Economics of technological change and the natural environment: how effective are innovations as a remedy for resource scarcity?

Lucas Bretschger

Working Paper 03/27
June 2004
Economics of technological change and the natural environment: how effective are innovations as a remedy for resource scarcity?

Lucas Bretschger*
ETH Zurich

Abstract

The paper aims to substantiate the importance of endogenous innovations when evaluating the compatibility of natural resource use and economic development. It explains that technological change has the potential to compensate for natural resource scarcity, diminishing returns to capital, poor input substitution, and material balance restrictions, but is limited by various restrictions like fading returns to innovative investments and rising research costs. It also shows how innovative activities are fostered by accurate price signals and research-favouring sectoral change. The simultaneous effects of increasing technical knowledge, decreasing resource inputs, and increasing world population largely determine the chances of long-run sustainable development. Consequently, future research has to be directed at a more thorough understanding of the mechanisms driving innovations in the presence of natural resource scarcity.

Keywords: endogenous technological change, environment, natural resources, sustainability

JEL-Classification: Q20, Q30, O41, O33

* WIF – Institute of Economic Research, ETH Zentrum, WET D5, 8092 Zurich, Switzerland, Tel. +41 1 632 21 92, Fax +41 1 632 13 62, Email: bretschger@wif.gess.ethz.ch

The author wishes to thank Hannes Egli, Christian Groth, Karen Pittel, Thomas Steger and three anonymous referees for valuable comments.
1. Introduction
In many world regions, the adoption of superior technologies has allowed to remarkably improve the state of the natural environment. For instance, in developed countries, local air and water qualities are much better than they used to be some decades ago. Technology has also helped to significantly increase efficiency in energy use in the past. Regarding current pollution and resource scarcity problems, there is little doubt that technical change has the potential to substantially contribute to new solutions. As a prominent example, it has been suggested that lower carbon use is compatible with high and rising energy use when we shift to alternative technologies, see IPPC (2001). On the other hand, shifts in technology and the subsequent sectoral composition of the economy have to be fostered by appropriate price signals, which cannot be expected under general market conditions.

In macroeconomic terms, technology plays a major part in the steady accumulation of capital which is necessary to increase living standards in the long run. The prediction of a further increase in world population and the fast economic development in world regions like China and India accentuates the problems of global resource scarcity and the need for technical progress. The de-linking of economic development from the use of the natural environment, as formalised in the Environmental Kuznets Curve, largely hinges on the availability of clean and resource-saving technologies.

However, while it is widely accepted that technology determines the relationship between economic growth and the environment, see Jaffe, Newell and Stavins (2003, p. 463), it is not unanimously believed that technical progress is powerful enough to solve current and future environmental problems. In ecological economics, technical progress plays a much less prominent role than in neo-classical contributions, see the assessment of Pearce (2002, p. 76). If, as many ecologists argue, clean and resource-saving technologies are not (or inadequately) available in the future, pollution will not be de-linked from income and decreases of natural resource use may entail limits to economic activities in the long run. It is therefore a major task for current economic research to identify the basic mechanisms driving technology improvements and to adequately depict the effects of innovation on future economic development.

With a few exceptions, until the 1990s the rate and the direction of technical change were treated as exogenous variables in macroeconomics and, subsequently, in environmental and resource economics. Most prominently, in neo-classical growth theory, long-run growth is determined by exogenous technical progress. When addressing resource scarcity problems with this framework, the focus of possible conclusions becomes limited. Regarding technology, it can be asked how much technical progress is needed to prevent income growth from becoming negative, see Nordhaus (1992). It can also be calculated under what conditions physical capital accumulation is strong enough to compensate for fading resource inputs. The results of this strand of research are assembled in the well-known RES symposium issue, see Solow (1974a), Stiglitz (1974) and Dasgupta and Heal (1974). According to this literature, living standards can be sustained even without technical progress, assuming high enough elasticities of substitution in production and sufficiently high saving rates. However, these assumptions are not in line with ecological economics. It has been criticised that the
assumed elasticities are unrealistically high, material balance constraints entirely neglected, and savings rates related to hypothetical planner solutions, see Cleveland and Ruth (1997).

This critique has to be taken seriously, which means that we have to introduce technology as an endogenous variable into theory and thoroughly scrutinise the incentives for innovation under different market and policy conditions. Realistic elasticities of substitution, material balance rules and savings rates should be among the determinants of modelling. The consequences of innovations and of current environmental problems for economic growth and natural resource use should be predictable from theory. This enables us at the same time to evaluate the usefulness of policies aimed at decreasing resource use and/or increasing growth rates, as well as the chances of raising income with declining natural resource use in the future.

