1 D_5 - D_6 D_1 = \frac{(D_1 D_2 - D_3) D_5 - D_6 D_1^2}{D_1} = \frac{(3\theta D_2 - D_3) D_5 - 9\theta^2 D_6}{D_1} < 0.$$

Thus $b_2 = \frac{a_1 D_5 - D_6 D_1}{a_1} > 0$.

If $c_1 < 0, d_1 < 0$, we should have $b_1 a_2 > a_1 b_2$ and $c_1 b_2 > b_1 D_6$. But this implies $b_2 < 0$. A contradiction.

Let us consider the sign of the first column of the Routh's table:

$$\begin{array}{ccccccc} 1 & D_1 & a_1 & b_1 & c_1 & d_1 & D_6 \\ + & + & - & - & \pm & \pm & - \end{array}$$

Since either c_1 or d_1 is positive, we only have *three* number of changes in sign.

So $P(\Lambda) = 0$ has *three* negative roots. ■

The local stability and bifurcation of the dynamic system are summarized in Figure 4. When the birth rate b is very high, i.e. greater than $\overline{\alpha} - \underline{\gamma}$, there is only a disease-free steady state which is locally stable. This is the case where disease is eradicated in the long run. When b decreases to exactly $\overline{\alpha} - \underline{\gamma}$, the stable disease-free equilibrium goes through a transcritical bifurcation to two equilibria: one is the unstable disease-free steady state and the other is the stable endemic steady state without health expenditure as $\hat{\theta}(b)$ is equal to 0 at $\overline{\alpha} - \underline{\gamma}$. To the left of this point, $\hat{\theta}(b)$ is a decreasing function. Below this function, when θ is relatively low, there is the endemic steady state with positive health expenditures, but above the function, only the endemic steady state without health expenditures exists. These steady states are stable. The disease free steady state continues to exist but is unstable.

6 Comparative Statics

We now explore how the steady state properties of the model change as the parameters are varied. The goal of comparative statics is twofold: (1) We show as parameters vary, there is a nonlinearity

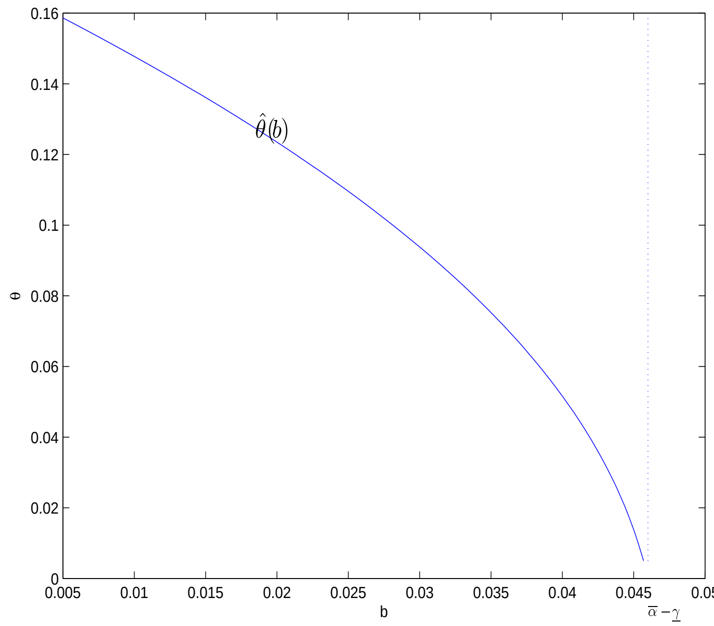

Figure 5: $\hat{\theta}(b)$

in steady state changes due to the switches in the steady state and the role played by the endogenous changes in health expenditures. (2) We study the endogenous relationship between health expenditure (as percentage of GDP) and GDP. This can help understanding the changing share of health expenditures in many countries. The first points out that non-linearities in equilibrium outcomes, which are often assume away, may be very important in understanding aggregate behavior. While we are unable to study global dynamics as it is difficult in the system to derive policy functions (unlike Goenka and Liu (2010) where the system is only two-dimensional) and thus, are unable to study the full range of dynamics, the results point out that even steady states may change in a non-linear way. For the second, it should be emphasized that while we are looking at only public health expenditures on infectious diseases this methodology can be extended to incorporate non-infectious diseases.

We specify the following functional forms: $f(k,l) = Ak^a l^{1-a}$; $\quad g(m) = \phi_3(m+\phi_1)^{\phi_2} - \phi_3\phi_1^{\phi_2}$; $\quad \alpha(h) = \alpha_1 + \alpha_2 e^{-\alpha_3 h}$; $\quad \gamma(h) = \gamma_1 - \gamma_2 \exp^{-\gamma_3 h}$. The parameter values are chosen as follows: $A = 1, a = 0.36, \delta = 0.05, \theta = 0.05$ and $b = 0.02, d = 0.005$ by convention. Since there are no counterparts for health related functions in the economic literature we choose the following parameters which satisfy assumptions *A.1-A.9* :$\phi_1 = 2, \phi_2 = 0.1, \phi_3 = 1, \alpha_1 = \alpha_2 = 0.023, \alpha_3 = 1, \gamma_1 = 1.01, \gamma_2 = \gamma_3 = 1$. So we have $\overline{\alpha} = 0.046$ and $\underline{\gamma} = 0.01$. $\hat{b} = \overline{\alpha} - \underline{\gamma} = 0.036$ and if $b > 0.036$ only disease free steady state exists. The function $\hat{\theta}(b)$ is given in Figure 5 for this economy. While sufficient conditions for stability (*A.10*) may not be satisfied as the parameters are varied, we check that the stability properties continue to hold in the parameter range of interest.

6.1 The discount rate θ

As θ is varied, in the endemic steady state without health expenditure,

$$\frac{dk^*}{d\theta} = \frac{1}{f_{11}} < 0, \quad \text{and} \quad \frac{dc^*}{d\theta} = \frac{\theta}{f_{11}} < 0.$$

The disease prevalence $l^* = \frac{\gamma+b}{\overline{\alpha}}$ remains unchanged.

In the endemic steady state with health expenditure, we have $\frac{\partial m}{\partial h} = \frac{\delta+(b-d)}{g'(m)} > 0$ and $\frac{\partial l'_\theta(h)}{\partial \theta} = \frac{-l'_\theta(h)}{\alpha(h)-(\gamma+b)+\theta} < 0$. Let $\Psi = f_{11}(f_{22}g'(m)l'l'_\theta + f_2 g'(m)l''_\theta + f_2 g''(m)\frac{\partial m}{\partial h}l'_\theta) - f_{12}l'f_{21}l'_\theta g' > 0$. By the

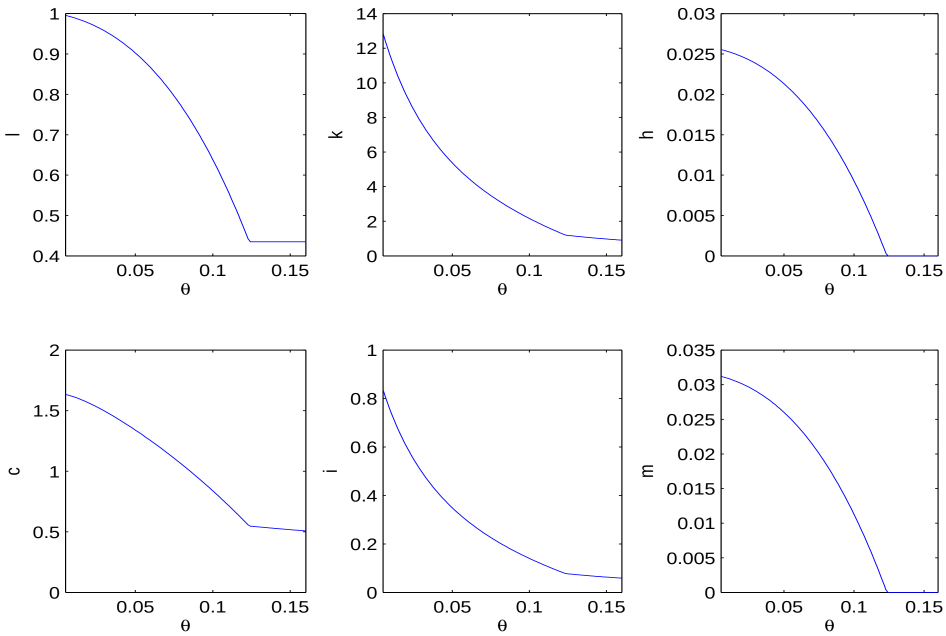

Figure 6: Varying θ

multi-dimensional implicit function theorem , we have:

