THE UNIVERSITY of York

Discussion Papers in Economics

No. 2000/48

Positive and Normative Issues of Economic Growth with Infectious Disease

by

Doriana Delfino and Peter J. Simmons

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Positive and Normative Issues of Economic Growth with Infectious Disease

D. Delfino*and P.J. Simmons†
Department of Economics and Related Studies, University of York.

November 17, 2000

Abstract

This paper uses a variant of the Lotka-Volterra system explaining the dynamic interaction between populations of infected and healthy individuals in which the demographic and epidemiological parameters (the net healthy birth rate, the death rate of the infected and the infection rate) are functions of economic variables and some simple economic growth models to examine deterministic growth paths of the system with an exogenous savings rate. Demographic-epidemiological parameters depend on productive capital which combined with healthy workers produces output. We find that there are generally multiple steady states. The system usually converges to a steady state in which the economy moderates the disease. If capital accumulation is set optimally to maximise welfare then there may be multiple steady states and optimal growth paths generally display four dimensional saddle point properties. Extensions of the framework to allow for density dependent infection, recovery from the disease and alternative social welfare functions are analysed.

Keywords: economic growth, infectious disease, dynamic optimal control

JEL Nos: O11, O41, I10, I18

*dd109@york.ac.uk
†ps1@york.ac.uk

In various historical periods there have been epidemics of different global infectious diseases which have had dramatic effects on population and indeed society of the time e.g. the Plague in Western Europe and the Middle East over the 500 year period from 1350; the nineteenth century and earlier examples of Tuberculosis; small pox and cholera in both Western Europe and Asia (Watts, 1997). There are contemporary worries about the epidemics of HIV/AIDS (WHO, 1996; WHO, 1997; WHO, 2000). Some of these diseases are no longer endemic in the population. Sometimes this control has been due to improvements in medical knowledge and technology e.g. the development of public health responsibilities of governments together with increasing awareness of the transmission mechanism of the disease in question (the control of cholera in India has largely come through this route). Sometimes it has been due to specific targeted regulatory action against the disease e.g. in European and especially Italian responses to the plague and to leprosy until the 18th century, there was an enforced isolation policy of infected individuals. Sometimes the increase in knowledge that followed economic growth led to an innovation which was targeted on the disease e.g. discovery in the 18th century in the U.K. of the merits of vaccination against the plague. Often control of the disease has been in major part due to education of the population at large about preventive public health measures e.g. cholera control in India. Historically, apart from segregation policy, economic growth has been one of the major stimuli to control the diseases. In this sense it could be argued that direct policy against disease has not been particularly successful over the broad historical picture (Alvi et al, 1998; De Cock and Chaison, 1999; Kimerling et al., 1999; Netto, 1999). Despite this, contemporary concerns with disease control especially in the Third World are usually framed in terms of specific targeted policies rather than on trying to promote the general level of prosperity to control the disease e.g. the TB-WHO programme (the DOTS strategy).

The causal direction outlined above concerns the effect of the economy on the population structure. However, there are also severe effects of the population structure on the economy e.g. the import of smallpox to Mid-America in the 16th century amongst the local population who had no immune resistance to the disease led to their decimation which in turn eliminated the local labour force for working the silver mines; the plague in the Middle East that had the effects of reducing the productive work force so heavily that localised famines emerged (Bartel and Taubman, 1979; Ettner, 1996; Lee, 1982; Luft, 1975; Watts, 1997). The interaction between the economy and the epidemiology of the population is thus two way with causal links in both directions.

Historically, when these infectious diseases were endemic they often generated epidemic cycles. For example a common pattern in medieval England was for a geographical area to succumb to an outburst of plague over a period of five months or so often concentrated at particular times of year but then the disease would die away, subsequently breaking out again. In the population and epidemiological literature, there are various frameworks for understanding the intrinsic dynamics of such diseases and their effects on the population structure. One of the most well researched of these is the Lotka-Volterra type

model which divides the population into the healthy and the infected individuals and then explores the dynamic interaction between these two groups. This dynamic pattern matches the cycles that emerge from the interaction represented in predator-prey models where infected individuals act as predators of susceptible prey. By contrast, in the economic literature models which explore the effects of economic variables on population growth and the stock of human capital do so in a rudimentary reduced form fashion e.g. the single sector growth model can accommodate endogenous growth of labour in efficiency units either through human capital effects or through effects on the birth and death rates of the population in this way but treats everyone as identical in terms of health status.

In order to model the two way interaction between the infectious disease and prosperity, this paper combines the economic and demographic-epidemiological approaches to examine how economic effects on birth and mortality rates and on the rate of infection of healthy by infected people determine the population structure and hence also the size of the labour force. The latter then determines the level of prosperity. The paper also examines the form of optimal economic policy to control the disease.

We start by reviewing the transmission mechanism of disease implicit in the Lotka-Volterra type model of May and Anderson (1989). This assumes random mixing of the two population groups and leads to dynamic population equations that are homogenous of degree one in the number of healthy and infected individuals.

Next we consider a variety of frameworks in which the disease dynamics are embedded in the economy. A basic assumption is that healthy individuals are productive in the economy but infected individuals are not. Generally the demographic-epidemiological parameters (the net growth rate of the healthy population, the death rate of the infected and the infection rate between healthy and infected) are functions of economic variables. Economic effects on demographic-epidemiological parameters work through the level of general economic prosperity as determined by the accumulation of productive capital. When combined with the productive system of the economy, the homogeneity of the demographic-epidemiological equations allows the existence of balanced growth paths in which capital stock, healthy and infected population groups all grow at a common rate. We consider both a descriptive epidemiological-economic model with an exogenous savings rate and the form that optimal accumulation policy would take in a centrally planned economy. For models with a central planner who can control the system, we choose a time additive social welfare function which depends positively on the level of consumption per capita of the whole population and negatively on the change in the prevalence of infected individuals. There is obviously some room for debate here. In the last section we outline the effects of changing the objective function. We do not consider microeconomic individual responses to the prevalence of the disease in this paper. There is evidence that such responses are important (e.g. in the case of the plague the wealthy could often move away from infected areas at least during periods of infection) and are mostly concentrated on avoiding dangerous

contact with infected people. However, we indicate how these effects could be added.

Our results indicate that the economic-epidemiological interaction has important structural effects. In a descriptive scenario, generally there are either two or three steady states depending on how powerful general prosperity is in controlling infection and on how productive the economy is. In both cases one trivial steady state is at the origin: the capital/healthy labour ratio is zero and population is extinct. This is also the only steady state of the pure demographic-epidemiological system. The other steady states have a functioning economy. When there are two other steady states, in both the population structure is constant but in only one of them are there no infected people. Moreover locally eradication of the disease is unstable and the system instead tends to converge (with cycles in the population structure and economic prosperity along the path of convergence) to the steady state with a constant proportion of infected. Where there is just one other steady state then it has no infected people and this steady state is globally stable. These dynamic patterns are relatively robust to assumptions made about the economic environment so long as the demographic-epidemiological variables vary with the level of prosperity measured by the productive capital/productive labour ratio. In the centrally planned framework we find that, in the general prosperity model, there are usually two steady states one with no infected individuals and one in which there is a constant prevalence of the disease. The steady state which has no infection is locally a four dimensional saddle point. The local dynamics of the other steady state are more complex. Depending on parameter values it may be a four dimensional saddle point with or without oscillatory solutions, a centre or a focus.

The paper is organised as follows. Section 1 introduces the demographic-epidemiological models used and explores their equilibrium points and dynamics. Section 2 explores the general economic prosperity model, its steady states and dynamics. We provide some numerical integrations for a particular example. Section 3 outlines extensions including individual reaction to disease prevalence, the possibility of recovery from the disease and different forms of welfare function.

1 Demographic-epidemiological model

At instant t there are x_t healthy and y_t infected individuals. Healthy individuals are in the labour force but infected individuals are not capable of work. All births are of healthy individuals but by interacting with an infected person, a healthy individual may become infected. There is a nondisease driven natural death rate of healthy individuals; infected individuals may die of the disease or from other causes. The three basic demographic-epidemiological parameters are α, the net growth rate of the healthy; β, the infection rate between the healthy and infected; and ω, the death rate of the infected group.

The physical interpretation of the transmission mechanism is that a given

healthy person has a chance of infection given by

$$\begin{aligned} &Pr(becoming\ infected) = \\ &= Pr(meeting\ an\ infected\ person\ \&\ the\ meeting\ leads\ to\ infection) = \\ &\quad = Pr(meeting\ an\ infected\ person) Pr(the\ meeting\ leads\ to\ infection) \end{aligned}$$

since these events are generally taken as independent. The chance of infection from an encounter between an infected and a healthy person (β) is usually assumed to be a constant. For different diseases β may depend on individual characteristics (e.g. the age of the healthy person) or characteristics of the meeting (the plague often spread through fleas passing from one person to another in multioccupied beds). In common with the literature we abstract from this and assume homogeneity within population groups in terms of health status. Given x healthy people and this transmission mechanism, the average number of new cases of infection at t is then

$$Pr(meeting\ an\ infected\ person)\beta x$$

With equally likely meetings between any two people the probability of meeting an infected is $y/(x+y)$. This approach has been taken by May and Anderson (1989) in exploring the dynamics of HIV and AIDS. This gives

$$\begin{cases} \dot{x}_t = \alpha x_t - \frac{\beta x_t y_t}{(x_t + y_t)} \\ \dot{y}_t = \frac{\beta x_t y_t}{(x_t + y_t)} - \omega y_t \end{cases} \tag{1}$$

where α, β, ω are constants.

Define $p_t = y_t/x_t$ and rewrite the dynamic equations in terms of just p_t and x_t as

$$\begin{cases} \dot{p}_t = \frac{\beta p_t}{1+p_t} - \omega p_t - p_t(\alpha - \frac{\beta p_t}{1+p_t}) = (\beta - \alpha - \omega)p_t \\ \dot{x}_t/x_t = \alpha - \beta \frac{p_t}{1+p_t} \end{cases} \tag{2}$$

In this pure demographic-epidemiological system we can solve as follows. The equation in p has the solution $p_t = p_0 \exp[\beta - \alpha - \omega)t]$. Putting this back into the equation for x gives

$$\frac{\dot{x}}{x} = \alpha - \beta \frac{p_0 \exp((\beta - \alpha - \omega)t)}{1 + p_0 \exp((\beta - \alpha - \omega)t)} \tag{3}$$

which has solution

$$x_t = x_0 \exp(\alpha t)[1 + p_0 \exp((\beta - \alpha - \omega)t)]^{-\frac{\beta}{(\beta - \alpha - \omega) y_0}} \tag{4}$$

and so

$$y_t = p_0 x_0 \exp((\beta - \omega)t)[1 + p_0 \exp((\beta - \alpha - \omega)t)]^{-\frac{\beta}{(\beta - \alpha - \omega) y_0}} \tag{5}$$

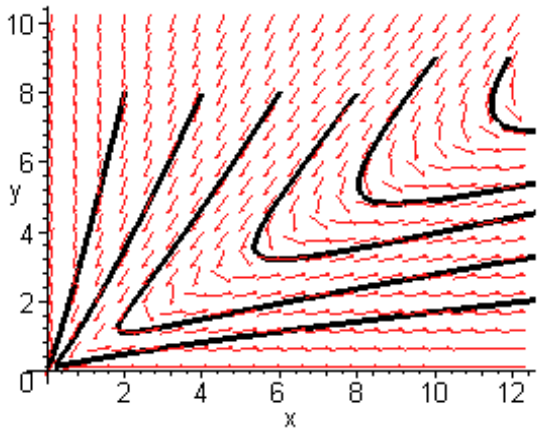

Figure 1: Demographic-epidemiological model, $\beta - \alpha - \omega < 0$ weak infection.