A natural and theoretically important distinction should be made between determinants affecting the supply of innovations and determinants influencing the demand for innovations. Moreover, the use of a specific innovation sector leads to formulating multi-sector models instead of the one-sector models used in the 1970s. It has been noted that for theories of “learning by doing, but perhaps even more important in the models of direct knowledge investment or research and development, a key issue may be the sectoral disaggregation of the model,” see Grubb, Köhler and Anderson (2002, p. 293). The representation of a more detailed sectoral structure aims at including all relevant substitution possibilities in an economy. It has been argued that structural change can act as a powerful additional substitution mechanism in the case of natural resources, see Bretschger (1999, p. 224). On the other hand, the result of Nordhaus (2002) of a low response of induced technological change can be criticised, because he only considers one aggregate energy-carbon sector with no opportunity for substitution toward non-carbon sources.

The present paper starts with a general approach to capital accumulation and resource use and explains the basic substitution problem when facing natural resource scarcity. It then introduces a simple framework of endogenous technology to focus on the different aspects of current theoretic modelling. The text builds on basic resource economics, see Dasupta and Heal (1979), and more recent contributions, see Bovenberg and Smulders (1995), Scholz and Ziemer (1999) and Groth and Schou (2002), and incorporates the model elements of new growth theory, see Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1998). In particular, the differences between one-sector models and multi-sector models, as widely used in endogenous innovation theory, are pointed out. The paper argues that the sectoral structure of the economy plays a crucial role. Specifically, it is demonstrated that poor input substitution and the use of an essential non-renewable resource in the innovation sector are not necessarily detrimental for economic development. On the other hand, long-run predictions of the model rely on the asymptotic properties of the used functions; these have to be carefully examined. Based on these fundamental findings, the text elaborates on seminal future lines of research in this field, including directed research, education of researchers, research policies, and resource reallocation costs.

By analysing the state of the art in modelling endogenous innovation, the paper seeks to contribute to our assessment of how effective innovations are as a remedy for natural resource
scarcity; that is, more generally, whether optimism or pessimism is adequate when predicting the effects of technological change on the environment and economic development. It is obvious that innovations \textit{per se} are not a panacea for dealing with resource scarcities. But if appropriate, the vision of abundant and customised knowledge in the future has to be upheld by corresponding results from economic theory. Current innovation models can be mathematically quite complex, and yet their economic content may be limited. The scope of these models and the consequences of certain modelling assumptions only become clear when the whole process is reasonably understood. By looking at these specific issues, the present text does not aim at covering all the subjects related to environment and technology, but rather supplements recent surveys, which include Azar and Dowlatabadi (1999), Grübler, Nakicenovic and Victor (1999), Weyant and Olavson (1999), Smulders (2000), Grübler, Nakicenovic and Nordhaus (2002), Grubb, Köhler and Anderson (2002), and Jaffe, Newell and Stavins (2003).

The remainder of the paper is organised as follows. Section 2 analyses economic dynamics in the presence of natural resources. In section 3, the basic innovation model with natural resources is presented. Sections 4 and 5 provide results for different variants of this model. In sections 6 and 7, more innovation types and additional elements affecting innovations are discussed. Section 8 concludes.

2. Scarcity and economic dynamics

2.1 Decomposing income growth

In general, improving technology is aimed at increasing productivity and consumer utility. We distinguish between (Hicks)neutral and biased technical change, which are both important when dealing with resource scarcity. Neutral technical change can sustain the growth in knowledge needed to improve efficiency in the use of all factors, including natural resources. In the context of natural resources, the role of technology can also be viewed as more specific: it should respond to scarcity problems a society faces with respect to the supply of natural resources, and lead to additional degrees of freedom when dealing with the use of the environment for different aims. In either case, all possibilities of substitution and, specifically, the effects technology exerts on promoting substitution, have to be studied. With this approach, it should then be possible to make predictions about future development.

To be more specific in the following arguments and to follow recent literature on environment and technology, it is useful to proceed in three steps. First, we introduce a general framework to show the importance of technology, without explaining how it is determined. We then adopt a well-established endogenous innovation model to implement the new modelling approaches in this field. Finally, we extend the framework in different directions to check its robustness and identify fields for further research.