$$\frac{dk^*}{d\theta} = \frac{1}{\Psi}(f_{22}g'(m)l'l'_\theta + f_2 g'(m)l''_\theta + f_2 g''(m)\frac{\partial m}{\partial h}l'_\theta - f_{12}l'(1 - f_2 g'\frac{\partial l'_\theta}{\partial \theta})) < 0,$$

$$\frac{dh^*}{d\theta} = \frac{1}{\Psi}(f_{11}(1 - f_2 g'\frac{\partial l'_\theta}{\partial \theta}) - f_{21}g'(m)l'_\theta) < 0,$$

$$\text{and, thus,} \quad \frac{dl^*}{d\theta} = l'\frac{dh^*}{d\theta} < 0,$$

$$\frac{dc^*}{d\theta} = (f_1 - \delta_k - (b-d))\frac{dk^*}{d\theta} + (f_2 l' - \delta_h - (b-d))\frac{dh^*}{d\theta} < 0.$$

Therefore, in the endemic steady state without health expenditure variations in the discount rate have no effect on the spread of infectious diseases, since without health expenditures the mechanism of disease spread is independent of individual's behavior. The smaller discount rate only leads to higher physical capital and consumption in exactly the same way as in the neo-classical model. In the endemic steady state with health expenditure, as the discount rate decreases, that is as the people become more patient, they spend more resources in prevention of infections or getting better treatment. The rise in health capital leads to a larger labor force, and both physical capital and consumption will increase. We can see from Figure 6 that the rate of investment in physical capital is increasing while that of health capital is decreasing as θ decreases. This leads to an initial increase in the share of health expenditure in GDP and then an eventual decrease. The intuition is that as people become more patient, they spend more on health. This has two effects. First, as the incidence of diseases is controlled the increase in the effective labor force increases the marginal product of capital which leads to the increasing rate of physical capital investment. Second, as the incidence of diseases decreases, due to the externality in disease transmission the fraction of infectives decreases. This decreases the rate of investment in health expenditures. This leads to a non-monotonicity in the share of health expenditures, see Figure 7. This should hold for a cross section of countries when we consider the expenditure on a given infectious disease.

6.2 The birth rate b

The other two exogenous parameters are the death rate d and the birth rate b. As they enter the model only in difference, we look at variations in b holding d constant. Thus, increasing b is similar to increasing the net birth rate (taking into account migration). In the endemic steady state without health

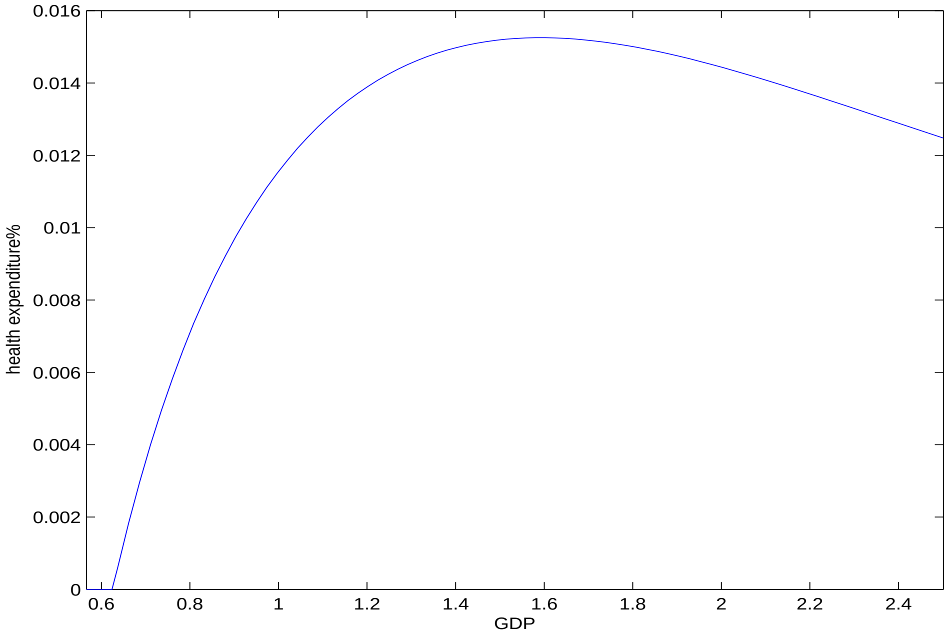

Figure 7: Share of health expenditure as θ is varied

expenditure,

$$\frac{dl^*}{db} = \frac{1}{\alpha} > 0, \quad \frac{dk^*}{db} = \underbrace{\frac{1}{f_{11}}}_{-} + \underbrace{\frac{f_{12}}{-\bar{\alpha} f_{11}}}_{+} \quad ?, \quad \text{and} \quad \frac{dc^*}{db} = \underbrace{\frac{\theta - k f_{11}}{f_{11}}}_{-} + \underbrace{\frac{\theta f_{12} - f_2 f_{11}}{-\bar{\alpha} f_{11}}}_{+} \quad ?.$$

This is because a rise of the birth rate has two effects. First, it has a negative effect as more needs to be invested to maintain the same capital per capita. Second, there is a positive effect: The proportion of healthy people increases due to more healthy newborns, and thus a higher labor force leads to higher physical capital and consumption. Hence, the two effects are offsetting and the net effect is unclear in general.

In the endemic case with health expenditure, by the implicit function theorem

$$\frac{dk^*}{db} = \underbrace{\frac{1}{\Psi}\left(f_{22} g' l' l'_\theta + f_2 g' l''_\theta + f_2 g'' \frac{\partial m}{\partial h} l'_\theta - f_{12} l'\right)}_{-} + \underbrace{-\frac{1}{\Psi} f_2 f_{12} g' \frac{1}{\alpha} l''_\theta}_{+} + \underbrace{\frac{1}{\Psi} f_2 f_{12} g' l' \frac{\partial l'_\theta}{\partial b}}_{?} \quad ?$$

$$\frac{dh^*}{db} = \underbrace{\frac{1}{\Psi}\left(f_{11} - f_{21} g' l'_\theta\right)}_{-} + \underbrace{\frac{1}{\Psi}\frac{1}{\alpha} g' l'_\theta \left(f_{21} f_{12} - f_{11} f_{22}\right)}_{-} + \underbrace{\frac{1}{\Psi}\left(-f_{11} f_2 g' \frac{\partial l'_\theta}{\partial b}\right)}_{?} \quad ?$$

$$\text{and then} \quad \frac{dl^*}{db} = \frac{1}{\alpha} + l'(h) \frac{dh^*}{db} \quad ?$$

where $\frac{\partial l'_\theta}{\partial b} = -\frac{\alpha'}{\alpha^2} + \frac{\theta(\theta\alpha' + \alpha(\alpha' - \gamma'))}{\alpha^2(\alpha - (\gamma + b) + \theta)^2}$.[11]

Therefore, the effect of a rise in birth rate is ambiguous. The basic reasoning is similar to the endemic case without health expenditure above, but here it becomes more complex by involving changes in health capital. First, there is a negative effect: The marginal cost of physical capital and health capital will increase which leads to lower physical capital and health capital. Second, since people are born healthy the labor force is increasing, which means that the marginal productivity of physical capital is increasing and hence, physical capital increases. On the other hand the higher labor force causes marginal productivity of labor to decline and hence, health capital decreases. Third, because of more healthy newborns, the marginal benefit of health is changing. The marginal benefit of health to labor

[11] Note $\partial l' / \partial b = -\alpha'^2 > 0$, but it is not clear $\partial l'_\theta / \partial b$ takes the positive sign or the negative sign.