In the pure demographic-epidemiological system the ratio p either rises exponentially so that eventually the healthy are eliminated or falls to zero so that the infected are eliminated depending on whether the infection rate is greater than the combined net growth rate of the healthy and death rate of the infected. There are no stationary or steady states of the system other than at the origin. However, the number of healthy individuals may not be monotonic: the growth rate $\frac{\dot{x}}{x}$ from (3) may vary in sign starting negative and then becoming positive if $\alpha < \beta p_0/(1+p_0)$. A typical example is shown in the phase diagrams of Figure 1 in which we take $\beta - \alpha - \omega < 0$ and Figure 2 in which $\beta - \alpha - \omega > 0$. This extreme dynamic behaviour is consistent with history. The effects of the plague in Egypt, England and Italy or of smallpox in mid-America were that whole villages were eliminated while other communities emerged relatively unscathed from an initial infection. We use this basic demographic-epidemiological structure in the sequel.

2 The economic-epidemiological model

Here α, β and ω are in general functions of economic prosperity. The healthy individuals work to produce output from productive capital. Output can either be consumed or invested to add to productive capital. We measure the level of economic prosperity by the ratio of productive capital to the number of healthy workers at time t. This is consistent with either an economic infrastructure interpretation (the higher the capital-labour ratio, the more mechanised the society) or, as consumption per healthy worker is a function of capital per healthy worker, consistent with an individualistic interpretation in which demographic-epidemiological parameters depend on consumption per healthy worker. We

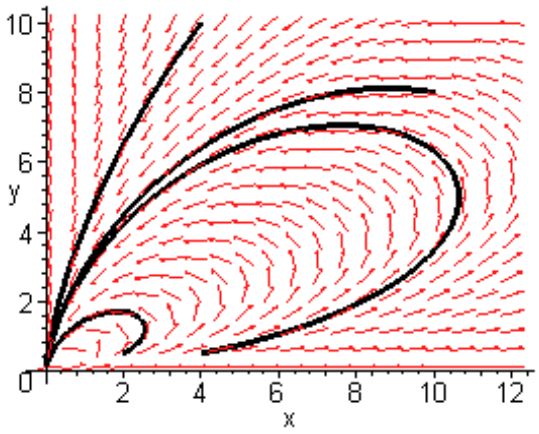

Figure 2: Demographic-epidemiological model, $\beta - \alpha - \omega > 0$ strong infection.

consider both the case in which savings is exogenous[1] and the case of optimally selected savings which takes into account the effects of productive capital on the future demographic-epidemiological parameters.

The basic technological assumptions on the economy are that there is a single good which can be consumed or invested. At any instant t output of the single good is given by $F(k_t, x_t)$ where k_t is the existing capital stock. The production function satisfies the usual neoclassical properties ($F(k_t, x_t)$ is increasing, homogenous of degree one and has diminishing marginal productivity of each input). In addition we assume that each input is essential in that $F(0, x) = F(k, 0) = F(0, 0) = 0$. We also impose that $\partial F/\partial k \to \infty$ as the ratio $k/x \to 0$ and that $\partial F/\partial k \to 0$ as $k/x \to \infty$.

Capital depreciates linearly at a constant rate of ϕ. The interaction between economics and the health structure of the population comes through the facts that α, β and ω are functions of k_t/x_t and that only healthy individuals are productive in technology.

2.1 Exogenous Savings Behaviour

In this simple scenario the only policy instrument is the division of output at each instant between consumption and investment. With this being exogenous,

[1] The exogenous savings case could be rationalised through decentralised individual savings decisions. With competitive input markets, worker per capita income is $w + rk_x$ where the real wage (w) is equal to the marginal product of labour and the rental rate on capital r is equal to the marginal product of capital. Capital stock is equally owned by all individuals so that k_x is capital stock owned per worker. An infected person only receives rental income so total income of the infected group is rk_y where k_y is the aggregate capital stock. From homogeneity of degree one of the production function aggregate income is $wx + r(k_x x + k_y) = F$. If every individual saves a constant share s of their income then aggregate savings sF is proportional to aggregate income.

we assume a proportional savings function, with rate of savings s. Hence, capital accumulates according to

$$\dot{k}_t = sF(k_t, x_t) - \phi k_t \tag{6}$$

and per capita consumption of the healthy is

$$c_t = (1-s)F(k_t, x_t)/x_t \tag{7}$$

At some stages in the sequel we select functional forms for the production function and α, β, ω. Initially, we just assume $\alpha^{'}(.) \geq 0$; $\beta^{'}(.) \leq 0$; $\omega^{'}(.) < 0$; $\beta'(.) - \alpha'(.) - \omega'(.) \leq 0$ and $\beta''(.) - \alpha''(.) - \omega''(.) \geq 0$ everywhere, $\alpha(0)$, $\beta(0)$, $\omega(0) > 0$ and that $\beta(0) - \alpha(0) - \omega(0) > 0$ so that in the absence of productive capital the prevalence of the disease increases without bound. It is also natural to assume that α, β, ω are bounded above by $\overline{\alpha}$, $\overline{\beta}$, $\overline{\omega}$ respectively and below by zero (dealing with a growing rather then declining susceptible population).

We have

$$\begin{cases} \dot{x}_t = \alpha x_t - \frac{\beta x_t y_t}{x_t + y_t} \\ \dot{y}_t = \frac{\beta x_t y_t}{x_t + y_t} - \omega y_t \\ \dot{k}_t = sF(k_t, x_t) - \phi k_t \end{cases} \tag{8}$$

where α, β, ω are functions of k_t/x_t. Given that $F(.)$ is homogeneous of degree one, we can define $z_t = k_t/x_t$, $\quad p_t = y_t/x_t$ and rewrite the dynamic equations in terms of just z_t and p_t as

$$\begin{cases} \dot{p}_t = \frac{\beta p_t}{1+p_t} - \omega p_t - p_t(\alpha - \frac{\beta p_t}{1+p_t}) = (\beta - \alpha - \omega)p_t \\ \dot{z}_t = sf(z_t) - \phi z_t - z_t(\alpha - \frac{\beta p_t}{1+p_t}) \end{cases} \tag{9}$$

where β, α, ω are functions of z_t .

This system has steady states in which z and p are constant. Generically there are three steady states:

(i) $p^* = 0$ and z^* such that $sf(z^*) - \phi z^* - \alpha z^* = 0$.

Since α is constant in steady state, here we have the steady state of a standard Solow-Swan growth model in which there are no infected people. All the standard analysis of golden rule savings ratios for this steady state would apply. Overall consumption per capita is given by $c_t/(1+p_t)$. In this steady state since $p^* = 0$, $c^*/(1+p^*) = f(z^*) - \phi z^* - \alpha z^*$ which is maximised over z^* at $f'(z^*) - \phi - \alpha' z^* - \alpha = 0$. The usual Golden Rule equates the net marginal product of capital to the population growth rate. Here it is equated to the marginal population growth rate allowing for the effects of z on population growth.

(ii) $p^* = 0$ and $z^* = 0$.

Due to the assumptions on technology and on boundedness of the demographic-epidemiological parameters this is always a steady state in which the healthy population grows at the rate $\alpha(0)$, there is no disease and no output or capital.

(iii) The most interesting cases are steady states where $p^* \neq 0$ and $z^* \neq 0$. These must solve

$$(\beta - \alpha - \omega) = 0 \tag{10}$$

$$sf(z^*) - \phi z^* - (\alpha - \frac{\beta p^*}{1+p^*})z^* = 0 \tag{11}$$

There must be a root z^* to (10) if $\beta(0) - \alpha(0) - \omega(0) > 0$ and $\beta(\infty) - \alpha(\infty) - \omega(\infty) < 0$. The root must be unique if $\beta'(.) - \alpha'(.) - \omega'(.) < 0$. In this steady state the infected/healthy structure of the population is constant through time with new healthy births just balancing the new infections net of infected deaths. The capital/healthy labour ratio is constant with new investment just matching the depreciation of capital and the growth in the number of healthy individuals net of attrition through infection.

Solving (11) for p we have

$$p^* = -\frac{sf - \phi z^* - \alpha z^*}{sf - \phi z^* - \alpha z^* + \beta z^*} \tag{12}$$

so long as the denominator is nonzero. Since we need $p^* \geq 0$, there can only be feasible steady states of this form when the numerator and denominator of this expression have opposite signs. As $\beta z^* > 0$ this needs the numerator to be negative. However for low z^* the numerator is positive since $f'(0)$ is high and $f(0) = 0$. For sufficiently high values of z^* the numerator is negative since $f' > 0$ and concavity of f, under the Inada conditions, means that eventually the linear term in depreciation dominates. We also know the numerator is zero at $z^* = 0$; hence as z^* increases from 0 the numerator is first positive and rising and then falls eventually becoming negative. Where it vanishes we have the steady state ($p^* = 0$, $z^* \neq 0$). Defining z_2 as the root of $(\beta - \alpha - \omega) = 0$, z_1 as the root of $sf - \phi z - \alpha z = 0$, and z_3 as the root of $sf - \phi z - \alpha z + \beta z = 0$, a steady state with $p^* \neq 0$ and $z^* \neq 0$ exists if $z_1 < z_2 < z_3$ but fails to exist if $z_1 > z_2$ as in Figure 3 and Figure 4, respectively. Generally $z_1 > z_2$ is more likely when the level of $(\beta - \alpha - \omega)$ is relatively low and when it tends to fall relatively fast. Alternatively, when f is not very concave, s is high and α, ϕ are low.

