Environment and Development Economics 9: 591–612 © 2004 Cambridge University Press
DOI: 10.1017/S1355770X04001524 Printed in the United Kingdom

Negative externalities, defensive expenditures, and labour supply in an evolutionary context

ANGELO ANTOCI
Dipartimento di Economia, Impresa e Regolamentazione, University of Sassari, via Sardegna 58, 07100 Sassari, Italy.
E-mail: angelo.antoci@virgilio.it

STEFANO BARTOLINI
Dipartimento di Economia Politica, University of Siena, Piazza. S. Francesco 7, 53100 Siena, Italy. E-mail: bartolinist@unisi.it

ABSTRACT. In this model, well-being depends on leisure, on an environmental renewable resource, and on a non-storable output, which can substitute for the environmental resource or can satisfy needs different from those satisfied by the resource. Individuals have free access to the environmental resource, which is subject to negative externalities: that is, is depleted by the production and consumption of the output. Individuals react to negative externalities by increasing their labour supply in order to produce substitutes for the diminishing resource. The increase in production and consumption that ensues generates further deterioration of the future quality or quantity of the free resource, thus giving rise to a self-reinforcing process. Multiple equilibria and 'critical mass effects' are consistent with the functioning of this economy and the resulting level of aggregate production may be higher than is socially desirable.

1. Introduction

Our main argument is that negative externalities can be a cause of growth in per capita output. Indeed, the evolutionary model presented here suggests that negative externalities may generate an increase in per capita output due to an increase in the labour supply.

We present a three-goods model. Economic agents' well-being depends on leisure, a free environmental renewable resource, and a non-storable output, which may be used as a substitute for the resource or employed to satisfy needs different from those satisfied by the resource. The environmental resource is caused to deteriorate by the production and consumption of the output. Faced with a reduction of the resource, agents may react by increasing the labour supply in order to produce and consume private substitutes for the diminishing resource, i.e. they may raise their defensive expenditures. By doing so, each of them contributes to raising aggregate production. The detrimental impact of each individual's activity

We are especially indebted for their comments to A. Leijonhufvud, R. Lopez, U. Pagano, L. Punzo, and three anonymous referees. The usual caveats apply.

on the resource is negligible, but the aggregate effect is remarkable; hence the increase in output generates a further deterioration of the environmental resource, thus giving rise to a self-feeding output growth process. During this process we observe a shift of consumption habits from free consumptions (the environmental resource) to costly ones (private consumption).

Given that the growth of output is negatively correlated with the other two goods – leisure and the environmental resource – in an economy of this kind, using output as an indicator of well-being leads to systematic overestimation of it. The extent of this overestimation is such that, in many of the dynamics of our model, the increase in output takes place at the price of a worsening of individuals' well-being: the uncoordinated efforts of individuals to defend themselves against negative externalities may push the economy along trajectories in which an increase in income does not off-set the loss of well-being due to the decline of environmental assets and the increase in work effort. In this case, output growth is both the consequence and the cause of a diminution in well-being. Not surprisingly, in this context, coordination failures occur. The novel feature of this approach is that it views coordination failure as resulting in the growth of the economy's aggregate output.

In our model, individuals are forced increasingly to satisfy their needs by relying on market goods in order to off-set a diminution in their well-being due to the decline in free resources; that is, individuals must make (private) defensive expenditures to counterbalance environmental deterioration.

There are many examples of this mechanism. Imagine that if the quality of water where people can swim for free (e.g. the sea or the river close to home) is spoilt by pollution, agents may decide to buy a substitute, e.g. a swimming pool or a holiday in some tropical resort. An example of this kind seems to be only a paradigmatic case of much more general considerations. The quality of the environment that used to be available for free until a generation ago in the rich countries is now only available at high costs. Holiday travel, second homes on the coast or in the countryside, etc., are the principal means to escape from the congested and polluted environments in which an enormous proportion of the population lives.[1] Considerable and increasing resources are devoted to pursuit of this escape. A crucial resource for many developing countries has become tourism from the rich countries, while the latter experience periodic mass migrations called summer holidays.

As regards leisure, urbanization seems to have been accompanied by an increase in the costs of its enjoyment connected with the progressive substitution of free goods with costly ones. According to Cross (1993), the pattern of leisure consumption in the industrial and post-industrial societies has evolved towards the greater use of costly goods at the expense of communal and collective leisure activities. Cities are places built for work,

[1] The observation that urban living generates a need to escape from pollution, noise, and congestion is extremely ancient and dates back to the first metropolises. For example, Juvenal complained that it was only possible to find peace and quiet in ancient Rome by possessing a country house outside the city (cited by Hueting, 1980).

where low-cost opportunities for leisure are extremely rare, beginning with the scarcity of places where people can meet. This is evidenced by the distress of the categories of the population enjoying more leisure. From the point of view of leisure, cities have the advantage of offering a wide variety of costly entertainment, and the symmetrical disadvantage that cheap entertainment is difficult to find (see e.g. Hueting, 1980). In this light, for instance, the massive growth of 'home entertainment' may be interpreted as a reaction against the difficulty of finding low-cost places to meet in an urban environment.

In any case, the prolonged enjoyment of leisure (weekends, holidays) takes place as much as possible outside the city. The city induces a separation between the places of work and leisure which creates costs. What Polanyi (1968) considered as a separation between the time of life and the time of work is an essentially urban phenomenon.

A paradigmatic example of the mechanism of this model – that is, the increase in output raises (through negative externalities) the demand for output – is the 'air conditioners trap' in Tokyo. The temperature of the city, which is very hot in summer, is considerably increased by the air conditioners in general and constant use. They cool the interior of buildings but emit heat to the exterior. This is a trap in the sense of a self-reinforcing mechanism: the increase in the use of air conditioners increases the demand for air conditioners, because people are forced to buy air conditioners by their widespread use.

The concept of defensive expenditures was introduced into the economic literature by Hirsch (1976) to refer to all those consumption choices made by individuals to defend themselves against the negative externalities due to economic growth. Many of Hirsch's ideas anticipated themes treated here.

The notion of defensive expenditure has subsequently become known in the environmental literature, where it has been applied in an attempt to correct GNP into a more reliable indicator of well-being. Under this approach, defensive expenditures should be subtracted from GNP because they do not increase the net availability of goods. Yet this literature has found it enormously difficult to identify defensive expenditures. In fact, expenditures which are intrinsically defensive are rare. Many of the canonical examples in the literature have a clearly defensive nature: double glazing is certainly a form of defence against noise; the use of mineral water is a substitute for tap water; expenditures for pollution abatement or prevention, for the treatment of illnesses caused by pollution, or for soil restoration are a direct response to environmental degradation. But the feature shared by the examples given at the beginning of this section is that these expenditures may or may not be defensive according to the motives for undertaking them. And this makes their certain identification very difficult. The purchase of a swimming pool may be defensive if it is a response to the deterioration of the local water, but it may not be defensive if it is prompted *ceteris paribus* by an increase in income.[2]

[2] The difficulty of identifying defensive expenditures due to the problem of motives has been noted since 1980 by Hueting (1980, pp. 177–178).

Understandably, econometric estimates of defensive expenditure have concentrated on a rather restrictive identification in order to avert the criticism that they unjustifiably inflate defensive expenditures. The result is probably a substantial underestimation of the phenomenon.[3]

The role of negative externalities as a factor contributing to growth is an under-investigated topic. In fact, neither the literature on endogenous growth nor growth models with environmental resources attribute any role to negative externalities as a possible cause of growth.

With regard to the theory of endogenous growth, this has concentrated entirely on the role of positive externalities as the engine of growth.