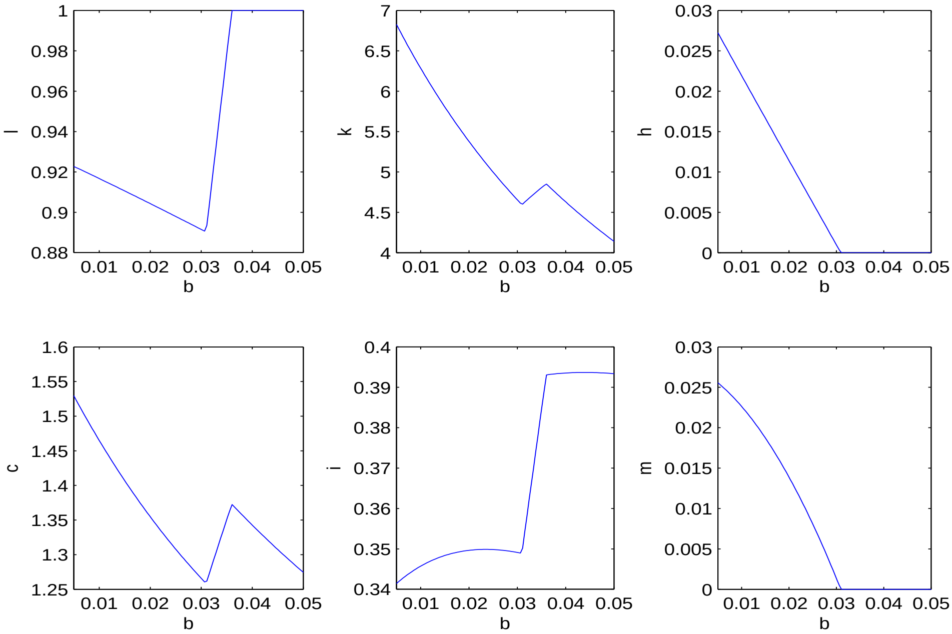

Figure 8: Varying b

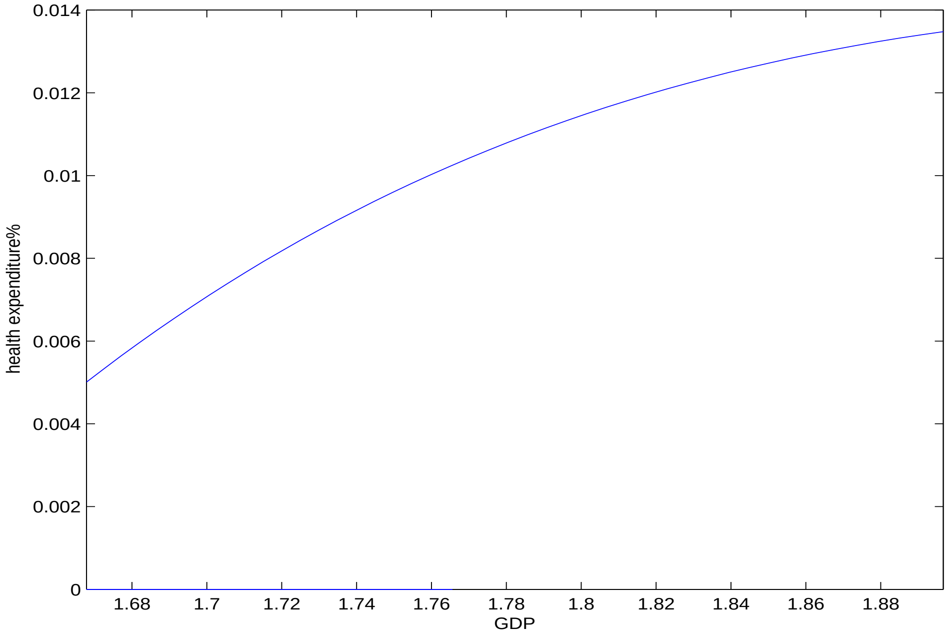

Figure 9: Share of health expenditure as b is varied

force is increasing ($\partial l'/\partial b > 0$), whereas the discounted marginal benefit of health to labor force $\partial l'_\theta/\partial b$ is unclear.

We vary b from 0.5% to 5%, which is the range of birth rates for countries in the world. So if $b \in [3.6\%, 5\%]$ there is only a disease free steady state, if $b \in [3.1\%, 3.6\%]$ there is an endemic steady state without health expenditure and if $b \in [0.5\%, 3.1\%]$ there is an endemic steady state with health expenditure. We can see from the Figure 8 that as b decreases, from the disease free steady state, the endemic steady state with no health expenditure emerges, and if it decreases further the endemic steady state with positive health expenditure emerges. The capital stock decreases, and as b decreases, it starts increasing due to the increasing health expenditures. This is mirrored in the effect on consumption. One of the interesting implications of this is that there will be a positive relationship between capital and hence, output and health capital, and consumption and health capital. Thus, one may be led to think that there is a causal relationship between income and health capital - that health is a luxury good. However, the link is through the birth rate. If we were to look at the relationship between net birth rate and health expenditure there would be the negative relationship which drives the link between income and health capital. The intuition is that as the net birth rate falls the cost of the marginal worker falling ill becomes higher and this leads to an increase in health expenditure and hence health capital.

In the literature (see Hall and Jones (2007)) an increase in longevity is interpreted as a decrease in the discount rate θ. Thus, the comparative statics exercise we do can be interpreted as studying the effect of increases in longevity on optimal health expenditures. We find that increases in longevity alone cannot explain the observed rising shares of health expenditures. However, when we look at the changes in the net birth rate (increases in b) we get the endogenous positive relationship between GDP and the share of health expenditures, see Figure 9. This is similar to the finding of Hall and Jones (2007). However, unlike their model we do not have to introduce a taste for health. They need to assume that the marginal utility of life extension does not decline as rapidly as that of consumption declines as income increases, i.e. there is a more rapid satiation of consumption than life extension. The mechanism in our model is more direct. Decreases in the net birth rate increases the marginal cost of an additional worker falling ill. The optimal response is to have increases in health expenditure, i.e. a more aggressive strategy to control the incidence of the disease. This interacts with the rising per capita capital stock and the increasing marginal product of capital which cause the GDP to rise as well.

7 The Conclusion

This paper develops a framework to study the interaction of infectious diseases and economic growth by establishing a link between the economic growth model and epidemiology models. We find that there are multiple steady states. Furthermore by examining the local stability we explore how the equilibrium properties of the model change as the parameters are varied. Although the model we present here is elementary, it provides a fundamental framework for considering more complicated model. It is important to understand the basic relationship between disease prevalence and economic growth before we go even further to consider more general models. The model also points the link between the health expenditures and income - both of which are endogenous - may be driven by parameters of population - as the birth rate drops the cost of a marginal worker becoming ill increases which leads to a negative relationship between population growth and health expenditures (controlling for disease induced mortality). An epidemiology model including control procedures, such as screening, tracing infectors, tracing infectives, post-treatment vaccination and general vaccination can be used to study the economic cost and benefit analysis of disease control. Moreover, the prevalence for many diseases varies periodically because of seasonal changes in the epidemiological parameters. It may also be one of the reasons of economic fluctuations. In addition the parameters can potentially be estimated and used to analyzed the economic effects of some specific infectious diseases in detail.