If $z_2 > z_1$ there is a range of initial capital labour ratios bounded above by z_1, such that starting from any capital labour ratio in that range, capital accumulates and the prevalence of the disease initially falls and then rises as the population grows but the existing capital is spread too thinly between workers. The continuing capital accumulation and the reduction of the work force due to the increase in infection sufficiently raises z to improve the epidemiological situation so that prevalence reaches a maximum and then starts falling again. Thereafter p and z cycle in a convergent way around the steady state with

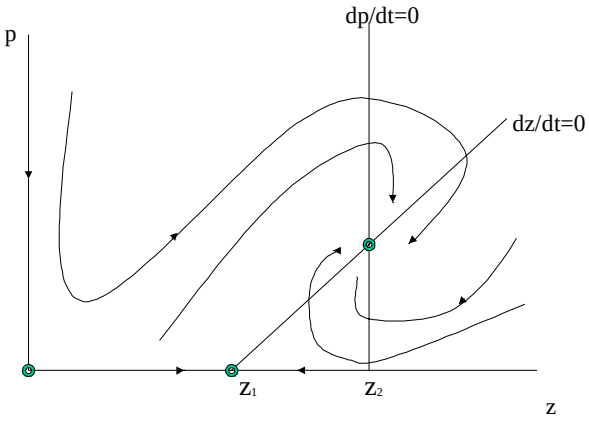

Figure 3: Productive capital model, $z_1 < z_2 < z_3$.

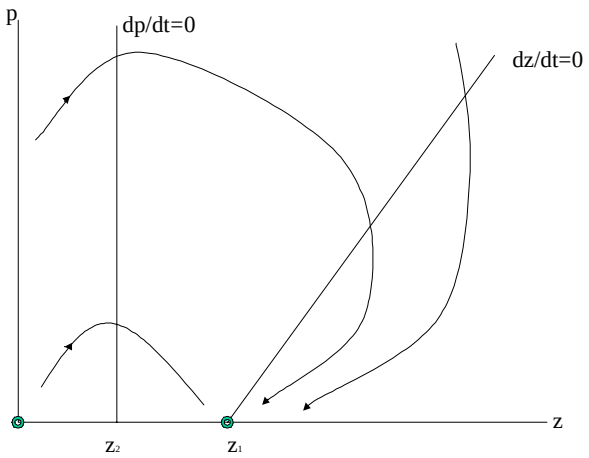

Figure 4: Productive capital model, $z_1 > z_2$.

$p^* \neq 0$ and $z^* \neq 0$. The initial reduction in prevalence may not occur on paths with a high initial capital labour ratio or a low prevalence; instead the system goes directly into convergent cycles.

Conversely if $z_2 < z_1$ then sufficiently high initial prevalence of the disease leads to a time path of increasing prevalence and accumulation until prevalence reaches a maximum; then prevalence falls while capital accumulates until the labour force has risen so much with decreasing infection and population growth to lead to a fall in z. On the last part of such a path, both p and z fall towards the steady state with $p^* = 0$ and $z^* \neq 0$.

The global phase spaces in Figure 3 and Figure 4 are consistent with the local stability properties of the different steady states. Linearising the equations

$$\begin{bmatrix} \dot{p}_t \\ \dot{z}_t \end{bmatrix} = \begin{bmatrix} \beta - \alpha - \omega & (\beta' - \alpha' - \omega')p \\ z\beta/(1+p)^2 & sf' - sf/z - z(\alpha' - \beta' p/(1+p)) \end{bmatrix} \begin{bmatrix} p \\ z \end{bmatrix} \tag{13}$$

Around the steady state $p_1^* = 0$, z_1^* solving $sf(z^*) - \phi z^* - \alpha(z^*)z^* = 0$ the Jacobian becomes

$$\begin{bmatrix} \beta - \alpha - \omega & 0 \\ z\beta & sf' - sf/z - z\alpha' \end{bmatrix} \tag{14}$$

At this steady state if $z_1^* < z_2^* < z_3^*$ (so we are in the case with three steady states) we know that $\beta - \alpha - \omega > 0$ since $\beta'(.) - \alpha'(.) - \omega'(.) < 0$. The determinant of the matrix is negative and we have a local saddle point. Alternatively if $z_1^* > z_2^*$ (i.e. there are only two steady states), then $\beta - \alpha - \omega < 0$. The trace is negative whilst the determinant is positive so that we have two roots with negative real parts and local stability (node or focus).

Around the steady state ($p^* \neq 0$, $z^* \neq 0$) when it exists the Jacobian becomes

$$\begin{bmatrix} 0 & (\beta' - \alpha' - \omega')p \\ z\beta/(1+p)^2 & sf' - \alpha - \phi - z(\alpha' - \beta' p/(1+p)) \end{bmatrix} \tag{15}$$

The trace is negative and the determinant is positive: the real parts of the two roots are both negative and we have local stability (node or focus).

Around the steady state ($p^* = 0, z^* = 0$) the Jacobian becomes

$$\begin{bmatrix} \beta - \alpha - \omega & 0 \\ 0 & sf' - \phi - \alpha \end{bmatrix} \tag{16}$$

The diagonal elements, which coincide with the eigenvalues, are both positive since at $z^* = 0$, $f'(z^*)$ is high. So we have local instability.

Numerical integration is useful to show the speed and amplitude of the dynamics although it requires specialization of functional forms. As an example we take functions of the form

$$f(z) = gz^a/a \qquad 0 < a < 1 \tag{17}$$

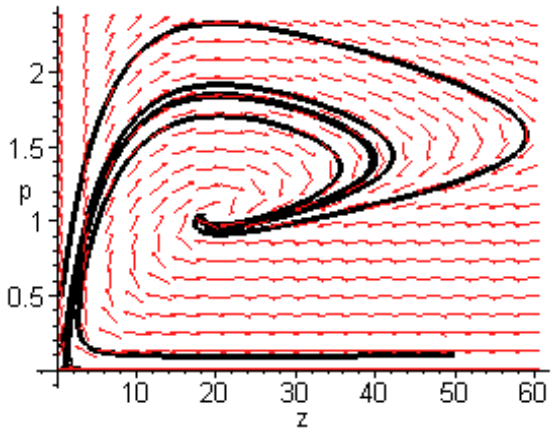

Figure 5: Productive capital model, global phase space for $g = 0.25$.

$$\alpha(z) = \alpha_0 - \alpha_1 \exp(-\alpha_2 z) \qquad \alpha > 0 \tag{18}$$

$$\beta(z) = \beta_0 + \beta_1 \exp(-\beta_2 z) \qquad \beta > 0 \tag{19}$$

$$\omega(z) = \omega_0 + \omega_1 \exp(-\omega_2 z) \qquad \omega > 0 \tag{20}$$

We select values $\alpha_0 = 0.04$, $\alpha_1 = 0.02$, $\alpha_2 = 0.5$, $\beta_0 = 0.2$, $\beta_1 = 0.2$, $\beta_2 = 0.5$, $\omega_0 = 0.2$, $\omega_1 = 0.1$, $\omega_2 = 0.5$, $a = 0.5$, $\phi = 0.2$, $s = 0.2$ and two alternative values for the scale of output: $g = 0.25$ and $g = 1$. With these parameters, as z varies from 0 to ∞, the net growth rate of the healthy individuals varies from 2% to 4%, the infection rate and the total death rate of the infected individuals vary from 40% to 20% and from 20% to 10%, respectively.

Using these values when $g = 0.25$ we have three steady states in the case of an exogenous saving rule: $(p^* = 0, z^* = 0)$, $(p^* = 0, z^* = 2.197)$ and $(p^* = 1.646, z^* = 2.917)$. The phase space is shown in Figure 5. When $g = 1$ we have two steady states: $(p^* = 0, z^* = 0)$ and $(p^* = 0, z^* = 2.197)$. The phase space is shown in Figure 6.

Starting from an initial (p_0, z_0) the economy shows cycles in which p and z alternately move either together or in opposite directions. In an economic upswing the prevalence of disease is rising but then reaches a maximum and in the later part of the upswing the disease prevalence falls. In the early stage of the upswing there is such low prosperity that the infection rate is too high relative to the healthy birth rate and infected death rate to reduce the prevalence of the disease. Once prosperity has sufficiently increased, the infection rate has fallen far enough to make further prosperity increases reduce prevalence. The opposite effect occurs in the downswing. The dynamics shown here match those of Figure 3 and Figure 4 although in Figure 5 the region of falling prevalence at low values of z is compressed near the vertical axis.

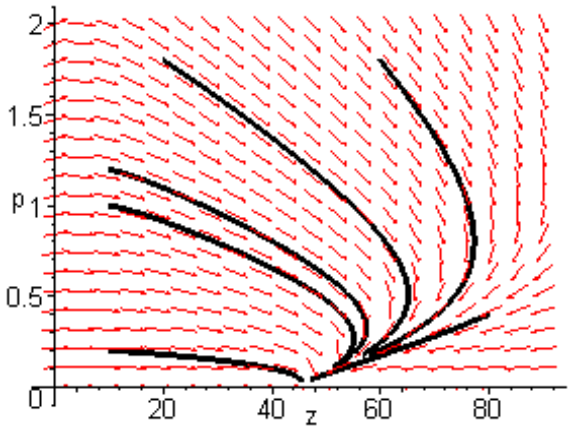

Figure 6: Productive capital model, global phase space for $g = 1$.

2.2 Optimal control of accumulation

Instead of taking the savings rate to be exogenous, there may be a centrally planned economy in which savings is chosen to balance the welfare effects of current and future consumption and the effects of capital accumulation on controlling disease[2].

2.2.1 The welfare function

In the centrally planned context defining intertemporal welfare is problematic: the population and its composition is changing. This raises normative issues that have been discussed in the optimal population literature (Dasgupta,1969; Meade 1955; Yaari, 1981). One issue to be beware of is the danger of maximising individual welfare of those alive by having just a single individual who owns all the capital stock. A further issue is that of time additivity. Firstly there may be intergenerational altruism within the private welfare functions of those alive at t. Second there is the issue of time discounting; generally in the body of the paper we do not allow for discounting essentially on ethical grounds (Dasgupta, 1974; Grout, 1982; Rawls, 1972) since there seems little reason to value generations differently. If we did use a time discounted additive welfare function the form of the necessary conditions are essentially unaltered, effectively we can absorb the discount rate in the depreciation rate ϕ and in the demographic-epidemiological functions α, β, ω. However discounting is also of assistance in demonstrating that an optimal solution exists. Apart from this at any t there is the difficulty

[2] A much more complex approach would model a decentralised system in which healthy individuals select labour supply and savings to optimise their lifetime utility taking into account the chance of becoming infected and of dying. Similarly the infected choose their savings out of any rental income taking into account their chance of death. Moreover new generations would continuously enter the picture.

of determining the relative benefit of the infected and healthy. If there is a time additive welfare function $\int u(.)\ dt$ with $u(.)$ reflecting the social welfare of the population at t, then it is plausible to take overall per capita consumption of the whole population $(c/(1+p))$ as one argument of $u(.)$ which implies that $u(.)$ is decreasing in p. This, however, gives an incentive to reduce p by whatever means possible ('population cleansing') which hardly seems acceptable. A preferable structure makes $u(.)$ also a decreasing function of $\dot{p}$ and generally we work with this formulation. At t welfare is given by $u(c/(1+p),\dot{p})$ where (using subscripts 1 and 2 for derivatives with respect to the first and second argument of $u(\frac{c}{(1+p)},\dot{p})$), $u_1 > 0$; $u_2 < 0$; $u_{11} < 0$; $u_{22} < 0$ and generally we take $u_{12} = 0$ together with $u_2(.,0) = 0$. We also generally assume that $\lim\ u_1 \rightarrow \infty$ as $c/(1+p) \rightarrow 0$.