In a one-sector framework, production possibilities of an economy can be described by an aggregate production function with the inputs capital, natural resource use and labour. If technology is neutral, enhancing the productivity of inputs in a symmetric way, we have:
where $Y$ is aggregate output, $A$ technical knowledge, $K$ physical capital, $R$ natural resources and $L$ labour. Natural resource input can also be interpreted as pollution when production pollutes the environment. Expressing (1) in growth rates, denoted by hats, yields:

$$\hat{Y} = \hat{A} + \theta_k \hat{K} + \theta_r \hat{R} + \theta_l \hat{L}$$

with the $\theta$s being the partial elasticities of production. Forces raising $\hat{Y}$ are the increases of physical capital and knowledge capital, i.e. $\hat{A}, \hat{K} > 0$. Let us assume that the effective resource input decreases over time so that we have $\hat{R} < 0$ in (2). This is predicted to happen when the resource is non-renewable, like oil, and its price obeys the well-known Hotelling rule. When resources are renewable, it might be the case that resource use exceeds natural regeneration, as happens in fisheries or rain-forests, or pollutes the environment. Then, for a social optimum, a decrease in the use of the resource is indicated, which has to be achieved by policy measures.

Population growth $\hat{L}$ has a positive impact on total production and increases resource use ceteris paribus. In relation to capital (and knowledge), the role of labour is more complex. When capital is a private good and is produced with the same production technique as final output (like in the neo-classical growth model), an increasing population leads ceteris paribus to a lower capital intensity and a lower per capita income. If, on the other hand, capital is a public good, like public knowledge, and the production of capital is intensive in the use of labour, we have the opposite effect of labour supply on per capita income.

As one might suspect, theoretical predictions on how powerful $\hat{A}$ and $\hat{K}$ effectively are and how they are affected by certain conditions hinge on a number of important assumptions. Specifically, $\hat{A}$ and $\hat{K}$ depend on the assumed returns to physical and knowledge capital, the scarcity of materials used for $K$-accumulation, the scarcity of resources used for $A$-accumulation, and population growth $\hat{L}$. At the same time, the effect of $\hat{R}$ on $\hat{A}$ and $\hat{K}$ depends on whether we adopt a one-sector or a multi-sector economy. Moreover, the impact of technology $\hat{A}$ on resource use $R$ depends on relative input prices, and - in a more general case - on the direction of endogenous innovations.

From this first approach it can be concluded, that - in order to ensure constant income or income growth in the long run - $\hat{K}$ and $\hat{A}$ must be sufficiently large to offset the drag of a negative $\hat{R}$. In per-capita terms, input substitution and productivity increases must also be big enough to take into account a rising labour force. Finally, in the long run, per-capita income is also determined by the asymptotic properties of production functions like $F$ from (1) when $R \to 0$ (combined with $\hat{A}$ and $\hat{K}$ taking much higher values than today), which applies for non-renewable resources. Thus, in all adopted models, input combinations attained in the future may lie outside the range of the past. Put differently, predictions are – at least partially
– based on assumptions on the production function which have not been tested empirically up to now. This will be the focus of section 5.

2.2 Input substitution and beyond

In the one-sector model of a resource-using economy, the results summarised by Solow (1974b, p.11) apply: “if the elasticity of substitution between exhaustible resources and other inputs is unity or bigger, and if the elasticity of output with respect to reproducible capital exceeds the elasticity of output with respect to natural resources, then a constant population can maintain a positive constant level of consumption per head forever.” So in terms of equation (2), it has been argued that, even with \( \hat{A} = 0 \), elasticities between inputs that are larger than unity prevent \( \hat{Y} \) from becoming negative with a negative \( \hat{R} \).

However, recent contributions point to necessary extensions of this result. First, elasticities might well lie below unity in the real world. Furthermore, it has been argued that \( \hat{K} \) cannot be indefinitely positive because physical capital is built out of material which is bounded due to material balance restrictions. In addition, the impact of labour \( L \) on \( \hat{A} \) and \( \hat{K} \) has to be clarified. Moreover, the \( \theta \)s depend on the type of the production function and are not necessarily constant in the general case. Furthermore, \( \hat{K} \) depends on the savings behaviour of households. For instance, the well-known Hartwick rule (Hartwick 1977) requires a constant savings rate, which is – with unitary elasticities – not obtained under normal utility discounting; only maximin-preferences lead to such an outcome, see Asheim and Withagen (1998). In addition, long-run individual decisions are biased because of various externalities. While part of the benefits of increasing \( A \) are appropriated by people who do not pay for them, negative externalities like stock pollution or over-exploitation of resources have a negative effect on the welfare of future generations, without any counterbalancing compensation.