In a companion paper, Goenka and Liu (2010) we examine a discrete time formulation of a similar model. In that paper, however, there is only a one way interaction between the disease and the economy. The disease affects the labor force as in this model, but the labor supply by healthy individuals is endogenous and the epidemiology parameters are treated as biological constants. We find that under standard assumptions the dynamics of the model with and without endogenous labor are topologically conjugate. Thus, there may be no loss in generality in using an exogenous labor-leisure choice as in this paper. Under the simplifying assumption of a one-way interaction, the dynamics become two-dimensional and we can study the global dynamics. The key result is that as the disease becomes more infective, cycles and then eventually chaos emerges. Here, we endogenize the epidemiology parameters. Thus, it is a framework to study optimal health policy. However, the dynamical system becomes six dimensional and we have to restrict our analysis to local analysis of the steady state. In Goenka and Liu (2009) we incorporate learning-by-doing into a similar model as the current paper. We find that the growth rate is reduced by disease incidence. However, unlike Lucas (1988) the growth rate depend on all the economic parameters of the model as the human and physical capital choice depends on these. Thus, even small differences in the disease prevalence or in the economic fundamentals can have long run effects.

8 Appendix A: Existence of Optimal Solution

For the proof we also recall Mazur's Lemma (Renardy and Rogers (2004)) and the reverse Fatou's Lemma as follows.

Let F be a family of scalar measurable functions on a finite measure space (Ω, Σ, μ), F is called uniformly integrable if $\{\int_E |f(t)|\, d\mu,\ f \in F\}$ converges uniformly to *zero* when $\mu(E) \to 0$.

Dunford-Pettis Theorem: *Denote $L^1(\mu)$ the set of functions f such that $\int_\Omega |f|\, d\mu < \infty$ and K be a subset of $L^1(\mu)$. Then K is relatively weak compact if and only if K is uniformly integrable.*

When applying Fatou's Lemma to the non-negative sequence given by $g - f_n$, we get the following reverse Fatou's Lemma .

Fatou's Lemma: *Let f_n be a sequence of extended real-valued measurable functions defined on a measure space (Ω, Σ, μ). If there exists an integrable function g on Ω such that $f_n \le g$ for all n, then* $\limsup_{n\to\infty} \int_\Omega f_n d\mu \le \int_\Omega \limsup_{n\to\infty} f_n d\mu$.

Mazur's lemma shows that any weakly convergent sequence in a normed linear space has a sequence of convex combinations of its members that converges strongly to the same limit. Because strong convergence is stronger than pointwise convergence, it is used in our proof for the state variables to converge pointwise to the limit obtained from weak convergence.

Mazur's Lemma: *Let $(X, ||\ ||)$ be a normed linear space and let $(u_n)_{n\in N}$ be a sequence in X that converges weakly to some u^* in X. Then there exists a function $\mathcal{N} : N \to N$ and a sequence of sets of real numbers $\{\omega_{i(n)} \mid i = n, ..., \mathcal{N}(n)\}$ such that $\omega_{i(n)} \ge 0$ and $\sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} = 1$ such that the sequence $(v_n)_{n\in N}$ defined by the convex combination $v_n = \sum_{i=n}^{\mathcal{N}(n)} \omega_{i(n)} u_i$ converges strongly in X to u^**,i.e.,$||v_n - u^*|| \to 0$ as $n \to \infty$.

Proof of Lemma 1
Proof. Since $\lim_{k\to\infty} f_1(k,l) = 0$, for any $\zeta \in (0, \theta)$ there exist a constant A_0 such that $f(k,1) \le A_0 + \zeta k$. Hence we have

$$f(k,l) \le f(k,1) \le A_0 + \zeta k. \tag{27}$$

Since $\dot{k} = f(k,l) - c - m - k(\delta + b - d)$, it follows that

$$\dot{k} \le f(k,l) \le A_0 + \zeta k.$$

Multiplying by $e^{-\zeta\tau}$ we get $e^{-\zeta\tau}\dot{k} - \zeta k e^{-\zeta\tau} \le A_0 e^{-\zeta\tau}$. Thus,

$$e^{-\zeta t} k = \int_0^t \frac{\partial(e^{-\zeta\tau} k)}{\partial\tau} d\tau + k_0 \le \int_0^t A_0 e^{-\zeta\tau} d\tau = \frac{-A_0 e^{-\zeta t}}{\zeta} + \frac{A_0}{\zeta} + k_0.$$

This implies $k \le \frac{-A_0}{\zeta} + \frac{(A_0 + k_0\zeta)e^{\zeta t}}{\zeta}$. Thus, there exists a constant A_1 such that

$$k \le A_1 e^{\zeta t}. \tag{28}$$

Therefore, note that $\zeta < \theta$, $\int_0^\infty k e^{-\theta t} dt \le \int_0^\infty A_1 e^{(\zeta-\theta)t} dt < +\infty$.

Moreover, since $-\dot{k} \le \kappa k$ and $\dot{k} \le A_0 + \zeta k \le A_0 + \zeta A_1 e^{\zeta t}$ there exists a constant A_2 such that $|\ \dot{k}\ | \le A_2 e^{\zeta t}$. Thus

$$\int_0^\infty |\ \dot{k}\ |\ e^{-\theta t} dt < \int_0^\infty A_2 e^{(\zeta-\theta)t} dt < +\infty.$$

Because $-\dot{k} \leq \kappa k$ and $c = f(k,l) - \dot{k} - m - \delta k - k(b-d)$, it follows from (27) and (28) that

$$\begin{aligned} c &\leq f(k,l) + k(\kappa - \delta - b + d) \\ &\leq A_0 + (\kappa - \delta - b + d + \zeta)k \\ &\leq A_0 + (\kappa - \delta - b + d + \zeta)A_1 e^{\zeta t}. \end{aligned}$$

Thus, we can choose a constant A_3 large enough such that $c \leq A_3 e^{\zeta t}$ which implies

$$0 \leq \int_0^\infty c e^{-\theta t} dt \leq \int_0^\infty A_3 e^{(\zeta - \theta)t} dt < +\infty.$$

Similarly there exists A_4 such that $m \leq A_4 e^{\zeta t}$ and $m \in L^1(e^{-\theta t})$.

Now we prove $|\dot{h}|, h$ belong to the space $L^1(e^{-\theta t})$.

Since $g(m) \leq \pi_0 m$[12], there exists a constant B_1 such that $\dot{h} \leq g(m) \leq B_1 e^{\zeta t}$.

Clearly $h = \int_0^t \dot{h} d\tau + h_0 \leq \int_0^t B_1 e^{\zeta\tau} d\tau + h_0 = \frac{B_1}{\zeta} e^{\zeta t} - \frac{B_1}{\zeta} + h_0$ which means there exist B_2 such that $h \leq B_2 e^{\zeta t}$ or $h \in L^1(e^{-\theta t})$. Moreover $-\dot{h} \leq (\delta + b - d)h$ because $g(m) \geq 0$. Therefore $-\dot{h} \leq (\delta + b - d)B_2 e^{\zeta t}$. So$|\dot{h}| \leq B_3 e^{\zeta t}$ with $B_3 = \max\{B_1, (\delta + b - d)B_2\}$. Thus $|\dot{h}| \in L^1(e^{-\theta t})$.

Obviously, $l \in L^\infty$ and $\lim_{t\to\infty} l e^{-\theta t} = 0$. It follows that

$$\int_0^\infty \dot{l} e^{-\theta t} dt = -l_0 + \theta \int_0^\infty l e^{-\theta t} dt \leq -l_0 + \theta \int_0^\infty e^{-\theta t} dt < +\infty.$$

Finally, we will prove that $|\dot{l}| \in L^1(e^{-\theta t})$. Since $0 \leq l \leq 1$ and $\alpha(h)$ is decreasing, we have

$$\begin{aligned} |\dot{l}| &\leq b + |\gamma(h)| + |\alpha(h)| \\ &\leq b + |\gamma(h)| + |\alpha(0)| \\ &= \gamma(h) + b + \alpha(0). \end{aligned}$$

Since $\lim_{h\to\infty} \gamma'(h) \to 0$, there exists a constant B_4 such that $\gamma(h) \leq B_4 + \zeta h \leq B_4 + \zeta B_2 e^{\zeta t}$. Thus, there exists B_5 such that$|\dot{l}| \leq B_5 e^{\zeta t}$. This implies $|\dot{l}| \in L^1(e^{-\theta t})$. We have proven that $\mathcal{K}$ is uniformly bounded on $L^1(e^{-\theta t})$.