2.2.2 Necessary conditions for optimality

The policy problem then has control variable c_t and state variables p_t, z_t and is to

$$max \int u(c/(1+p),\dot{p})dt \tag{21}$$

subject to

$$\begin{array}{l} \dot{p} = (\beta - \alpha - \omega)p \\ \dot{z} = f(z) - \phi z - (\alpha - \beta\frac{p}{1+p})z - c \end{array} \tag{22}$$

The Hamiltonian is

$$H = u(c/(1+p),\dot{p}) + \lambda_1 \left[f(z) - \phi z - (\alpha - \beta\frac{p}{1+p})z - c \right] + \lambda_2(\beta - \alpha - \omega)p \tag{23}$$

Optimising H over c:

$$u_1 = \lambda_1(1+p) \tag{24}$$

The equations of motion for the costate variables are

$$\begin{array}{rcl} \dot{\lambda_1} & = & -(\lambda_2 + u_2)(\beta' - \alpha' - \omega')p \\ & & -\lambda_1[f' - \phi - (\alpha - \beta p/(1+p)) - z(\alpha' - \beta' p/(1+p))] \end{array} \tag{25}$$

$$\dot{\lambda_2} = u_1 c/(1+p)^2 - \lambda_1 z\beta/(1+p)^2 - (\lambda_2 + u_2)(\beta - \alpha - \omega) \tag{26}$$

Here (24) has the interpretation that the marginal utility of consumption per overall capita divided by $(1+p)$ which corrects for the unproductive workers should be equated to the marginal value of investment per healthy worker. Next (25) modifies the usual Euler equation in two respects. First the social marginal productivity of capital deepening is reduced because it raises the healthy birth rate via α and it reduces the infection rate via the transmission rate β. Second because capital deepening alters the growth of infection and of p, it changes the marginal value of the population structure in the future. (26) indicates that the shadow value of the prevalence of the disease for the future rises because of its effects in reducing the instantaneous utility but falls because of its effects on the future value of investment and the future growth of prevalence. Notice that these conditions imply that on one margin higher prevalence is socially desirable as it increases the capital deepening effect of a given amount of investment by having fewer healthy workers available to use it. By contrast, it reduces the value of current $c/(1+p)$. (22), (24) - (26) are necessary conditions for optimality. Subsequently we consider their sufficiency and also whether a solution exists to the problem. In the case of time discounting the necessary conditions are modified[3] but essentially the analysis below remains valid.

2.2.3 Dynamic analysis

To analyse the dynamic properties of this solution we start by examining the steady states. There are two in general. The first has $p^* = 0$ and corresponds to the Golden Rule of a growth model with endogenous net population growth at rate α and conditional on the absence of infection ($\beta = 0$). Knowing $p^* = 0$, consider the steady state capital-healthy-labour-ratio z which maximises steady state utility $u(c,0)$. This solves

$$f' - \phi - \alpha - \alpha' z^* = 0 \tag{30}$$

[3] If the welfare function is

$$\int e^{-rt} u(c/(1+p), \dot{p}) \tag{27}$$

the present value Hamiltonian is

$$H = e^{-rt}\left(u(\frac{c}{1+p}, \dot{p}) + \lambda_1(f(z) - \phi z - (\alpha - \beta\frac{p}{1+p})z - c) + \lambda_2((\beta - \alpha - \omega)p)\right) \tag{28}$$

and the necessary conditions become

$$u_1 = \lambda_1(1+p)$$

$$\dot{\lambda}_1 = -\lambda_1(f' - \phi - r - z(\alpha' - \beta'\frac{p}{1+p}) - (\alpha - \beta\frac{p}{1+p})) - \lambda_2(\beta' - \alpha' - \omega')p$$

$$\dot{\lambda}_2 = \lambda_1[c(1+p) - \beta z] - \lambda_2(\beta - \alpha - \omega - r) \tag{29}$$

and the dynamic equations for the state variables. This affects the steady state values of the costate variables and adds a growth term to their dynamics.

Putting these values into the costate equations (25) and (26) at the steady state gives

$$\begin{aligned} 0 &= -\lambda_1^*[f' - \phi - (\alpha - \beta p^*/(1+p^*)) - z^*(\alpha' - \beta' p^*/(1+p^*))] \\ &\quad -\lambda_2^* p^*(\beta' - \alpha' - \omega') \end{aligned} \tag{31}$$

(which is automatically zero since $p^* = 0$ and $f' - \phi - \alpha - \alpha' z^* = 0$) and

$$0 = u_1 c^* - (\lambda_2^* + u_2)(\beta - \alpha - \omega) - \lambda_1^* z^* \beta \tag{32}$$

or

$$0 = \lambda_1^*(c^* - z^*\beta) - \lambda_2^*(\beta - \alpha - \omega) \tag{33}$$

since at $p^* = 0, u_1^* = \lambda_1^*$ and $u_2^* = 0$. So one steady state is $p^* = 0$; z^* such that $f' - \phi - \alpha - \alpha' z^* = 0$; u_1 is evaluated at $c^* = f(z^*) - \phi z^* - \alpha z^*$; λ_1^* solves $u_1^* = \lambda_1^*$; λ_2^* satisfies $\lambda_1^*(c - z^*\beta) - \lambda_2^*(\beta - \alpha - \omega) = 0$.

The sign of λ_2^* is ambiguous. Since $\partial \dot{z}/\partial p > 0$ the model does not satisfy the most common condition imposed for negativity of λ_2 (Leonard, 1981). However, as we see below $\lambda_2 < 0$ forms part of a set of sufficiency conditions for optimality of the solutions to these equations.

In general there is also a second steady state with $p^* \neq 0$. This satisfies the equations

$$u_1 = \lambda_1^*(1 + p^*) \tag{34}$$

$$\beta - \alpha - \omega = 0 \tag{35}$$

$$0 = f(z^*) - \phi z^* - z^*(\alpha - \beta p^*/(1+p^*)) - c^* \tag{36}$$

$$\begin{aligned} 0 &= -(\lambda_2^* + u_2)(\beta' - \alpha' - \omega')p^* - \lambda_1^*[f' - \phi - (\alpha - \beta p^*/(1+p^*)) \\ &\quad -z^*(\alpha' - \beta' p^*/(1+p^*))] \end{aligned} \tag{37}$$

$$0 = [u_1 c^* - \lambda_1^* z^* \beta]/(1+p^*)^2 = \lambda_1^*[c^*(1+p^*) - z^*\beta]/(1+p)^2 \tag{38}$$

implying either $\lambda_1^* = 0$ which is impossible under the assumptions on u_1, or $(1+p^*)c^* = z^*\beta$. Using this in (36) we can solve for p and c:

$$p^* = -\frac{f - \phi z^* - \alpha z^*}{f - \phi z^* - \alpha z^* + \beta z^*} \tag{39}$$

$$c^* = f - \phi z^* - \alpha z^* + \beta z^* \tag{40}$$

and λ_1^* is given by

$$\lambda_1^* = \frac{u_1(\beta z^*/(1+p^*)^2)}{(1+p^*)} \tag{41}$$

Here (35) sets z^* and (37) sets λ_2^*.

This second steady state also has a Golden Rule interpretation. Consider the problem of selecting steady state values of p and z which maximise $c/(1+p)$ and satisfy the steady state conditions

$$(\beta - \alpha - \omega)\, p = 0 \tag{42}$$

$$c/(1+p) = \left[f(z) - \phi z - (\alpha - \beta \frac{p}{1+p}) z \right] /(1+p) \tag{43}$$

There are two possibilities. First z^* is set to satisfy (42) and p^* maximises (43) and so solves[4]

$$f(z^*) - \phi z^* - \alpha z^* = \beta z^* \frac{(1-p^*)}{(1+p^*)} \tag{44}$$

yielding $c^*/(1+p^*) = \beta z^*/(1+p^*)^2$ and $p^* = -[f - \phi z^* - \alpha z^*]/[f - \phi z^* - \alpha z^* + \beta z^*] > 0$ if $f - \phi z^* - \alpha z^* < 0$. Second we could have $p^* = 0$ which then returns us to the first steady state satisfying (30). Both these solutions are local maxima so that the global Golden Rule could be either.

To determine the local dynamics, in the Appendix (A1) we calculate the Jacobian of the linearised system and its value at the different steady states. Around the steady state with $p^* = 0$ the Jacobian has a zero trace so that, unless all real parts of roots are zero, there are elements of both local stability and instability. We can show that the matrix has two pairs of real roots, each pair consisting of a real number of opposite sign. To see this rewrite the matrix as

$$\begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & 0 & a_{23} & 0 \\ a_{31} & a_{32} & 0 & 0 \\ a_{41} & a_{31} & -a_{21} & -a_{11} \end{bmatrix} \tag{45}$$

Its characteristic polynomial is $(X - a_{11})\left(X^3 + X^2 a_{11} - a_{32}a_{23}X - a_{32}a_{23}a_{11}\right)$. There are four real roots to this matrix in pairs of opposite sign: $\pm a_{11}, \pm\sqrt{a_{32}a_{23}}$. Note that $a_{32} > 0$ and $a_{23} > 0$ since $\alpha'' > 0$. Locally the steady state is a four dimensional saddle point.

[4] This is actually a local maximum of $c/(1+p)$ in p since at this value $\frac{\partial^2(c/(1+p))}{\partial p^2} = -\frac{2\beta z}{(1+p)^3} < 0$.

Around the steady state with $p^* \neq 0$ the Appendix (A1) shows that the Jacobian matrix has a zero trace so the sums of the real parts of the roots are zero. The Jacobian has the form

$$\begin{bmatrix} 0 & a & 0 & 0 \\ b & c & d & 0 \\ f & e & -c & -a \\ g & f & -b & 0 \end{bmatrix} \tag{46}$$

so that the characteristic equation is $X^4 - 2baX^2 - X^2c^2 - edX^2 + b^2a^2 + ga^2d$ which can be reparametrised as $X^4 + AX^2 + B$. If $(A^2 - 4B)$ is positive and $(-A - \sqrt{(A-4B)})$ is positive there are two pairs of real roots of equal absolute value but opposite sign which represents a 4-D saddle point. When $(A^2 - 4B)$ is positive and $(-A - \sqrt{(A-4B)})$ is negative then there are two pairs of imaginary roots. When $(A^2 - 4B)$ is positive and $(-A + \sqrt{(A-4B)})$ is positive and $(-A - \sqrt{(A-4B)})$ is negative then there is one pair of pure imaginary and a pair of real roots of opposite sign. However, if $(A^2 - 4B)$ is negative we have two pairs of complex conjugate roots with the common real part of each pair being of equal absolute value but opposite sign. We can think of this as a combination of a 4-D saddle point and locally oscillatory solutions. Summarising the possibilities are shown in Table 1.