As far as the literature on sustainable development is concerned, this tends to view economic growth as limited by the finiteness of environmental resources. This literature, in fact, has been generated by doubts about the existence of 'limits to growth' due to the limits of the environment. As a consequence of this 'imprinting', growth models including environmental resources have been used to define what those limits are. They have concentrated, that is to say, on identifying the conditions that a process of economic growth must satisfy to be 'sustainable', i.e. on the definition of sustainability. The question to which an answer is sought is this: how extensive is the limit represented by the finiteness of resources? How stringent are the conditions that an economy must respect in order not to exceed that limit?

Hence, this literature has never explored the possibility that the deterioration of resources may be a cause of growth as well as an effect.[4] By contrast, we argue that there are reasons to believe that negative externalities may be a factor which contributes to growth, and that there are matters that can be explained by this consideration.

Since our model does not include capital accumulation, it does not lend itself to description of the long-period income dynamic; the conclusions of the model refer to a limited time horizon. However, the analysis performed can be extended to models with accumulation. In fact, the substitution mechanism described has been introduced into growth models with

[3] See Leipert and Simonis (1989) for estimates concerning the German economy and Cullino (1993) for the Italian one.

[4] Shogren and Crocker (1991) offer a partial treatment of the mechanism described here. They point out that, 'most environmental policy does not resolve environmental problems ... While continuing to allow the mass of waste to flow into the environment, it simply transfers through time and across space' (pp. 195–6). Shogren and Crocker use a static model to show that if defensive expenditures transfer externalities, then individuals protect themselves to an extent higher than the socially optimal level. However, they only consider pollution abatement policies which directly transfer the pollution to other subjects. Although our model is able to capture this particular case, in it the individual contribution to externalities can be seen as negligible even if the aggregate effect is substantial; hence the increase in externalities is due to the general and indirect effect of an increase in income. This enables us to describe a self-reinforcing mechanism which is intrinsically dynamic in nature.

optimizing agents with results that confirm and reinforce those presented here, allowing them to be extended to the long period.[5]

The paper is organized as follows. In section 2 we present the model; in sections 3, 4, and 5 we analyse it. Section 6 concludes.

2. The model

Let us consider a population of individuals (a continuum) of size N; individuals' well-being depends on three goods: leisure, a free environmental resource E, and a non-storable output Y that can be used to satisfy needs different from those satisfied by the environmental resource, Y_1, or as a substitute for the free resource, Y_2. For simplicity, we may consider Y as a homogeneous good which can be used for different purposes. However, Y may also be interpreted as an aggregate measure of heterogeneous goods.

2.1. Strategies: little time or little money

Let us assume that, *in each instant of time* t, individuals have to choose between the following two options (*strategies*):

(1) *Strategy* (l): They produce the flow $\bar{Y}_1$ of the output. We assume that, to obtain $\bar{Y}_1$, they must work at a rate $L = L^l$ (they work 'little').
(2) *Strategy* (h): They produce the flow $\bar{Y}_1 + \bar{Y}_2$ of the output. To produce $\bar{Y}_1 + \bar{Y}_2$, they must work at a rate $L = L^h$, where $1 > L^h > L^l > 0$ (they work 'hard').

$\bar{Y}_1, \bar{Y}_2, L^l$, and L^h are strictly positive parameters of the model.

$\bar{Y}_1$ and $\bar{Y}_1 + \bar{Y}_2$ can be interpreted as the wage-goods obtained if the labour supplied is respectively L^l and L^h. Roughly speaking, the two alternative strategies depict a context in which individuals can choose between having little time or having little money.

The parameter $\bar{Y}_2$ simply indicates the additional output with respect to $\bar{Y}_1$ obtainable from increasing work effort from L^l to L^h.

2.2. Environmental impact of production and consumption

All individuals have free access to the environmental resource in every instant of time t; no individual is excluded from the consumption of the resource. Let $E(t)$ be the stock of the free resource at the instant of time t. Let us assume that the time derivative $\dot{E}(t)$ of $E(t)$ – without the negative impact of individuals' economic activity – is given by the usual logistic equation

$$\dot{E}(t) = \beta E(t)[\bar{E} - E(t)] \tag{1}$$

where $\bar{E}$ and β are strictly positive parameters; $\bar{E}$ represents the value to which $E(t)$ would tend in the absence of production and of consumption; $\bar{E}$ can therefore be interpreted as the economy's 'endowment' of the

[5] This model is a development of Antoci and Bartolini (1997, 1999), which are the first models in which growth of per capita output is fed by negative externalities. This idea has been latterly transferred to a world with optimizing agents in Bartolini and Bonatti (2002). In a model *à la* Solow–Ramsey they confirm and reinforce the results obtained with evolutionary choice mechanisms.

environmental good. The parameter β measures the speed of convergence of E to $\bar{E}$.

Let $x(t)$ be the proportion of individuals adopting strategy (l) at the instant of time t, $0 \leq x(t) \leq 1$. Consequently, $1 - x(t)$ is the proportion of individuals choosing strategy (h). We assume that the aggregate negative impacts on $\dot{E}(t)/E(t)$ by individuals choosing strategies (l) and (h) are respectively $\gamma \bar{Y}_1 x N$ and $\gamma(\bar{Y}_1 + \bar{Y}_2)(1 - x)N$, where γ is a strictly positive parameter that measures the environmental impact of the production and consumption of the output.

Augmenting equation (1) by the negative impact of individuals' activity we obtain

$$\begin{aligned} \dot{E} &= E[\beta(\bar{E} - E) - \gamma \bar{Y}_1 x N - \gamma(\bar{Y}_1 + \bar{Y}_2)(1 - x)N] \\ &= E[\beta(\bar{E} - E) - \gamma N(\bar{Y}_1 + \bar{Y}_2(1 - x))] \end{aligned} \tag{2}$$

We assume that production and consumption activities freely dispose of their polluting waste because of the absence of property rights on the natural resource. Although a single agent's productive activity has a negligible impact on environmental quality, the aggregate effect of individuals' production in instant t is not negligible and depends on the technological parameter γ, on N and $x(t)$. In fact, according to (2), $\dot{E}(t)$ is an increasing function of x, the proportion of individuals who work little. Furthermore, given the distribution of strategies across the population, $\dot{E}(t)$ decreases if the values of the technological parameter γ or the population size N increase.

2.3. Payoffs

We assume that all individuals have the same payoff function and that it is of the Cobb–Douglas type. Furthermore, we assume that E and $\bar{Y}_2$ are perfect substitutes with a marginal rate of substitution equal to the parameter $d > 0$. In particular, the payoffs from strategies (l) and (h) are respectively

$$\begin{aligned} U_l(E) &\equiv E^a(1 - L^l)^b(\bar{Y}_1)^{1-a-b} \\ U_h(E) &\equiv (E + d\bar{Y}_2)^a(1 - L^h)^b(\bar{Y}_1)^{1-a-b} \end{aligned}$$

where $a, b > 0$ and $a + b < 1$; $1 - L^i (i = l, h)$ represents leisure.

Note that the output, which is a homogeneous good, can be used to substitute the environmental resource $(\bar{Y}_2)$ or to satisfy needs different from those satisfied by the resource $(\bar{Y}_1)$.[6]

Since the consumption of $\bar{Y}_2$ is not affected by negative externalities, unlike E, strategy (h) provides individuals with a self-protection device against negative externalities based on the substitution for a common consumption, E, by a private one, $\bar{Y}_2$. This costly substitute must be financed by working more.

[6] In macroeconomic models it is quite common practice to treat a homogeneous good as utilizable for different purposes (for example, consumption or accumulation).

The hypothesis of perfect substitutability between E and $\bar{Y}_2$ can be relaxed by assuming only 'imperfect' substitutability, obtaining similar results[7] (see appendix 1). The unique feature of this function playing a central role in our analysis is that the payoff of individuals consuming $\bar{Y}_1 + \bar{Y}_2$ decreases less than that of others when the stock E of environmental good decreases.