Moreover, $\lim_{a\to\infty} \int_a^\infty k e^{-\theta t} dt \leq \lim_{a\to\infty} \int_a^\infty A_1 e^{(\zeta-\theta)t} dt = 0$. This property is true for other variables in $\mathcal{K}$. Therefore $\mathcal{K}$ satisfies Dunford-Pettis theorem and it is relatively compact in the weak topology $\sigma(L^1(e^{-\theta t}), L^\infty)$. ■

9 Appendix B: Center Manifold Calculation

Here, we introduce the procedure of calculating center manifold instead of the calculation part itself. We use $\dot{x} = g(x,b)$ to denote the dynamic system, where $x = (k,h,l,c)^T \in \Re_+^4$, and $g : \Re_+ \times \Re_+^4 \to \Re_+^4$ is the vector field. Moreover, we use x^* to denote its equilibrium point, and so $g(x^*, b) = 0$. Bifurcation occurs when $b^* = \overline{\alpha} - \underline{\gamma}$. We assume $g(x,b)$ to be at least C^5. We follow the procedure given by Wiggins (2003) and Kribs-Zaleta (2002):

1. Using $\tilde{x} = x - x^*$ and $\tilde{b} = b - b^*$, we transform the dynamical system into $\dot{\tilde{x}} = g(\tilde{x} + x^*, \tilde{b} + b^*)$ with the equilibrium point $\tilde{x}^* = 0$ and bifurcation point $\tilde{b}^* = 0$. Then we linearize the system at point 0 to get $\dot{\tilde{x}} = D_x g(x^*, b^*)\tilde{x} + D_b g(x^*, b^*)\tilde{b} + R(\tilde{x}, \tilde{b})$, where $R(\tilde{x}, \tilde{b})$ is the high order term;

2. Let $A = D_x g(x^*, b^*)$, $B = D_b g(x^*, b^*)$ and calculate matrix A's eigenvalues, corresponding eigenvectors matrix TA (placing the eigenvector corresponding to zero eigenvalue first) and its inverse TA^{-1}. By transformation $\tilde{x} = TA \cdot y$, we get $y = TA^{-1} \cdot A \cdot TA \cdot y + TA^{-1} \cdot B \cdot \tilde{b} + TA^{-1} \cdot R(TA \cdot y, \tilde{b})$, where $TA^{-1} \cdot A \cdot TA$ is its Jordan canonical form;

[12] If $\lim_{m_t\to\infty} g' = 0$ holds, there exists a constant B_0 such that $g(m) \leq B_0 + \zeta m$ where $\zeta \in (0, \theta)$.Thus, $\dot{h} \leq g(m) \leq \zeta A_4 e^{\zeta t}$.

3. We separate y into two vectors y_1, the first term, and y_2, the rest terms, and then we can rewrite the system as:

$$y_1' = \Gamma_1 y_1 + \tilde{R}_1(TA \cdot y, \tilde{b})$$
$$y_2' = \Gamma_2 y_2 + \tilde{R}_2(TA \cdot y, \tilde{b});$$

Since $TA^{-1} \cdot B \neq 0$, we separate it into two vectors Δ_1 with only one element, and Δ_2 with the rest, and form a system as:

$$\begin{pmatrix} y_1 \\ \tilde{b} \\ y_2 \end{pmatrix}' = \underbrace{\begin{pmatrix} \Gamma_1 & \Delta_1 & 0 \\ 0 & 0 & 0 \\ 0 & \Delta_2 & \Gamma_2 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} y_1 \\ \tilde{b} \\ y_2 \end{pmatrix}}_{y_b} + \underbrace{\begin{pmatrix} \tilde{R}_1(TA \cdot y, \tilde{b}) \\ 0 \\ \tilde{R}_2(TA \cdot y, \tilde{b}) \end{pmatrix}}_{\tilde{R}_b(TA \cdot y, \tilde{b})};$$

4. In order to put matrix C into Jordan canonical form, we make another linear transformation $y_b = TC \cdot z$, and get $\dot{z} = TC^{-1} \cdot C \cdot TC \cdot z + TC^{-1} \cdot \tilde{R}_b(TA \cdot TC \cdot z, \tilde{b})$, where $z = (z_1, \tilde{b}, z_2, z_3, z_4)$. Therefore, we can now write the system as:

$$\begin{aligned} z_1' &= \Pi_1 z_1 + \hat{R}_1(z_1, z_2, z_3, z_4, \tilde{b}) \\ z_2' &= \Pi_2 z_2 + \hat{R}_2(z_1, z_2, z_3, z_4, \tilde{b}) \\ z_3' &= \Pi_3 z_3 + \hat{R}_3(z_1, z_2, z_3, z_4, \tilde{b}) \\ z_4' &= \Pi_4 z_4 + \hat{R}_4(z_1, z_2, z_3, z_4, \tilde{b}) \\ \tilde{b}' &= 0; \end{aligned}$$

5. Take $z_i = h_i(z_1, \tilde{b}) \quad (i = 2, 3, 4)$ as a polynomial approximation to the center manifold, and differentiate both sides w.r.t. t:

$$\Pi_i z_i + \hat{R}_i(z_1, h_2, h_3, h_4, \tilde{b}) = D_{z_1} h_i(z_1, \tilde{b})[\Pi_1 z_1 + \hat{R}_1(z_1, h_2, h_3, h_4, \tilde{b})].$$

   And then solve for the center manifold by equating the coefficient of each order;

6. Finally, we write the differential equation for the dynamical system on the center manifold by substituting $h_i(z_1, \tilde{b})$ in $\hat{R}_1(z_1, z_2, z_3, z_4, \tilde{b})$, and get the system:

$$\begin{aligned} z_1' &= \Pi_1 z_1 + \hat{R}_1(z_1, h_2(z_1, \tilde{b}), h_3(z_1, \tilde{b}), h_4(z_1, \tilde{b}), \tilde{b}) \\ \tilde{b}' &= 0. \end{aligned}$$

However, in our economic epidemiology model as dynamics of l and h is independent of the rest of system dynamics, we could just simply calculate their dynamics on the center manifold, which is given by:

$$\dot{z_1} = \overline{\alpha} z_1 (z_1 - \frac{1}{\overline{\alpha}} \tilde{b}).$$

10 Appendix C: Stability Analysis

For the determinants calculation, we have:

$$\begin{aligned} a_{11} &= \theta, a_{13} = f_2^*, a_{14} = a_{15} = -1, a_{22} = -\delta - (b - d), a_{25} = g'^*, a_{32} = (1 - l^*)(\gamma'^* - \alpha'^* l^*) \\ a_{33} &= b + \gamma^* - \alpha^*, a_{41} = c^* f_{11}^*, a_{43} = c^* f_{12}^*, a_{51} = -f_{11}^* \frac{g'^*}{g''^*}, a_{52} = \frac{f_1^*(\gamma''^* - \alpha''^* l^*)}{\gamma'^* - \alpha'^* l^*} \frac{g'^*}{g''^*} \\ a_{53} &= \left(\frac{f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*)}{(1 - l^*)(\gamma'^* - \alpha'^* l^*)} - f_{12}^* \right) \frac{g'^*}{g''^*}, a_{54} = \frac{f_1^*}{c^*} \frac{g'^*}{g''^*}, a_{55} = f_1^*, a_{56} = \frac{f_1^*}{\lambda_3^*} \frac{g'^*}{g''^*} \\ a_{61} &= -\frac{f_{12}^*}{c^*}, a_{62} = -\lambda_3^*(2\alpha'^* l^* - \gamma'^* - \alpha'^*), a_{63} = -\frac{f_{22}^*}{c^*} - 2\lambda_3^* \alpha^*, a_{64} = \frac{f_2^*}{c^{*2}}, a_{66} = \frac{f_2^*}{c^* \lambda_3^*}. \end{aligned}$$

Let us denote $X = \delta + b - d, Y = \alpha - \gamma - b$, we have the following relation will be used in the calculation.