$(A^2 - 4B)$	+	+	+	−
$(-A - \sqrt{(A-4B)})$	+	−	−	
$(-A + \sqrt{(A-4B)})$		+	−	
	4-D saddle	cycles	focus	4-D saddle and local oscillations

Table 1: Optimal control of productive capital model, local dynamics

Under our assumptions both of these steady states may not always exist and in particular the steady state with $p^* \neq 0$ generally requires a relatively high infection rate. For example, suppose we confine attention to isoelastic utility so that $u(c/(1+p), \dot{p}) = (c/(1+p))^{(1-b)}/(1-b)$ where $b > 0$. Putting together the optimal control conditions and the conditions for $\dot{\lambda}_2 = 0$, $\dot{z} = 0$ gives us the equations

$$\lambda_1^{\frac{1}{2b-1}} = 2(z\beta)^{\frac{b-1}{2b-1}}/(f - \phi z - \alpha z + \beta z) \tag{47}$$

$$c = (f - \phi z - \alpha z + \beta z)/2 \tag{48}$$

$$p = \frac{2z\beta}{(f - \phi z - \alpha z + \beta z)} - 1 \tag{49}$$

so that $p^* > 0$ requires $\beta > f/z - \phi - \alpha$ given that $f - \phi z - \alpha z + \beta z > 0$.

2.2.4 Numerical analysis

To understand the connections between the local stability properties of the alternative steady states and also the qualitative amplitude of any oscillations, numerical integration of the four dimensional non-linear system is illuminating. To display the results of this geometrically whilst remaining in at most three dimensions, we integrate around contours of the Hamiltonian. Since along any solution path for $[p(t), z(t), \lambda_1(t), \lambda_2(t)]$, the Hamiltonian is identically constant (Dankowicz, 1997) we can use the equation

$$\begin{aligned} \bar{H} &= u\left(\frac{c(\lambda_1(t),p(t))}{(1+p(t))}\right) + \lambda_2(t)\left[(\beta-\alpha-\omega)p(t)\right] \\ &\quad +\lambda_1(t)\left[f(z(t)) - \phi z(t) - z(t)(\alpha - \beta\frac{p(t)}{1+p(t)}) - c(\lambda_1(t),p(t))\right] \end{aligned} \quad (50)$$

to solve for $\lambda_2(t)$

$$\begin{aligned} \lambda_2(t) &= \frac{\left[\bar{H} - u\left(\frac{c(\lambda_1(t),p(t))}{(1+p(t))}\right)\right]}{\left[(\beta-\alpha-\omega)p(t)\right]} - \\ &\quad \frac{\lambda_1(t)\left[f(z(t)) - \phi z(t) - z(t)(\alpha - \beta\frac{p(t)}{1+p(t)}) - c(\lambda_1(t),p(t))\right]}{\left[(\beta-\alpha-\omega)p(t)\right]} \end{aligned} \quad (51)$$

where $\bar{H}$ is the constant value of the Hamiltonian on the surface. We can then substitute this expression for $\lambda_2(t)$ into the differential equations for the other variables $p(t)$, $z(t)$, $\lambda_1(t)$ to reduce the system to three dimensions without losing any information. Notice that because the system has the Hamiltonian conservation property, close to a stationary position of $p(t)$, $\lambda_2(t)$ becomes unbounded.

For the welfare function we take the isoelastic case and add $b = 2.0$ to the earlier set of parameters. With these parameter values when $g = 0.25$ the control model has two steady states ($p^* = 0$, $z^* = 1.131$, $\lambda_1^* = 13.403$, $\lambda_2^* = -38.818$) and ($p^* = 0.439$, $z^* = 2.197$, $\lambda_1^* = 8.678$, $\lambda_2^* = -12.850$). The eigenvalues at the first steady state where $p^* = 0$ are ± 0.028 and ± 0.124. At the second steady state when $p^* \neq 0$ the eigenvalues are ± 0.093 and ± 0.027. Figure 7 shows the three dimensional phase space in p, z and λ_1 around a level surface of the Hamiltonian corresponding to $\bar{H} = -3.626$. Figures 8 and 9 show magnifications of this in the vicinity of each of the two steady states, respectively. Since λ_2 only enters the differential equation for λ_1, the constraint that the 3-D solution lies on the Hamiltonian surface only bites in that equation. Hence, in the $\dot{\lambda}_1$ equation (25) the term in λ_2 is replaced by (51). If the system approaches one of the steady states, then on the Hamiltonian surface this makes λ_1 diverge to either $\pm\infty$ depending on whether z converges to the steady state faster or slower than p. In the diagram, since λ_1 diverges to ∞ along some trajectories, on these paths p must be converging more rapidly than z. On other trajectories

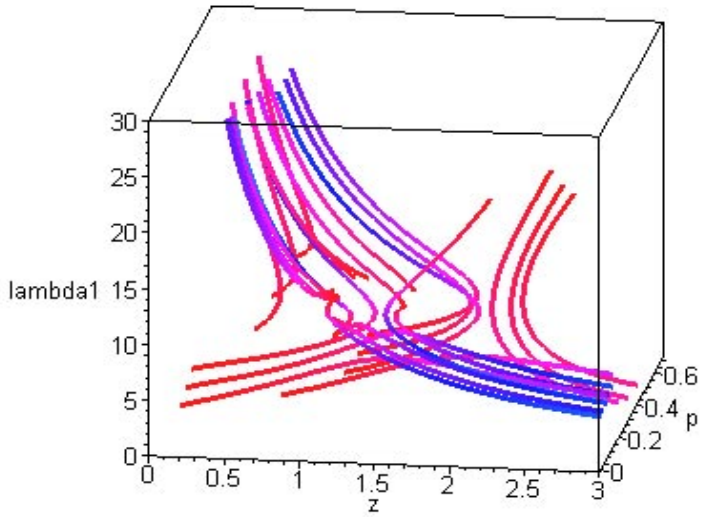

Figure 7: Optimal control of productive capital model, global phase space for g=0.25.

λ_1 is converging to zero. With the chosen parameters the two steady states both show dual instability of a saddle point type and there are no oscillations; this reflects the fact that all the eigenvalue are real. Careful study of Figures 7-9 reveals that there are six patterns of paths: along four of these λ_1 eventually grows without bound and along two λ_1 converges to zero. Typically z is not monotonic but initially increases, subsequently falls. Starting near $p^* = 0.4$ the movement of prevalence is very slow but starting near $p^* = 0$ there is a relatively rapid movement of prevalence towards the steady state with $p^* \neq 0$. As a whole the system spends most its time close to the steady state in which p and z are nonzero and away from the origin.

To show that the system sometimes loses the second steady state with $p^* \neq 0$ we also take the above example with $g = 1$. Here the control model has a single steady state: ($p^* = 0$, $z^* = 17.357$, $\lambda_1^* = 0.058$, $\lambda_2^* = -1.001$). The eigenvalues at the unique steady state where $p^* = 0$ are ± 0.0400 and ± 0.120. So again we have a 4-D saddle point.

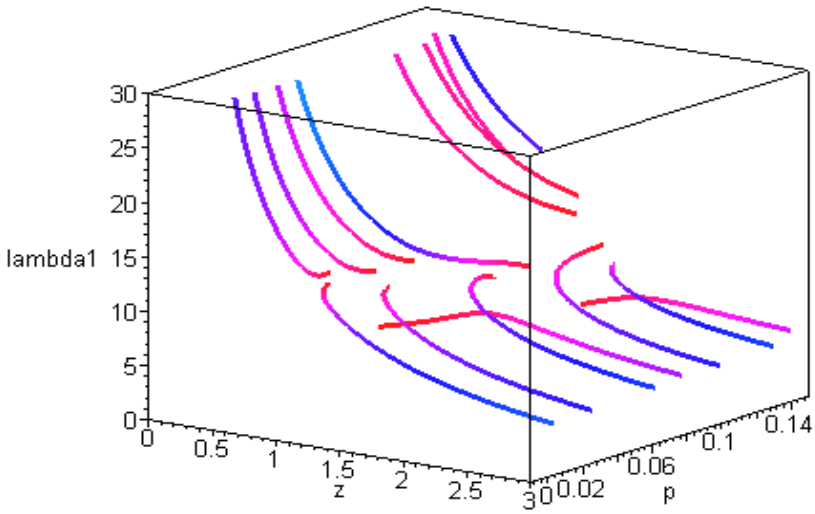

Figure 8: Optimal control productive capital model, magnified view of the global phase space near $p = 0$.

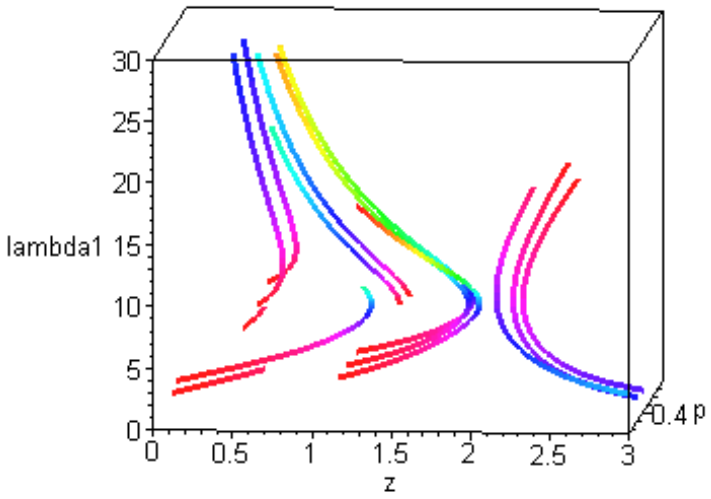

Figure 9: Optimal control productive capital model, magnified view of the global phase space for values near $p = 0.4$.

2.2.5 Sufficiency conditions for optimality

Consider the sufficiency of the necessary conditions for optimal control. First take the case in which α, β and ω are constants. The control problem is then the usual one of trading off consumption against investment. In this case $p(t)$ is like an exogenous time trend, growing or falling at rate $\gamma = (\beta - \alpha - \omega)$, which affects the accumulation process of z and preferences. We can apply the usual sufficiency argument to show that if $u(.)$ and $f(.)$ are concave then the necessary conditions for optimality are sufficient so long as $\lim_{t\to\infty} \lambda_1(t)z(t) = 0$. (See Appendix A2).