An important issue is that economic development \( \hat{Y} \) is also affected by structural change. For instance, when labour moves from resource intensive and knowledge extensive sectors to resource extensive and knowledge intensive sectors, \( \hat{Y} \) is positively affected through an increasing \( A \) while resource use decreases. Interestingly, reallocating labour to dynamic sectors can act as a counterforce to decreasing returns in knowledge accumulation. Thus, the main result from one-sector analysis, as stated above, does not carry over to a multi-sector framework. The effects of input substitutability can, under certain conditions, even be reversed in multi-sector models, which is explained in more detail in the model of section 3.

Capturing the effects of structural change is not only useful for building a robust fundament for theoretic modelling, but also for providing a reliable guideline for the interpretation of sectoral empirical results. Empirical estimations for specific sectors have to be assessed under the premise that the different sectors do not contribute uniformly to resource use and economic dynamics. For example, it was found that pollution abatement crowds out productive investment almost entirely in the pulp and paper industry, see Gray and Shadbegian (1998), that elasticities of substitution between natural resources and capital are small in several sectors of the economy, as argued by Cleveland and Ruth (1997), and that the analysis of regional material flows shows modest recycling in developed regions, see Binder
et al. (2001). But when the different sectors of an economy behave differently and interact with each other, the sectoral results do not necessarily carry over to the aggregate level. In the same way, it is not useful to use a sector-by-sector sustainability definition to assess whether a society is able to keep productive capacities at least constant.

In order to find the impact of sectoral results in general equilibrium, we need appropriate multi-sector models (of closed and open economies) with endogenous technology. It will then become possible to show that the aggregate effects of substitution possibilities in a certain sector depend on specific sector characteristics.

To conclude, technology can have an important role in overcoming poor input substitution and material balance constraints, but it has to be further specified whether and how this can be the case. Assuming new technology to be the output of a specific research sector, sectoral allocation of inputs in the economy and the corresponding market incentives, as well as total resource availability, decide on how much knowledge is produced.

2.3 Introducing technology
In 1962, Arrow expressed that “.. a view of economic growth that depends so heavily on an exogenous variable, let alone one so difficult to measure as the quantity of knowledge, is hardly intellectually satisfactory.” What is true for economic dynamics in general is especially true for theory dealing with natural resource scarcity. So today, given the recent advances in endogenous growth theory, assuming $\dot{A}$ to be equal to zero or exogenous is no longer an adequate procedure. A comprehensive theory for $\dot{A}$ has to be supplemented, and the causal relationship between $\dot{A}$ and $\dot{K}$ should be specified. The case of biased technical change also has to be considered, which means that $\dot{A}$ has to be disaggregated and analysed at the input level.

In economic models, technical progress is usually assumed to increase an aggregate stock variable called “technical” or “public” knowledge, which raises (symmetrically or asymmetrically) the productivity of inputs into production. There are similarities between knowledge and physical capital accumulation, because savings are needed to finance both types of investments. In either case, there is no immediate (direct) compensation through the market. Is it useful to conduct the analysis similar to physical capital accumulation? The answer is no, for several reasons.

First, the production technique in the innovation sector is significantly different from capital and final goods production. In particular, innovative activities are normally assumed to be especially (skilled) labour and knowledge intensive and might, under certain conditions, be resource extensive. Second, in most applications the market form of the sectors using technical knowledge is not perfect competition, which is obvious in the case of product innovations. Third, positive externalities of innovative activities have been emphasised in theory and empirics, see Arrow (1962), Griliches (1998) and Baumol (2002). The inherent connection of these so-called “spillovers” to knowledge accumulation is also reflected in the Schumpeterian distinction between invention (a technical event), innovation (a market event) and diffusion, corresponding to the positive externalities. A differentiation between learning by doing effects and direct knowledge investments, as e.g. used by Grübler, Nakicenovic and
Victor (1999), becomes somewhat artificial in this context, because the two effects are closely linked to each other (the “doing” means “doing research” here). Positive spillovers raise the returns to aggregate capital, which can be assumed to be decreasing as in the neo-classical growth model, see Jones (1995), or constant, see Scholz and Ziemes (1999). In the former model, long-run growth is only possible with exogenous population growth; the latter case, however, provides a constant endogenous growth rate. Including non-renewable natural resources can, under certain conditions, make redundant the need for exactly constant returns to get constant growth rates, see Groth and Schou (2002).