$$\lambda_3^* = \frac{f_2^*}{c^*(\theta - 2\alpha^* l^* + b + \gamma^* + \alpha^*)} = \frac{f_2^*}{c^*(\theta + Y)}$$

$$\begin{aligned}
a_{22} &= -X, a_{33} = -Y & (29)\\
a_{55} &= f_1^* = \theta + (\delta + b - d) = \theta + X & (30)\\
a_{66} &= \frac{f_2^*}{c^* \lambda_3^*} = \theta - b - \gamma^* + \alpha^* = \theta + Y & (31)\\
a_{66} a_{54} &= a_{56} a_{64} = \frac{f_1^* f_2^*}{\lambda_3^*} \frac{g'^*}{g''^*} \frac{1}{c^{*2}} & (32)\\
-a_{41} a_{54} &= c^* f_{11}^* \frac{f_1^*}{c^*} \frac{g'^*}{g''^*} = f_{11}^* \frac{g'^*}{g''^*} f_1^* = a_{51} a_{55} = a_{51}(X + \theta) & (33)
\end{aligned}$$

Since

$$f_1^* = f_2^* g'^* \frac{(1 - l^*)(\gamma'^* - \alpha'^* l^*)}{\theta + \alpha^* - b - \gamma^*} = \frac{a_{13} a_{25} a_{32}}{a_{66}}$$

we also get

$$\begin{aligned}
a_{55} a_{66} &= a_{13} a_{25} a_{32} = (\theta + X)(\theta + Y). & (34)\\
a_{41} a_{56} a_{64} &= a_{41} a_{54} a_{66} = -a_{51}(X + \theta)(Y + \theta). & (35)
\end{aligned}$$

As $\lambda_3^* c^* = \frac{f_2^*}{a_{66}}$, we have

$$\begin{aligned}
a_{56} a_{61} &= \frac{f_1^*}{\lambda_3^*} \frac{g'^*}{g''^*} \left(-\frac{f_{12}^*}{c^*}\right) = \left(\frac{-g'^* f_{12}^*}{g''^*}\right) . \frac{(X + \theta)}{\lambda_3^* c^*} = \\
&= \left(\frac{-g'^* f_{12}^*}{g''^*}\right) . \frac{a_{55} a_{66}}{f_2^*} = \left(\frac{-g'^* f_{12}^*}{g''^*}\right) . \frac{f_2 a_{25} a_{32}}{f_2^*} = \left(\frac{-g'^* f_{12}^*}{g''^*}\right) .a_{25} a_{32}.
\end{aligned}$$

Thus,

$$\begin{aligned}
&a_{25} a_{56} a_{62} + a_{25} a_{32} a_{53} - a_{56} a_{61} = a_{25} a_{56} a_{62} + a_{25} a_{32} a_{53} + \frac{g'^* f_{12}^*}{g''^*} a_{25} a_{32}\\
= \; & a_{25}[-\frac{f_1^*}{\lambda_3^*} \frac{g'^*}{g''^*} \lambda_3^* (2\alpha'^* l^* - \gamma'^* - \alpha'^*) + (1 - l^*)(\gamma'^* - \alpha'^* l^*) \left(\frac{f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*)}{(1 - l^*)(\gamma'^* - \alpha'^* l^*)} - f_{12}^*\right) \frac{g'^*}{g''^*}\\
&+ \frac{g'^* f_{12}^*}{g''^*} (1 - l^*)(\gamma'^* - \alpha'^* l^*)].
\end{aligned}$$

Hence,

$$a_{32} a_{25} a_{53} + a_{56} a_{25} a_{62} - a_{56} a_{61} = 0. \qquad (36)$$

$$a_{54} a_{43} a_{25} a_{32} = \frac{f_1^*}{c^*} \frac{g'^*}{g''^*} c^* f_{12}^* a_{25} a_{32} = (-f_1^*)(\frac{-g'^* f_{12}^*}{g''^*}) a_{25} a_{32} = -(X + \theta) a_{56} a_{61} \qquad (37)$$

The characteristic equation, $|\Lambda I - \mathcal{J}_3| = 0$ can be expanded and written as a polynomial of λ as

$$P(\Lambda) = \Lambda^6 - D_1 \Lambda^5 + D_2 \Lambda^4 - D_3 \Lambda^3 + D_4 \Lambda^2 - D_5 \Lambda + D_6 = 0$$

where the D_i are the sum of the i-th order minors about the principal diagonal of $\mathcal{J}_3$.

Thus, for D_1 we have

$$D_1 = a_{11} + a_{22} + a_{33} + a_{44} + a_{55} + a_{66}$$

which are first order minors about the diagonal.

$$\begin{aligned}
D_2 &= a_{11}(a_{22} + a_{33} + a_{55} + a_{66}) + a_{41} + a_{51}\\
&\quad + a_{22}(a_{33} + a_{55} + a_{66}) - a_{52} a_{25} + a_{33}(a_{55} + a_{66}) + a_{55} a_{66}
\end{aligned}$$

Replace

$$\begin{aligned}
a_{22}+a_{33}+a_{55}+a_{66} &= 2\theta \\
a_{33}+a_{55}+a_{66} &= 2\theta+X \\
a_{55}+a_{66} &= 2\theta+(\delta+b-d)+(\alpha-\gamma-b) \\
&= 2\theta+X+Y
\end{aligned}$$

we get

$$\begin{aligned}
D_2 &= 2\theta^2 - X(2\theta+X) - Y(2\theta+X+Y) + (\theta+X)(\theta+Y) + a_{41} + a_{51} - a_{52}a_{25} \\
&= 3\theta^2 - \theta(X+Y) - X^2 - Y^2 + a_{41} + a_{51} - a_{52}a_{25}.
\end{aligned}$$

$$\begin{aligned}
D_3 = \; & a_{11}a_{22}a_{33} + a_{22}a_{41} + a_{11}a_{22}a_{55} + a_{22}a_{51} - a_{11}a_{25}a_{52} + a_{11}a_{22}a_{66} \\
& + a_{33}a_{41} + a_{11}a_{33}a_{55} + a_{33}a_{51} + a_{11}a_{33}a_{66} - a_{41}a_{54} + a_{41}a_{55} \\
& + a_{41}a_{66} + a_{11}a_{55}a_{66} - a_{56}a_{61} + a_{51}a_{66} \\
& + a_{22}a_{33}a_{55} + a_{32}a_{25}a_{53} - a_{25}a_{52}a_{33} + a_{22}a_{33}a_{66} \\
& + a_{22}a_{55}a_{66} + a_{25}a_{56}a_{62} - a_{25}a_{52}a_{66} + a_{33}a_{55}a_{66}
\end{aligned}$$

We keep only $a_{41}, a_{25}a_{52}, a_{25}a_{56}a_{62}, a_{25}a_{32}a_{53}, a_{56}a_{61}$ in the expression and replace $a_{11}, a_{22}, a_{55}, a_{66}$ via X, Y, θ from (29)-(31) and using (33)

$$\begin{aligned}
D_3 = \; & \theta[\theta^2 - 2\theta(X+Y) - 2(X^2+Y^2)] + 2\theta a_{41} + 2\theta a_{51} - 2\theta a_{25}a_{52} \\
& + a_{32}a_{25}a_{53} + a_{25}a_{56}a_{62} - a_{61}a_{56}.
\end{aligned}$$

It follows from (36) we have $a_{32}a_{25}a_{53} + a_{25}a_{56}a_{62} - a_{61}a_{56} = 0$.

Hence

$$D_3 = \theta[\theta^2 - 2\theta(X+Y) - 2(X^2+Y^2)] + 2\theta a_{41} + 2\theta a_{51} - 2\theta a_{25}a_{52}.$$

$$\begin{aligned}
D_6 = \; & a_{66}[a_{55}a_{22}a_{33}a_{41} - a_{25}a_{32}a_{43}a_{51} - a_{25}a_{33}a_{41}a_{52} + a_{25}a_{32}a_{41}a_{53} \\
& - a_{25}a_{54}a_{11}a_{32}a_{43} + a_{25}a_{54}a_{13}a_{32}a_{41} - a_{54}a_{22}a_{33}a_{41}] + a_{56}a_{64}a_{22}a_{33}a_{41} \\
& + a_{56}a_{25}[a_{64}a_{11}a_{32}a_{43} - a_{64}a_{13}a_{32}a_{41} + a_{32}a_{43}a_{61} + a_{33}a_{41}a_{62} - a_{32}a_{41}a_{63}]
\end{aligned}$$

D_4, D_5 are explicitly computed and the signs depend on X, Y, θ and $a_{63} = -\frac{f_{22}^*}{c^*} - 2\lambda_3^*\alpha^*$.