A similar but more complex argument holds if we allow for the effects of z on the demographic-epidemiological parameters. With a finite time horizon and nonnegative terminal constraints on the state variables, the transversality conditions

$$\lambda_1(T)z(T) = 0; \qquad \lambda_2(T)p(T) = 0 \tag{52}$$

form part of the necessary conditions for optimality. Then we can use the sufficiency theorem that if the maximised Hamiltonian is concave in the state variables p, z (Chiang, 1992; Mangarasian, 1966) any path satisfying the necessary conditions is optimal. If the time horizon is infinite then the transversality conditions are not generally necessary but we can use an adaptation of a similar sufficiency theorem (Theorem 9.3.1 in Leonard and Long, 1992) that if a form of transversality condition holds

$$\lim_{T\to\infty} [\lambda_1(T)z(T) + \lambda_2(T)p(T)] = 0$$

and the maximised Hamiltonian is concave in the state variables p, z then any path satisfying the necessary conditions is optimal in the catching up sense. Next we check the concavity of the maximised Hamiltonian. If $H^* = \max_c H$ then by the envelope theorem

$$\partial H^*/\partial p = \partial H/\partial p = -u_1 c/(1+p)^2 + \lambda_1\beta/(1+p)^2 - (u_2 + \lambda_2)(\beta - \alpha - \omega) \tag{53}$$

so

$$\partial^2 H^*/\partial p^2 = u_1/(1+p)^4\,[c(1+p) - \beta z] - [1 + 1/\eta]\,c/(1+p)^2 + u_{22}(\beta - \alpha - \omega)^2 \tag{54}$$

where η is the elasticity of the marginal utility of overall per capita consumption. Sufficient conditions for $\partial^2 H^*/\partial p^2$ to be negative are that $[1 + 1/\eta] > 0$, $[c(1+p) - \beta z] < 0$ and $u_{22} < 0$.

Similarly

$$\partial^2 H^*/\partial p \partial z = \lambda_1(z\beta' + \beta)/(1+p)^2 + (u_{22} + \lambda_2)(\beta' - \alpha' - \omega') \tag{55}$$

which is positive if the elasticity of β is greater than -1. Similarly

$$\partial^2 H^*/\partial z^2 = u_{22}(\beta' - \alpha' - \omega')^2 p^2 + (u_2 + \lambda_2)(\beta'' - \alpha'' - \omega'')p$$
$$+\lambda_1(f'' - 2(\alpha' - \beta'\frac{p}{1+p}) - z(\alpha'' - \beta''\frac{p}{1+p})) \quad (56)$$

This is negative if $\lambda_2(t) < 0$ and $\beta'' - \alpha'' - \omega'' > 0$. There is nothing in the theory to constrain the sign of $\lambda_2(t)$ a priori. Interpreting $\lambda_2(t)$ at time t as the shadow cost of a marginal increase in prevalence of the disease, we might expect $\lambda_2(t)$ to be negative since this has a cost in reducing instantaneous utilty. However, the increase in $p(t)$ raises the capital per productive worker ratio for a given capital stock by switching more of the population into the infected group (i.e. in the $\dot{z}$ equation the effect of $p(t)$ is positive) which is welfare improving. In the numerical examples presented, $\lambda_2(t)$ is negative for all t and at each steady state the other sufficient conditions for concavity are fulfilled.

3 Extensions and Alternatives

There are a variety of points at which alternative or more general assumptions could be made. We have excluded any individual adaptive behaviour to the prevalence of the disease. We have not allowed for the possibility of cure of the disease. There is a lot of scope for consideration of alternative social welfare functions. We have considered, economic factors as acting in a public way on all demographic-epidemiological parameters. That is there are no individual property rights in preventive health and all the demographic parameters are affected by an investment policy. There are also interesting special cases of our framework. For example if $\alpha' = \omega' = 0$ identically then the model can be thought of as finding the optimal targeted investment policy for controlling the spread of infection. Similar interpretations attach to the special cases in which only either α or ω vary with z. They correspond to net population growth or cure of the disease policies.

3.1 Individual Adaptive Behaviour

The most important area for allowing adaptive behaviour by individuals in the population is through β the infection rate e.g. in accounts of the plague in England the rich could take preventive segregation measures by physically moving out of a city which had plague outbreaks. To allow for this we take $\beta = \beta(z,p)$ with $\beta_z < 0, \beta_p < 0$. From an epidemiological perspective this is often considered as density dependent effects on the interaction process between the two subpopulations. In the descriptive economic model this has the effect of changing the nature of the phase space although the general pattern of steady states remain unchanged. The main effect on the $\dot{p} = 0$ locus is to generate additional nonlinearity[5] in the locus as in Figure 13 which shows two possible

[5] Here when $(\beta - \alpha - \omega) = 0$, $\frac{\partial p}{\partial z} = -\left(\beta' - \alpha' - \omega'\right)/\beta_p < 0$.

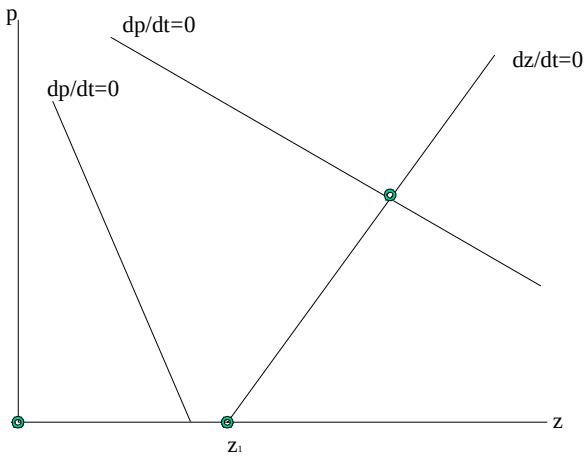

Figure 10: Individual adaptive behaviour

configurations.

The local stability properties of the first and third steady states are unchanged but they may change for the second steady state where $p^* \neq 0$ and $z^* \neq 0$ if it exists. The linearisations of (13) become

$$\begin{bmatrix} \dot{p}_t \\ \dot{z}_t \end{bmatrix} = \begin{bmatrix} \beta_p p + \beta - \alpha - \omega & (\beta_z - \alpha' - \omega')p \\ z\beta/(1+p)^2 + z\beta_p p/(1+p) & sf' - sf/z - \phi \\ & -z(\alpha' - \beta_z p/(1+p)) \end{bmatrix} \begin{bmatrix} p \\ z \end{bmatrix} \tag{57}$$

and evaluating the roots of this matrix at the first or third steady states leads to no change in sign pattern.

Around the second steady state when it exists we have the Jacobian

$$\begin{bmatrix} \beta_p p & (\beta_z - \alpha' - \omega')p \\ z\beta/(1+p)^2 + z\beta_p p/(1+p) & sf' - sf/z - \phi - z(\alpha' - \beta_z p/(1+p)) \end{bmatrix} \tag{58}$$

Here the trace is negative but the determinant may conceivably become negative leading to a saddle point. An example of the phase space is given in Figure 14 and Figure 15 where we have chosen

$$\beta = \beta_0/(1+p(t)) + \beta_1 \exp(-\beta_2 z(t))$$

together with the earlier set of parameters except that in the first diagram $g = 0.25$ and in the second diagram $g = 3$. When $g = 0.25$ the steady states are at $(p^* = 0.369, z^* = 0.500)$ and $(p^* = 0, z^* = 0.0203)$. The eigenvalues are then respectively -0.044, -0.280 for $p^* = 0$ and -0.313, 0.680 for $p^* \neq 0$. When $g = 3$ there is a single steady state $p^* = 0, z^* = 25$ which has eigenvalues -0.04, -0.32.

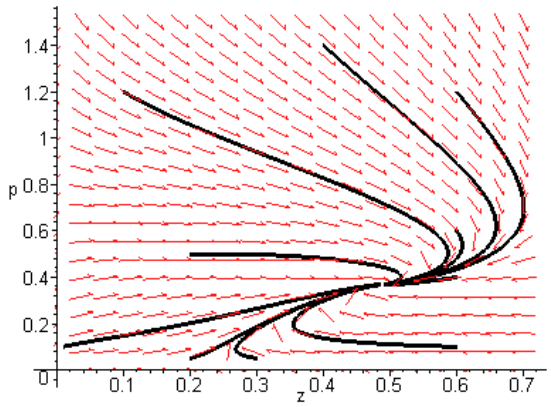

Figure 11: Individual adaptive behaviour, global phase space for $g = 0.25$.

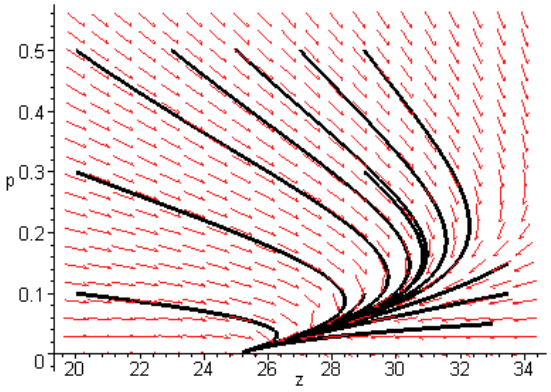

Figure 12: Individual adaptive behaviour, global phase space for $g = 3.0$.

3.2 Cure of the Disease

Recovery from the disease may take alternative forms when it is possible at all. For some diseases, like TB, individuals who recover from the disease develop immunity. In this case we would have three groups in the population: those who are healthy but susceptible to infection, those who are immune and those who are infected. The immune individuals and the healthy individuals can both work. The basic equations of the system thus become

$$\begin{cases} \dot{x}_t = \alpha x_t - \frac{\beta x_t y_t}{x_t+y_t+r_t} \\ \dot{y}_t = \frac{\beta x_t y_t}{x_t+y_t+r_t} - \omega y_t - \rho y_t \\ \dot{k}_t = sF(k_t, x_t + r_t) - \phi k_t \\ \dot{r}_t = \rho y_t - \gamma r_t \end{cases} \tag{59}$$

where ρ is the recovery rate of the infected and γ is the net death rate of the recovered individuals. We do not analyse this case here. Alternatively if the cured individuals do not develop immunity then the equations can be written as

$$\begin{cases} \dot{x}_t = \alpha x_t - \frac{\beta x_t y_t}{x_t+y_t} + \rho y_t \\ \dot{y}_t = \frac{\beta x_t y_t}{x_t+y_t} - \omega y_t - \rho y_t \\ \dot{k}_t = sF(k_t, x_t) - \phi k_t \end{cases} \tag{60}$$

Transforming the equations to p and z

$$\begin{cases} \dot{p} = (\beta - \omega - \alpha - \rho)p - \rho p^2 \\ \dot{z}_t = sf(z_t) - \phi z_t - z_t(\alpha - \beta p_t/(1+p_t) + \rho p_t) \end{cases} \tag{61}$$

The pure demographic-epidemiological system represented by the $\dot{p}$ equation now has the solution

$$p_t = p_0/(\frac{\rho}{\beta - \omega - \alpha - \rho}p_0 + (1 - \frac{\rho}{\beta - \omega - \alpha - \rho}p_0)e^{-(\beta-\omega-\alpha-\rho)t}) \tag{62}$$