Note that all individuals have access to the same stock $E(t)$ of the environmental resource, which is both an individual and an aggregate endowment. That is, there is no rivalry in the consumption of this resource among individuals, for whom it is a pure public good.

2.4. Economic dynamics

The evolution over time of the stock E of the environmental resource will be analysed under two alternative hypotheses concerning the adoption of strategies (l) and (h). We shall first analyse the dynamic of E assuming that the adoption process of the two strategies is characterized by a certain amount of inertia. That is to say, at a given instant of time, only a small proportion of the population considers the possibility of changing its strategy, and does so if it perceives that the alternative strategy is more remunerative. Under this hypothesis, it may happen that some members of the population adopt the less remunerative strategy; the more remunerative one (given the value of E) being adopted only as the final result of an *adaptive* transition dynamic.

We shall then consider the alternative hypothesis that each individual, at each instant of time, adopts the strategy that ensures the highest payoff (best response dynamic). In this case, the variable x becomes a function of E and does not vary with continuity. In particular, given the value of E, it 'jumps' from value 0 to value 1, or vice versa, according to which of the two strategies is more remunerative. We shall show that in this case the dynamics of the economy can be described by trajectories which are also present among those of the adaptive dynamic: suffice it to consider the initial value of x as equal to 0 or to 1 as the case may be. We shall also show that the best response dynamic and the adaptive dynamic have the same attractive fixed points.

In what follows we analyse the adaptive adoption process of the two strategies outlined above, assuming that this process is described by the so-called 'replicator dynamic' (see e.g. Weibull, 1995), according to which those behaviours that yield an above-average payoff will spread more rapidly at the expense of less rewarding ones. In particular, we assume that the dynamic of x is given by

$$\dot{x} = x[U_l(E) - \bar{U}(E, x)]$$

where $\dot{x}$ is the time derivative of $x(t)$ and $\bar{U}(E, x) \equiv U_l(E)x + U_h(E)(1 - x)$ is the average payoff.

Replicator dynamic may be generated by several individual and social learning mechanisms (see e.g. Borgers and Sarin, 1997; Schlag, 1998). It

[7] Furthermore, the results concerning well-being analysis hold *a fortiori* in the latter case.

https://doi.org/10.1017/S1355770X04001524 Published online by Cambridge University Press

is generally used to model the adoption process of strategies in contexts in which individuals play strategies on the occasion of random pair-wise encounters. However, it is also possible to find rationales for replicator dynamic when, as in our case, the payoff of each strategy depends on the choices of *all* individuals in the population; that is, outside the random matching paradigm (see, e.g. Sacco, 1994; for an application of replicator dynamic in a context similar to ours, see Sethi and Somanathan, 1996).

It is evident that replicator dynamic can be rewritten in the following form

$$\dot{x} = x(1-x)\Delta U(E) \tag{3}$$

where $\Delta U(E) \equiv U_l(E) - U_h(E)$ is the payoff difference between strategies (l) and (h).

In the context analysed here (where only two strategies are present), replicator dynamic has the same attractive fixed points as any other dynamic of adaptive type which satisfies the following condition

$$\dot{x} > 0 \text{ if } \Delta U(E) > 0, \quad \dot{x} < 0 \text{ if } \Delta U(E) < 0, \quad \dot{x} = 0 \text{ if } \Delta U(E) = 0,$$

for every x such that $0 < x < 1$;[8] in fact, all the results in the appendices are obtained using this property alone.

3. Classification of dynamics

Since there is no accumulation of assets in our economy, the aggregate level of production and (private) consumption is $\bar{Y}_1 x N + (\bar{Y}_1 + \bar{Y}_2)(1-x)N$, which is a decreasing function of x. Thus, the growth of the activity level is represented by an increase in the proportion of individuals that choose to work hard.

This section provides a classification of the dynamic regimes under equations (2) and (3).[9] We shall use the following terminology:

L^l-dominance: we shall classify a dynamic regime as being of the L^l-dominance type if there exists only one single attractive fixed point and if within it $x = 1$; that is, all the individuals adopt strategy (l);

L^h-dominance: we shall classify a dynamic regime as being of the L^h-dominance type if there exists only one single attractive fixed point and if within it $x = 0$; that is, all the individuals adopt strategy (h);

Bi-stable dynamics: we shall classify a dynamic regime as being of the bi-stable type if there exist only two attractive fixed points; one with $x = 0$ and the other with $x = 1$.

[8] Besides the usual 'boundary' conditions which ensure that x does not become negative or greater than 1.

[9] For the sake of simplicity, we shall not consider 'non-robust' cases, that is, those corresponding to equality conditions on parameters.

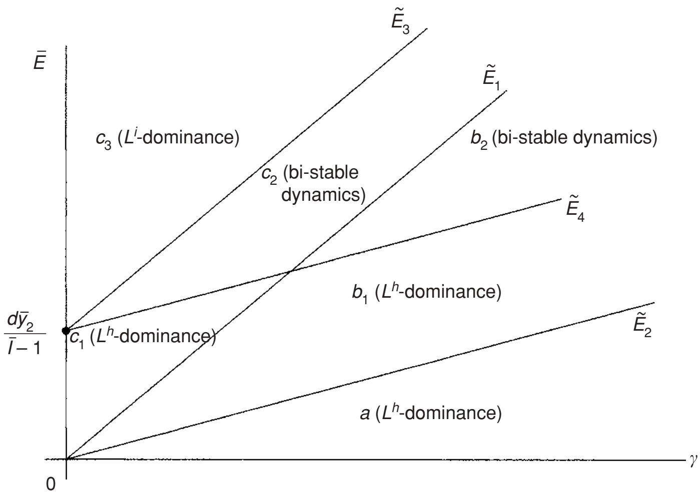

Figure 1. *Collocation of dynamic regimes in the plane* $(\gamma, \bar{E})$

Let us consider the following straight lines in the plane $(\gamma, \bar{E})$ (see appendix 1) (see figure 1)[10]

$$\bar{E} = \tilde{E}_1(\gamma) \equiv \frac{N(\bar{Y}_1 + \bar{Y}_2)}{\beta}\gamma \qquad \bar{E} = \tilde{E}_2(\gamma) \equiv \frac{N\bar{Y}_1}{\beta}\gamma$$

$$\bar{E} = \tilde{E}_3(\gamma) \equiv \frac{d\bar{Y}_2}{\bar{l}-1} + \frac{N(\bar{Y}_1 + \bar{Y}_2)}{\beta}\gamma \qquad \bar{E} = \tilde{E}_4(\gamma) \equiv \frac{d\bar{Y}_2}{\bar{l}-1} + \frac{N\bar{Y}_1}{\beta}\gamma$$

where $\bar{l} \equiv (\frac{1-L^l}{1-L^h})^{\frac{b}{a}} > 1$.

The details of the mathematical analysis of equations (2) and (3) are given in appendix 1. The results of such analysis can be summarised as follows:

Case (a): $\bar{E} < \tilde{E}_2(\gamma)$ — **Dynamics**: L^h-dominance (figure 2)
Case (b): $\tilde{E}_2(\gamma) < \bar{E} < \tilde{E}_1(\gamma)$
- **Sub-case (b.1)**: $\bar{E} < \tilde{E}_4(\gamma)$ — **Dynamics**: L^h-dominance (figure 3)
- **Sub-case (b.2)**: $\bar{E} > \tilde{E}_4(\gamma)$ — **Dynamics**: Bi-stable (figure 4)

Case (c): $\bar{E} > \tilde{E}_1(\gamma)$
- **Sub-case (c.1)**: $\bar{E} < \tilde{E}_4(\gamma)$ — **Dynamics**: L^h-dominance (figure 5)
- **Sub-case (c.2)**: $\tilde{E}_4(\gamma) < \bar{E} < \tilde{E}_3(\gamma)$ — **Dynamics**: Bi-stable (figure 6)
- **Sub-case (c.3)**: $\bar{E} > \tilde{E}_1(\gamma)$ — **Dynamics**: L^l-dominance (figure 7)

Figure 1 shows the subsets of the plane $(\gamma, \bar{E})$ corresponding to each dynamic regime. In figures 2–7, attractive fixed points are represented by full dots •, repulsive fixed points by open dots ∘, and saddle points by tracing only the trajectories converging and diverging from them (i.e. their stable and unstable manifolds).