By replacing $a_{11}, a_{22}, a_{55}, a_{66}$ via X, Y, θ and using (29)-(37) we have :

$$\begin{aligned}
D_4 = \; & a_{41}[2XY + a_{25}a_{52} + X^2 + 3X\theta - Y\theta - Y^2 + \theta^2] + a_{51}[-2XY - 2\theta X - 3\theta Y - \theta^2 - Y^2 + X + \theta] \\
& + 2XY(X+\theta)(Y+\theta) + \theta XY(X+\theta) + (\theta+Y)\theta XY - \theta Y(Y+\theta)(X+\theta) \\
& + [-\theta^2 + \theta Y + Y^2]a_{25}a_{52} + (\theta - Y)a_{56}a_{25}a_{62} + (Y+2\theta)a_{25}a_{32}a_{53} + (2X+Y+\theta)a_{56}a_{61} \\
& - a_{56}a_{25}a_{32}a_{63}.
\end{aligned}$$

Using (29)-(37) we get

$$\begin{aligned}
D_5 = \; & a_{41}[-X\theta(X+\theta) - Y\theta(Y+\theta)] - a_{51}(X+Y+2\theta)((X+Y)\theta+\theta^2)) + a_{51}X(Y+\theta)(X+\theta) \\
& + (Y+\theta)\theta XY(X+\theta) + [(Y+\theta)\theta - \theta a_{41}]Ya_{25}a_{52} \\
& + (Y+\theta)\theta a_{25}a_{32}a_{53} + [(X+Y)\theta+\theta^2) + \theta(X+\theta)]a_{56}a_{61} - \theta Y a_{56}a_{25}a_{62} \\
& + a_{41}(a_{25}a_{32}a_{53} + a_{56}a_{25}a_{62}) - a_{25}a_{32}a_{43}a_{51} \\
& - a_{56}\theta a_{25}a_{32}a_{63}
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
a_{41}(a_{25}a_{32}a_{53} + a_{56}a_{25}a_{62}) - a_{25}a_{32}a_{43}a_{51} &= \\
a_{41}a_{56}a_{61} - a_{25}a_{32}a_{43}a_{51} &= \\
\frac{a_{54}a_{41}a_{56}a_{61} + (X+\theta)a_{56}a_{61}a_{51}}{a_{54}} &= 0
\end{aligned}$$

Thus $D_5 = A_0 - a_{56}\theta a_{25}a_{32}a_{63}$ where $A_0 > 0$ but the sign of D_5 is ambiguous since we do not know sign of a_{63}.

Proof of Proposition 4
Proof.

By using (32) we rewrite D_6 as follows :

$$\begin{aligned} D_6 &= a_{66}(a_{55}a_{22}a_{33}a_{41} - a_{25}a_{33}a_{41}a_{52}) + a_{56}a_{25}(a_{32}a_{43}a_{61} + a_{33}a_{41}a_{62} - a_{32}a_{41}a_{63}) + a_{66}a_{25}a_{32}(a_{41}a_{53} - a_{43}a_{51}) \\ & a_{11}a_{32}a_{43}a_{25}(a_{56}a_{64} - a_{66}a_{54}) + a_{13}a_{32}a_{41}a_{25}(a_{66}a_{54} - a_{56}a_{64}) + a_{22}a_{33}a_{41}(a_{56}a_{64} - a_{66}a_{54}) \end{aligned}$$

$$= a_{66}(a_{55}a_{22}a_{33}a_{41} - a_{25}a_{33}a_{41}a_{52}) + a_{56}a_{25}[a_{32}(a_{43}a_{61} - a_{41}a_{63}) + a_{33}a_{41}a_{62}] + a_{66}a_{25}a_{32}(a_{41}a_{53} - a_{43}a_{51}).$$

Obviously $a_{66}(a_{55}a_{22}a_{33}a_{41} - a_{25}a_{33}a_{41}a_{52}) < 0$.

Note that $a_{56}a_{25} < 0$, $a_{33} = (b + \gamma^* - \alpha^*) = \alpha^*(l^* - 1)$ and by concavity of f, $f_{12}^{*2} < f_{11}^* f_{22}^*$ we have

$$\begin{aligned} & a_{56}a_{25}[a_{32}(a_{43}a_{61} - a_{41}a_{63}) + a_{33}a_{41}a_{62}] \\ = & a_{56}a_{25}[a_{32}(-f_{12}^{*2} - c^* f_{11}^*(-\frac{f_{22}^*}{c^*} - 2\lambda_3^*\alpha^*) - a_{33}c^* f_{11}^*\lambda_3^*(2\alpha'^* l^* - \gamma'^* - \alpha'^*)] \\ = & a_{56}a_{25}[a_{32}(-f_{12}^{*2} + f_{11}^* f_{22}^* + 2c^* f_{11}^*\lambda_3^*\alpha^*) - a_{33}c^* f_{11}^*\lambda_3^*(2\alpha'^* l^* - \gamma'^* - \alpha'^*)] \\ < & a_{56}a_{25}[c^* f_{11}^*\lambda_3^*(2\alpha^* a_{32} - a_{33}(2\alpha'^* l^* - \gamma'^* - \alpha'^*))] \\ = & a_{56}a_{25}[c^* f_{11}^*\lambda_3^*(2\alpha^*(1 - l^*)(\gamma'^* - \alpha'^* l^*) - \alpha^*(l^* - 1)(2\alpha'^* l^* - \gamma'^* - \alpha'^*))] \\ = & a_{56}a_{25}[c^* f_{11}^*\lambda_3^*(\alpha^*(1 - l^*)(2\gamma'^* - 2\alpha'^* l^* + 2\alpha'^* l^* - \gamma'^* - \alpha'^*)] \\ = & a_{56}a_{25}c^* f_{11}^*\lambda_3^*(\alpha^*(1 - l^*)(\gamma'^* - \alpha'^*) \\ = & \frac{f_1^*}{\lambda_3^*}\frac{g'^{*2}}{g''^*}c^* f_{11}^*\lambda_3^*(\alpha^*(1 - l^*)(\gamma'^* - \alpha'^*) \end{aligned}$$

and

$$\begin{aligned} & a_{66}a_{25}a_{32}(a_{41}a_{53} - a_{43}a_{51}) \\ = & a_{66}a_{25}a_{32}[c^* f_{11}^*\frac{g'^*}{g''^*}(\frac{f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*)}{(1 - l^*)(\gamma'^* - \alpha'^* l^*)} - f_{12}^*) + c^* f_{12}^* f_{11}^*\frac{g'^*}{g''^*}] \\ = & a_{66}a_{25}c^* f_{11}^*\frac{g'^*}{g''^*}f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*) \\ = & \frac{f_2^*}{c^*\lambda_3^*}g'^* c^* f_{11}^*\frac{g'^*}{g''^*}f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*). \end{aligned}$$