The descriptive economic system still has a steady state at $p^* = z^* = 0$ and at $p^* = 0, sf(z^*) - \phi z^* - z^*\alpha = 0$. However the possibilities for the third steady state are more complex. Where $p^* \neq 0$ the $\dot{p} = 0$ locus solves $\beta - \omega - \alpha - \rho = \rho p$ and so has a negative slope under the assumption that $\beta' - \omega' - \alpha' - \rho' < 0$. In this case the $\dot{z} = 0$ locus is quadratic in p with a maximum: $sf(z) - \phi z - z\alpha + p(sf(z) - \phi z - z\alpha + \beta z - \rho z)) - \rho z p^2 = 0$. This then has implications for the existence of this steady state. The consequences on the local stability of the steady states are very similar to the case of adaptive individual behaviour: if $\rho = \rho(z)$ with $\rho'(z) > 0$ and $\beta(0) - \omega(0) - \alpha(0) - \rho(0) > 0$ the steady state at the origin is still locally unstable[6]. The steady state with $p^* = 0$ but $z^* \neq 0$ has

[6] At $p^* = z^* = 0$ the Jacobian is

$$\begin{bmatrix} \beta - \alpha - \omega - \rho & 0 \\ 0 & sf' - \phi - \alpha \end{bmatrix}$$

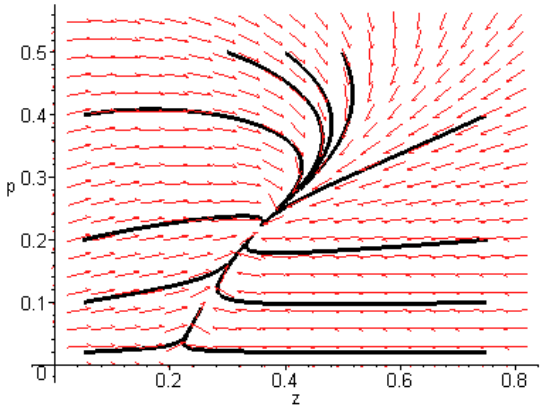

Figure 13: Allowing for recovery, global phase space for $g = 0.25$ and $\rho = 0.05$.

a negative trace and an ambiguous determinant and so is either a saddle point or a stable node[7]. The local stability of the third steady state where it exists is unclear[8]. In Figure 16 and Figure 17 for the usual parameter values when $g = 0.25$ the third steady state exists and is a stable node, whereas for $g = 3.0$ the third steady state fails to exist. In the numerical example when $g = 0.25$ and $\rho = 0.05$ the eigenvalues for the case $p^* \neq 0$, $z^* \neq 0$ are -0.067, -0.078 and when $p^* = 0$, $z^* \neq 0$ the eigenvalues are -0.119, 0.017.

3.3 Alternative Welfare Functions

One problem is to balance the interest of the current infected with those of future and present healthy susceptibles. From the viewpoint of output and consumption per capita the infected are a net drain on the economy; they do not work but under the assumption that consumption is equally distributed, they consume output thus "diluting" the consumption of the workers. They also infect healthy workers thus reducing the future labour force. This also raises the future capital worker ratio ceteris paribus which decreases the productivity of workers. If social welfare at one instant was taken to depend only on con-

[7] At $p^* = 0$, $z^* \neq 0$ the Jacobian is

$$\begin{bmatrix} \beta - \alpha - \omega - \rho & 0 \\ z\beta - \rho z & sf' - \phi - \alpha - z\alpha' \end{bmatrix}$$

[8] At $p^* \neq 0$, $z^* \neq 0$ the Jacobian is

$$\begin{bmatrix} \beta - \alpha - \omega - \rho - 2p\rho & (\beta' - \alpha' - \omega' - \rho')p - p^2\rho' \\ z\beta/(1+p)^2 - \rho z & sf' - \phi - (\alpha - \beta p/(1+p) + \rho p) - z(\alpha' - \beta' p/(1+p) + \rho') \end{bmatrix}$$

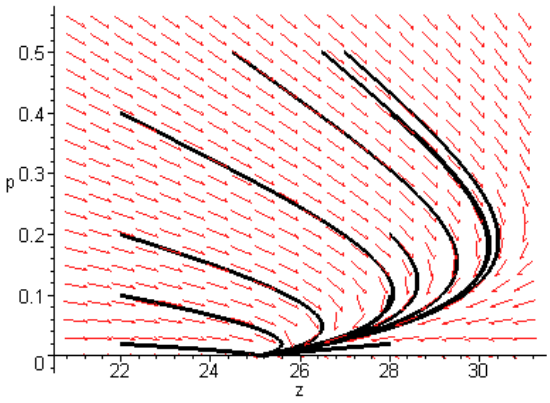

Figure 14: Allowing for recovery, global phase space for $g = 3.0$ and $\rho = 0.05$.

sumption per capita of the whole population then this would be maximised over the structure of the population by having zero infected. In a scenario in which resources can be allocated to alternative types of health care policy, there would be an incentive to minimise reductions in the death rate of the infected and to spend all resources on either net birth policy of the healthy or on preventive health care. With recovered individuals in the picture this is no longer so clear. There would then be an economic benefit from both trying to raise the recovery rate and reduce the death rate of the infected so that the chance of infected people recovering and then becoming productive workers is increased.

An obvious alternative welfare function is $\int u(c/(1+p),p)dt$ with $u(.,.)$ decreasing in its second argument. This gives some marginal differences in the optimal policy. The Hamiltonian becomes

$$H = u(c/(1+p),p) + \lambda_1(f(z) - \phi z - z(\alpha - \beta\frac{p}{1+p}) - c) + \lambda_2(\beta - \alpha - \omega)p \tag{63}$$

The equation for the optimal control is unchanged:

$$u_1 = \lambda_1(1+p) \tag{64}$$

The equations of motion for the costate variables are

$$\dot{\lambda_1} = -\lambda_1[f' - \phi - (\alpha - \beta p/(1+p)) - z(\alpha' - \beta' p/(1+p))] - \lambda_2 p(\beta' - \alpha' - \omega') \tag{65}$$

$$\dot{\lambda_2} = u_1 c/(1+p)^2 - u_p - \lambda_1 z\beta/(1+p)^2 - \lambda_2(\beta - \alpha - \omega) \tag{66}$$

Given that $u_2(.,0) = 0$ there are at least two steady states; one at $p^* = 0$ and with

$$f - \phi z^* - z^*(\alpha - \beta\frac{p^*}{1+p^*}) - c^* = 0 \tag{67}$$

$$f' - \phi - \alpha - z^*\alpha' = 0 \tag{68}$$

$$\lambda_1^*(c^* - z^*\beta) - \lambda_2^*(\beta - \alpha - \omega) = 0 \tag{69}$$

$$u_1 = \lambda_1^* \tag{70}$$

and any others being solutions of

$$\beta - \alpha - \omega = 0 \tag{71}$$

$$u_1 = \lambda_1^*(1 + p^*) \tag{72}$$

$$\lambda_1^*(c^*(1 + p^*) - z^*\beta) - u_2 = 0 \tag{73}$$

$$f - \phi z^* - z^*(\alpha - \beta\frac{p^*}{1+p^*}) - c^* = 0 \tag{74}$$

$$\begin{aligned}\lambda_1^*[f' - \phi - (\alpha - \beta p^*/(1+p^*)) - z^*(\alpha' - \beta' p^*/(1+p^*))] \\ +\lambda_2^* p^*(\beta' - \alpha' - \omega') = 0\end{aligned} \tag{75}$$

Here z^* is set by (71). If we invert (72) to get $c = u_1^{-1}(\lambda_1(1+p))$ and eliminate c, then it is quite possible that (73) and (74) have multiple solutions for λ_1^* and p^*. For any of these values for λ_1^* and p^*, λ_2^* is set by (75). Thus we may have more than two steady states. By contrast, linearising the dynamic equations around the steady states the only change in the Jacobians of (45) and (46) is in the first element of the last row which becomes $a_{41} = \frac{\delta(u_1 c)}{\delta p}\frac{1}{(1+p)^2} - \frac{2u_1 c}{(1+p)^3} + \frac{2\lambda_1 z\beta}{(1+p)^3} - u_{22}$. The four dimensional saddlepoint structure of the steady states is independent of the properties of the a_{41} term and so is preserved. The conditions for concavity of the maximised Hamiltonian in the state variables alter:

$$\partial^2 H^*/\partial p^2 = -2\lambda_1\beta/(1+p)^3 - u_{22} < 0 \qquad \text{if} \qquad u_{22} > 0 \tag{76}$$

Similarly

$$\partial^2 H^*/\partial p \partial z = \lambda_1 \beta^{'}/(1+p)^2 > 0 \tag{77}$$

So long as

$$\partial^2 H^*/\partial z^2 = \lambda_2(\beta'' - \alpha'' - \omega'')p + \lambda_1(f'' - 2(\alpha' - \beta'\frac{p}{1+p}) - z(\alpha'' - \beta''\frac{p}{1+p})) < 0 \tag{78}$$

under suitable conditions on the second derivatives of the demographic-epidemiological parameters.

4 Conclusions

We use the May and Anderson (1989) version of the Lotka-Volterra equations to model the interaction between healthy and infected individuals and then embed that process in an accumulation model of productive capital.

There is a two way interaction between the population structure and the economy. Infected individuals are unproductive and cannot work, the higher the state of economic development, the lower is the infection rate and the death rates of both the healthy and the infected. Taken on its own the population structure has no steady state (other than extinction when parameters have the appropriate magnitude). If there is an exogenous savings rule determining the rate of accumulation, we find that with both types of capital stock, there is a steady state in which the disease is eradicated and the capital-healthy labour ratio is constant. However, if the level of economic development required to keep the prevalence of the disease constant is high relative to the productivity of technology and the savings rate, then there is also a steady state in which prevalence of the disease is constant at a positive level of the disease. When this steady state exists it is globally stable and the disease is not eradicated by economic effects in the long run. However, if the savings rate is high or the technology sufficiently productive for this steady state to fail to exist, then the economy converges to the steady state in which the disease is eliminated and the capital-healthy labour ratio is constant. In any case the convergence to a long run equilibrium is generally oscillatory. Extinction of the population and a zero level of capital stock is also a steady state but it is unstable. We illustrate these qualitative results with examples of numerical integration using a Cobb-Douglas production function and 'plausible' functions for the way in which economic development impacts on the demographic-epidemiological parameters. The qualitative results are consistent with broad historical observation that control of major infectious diseases has come pari passu with economic development.

We also examine the situation in which a central planner can determine the process of accumulation and does so in order to optimise a measure of intertemporal welfare. When population is changing it is not easy to see how the latter should be defined; we take it to be a time additive function of the level of overall

per capita consumption of those alive and of the rate of change of prevalence. In normative terms prevalence has an ambiguous effect: on the one hand higher prevalence lowers instantaneous per capita consumption; on the other hand it makes the existing capital stock go further when shared between the lower number of healthy workers which impacts on the future infection rate and death rates by raising economic development. We find that the centrally planned system has two steady states: one of them is locally a 4-D saddle point in which the disease is eradicated; the other has a positive prevalence of the disease and a richer variety of possible local dynamic patterns depending on parameter values. With four nonlinear differential equations and two steady states it is difficult to get further qualitative information on the system; so we exploit a technique of numerically integrating around a contour of the maximised Hamiltonian which allows us to eliminate one dimension and to plot three dimensional phase spaces.