[10] Remember that the parameter γ measures the negative impact of economic activity on the environment while the parameter $\bar{E}$ can be interpreted as the 'endowment' of the environmental good in the economy.

https://doi.org/10.1017/S1355770X04001524 Published online by Cambridge University Press

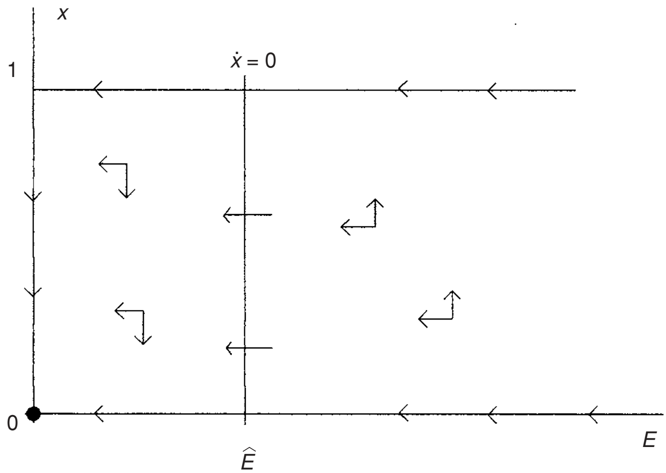

Figure 2. *Case (a)*

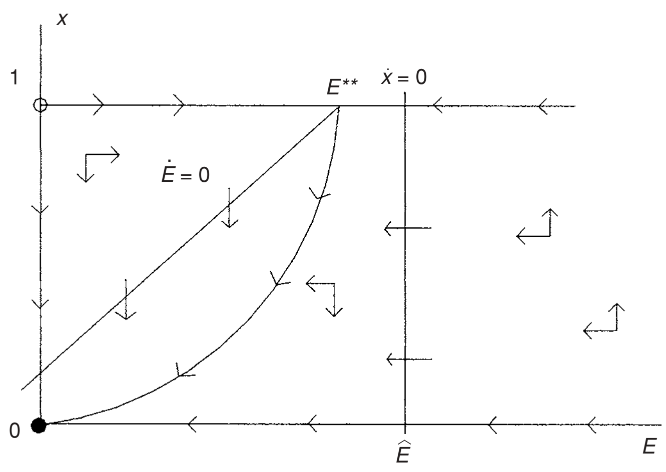

Figure 3. *Case (b), sub-case (b.1)*

In case (a), only the fixed points (0,0) and (0,1) exist (see figure 2); (0,1) is a saddle point while (0,0) is attractive, and all the trajectories, except for the one belonging to the $x = 1$ line, approach it. In this fixed point, the environmental resource is completely depleted and all individuals choose strategy (h) (i.e. the aggregate production of the economy reaches the highest possible level).

In figure 2, to the right of the straight line $E = \hat{E} \equiv d\bar{Y}_2/(\bar{l} - 1)$ (see appendix 1, expression 4), the better performing strategy is (l); vice versa to the left of it. Consequently, under the assumption that every individual adopts the best response strategy (given E) in any instant of time (best response dynamic), the economy follows the trajectory with $x = 1$ for $E > \hat{E}$

https://doi.org/10.1017/S1355770X04001524 Published online by Cambridge University Press

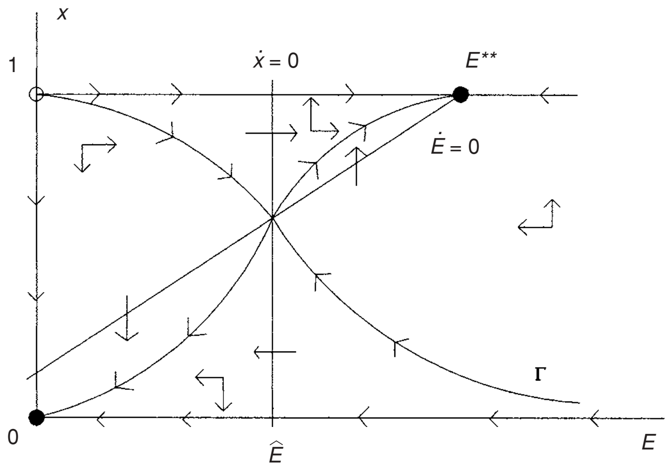

Figure 4. *Case (b), sub-case (b.2)*

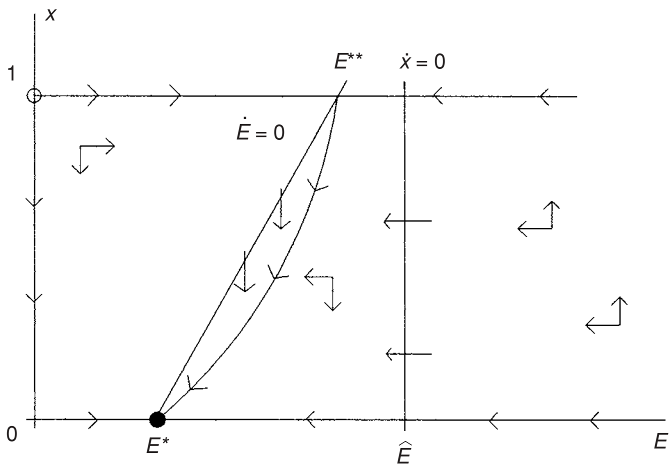

Figure 5. *Case (c), sub-case (c.1)*

and then jumps to the trajectory with $x = 0$ for $E < \hat{E}$.[11] The analysis of best response dynamics in cases (b) and (c) can be worked out in the same way. It is evident that these dynamics have the same attracting fixed points as dynamics (2), (3); only the attraction basins of fixed points may be different.

In case (b), there are two possible dynamic regimes. In sub-case (b.1), there are three fixed points (0,0), (0,1), and $(E^{**}, 1)$ $E^{**} > 0$; the first is attractive, the second is repulsive, and the third is a saddle point. As in case

[11] For $E = \hat{E}$, both strategies give the same payoff and the choice of strategies is indeterminate, that is, every x belonging to the interval [0, 1] might be observed.

https://doi.org/10.1017/S1355770X04001524 Published online by Cambridge University Press