Hence

$$\begin{aligned} & a_{56}a_{25}[a_{32}(a_{43}a_{61} - a_{41}a_{63}) + a_{33}a_{41}a_{62}] + a_{66}a_{25}a_{32}(a_{41}a_{53} - a_{43}a_{51}) \\ < & \frac{f_1^*}{\lambda_3^*}\frac{g'^{*2}}{g''^*}c^* f_{11}^*\lambda_3^*\alpha^*(1 - l^*)(\gamma'^* - \alpha'^*) + \frac{f_2^*}{c^*\lambda_3^*}g'^* c^* f_{11}^*\frac{g'^*}{g''^*}f_1^*(2\alpha'^* l^* - \alpha'^* - \gamma'^*) \\ = & \frac{g'^{*2}}{g''^*}f_1^* c^* f_{11}^*[\alpha^*(1 - l^*)(\gamma'^* - \alpha'^*) + \frac{f_2^*}{c^*\lambda_3^*}(2\alpha'^* l^* - \alpha'^* - \gamma'^*)] \\ = & \frac{g^{2\prime *}}{g''^*}f_1^* c^* f_{11}^*[\alpha^*(1 - l^*)(\gamma'^* - \alpha'^*) + (\theta - b - \gamma^* + \alpha^*)(2\alpha'^* l^* - \alpha'^* - \gamma'^*)]. \\ = & \frac{g^{2\prime *}}{g''^*}f_1^* c^* f_{11}^*[(\alpha^* - b - \gamma^*)(\gamma'^* - \alpha'^*) + (\theta - b - \gamma^* + \alpha^*)(\frac{2\alpha'^*(b + \gamma^*)}{\alpha^*} - \alpha'^* - \gamma'^*)] \\ < & 0 \text{ by A.10(i).} \end{aligned}$$

The proof is complete. ■

Proof of Lemma 4

Proof.

$$\begin{aligned} D_2 &= a_{11}(a_{22}+a_{33}+a_{55}+a_{66})+a_{41}+a_{51} \\ &\quad +a_{22}(a_{33}+a_{55}+a_{66})-a_{52}a_{25}+a_{33}(a_{55}+a_{66})+a_{55}a_{66} \\ &= 2\theta^2 - X(2\theta+X) - Y(2\theta+X+Y)+(\theta+X)(\theta+Y)+a_{41}+a_{51}-a_{52}a_{25} \\ &= [3\theta^2-\theta(X+Y)-X^2-Y^2]+a_{41}+a_{51}-a_{52}a_{25}. \end{aligned}$$

and

$$D_3 = \theta[\theta^2-2\theta(X+Y)-2(X^2+Y^2)]+2\theta a_{41}+2\theta a_{51}-2\theta a_{25}a_{52}$$

Thus

$$\begin{aligned} & D_1D_2-D_3 = 3\theta[3\theta^2-\theta(X+Y)-X^2-Y^2]-\theta[\theta^2-2\theta(X+Y)-2(X^2+Y^2)] \\ & +3\theta[a_{41}+a_{51}-a_{52}a_{25}]-[2\theta a_{41}+2\theta a_{51}-2\theta a_{25}a_{52}] \\ = \; & \theta[8\theta^2-\theta(X+Y)-X^2-Y^2]+[\theta a_{41}+\theta a_{51}-\theta a_{25}a_{52}] \end{aligned}$$

which is negative since $\theta a_{41}+\theta a_{51}-\theta a_{25}a_{52}<0$ and $8\theta^2-\theta(X+Y)-X^2-Y^2<0$ due to *A.10 (ii)*. Furthermore $D_2<0, D_3<0$ since

$$\theta^2-2\theta(X+Y)-2(X^2+Y^2) < 3\theta^2-\theta(X+Y)-X^2-Y^2 < 8\theta^2-\theta(X+Y)-X^2-Y^2 < 0.$$

■

References

1. Acemoglu, D. and Johnson, S. (2007) Disease and development: The effect of life expectancy on economic growth, *Journal of Political Economy*, 115: 925-985.

2. Anderson, R.M. and May, R.M. (1991) *Infectious Diseases of Humans*, Oxford: Oxford Science Publications.

3. Arrow, K.J. and Kurz, M. (1970) *Public investment, the rate of return, and optimal fiscal policy*, Baltimore: The Johns Hopkins Press.

4. Ashraf, Q.H., Lester, A., and Weil, D.N. (2009) When does improving health raise GDP? In D. Acemoglu, K. Rogoff, and M. Woodford (Eds.) *NBER Macroeconomics Annual 2008*, Chicago: The University of Chicago Press.

5. Barrett, S. and Hoel, M. (2004) Optimal disease eradication, mimeo. University of Oslo.

6. Bell, C., Devarajan, S. and Gersbach, H., (2003) The long-run economic costs of AIDS: theory and an application to South Africa, *Working paper*, The World Bank.

7. Bleakley, H. (2007) Disease and development: Evidence from hookworm eradication in the American South, *Quarterly Journal of Economics*, 122: 73-117.

8. Bloom, D.E., Canning, D. and Fink, G. (2009) Diseases and development revisited, NBER Working Paper No. 15137.

9. Brander, J.A. and Dowrick, S. (1994) The Role of Fertility and Population in Economic Growth: Empirical Results: From Aggregate Cross-National Data, *Journal of Population Economics*, 7: 1-25.

10. Chichilnisky, G. (1981) Existence and characterization of optimal growth paths including models with non-convexities in utilities and technologies, *Review of Economic Studies*, 48, 51-61.

11. d'Albis, H. and Augeraud-Véron, E. (2008) The optimal prevention of epidemics, Mimeo., LERNA, Toulouse.

12. d'Albis, H., Gourdel, P. and Le Van, C. (2008) Existence of solutions in continuous-time optimal growth models, *Economic Theory*, 37(2): 321-333.

13. Delfino, D. and Simmons, P.J. (2000) Positive and normative issues of economic growth with infectious diseases, Discussion Papers in Economics 2000/48, University of York.

14. Dunford, N. and J.T. Schwartz, (1967) *Linear operators*, New York: Interscience Publishers.

15. Geoffard, P.-Y. and Philipson, T. (1997) Rational epidemics and their rational control, *International Economic Review*, 37: 603-624.

16. Gersovitz, M. and Hammer, J.S. (2004) The economical control of infectious diseases, *The Economic Journal* : 1-27.

17. Goenka, A., and Liu, L. (2010) Infectious diseases and endogenous fluctuations, *Economic Theory*, Published online: August 2, 2010

18. Goenka, A., and Liu, L. (2009) Infectious diseases and endogenous growth, National University of Singapore, mimeo.

19. Hall, R.E. and Jones, C. (2007) The value of life and the rise in health spending, *Quarterly Journal of Economics*, 122 (1): 39-72.

20. Hethocote, H.W. (2000) The mathematics of infectious diseases, *SIAM Review* 42: 599-653.

21. Hethcote, H.W. (2009) The basic epidemiology models, in S. Ma and Y. Xia (Eds.) *Mathematical understanding of infectious disease dynamics*, Singapore: World Scientific.

22. Kribs-Zaleta, C.M. (2002) Center manifolds and normal forms in epidemic models, in C. Castillo-Chavez, *et al* (Eds.) *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction* : 269-86.

23. Kremer, M. (1996) Integrating Behavioral Choice into Epidemiological Models of the AIDS Epidemic, *Quarterly Journal of Economics*, 111: 549-573.

24. Kremer, M. and Miguel, E. (2004) Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, *Econometrica* 72: 159-217.

25. Le Van, C. and Vailakis, Y (2003) Existence of a competitive equilibrium in a one sector growth model with irreversible investment, *Economic Theory*, 22: 743-771.

26. Lucas, Robert E. (1988) On the mechanics of economic development, *Journal of Monetary Economics* 22: 3-42.

27. Philipson, T. (2000) Economic epidemiology and infectious diseases, in J. Newhouse and T. Culyer (Eds.) *Handbook of Health Economics*, New York: North-Holland.

28. Renardy, M and Rogers, R.C. (2004) *An introduction to partial differential equations*. Texts in Applied Mathematics 13 (Second edition ed.). New York: Springer-Verlag.

29. Romer, P. (1986) Cake eating, chattering, and jumps: existence results for variational problems, *Econometrica*, 54, 897-908.

30. Weil, D. (2007) Accounting for the effect of health on economic growth, *Quarterly Journal of Economics*, 122.

31. Wiggins, S. (2003) *Introduction to applied nonlinear dynamical systems and chaos*, New York: *Springer-Verlag*.

32. Young, A. (2005) The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations, *Quarterly Journal of Economics* 120: 243-266.