Various extensions are considered: to allow for recovery from the disease; to try to analyse sensitivity of the planned economy results to the form of intertemporal welfare function and to consider prevalence health action. Generally these extensions do not change the basic form of the dynamics of the system but they indicate that an important outstanding research question is to relate the aggregate macro view of the system that is used here to more explicit micro underpinnings.

A Appendix 1

A.1 The Jacobian for the centrally planned productive model

Linearising the dynamic equations (22), (25) and (26) gives a Jacobian of

$$\begin{bmatrix} (\beta-\alpha-\omega) & p(\beta'-\alpha'-\omega') & 0 & 0 \\ \frac{z\beta}{(1+p)^2}-\frac{\delta c}{\delta p} & f'-\phi-(\alpha-\beta\frac{p}{(1+p)}) & -\frac{\delta c}{\delta \lambda_1} & 0 \\ & -z(\alpha^{'}-\beta^{'}\frac{p}{(1+p)}) & & \\ & -\lambda_1[f^{''}-2(\alpha^{'}-\beta^{'}\frac{p}{(1+p)})] & & \\ -\frac{\lambda_1}{(1+p)^2}(\beta+z\beta^{'}) & +\lambda_1 z(\alpha^{''}-\beta^{''}\frac{p}{(1+p)}) & -(f^{'}-(\alpha-\phi-\beta\frac{p}{(1+p)}) & -p(\beta'-\alpha'-\omega') \\ -\lambda_2(\beta'-\alpha'-\omega') & -\lambda_2 p(\beta^{''}-\alpha^{''}-\omega^{''}) & -z(\alpha^{'}-\beta^{'}\frac{p}{(1+p)})) & \\ \frac{\delta(u_1 c)}{\delta p}\frac{1}{(1+p)^2} & -\frac{\lambda_1}{(1+p)^2}(\beta+z\beta^{'}) & \frac{\delta(u_1 c)}{\delta \lambda_1}\frac{1}{(1+p)^2} & \\ -\frac{2u_1 c}{(1+p)^3}+\frac{2\lambda_1 z\beta}{(1+p)^3_1} & -\lambda_2(\beta'-\alpha'-\omega') & -\frac{z\beta}{(1+p)^2} & -(\beta-\alpha-\omega) \end{bmatrix} \tag{79}$$

where

$$\frac{\delta c}{\delta p} = \lambda_1(1+p)/u_{11} + c/(1+p) \tag{80}$$

$$\frac{\delta(u_1 c)}{\delta p} = \lambda_1(1+p)^2/u_{11} + 2\lambda_1 c)] \tag{81}$$

$$\frac{\delta c}{\delta \lambda_1} = (1+p)^2/u_{11} \tag{82}$$

$$\frac{(u_1 c)}{\delta \lambda_1} = c(1+p) + u_1(1+p)^2/u_{11}$$

$$\frac{\delta(u_1 c)}{\delta z} = \frac{\delta c}{\delta z} = \frac{\delta c}{\delta \lambda_2} = 0 \tag{83}$$

Specialising this to the values around the steady state with $p^* = 0$ (using the steady state equations) gives

$$\begin{bmatrix} (\beta-\alpha-\omega) & 0 & 0 & 0 \\ z\beta-\lambda_1/u_{11}+c & 0 & -1/u_{11} & 0 \\ -\lambda_1(\beta+z\beta^{'}) & -\lambda_1[f^{''}-2\alpha^{'}-z\alpha^{''})] & 0 & 0 \\ -\lambda_2(\beta'-\alpha'-\omega') & & & \\ u_1^2/u_{11}+2\lambda_1 z\beta & -\lambda_1(\beta+z\beta^{'}) & c+u_1/u_{11}-z\beta & -(\beta-\alpha-\omega) \\ & -\lambda_2(\beta'-\alpha'-\omega') & & \end{bmatrix} \tag{84}$$

Around the steady state with $p^* \neq 0$ the Jacobian matrix becomes

$$\begin{bmatrix} 0 & p(\beta'-\alpha'-\omega') & 0 & 0 \\ -\lambda_1(1+p)/u_{11} & -\frac{\lambda_1}{\lambda_2}(\beta'-\alpha'-\omega') & -(1+p)^2/u_{11} & 0 \\ -\frac{\lambda_1}{(1+p)^2}(\beta+z\beta^{'}) & -\lambda_1[f^{''}-2(\alpha^{'}-\beta^{'}\frac{p}{(1+p)})] & & \\ -\lambda_2(\beta'-\alpha'-\omega') & +\lambda_1 z(\alpha^{''}-\beta^{''}\frac{p}{(1+p)}) & \frac{\lambda_1}{\lambda_2}(\beta'-\alpha'-\omega') & -p(\beta'-\alpha'-\omega') \\ & -\lambda_2 p(\beta''-\alpha''-\omega'') & & \\ \lambda_1/(1+p) & -\frac{\lambda_1}{(1+p)^2}(\beta+z\beta^{'}) & & \\ \left[u_1/u_{11}+2z\beta/(1+p)^2\right] & -\lambda_2(\beta'-\alpha'-\omega') & u_1/u_{11} & 0 \end{bmatrix} \tag{85}$$

A.2 Sufficiency of the necessary conditions for optimality with exogenous growth in prevalence

Let $z^*(t)$, $\lambda_1^*(t)$, $c^*(t)$ satisfy

$$u_1 = \lambda_1^*(1+p_0e^{\gamma t}) \tag{86}$$

$$\dot{z} = f(z^*) - \phi z^* - z^*(\alpha - \beta\frac{p_0e^{\gamma t}}{1+p_0e^{\gamma t}}) - c^* \tag{87}$$

$$\dot{\lambda_1} = -\lambda_1^*[f' - \phi - (\alpha - \beta p_0e^{\gamma t}/(1+p_0e^{\gamma t}))] \tag{88}$$

and let $z(t), c(t)$ be any feasible path which satisfies (87). Then

$$\int u(\frac{c}{1+p_0e^{\gamma t}}) - u(\frac{c^*}{1+p_0e^{\gamma t}})dt \leq \int u_1^*(\frac{c}{1+p_0e^{\gamma t}} - \frac{c^*}{1+p_0e^{\gamma t}})dt = \tag{89}$$

$$\int \lambda_1^*[(f(z) - \phi z - z(\alpha - \beta\frac{p_0e^{\gamma t}}{1+p_0e^{\gamma t}}) - \dot{z}) \tag{90}$$

$$-(f(z^*) - \phi z^* - z^*(\alpha - \beta\frac{p_0e^{\gamma t}}{1+p_0e^{\gamma t}}) - \dot{z}^*) \tag{91}$$

$$\leq \int \lambda_1^*[\{f'^* - \phi - (\alpha - \beta\frac{p_0e^{\gamma t}}{1+p_0e^{\gamma t}})\}(z-z^*) - (\dot{z}-\dot{z}^*)]dt \tag{92}$$

$$= \int [-\lambda_1^*(z-z^*) - \lambda_1^*(\dot{z-z}^*)]dt \tag{93}$$

$$= -\lim_{t\to\infty} \lambda_1^*(t)z(t) \leq 0 \tag{94}$$

References

[1] Alvi, A. R. et al. (1998). "Prevalence of pulmonary tuberculosis on the roof of the world", *International Journal of Tuberculosis and Lung Diseases*, 2 (11), 909-913.

[2] Bartel, A. and Taubman, P. (1979). "Health and labor market success: The role of various diseases", *Review of Economics and Statistics*, 61, 1-8.

[3] Chiang, A. (1992). *Elements of Dynamic Optimization*, MCGraw-Hill, New York.

[4] Dankowicz, H. (1997). *Chaotic Dynamics in Hamiltonian Systems*, World Scientific, Singapore.

[5] Dasgupta, P. (1969). "On the concept of optimal population", *Review of Economic Studies*, 36, 3, 295-318.

[6] Dasgupta, P. (1974). "Some alternative criteria for justice between generations", *Journal of Public Economics*, 3, 405-424.

[7] De Cock, K. M. and Chaisson, R. E. (1999). "Will DOTS do it? A reappraisal of tuberculosis control in countries with high rates of HIV infection", *International Journal of Tuberculosis and Lung Diseases*, 3 (6), 457-465.

[8] Ettner, S. L. (1996). "New evidence on the relationship between income and health", *Journal of Health Economics*, 15, 67-85.

[9] Grout, P. A. (1982). "Welfare Economics of Decision Making with Changing Preferences", *Review of Economic Studies*, 49, 83-90.

[10] Kimerling, M. E. et al. (1999). "Inadequacy of the current WHO retreatment regimen in a central Siberian prison: treatment failure and MDR-TB", *International Journal of Tuberculosis and Lung Diseases*, 3 (5), 451-453.

[11] Lee, L. F. (1982). "Health and wage: a simultaneous equations model with multiple discrete indicators", *International Economic Review*, 23, 192-222.

[12] Leonard, D. (1981). "The signs of the Costate Variables and Sufficiency Conditions in a Class of Optimal Control Problems", *Economic Letters*, 8, 321-325.

[13] Leonard, D. and Long, N. V. (1992). *Optimal Control Theory and Static Optimization in Economics*, Cambridge University Press, Cambridge.

[14] Luft, H. S. (1975). "The impact of poor health on earnings", *Review of Economics and Statistics*, 57, 43-57.

[15] Mangarasian, O. L. (1966). "Sufficient conditions for the Optimal Control of Nonlinear Systems", *Journal of SIAM Control*, 4, 139-152.

[16] May, R. M. and Anderson, R. M. (1989). "The Transmission Dynamics of Human Immunodeficiency Virus", in S. A. Levin, T. G. Hallam and L. J. Gross, Eds., *Applied Mathematical Ecology*, Springer-Verlag, Berlin.

[17] Meade, J. E. (1955). *Trade and Welfare*, Oxford University Press, Oxford.

[18] Netto, E. M. et al. (1999). "Progress in global tuberculosis control 1995-1996, with emphasis on 22 high-incidence countries", *International Journal of Tuberculosis and Lung Diseases*, 3 (4), 310-320.

[19] Rawls, J. (1972). *A theory of justice*, Oxford University Press, Oxford.

[20] Watts, S. (1997). *Disease, Power and Imperialism*, Yale University Press, New Haven and London.

[21] WHO, (1996). *HIV/AIDS: the global epidemic.*

[22] WHO, (1997). *Tuberculosis and AIDS.*

[23] WHO, (2000). *The World Health Report 2000, Health Systems: Improving Performance.*

[24] Yaari, M. E. (1981). "Rawls, Edgeworth, Shapley, Nash - Theories of distributive justice reexamined", *Journal of Economic Theory*, 24, 1, 1-39.