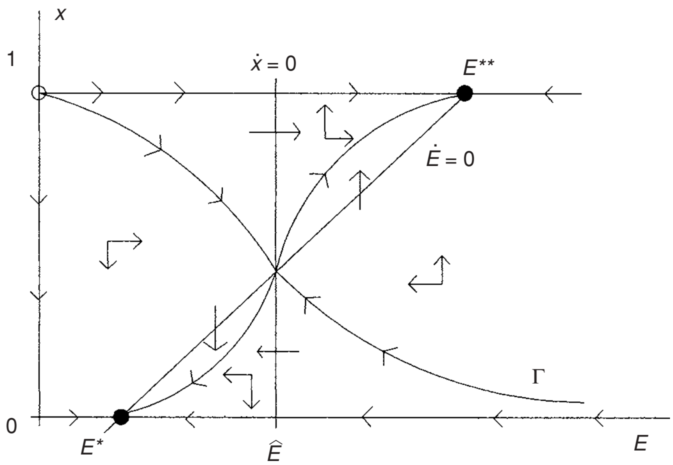

Figure 6. *Case (c), sub-case (c.2)*

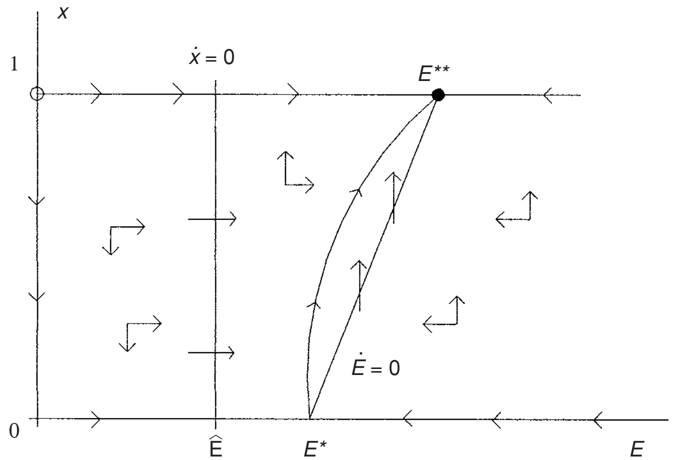

Figure 7. *Case (c), sub-case (c.3)*

(a), all the trajectories, except those belonging to the $x = 1$ line, approach (0,0) (figure 3). In sub-case (b.2), there are the fixed points (0,0), (0,1) and $(E^{**}, 1)$ plus a fixed point $(\tilde{E}, \tilde{x})$ with $\tilde{E} > 0$ and $0 < \tilde{x} < 1$. The latter is a saddle point and its stable manifold.[12] Γ separates the attraction basins of the two attractive fixed points (0,0) and $(E^{**}, 1)$. The fixed point (0,1) is repulsive. This sub-case is characterized by the presence of two attractive

[12] The stable manifold is the set made up of the two trajectories which converge at the saddle point and the saddle point itself.

fixed points: depending on the initial values of E and x, the economy may converge to the state in which the aggregate production is at the lowest possible level (all individuals choose strategy (l)) and $E = E^{**} > 0$, or that in which aggregate production is at the highest possible level (all individuals choose strategy (h)) and the stock of environmental resources is completely exhausted, $E = 0$ (figure 4).

Note that along the curve Γ the value of x increases as the value of E decreases. This means that, in order for the economy to converge at $(E^{**}, 1)$, the lower the initial value of E, the higher must be the initial proportion of individuals choosing strategy (l).

Case (c) admits three sub-cases. In sub-case (c.1), there are the four fixed points (0,0), (0,1), $(E^*, 0)$, and $(E^{**}, 1)$, $0 < E^* < E^{**}$. All the trajectories with $E > 0$ and $x < 1$ converge at the attractive fixed point $(E^*, 0)$, where the aggregate production reaches the highest possible level (figure 5). Nevertheless, unlike in case (a) and sub-case (b.1), at the attractive fixed point with $x = 0$ we have $E = E^* > 0$. This is due to the relatively low value of γ (the parameter which measures the environmental impact of productive activity and consumption) which determines the 'sustainability' of the maximum level of aggregate production. In sub-case (c.2), we have the four fixed points (0,0), (0,1), $(E^*, 0)$, and $(E^{**}, 1)$ as above, plus a fixed point $(\tilde{E}, \tilde{x})$ with $\tilde{E} > 0$ and $0 < \tilde{x} < 1$. The latter is a saddle point and its stable manifold Γ separates the attraction basins of the two attractive fixed points $(E^*, 0)$ and $(E^{**}, 1)$. The fixed point (0,1) is repulsive and the fixed point (0,0) is a saddle point (figure 6).

Like sub-case (b.2), this sub-case is also characterized by the presence of two attractive fixed points; however, in this case at the fixed point in which aggregate production is at the highest possible level the stock of environmental resources is not completely exhausted.

In sub-case (c.3), there are the four fixed points (0,0), (0,1), $(E^*, 0)$, and $(E^{**}, 1)$. All the trajectories with $E > 0$ and $x > 0$ approach the fixed point $(E^{**}, 1)$, where the aggregate production is at the lowest possible level (figure 7).

4. Predictions

It can be seen from figure 1 that, for $\bar{E} \leq d\bar{Y}_2/(\bar{l} - 1)$, L^h-dominance holds for every value of γ.[13] For $\bar{E} > d\bar{Y}_2/(\bar{l} - 1)$, if γ is sufficiently low then L^l-dominance holds; if γ increases (*ceteris paribus*), dynamics reach the bi-stable regime and finally the L^h-dominance regime.

This prediction illustrates the role of negative externalities as the engine of an increase in output. The greater the environmental impact of economic activity, the more likely it becomes that the economy will follow a path of increased output and work effort. The reason for this is that if agents substitute with produced goods ($\bar{Y}_2$) for the free good, whose availability is eroded by the negative externalities (E), they generate a further increase in the negative externalities and trigger a self-reinforcing mechanism which leads (immediately or gradually) the entire population to maximum

[13] Remember that the parameter γ measures the negative impact of the ecomomic activity on the environmental resource.

work effort and output. This is more likely to happen the greater the environmental impact of the economic activity γ, the reason being that the payoff from strategy L^l – that is, the strategy whose patterns of consumption relies on the environmental resource – lowers more with respect to the payoff from strategy L^h (which is reduced less by negative externalities), the higher is γ.

A similar effect is generated by a reduction of $\bar{E}$, the endowment of the environmental resource. In fact, figure 1 shows that, given γ, when $\bar{E}$ increases, dynamics pass from L^h-dominance to L^l-dominance via the bi-stable regime. Hence a lower endowment of $\bar{E}$ tends to increase aggregate production and labour supply. The economic intuition underlying this prediction is the same as the one on which the previous prediction rests: agents react to the scarcity of environmental resources by increasing output. Thus, any exogenous shock that reduces the endowment of free resources may trigger output growth. This and the previous prediction clash with the conventional environmentalist wisdom that the scarcity of resources is a limit to growth.

Comparative dynamics concerning the remaining parameters of the model can be worked out in a similar way. Note that the slopes of all the straight lines in figure 1 increase if the value of $N\bar{Y}_1/\beta$ increases. Consequently, the L^h-dominance region expands while the L^l-dominance region shrinks (see figure 1). With similar arguments we can show that the same holds if $d\bar{Y}_2$ increase or if $(1 - L^l)/(1 - L^h)$ decreases (i.e. L^h approaches L^l). To see the effects of parameters' variations on the bi-stable regime region, observe that $\tilde{E}_3(\gamma) - \tilde{E}_4(\gamma) = \gamma N\bar{Y}_2/\beta$; hence, the area (in correspondence to any given interval $[0, \bar{\gamma}]$) of the bi-stable dynamics region expands if $N\bar{Y}_2/\beta$ increases; note, however, that the expansion of this region is at the expense of the size of the L^l-dominance region only.

The foregoing analysis suggests that an (exogenous) increase of $\bar{Y}_1$ and $\bar{Y}_2$ (i.e. of the productivity of labour), an increase of N, a reduction of L^h (given L^l), and a reduction of β may stimulate aggregate output growth.

The prediction of this model that a population increase stimulates a growth of output is also implied by models of endogenous technological change (Grossman and Helpman, 1991; Aghion and Howitt, 1992; Kremer, 1993) in which innovations have a fixed cost made less onerous by the increased size of the markets connected with that of the population. However, our model emphasizes an entirely different reason for making this prediction: the increased pressure on resources which stimulates individuals to adopt consumption patterns based on private goods.

Note finally that the prediction that increased labour productivity raises output – a prediction that any model with an exogenous labour supply, like the majority of growth models, will obviously make – acquires entirely different significance in a model where the labour supply is endogenous. In this case, in fact, it is no longer obvious that increased labour productivity raises per capita output: agents may use the increase in productivity to augment their leisure rather than their output. In this model, instead, increased productivity tends to raise output because it induces individuals to work more, given that this improves their capacity to defend themselves against negative externalities.

5. Well-being

Let us see how the values of the payoffs of each strategy and of the average payoff $\bar{U}(E,x) = U_l(E)x + U_h(E)(1-x)$ evolve along the trajectories that the economy may follow. Before each of the cases (a)–(c) is analysed, some general considerations are in order. In figures 2–7, the line $\dot{x} = 0$ separates the set $\{E \geq 0, 0 \leq x \leq 1\}$ into two parts. To the right of $\dot{x} = 0$, i.e. for $E > \hat{E}$, one has $U^l(E) > U^h(E)$ and, consequently, $U^l(E) > \bar{U}(E,x)$. Therefore, given $E > \hat{E}$, the average payoff increases if all individuals adopt strategy (l). The opposite happens to the left of $\dot{x} = 0$, i.e. if $E < \hat{E}$; in this case $U^l(E) < U^h(E)$ and $\bar{U}(E,x)$ increases if all individuals adopt (h). In the analysis that follows, we shall see that in many cases the economy can reach a fixed point with a higher average payoff if individuals coordinate themselves on strategy (l). However, the normative implications of this result differ according to the initial position (E^0, x^0) from which the economy starts. If $E^0 \geq \hat{E}$ (see figures 3 and 5), along the trajectory that passes through $(E^0, 1)$ one has (at each instant of time) a payoff greater (for each individual) than that associated with the trajectory that passes through (E^0, x^0), with $x^0 < 1$. Consequently it is in the interest of individuals to coordinate themselves on strategy (l) independently of their inter-temporal preferences (i.e. independently of the discount factor used).

Different is the case in which, starting from $(E^0, 1)$, the economy is able to reach a better fixed point, but $E^0 < \hat{E}$. In this case, in fact, the value of $\bar{U}(E,x)$ along the trajectory that starts from $(E^0, 1)$ is initially less than along the trajectory that starts from (E^0, x^0), with $x^0 < 1$. Consequently, the choice of coordinating individuals on strategy (l) raises a problem analogous to that of the 'Golden rule' in Solow's model; that is, it will be 'optimal' to coordinate individuals on strategy (l) only if they are sufficiently 'patient' to wait for regeneration of the environmental good.

The results that follow should be interpreted bearing these considerations in mind. We use the usual symbology: $(\hat{E}, \hat{x}) \prec (\breve{E}, \breve{x})$ indicates that the fixed point $(\breve{E}, \breve{x})$ strictly Pareto-dominates $(\hat{E}, \hat{x})$. Well-being analysis is set out in appendix 2; the results of such analysis can be summarized as follows:

Case (a)

L^h-dominance (figure 2)	(0,1) $\prec$ (0,0).

Case (b)

(b.1): L^h-dominance (figure 3)	(0,1) $\prec$ (0,0) $\prec$ $(E^{**}, 1)$, if (10) holds (see appendix 2), and (0,1) $\prec$ $(E^{**}, 1)$ $\prec$ (0,0) if the opposite of (10) holds.
(b.2): Bi-stable dynamics (figure 4)	(0,1) $\prec$ (0,0) $\prec$ $(\tilde{E}, \tilde{x})$ $\prec$ $(E^{**}, 1)$.

Case (c)

(c.1): L^h-dominance (figure 5)	(0,1) $\prec$ (0,0) $\prec$ $(E^*, 0)$ $\prec$ $(E^{**}, 1)$, if (11) holds (see appendix 2), and (0,1) $\prec$ (0,0) $\prec$ $(E^{**}, 1)$ $\prec$ $(E^*, 0)$ if the opposite of (11) holds.

(c.2): Bi-stable dynamics (figure 6) $(0,1) \prec (0,0) \prec (E^*,0) \prec (\tilde{E},\tilde{x}) \prec (E^{**},1)$.

(c.3): L^l-dominance (figure 7) $(0,1) \prec (0,0) \prec (E^*,0) \prec (E^{**},1)$.

In case (a), the fixed point (0,0) Pareto-dominates the fixed point (0,1). In this sense, the process of aggregate production growth which brings the economy to the fixed point (0,0) is desirable.

In sub-case (b.1), even if $(E^{**},1)$ is unstable, it may Pareto-dominate the attracting fixed point (0,0). More specifically, this occurs if and only if condition (10) in appendix 2 holds; in this case, the individuals would be better off in $(E^{**},1)$, where the level of activity of the economy is lower. In the sub-case (b.2), dynamics are of the bi-stable type: both (0,0) and $(E^{**},1)$ are attractive. Individuals' well-being is always higher at the fixed point $(E^{**},1)$, where the level of aggregate production is lower. There would moreover be greater well-being in $(\tilde{E},\tilde{x})$ than in (0,0); however, the first is a saddle point and therefore only two trajectories converge to it.

Well-being properties of fixed points in case (c) are similar to those of case (b). In case (c.1), the fixed point $(E^{**},1)$ may Pareto-dominate the fixed point $(E^*,0)$ even if the former is not attracting (this is the case if condition (11) in appendix 2 is satisfied) and both fixed points dominate (0,1) and (0,0). In sub-case (c.2) there is a bi-stable regime where the attracting fixed point $(E^{**},1)$ always dominates the other attracting point, $(E^{**},0)$. In sub-case (c.3), there is only one attracting point, $(E^{**},1)$, which dominates all the others.

Let us summarize what has been said regarding cases (a), (b) and (c). We can state that the fixed point $(E^{**},1)$, when it exists, always Pareto-dominates the others if it is attracting. It may, moreover, Pareto-dominate the others even when it is not attractive. Therefore, in the L^l-dominance and bi-stable dynamic regimes the fixed point $(E^{**},1)$ always dominates all the others; moreover, even in the L^h-dominance dynamic regime, this fixed point may dominate all the others.

To conclude, the substitution of free goods with produced goods may easily generate output growth dynamics in which the positive impact on well-being of increased per capita output is more than off-set by the poorer quality of the environment. Hence this model suggests that environmental degradation may be the consequence and the cause of undesirable growth in output.

Finally, observe that best reply dynamics do not always ensure better outcomes than adaptive dynamics. See e.g. figure 3; if the initial value E^0 of E is such that $E^0 < \hat{E}$, best reply dynamics always approach the Pareto-dominated fixed point $(E,x) = (0,0)$, while adaptive dynamics will converge to the Pareto-dominant fixed point $(E,0) = (E^{**},1)$ (fixed E^0 as above) if the initial value of x, x^0, is high enough (that is, if the point (E^0,x^0) lies above the curve Γ).

6. Concluding remarks: industrial revolutions

A natural terrain for empirical verification of the relation between negative externalities and the enlargement of the goods and labour markets concerns what is probably the most spectacular process of increased labour supply: the 'human resouces mobilization' associated with all processes of industrial take-off. The mobilization of the labour market, with the explosion of working time, of rates of activity and dependent employment, is a stylized fact typical of every industrial revolution, from the British one to the most recent of them, for example the second-generation Asian Tigers.[14]

The mobilization of human resources, the connected phenomena of urbanization, and the decline of the rural share of total employment have long been at the centre of the theory of development which makes large use of dual models (modern/traditional sectors) based on 'pull' factors to explain them.[15]

By contrast, in our model individuals are 'pushed' into the modern sector by the negative externalities that it generates.[16] In fact this model can be viewed as a dual development model: strategy (h) represents participation in the modern (industrial) sector, while strategy (l) represents participation in the traditional (agricultural) sector. The former strategy implies a relatively high work effort, and its payoff is relatively less affected by environmental degradation. The payoff of strategy (l) relies crucially on the stock of the environmental resource. This assumption reflects the well-known fact that traditional agriculture is largely based on natural commons (common fields, fresh water, forests, fisheries, etc.). The environmental cleavages generated by the expansion of the modern sector undermine this basis, compelling individuals to satisfy their needs by increasing their participation in the market sector of the economy. This, in turn, leads to a further enlargement of the modern sector, generating a self-reinforcing mechanism. Hence the growth of the modern sector is fuelled by the negative externalities that it produces.

[14] The Industrial Revolution was associated with a dramatic increase in working hours. The eight-hour working day of medieval 'industry' was extended to the ten, 12, 14 hours of the Industrial Revolution, and it took around a century to return to the eight-hour day (Schor, 1991). As regards the labour supply in the second-generation Asian Tigers – where rates of activity and dependent employment have often doubled in the space of ten years – see for example Krugman (1995), who coined the term 'human resources mobilization'.

[15] For instance, in the 'classical' models of development theory, individuals are 'pulled' into the modern sector by wage differentials (Lewis, 1954; Todaro, 1969).

[16] The idea that push factors are important in generating urbanization is anything but new. For example, playing a major role in historiography on the Industrial Revolution is the idea that the decline of traditional agriculture was of decisive importance in pushing the rural population into urban migration (see for instance Polanyi, 1968). For a review of the debate on push/pull factors in contemporary development economics see Williamson (1995). Besides environmental degradation, also indicated as push factors are natural disasters and wars. These may be simply conceived in our model as shocks which reduce the stock of the environmental resource. The erosion of traditional institutions and behaviours due to modernization has also received attention as a push factor in a number of historical, sociological, anthropological, studies. Two attempts to formalize push mechanisms completely different from ours have been made by Shaw (1974) and Stark and Taylor (1989).

At first sight, the predictions of the model – that shocks which raise the impact of production on the environment, the population, and labour productivity, or which reduce the endowment of the natural resource, may trigger an increase in the labour supply – seem compatible with the evidence on certain stylized facts common to all industrial revolutions.

For instance the Industrial Revolution exhibits a shock on all the four factors engendering growth. Besides the population increase that preceded the Industrial Revolution, the effect of the technological shock due to the adoption of industrial technologies is also evident, since it determined an increase in both labour productivity and the environmental impact of production.

Since the Industrial Revolution, environmental devastation has been an empirical regularity in all industrialization processes. A large body of literature suggests that environmental cleavages may be among the factors responsible for the urbanization of large masses of peasants whose traditional agriculture relies largely on environmental resources. These masses of urbanized peasants, the object of the mobilization, furnish the low-cost industrial labour supply indispensable for take-off.

In China, the pollution and depletion of the water supply in the Yellow River basin due to the impetuous development of industry are only the spectacular replication on a Chinese scale of this mechanism. The degradation of the basin affects the lives of around 200 million poor peasants whose agriculture is based on the use of water from it, forcing them to swell the ranks of the urban-industrial labour force.[17]

As regards the endowment of resources, the shocks that affect it may be of an institutional nature. The 'enclosures' provide an example of this type of shock (Antoci and Bartolini, 1999; Bartolini and Bonatti, 2002): as widely acknowledged by historians of the Industrial Revolution, the enclosures caused a collapse in the endowment of commons which played an important role in forming the industrial labour supply. Indeed, the restriction imposed by the enclosures on access to common resources was a grievous blow to the poorest segments of the rural population, 'pushing' them to urbanization.

References

Antoci, A. and S. Bartolini (1997), 'Externalities and growth in an evolutionary game', Department of Economics Discussion Paper No. 10, University of Trento, Trento.

Antoci, A. and S. Bartolini (1999), 'Negative externalities as the engine of growth in an evolutionary context', Nota di lavoro 83.99, FEEM, Milano.

Bartolini, S. and L. Bonatti (2002), 'Environmental and social degradation as the engine of economic growth', *Ecological Economics* **41**: 1–16.

Bartolini, S. and L. Bonatti (2003), 'Undesirable growth in a model with capital accumulation and environmental assets', *Environment and Development Economics* **8**: 11–30.

[17] According to World Bank (1997), the damage caused by pollution amounts to almost 8 per cent of Chinese GDP. The main focus of the study is the aggregate cost af air pollution. For a study on the cost of water pollution in the urban–rural area of Chongqing see Yongguan, Seip, and Vennemo (2001).

Borgers, T. and R. Sarin (1997), 'Learning through reinforcement and replicator dynamic', *Journal of Economic Theory* **77**: 1–14.

Cross, G. (1993), *Time and Money: The Making of Consumer Culture*, London: Routledge.

Cullino, R. (1993), 'Una stima delle spese difensive e ambientali', in I. Musu and D. Siniscalco (eds), *Ambiente e Contabilità Nazionale*, Bologna: Il Mulino.

Grossman, G. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge, MA: MIT Press.

Hirsh, F. (1976), *Social Limits to Growth*, Cambridge, MA: Harvard University Press.

Hueting, R. (1980), *New Scarcity and Economic Growth. More Welfare Through Less Production*?, Amsterdam, New York, Oxford: North-Holland.

Krugman, F. (1995), *Pop Internationalism*, Cambridge, MA: MIT Press.

Leipert, C. and U.E. Simonis (1989), 'Environmental protections expenditures: the German example', *Rivista Internazionale di Scienze Economiche e Commerciali* **36**: 255–270.

Lewis, W.A. (1954), 'Economic development, with unlimited supply of labour', *Manchester School of Economic and Social Studies* **22**: 139–191.

Polanyi, K. (1968), *The Great Transformation*, Boston: Beacon.

Sacco, P.L. (1994), 'Selection mechanisms in economics', *Behavioral Science* **39**: 311–325.

Schor, J. (1991), *The Overworked American: The Unexpected Decline of Leisure in America*, New York: Basic Books.

Schlag, K. (1998), 'Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits', *Journal of Economic Theory* **78**: 130–156.

Sethi, R. and E. Somanathan (1996), 'The evolution of social norms in common property resource use', *American Economic Review* **86**: 766–788.

Shaw, P. (1974), 'Land tenure and rural exodus in Latin America', *Economic Development and Cultural Change* **23**: 123–132.

Shogren, J.F. and T.D. Crocker (1991), 'Cooperative and non-cooperative protection against transferable and filterable externalities', *Environmental and Resource Economics* **1**: 195–213.

Stark, O. and E. Taylor (1989), 'Relative deprivation and international migration', *Demography* **26**: 1–14.

Todaro, M.P. (1969), 'A model of labour, migration and urban unemployment in less developed countries', *American Economic Review* **59**: 138–148.

Weibull, J.W. (1995), *Evolutionary Game Theory*, Cambridge, MA: MIT Press.

Williamson, J.G. (1995), 'Migration and city growth during industrial revolutions', in H. Giersch (ed.), *Urban Agglomeration and Economic Growth*, Berlin: Springer Verlag.

World Bank (1997), *China* 2020, *Clear Water, Blue Sky*, Washington, DC: World Bank.

Yongguan, C., H.M. Seip, and H. Vennemo (2001), 'The environmental cost of water pollution in Chongqing, China', *Environment and Development Economics* **6**: 313–333.

Appendix 1. Basic mathematical results

The locus in which $\dot{x} = 0$

In this appendix we analyse the system of differential equations (2) and (3) which are defined in the set

$$\{(E, x) : E \geq 0, 0 \leq x \leq 1\}$$

From equation (3), it holds that $\dot{x}=0$ for $x=0, x=1$ and for $\Delta U(E)=0$. It is easy to check that $\Delta U(E)=0$ if and only if

$$E = \hat{E} \equiv \frac{d\bar{Y}_2}{\left(\frac{1-L^l}{1-L^h}\right)^{\frac{b}{a}} - 1} \tag{4}$$

where $(\frac{1-L^l}{1-L^h})^{\frac{b}{a}} > 1$ always. To the left (to the right) of the straight line (4), in the plane (E, x), it holds that $\dot{x} < 0$ (respectively, $\dot{x} > 0$).

The locus in which $\dot{E}=0$

We observe that $\dot{E}=0$ if $E=0$ and along the straight line

$$x = \frac{1}{\bar{Y}_2}\left(\frac{\beta}{\gamma N}E + \bar{Y}_1 + \bar{Y}_2 - \frac{\beta\bar{E}}{\gamma N}\right) \tag{5}$$

To the left of (5) it holds that $\dot{E} > 0$; to the right it holds that $\dot{E} < 0$. Moreover, for $x=0$, equation (5) gives

$$E = E^* \equiv \bar{E} - \frac{\gamma N}{\beta}(\bar{Y}_1 + \bar{Y}_2) \tag{6}$$

where $E^* > 0$ if and only if

$$\bar{E} > E_1(\gamma) \equiv \frac{N(\bar{Y}_1 + \bar{Y}_2)}{\beta}\gamma \tag{7}$$

For $x=1$, equation (5) gives

$$E = E^{**} \equiv \bar{E} - \frac{\gamma N}{\beta}\bar{Y}_1 \tag{8}$$

where $E^{**} > E^*$ always and $E^{**} > 0$ if and only if

$$\bar{E} > E_2(\gamma) \equiv \frac{N\bar{Y}_1}{\beta}\gamma \tag{9}$$

The fixed points of dynamics

Let us now consider the fixed points of the dynamics (2), (3); that is, the points (E, x) of the set $\{(E, x) : E \geq 0, 0 \leq x \leq 1\}$ in which $\dot{E}=\dot{x}=0$. We observe that the points

$$(E, x) = (0{,}0) \quad \text{and} \quad (E, x) = (0{,}1)$$

are always fixed points.

Moreover, the points of intersection of the straight line (5) with the lines $x=0$ and $x=1$ are also fixed points if $E>0$ holds at these points of intersection. Therefore, if (9) holds, then there exists the fixed point

$$(E, x) = (E^{**}, 1)$$

and if (7) holds, then there exists the fixed point

$$(E, x) = (E^*, 0)$$

Since $E^{**} > E^*$ (that is, (9) implies (7)), if the latter fixed point exists, then so does the former.

Finally, there exists a fixed point with $E > 0$ and $0 < x < 1$ when the straight lines (4) and (5) meet at a point with $E > 0$ and $0 < x < 1$. Therefore, for (4), (6), and (8), this fixed point exists if and only if

$$E^{**} > \hat{E} > E^*$$

Note that $\hat{E} > E^*$ if and only if

$$\bar{E} < \tilde{E}_3(\gamma) \equiv \frac{d\bar{Y}_2}{\left(\frac{1-L^l}{1-L^h}\right)^{\frac{b}{a}} - 1} + \frac{N(\bar{Y}_1 + \bar{Y}_2)}{\beta}\gamma$$

and that $\hat{E} < E^{**}$ if and only if

$$\bar{E} > \tilde{E}_4(\gamma) \equiv \frac{d\bar{Y}_2}{\left(\frac{1-L^l}{1-L^h}\right)^{\frac{b}{a}} - 1} + \frac{N\bar{Y}_1}{\beta}\gamma$$

The analysis of the stability of the fixed points of dynamics (2), (3) is straightforward and is therefore omitted.

We can relax the hypothesis of perfect substitutability between E and $\bar{Y}_2$; the same dynamic regimes can be obtained by simply assuming that the value of the payoff $U_l(E, 1 - L^l, \bar{Y}_1)$ of strategy (l) decreases less than that of the payoff $U_h(E, 1 - L^l, \bar{Y}_1, \bar{Y}_2)$ of strategy (h) if E decreases; that is, the value of the partial derivative w.r.t. E of U_l is greater (*ceteris paribus*) than that of the corresponding partial derivative of U_h. Under this hypothesis, the locus $\dot{x} = 0$ is still a vertical straight line with $\dot{x} < 0$ (respectively, $\dot{x} > 0$) to the left (right) of it. Consequently, the dynamic regimes of figures 2–7 are obtained also under this less restrictive assumption. Furthermore, it is easy to check that the results on the undesirability of the increase of output (appendix 2) hold *a fortiori* under imperfect substitutability.

Appendix 2. Well-being analysis

Well-being analysis for case (a)

In case (a) we have only two fixed points: the attracting fixed point $(E, x) = (0,0)$ and the saddle point $(E, x) = (0,1)$. In the former, the average well-being of the population $\bar{U}(E, x) = U_l(E)x + U_h(E)(1 - x)$ coincides with $U^h(0) = (d\bar{Y}_2)^a(1 - L^h)^b \bar{Y}_1^{1-a-b}$ and is strictly greater than the average well-being calculated at the second fixed point, that is, $\bar{U}(0,1) = U^l(0,1) = 0$. Therefore the fixed point (0,0) Pareto-dominates the fixed point (0,1).

Well-being analysis for case (b)

Let us first consider the sub-case (b.1). In (b.1) we have the fixed points: $(E, x) = (0,0), (E, x) = (0,1), (E, x) = (E^{**}, 1)$; only the first is attractive. In this case too, the average well-being in (0,0) is greater than in (0,1), while the average well-being in $(E^{**}, 1), \bar{U}(E^{**}, 1) = U^l(E^{**})$, is strictly greater than that in (0,0) if and only if[18]

$$\bar{E} - \frac{\gamma N}{\beta}\bar{Y}_1 > \frac{d\bar{Y}_2}{\left(\frac{1-L^l}{1-L^h}\right)^{\frac{b}{a}}} \tag{10}$$

We observe that the left side of (10) coincides with E^{**} (see (8)) and that the right side is strictly less than $\hat{E}$ (see (4)). This implies that even if $(E^{**}, 1)$ is unstable, it can Pareto-dominate the attracting fixed point (0,0). More specifically, this occurs if and only if (10) holds.

In the sub-case (b.2), where $\hat{E} < E^{**}$ holds, condition (10) is always satisfied; therefore $(E^{**}, 1)$ Pareto-dominates (0,0). In this sub-case, there is also the fixed point $(\tilde{E}, \tilde{x})$ with $\tilde{E} > 0$ and $0 < \tilde{x} < 1$. At this point, $\bar{U}(\tilde{E}, \tilde{x}) = U^l(\tilde{E}) = U^h(\tilde{E})$ holds; it is consequently easy to verify that $(\tilde{E}, \tilde{x})$ Pareto-dominates (0,0) and that it is Pareto-dominated by $(E^{**}, 1)$.

Well-being analysis for case (c)

Let us compare the average payoff at points $(E^{**}, 1)$ and $(E^*, 0)$; it is easy to verify that

$$\bar{U}(E^{**}, 1) = U^l(E^{**}, 1) > \bar{U}(E^*, 0) = U^h(E^*, 0)$$

if and only if

$$E^{**} > \hat{E} - \frac{\frac{\gamma N}{\beta}\bar{Y}_2}{\left(\frac{1-L^l}{1-L^h}\right)^{\frac{b}{a}} - 1} \tag{11}$$

Condition (11) shows that the fixed point $(E^{**}, 1)$ Pareto-dominates the fixed point $(E^*, 0)$ in subcases (c.2) and (c.3) (see figures 5–6) and that it may also dominate it in case (c.1) (see figure 4). The other fixed points can be Pareto-ordered as in case (b).

[18] To obtain condition (10), it is sufficient to observe that $U^l(E^{**}) = (\bar{E} - \gamma N\bar{Y}_1/\beta)^a (1-L^l)^b \bar{Y}_1^{1-a-b}$ and rewrite the inequality $U^l(E^{**}) > U^h(0)$ in the form (10).