Fourth, material balance principles do not apply for knowledge capital as is the case for physical capital. It is, however, an open issue whether the total amount of knowledge is unlimited in the very long run. But the consensus seems to be that the possibility to increase knowledge will not cease for a very long period of time. Moreover, technical knowledge is often embodied in new physical capital, so that the two stocks are by no means independent of each other. Fifth, unlike capital investments, R&D investments generate a specialised, sunk and intangible asset with the specific characteristics of a high variance and a skewed distribution of expected returns, see Jaffe, Newell and Stavins (2003, p. 471). The view that R&D can, under these circumstances, be modelled as a profit-motivated activity has been seriously challenged by evolutionary economics, see Nelson and Winter (1982). These authors argue that, when determining the amount spent for R&D, firms usually apply rules of thumb or certain routines. Nevertheless, such routines might come closer to profit-maximising behaviour, once market signals have an impact on firm behaviour. Moreover, even with a fixed budget for R&D (or a fixed share of sales spent for R&D), the quantity of R&D varies with the price of innovations, that is with the cost in the research lab, so that even in this case, supply determinants of innovation play a role, as in section 3 below.

In the next section, we present the central equations from a well-known type of dynamic resource model. This will show the important mechanisms when simultaneously dealing with resource scarcity and endogenous knowledge formation. It will turn out that certain critical conditions of the one-sector capital accumulation approach can be relaxed by introducing endogenous technology, but that new challenges arise for predictions based on theoretical modelling.

3. Modelling endogenous technology

In order to explain the term $\hat{A}$ in (2) via a theoretical approach, three major aspects must be covered, which are (i) finding appropriate input/output relationships for the research sector, (ii) specifying the incentives to invest in research under market and optimal conditions, and (iii) analysing the various effects of technology on production and consumption possibilities. In order to do so, it is helpful to distinguish two aspects of $\hat{A}$ in (2). Successful innovations yield a private return, e.g. profits after the invention of a new product. In this case, the newly invented goods are a measure of technological progress. On the other hand, innovation provides benefits to the public through positive spillovers. According to endogenous
innovation theory, public knowledge capital is accumulated through such positive externalities.

It is thus useful to look at a basic model which follows recent resource economics. This is done also to show in detail the various issues where theory can be improved in future research. The approach considers the invention of new goods varieties; more aspects of technical change are treated in section 6.

Assume research to be directed at inventing new goods varieties which we label intermediate goods. At every point in time, $n$ goods are available. R&D firms use labour $L_n$ as a rival and knowledge $\kappa$ as a non-rival input to invent new designs for new intermediate goods, according to:

$$\dot{n} = b \cdot L_n \cdot \kappa$$

(3)

where the dot denotes the derivative with respect to time and $b$ is a productivity parameter for research. The sectoral labour input is determined by $L_n = L - L_X$ where $L$ is total labour supply and $L_X$ is labour used in the rest of the economy. Natural resources are not included as an input in (3); this will be discussed in section 4.2. The spillover (which is a positive externality) to public knowledge is captured by:

$$\kappa = n^\eta$$

(4)

where $\eta$ represents the intensity of spillovers and (3) can be rearranged to yield the innovation growth rate $g$ as:

$$\frac{\dot{n}}{n} \equiv g = b \cdot L_n \cdot n^{\eta-1}$$

(3')

There are several implications following from (3'). First, the size of $\eta$ is important for the returns from knowledge accumulation and for innovation growth. With proportional spillovers from R&D to public knowledge, i.e. $\eta = 1$, we get $n = \kappa$. Then the growth rate of knowledge $g$ simply becomes $g = b \cdot L_n$, which means that returns to knowledge are constant and a given labour input produces a constant innovation growth rate. Second, there is a clear trade-off between using labour in research and the rest of the economy. Put differently, the opportunity cost of knowledge accumulation can be expressed in terms of the model. Third, a higher employment of labour in research leads to a higher innovation growth rate. This productive effect of labour is fostered by the intensive use of labour in research and the public good-property of the produced capital, which is knowledge.

To yield meaningful conclusions from (3), incentives to use labour in R&D – that is to undertake research – have to be carefully studied. In particular, it has to be established how the incentives depend on natural resource use, which definitely requires saying something about the rest of the model economy. In general, we have to distinguish between primary
(raw) inputs, produced inputs (intermediate goods) and final goods which are consumed. Labour and natural resources are the important primary inputs. Natural resources at the same time represent material inputs in this simple approach. Intermediate goods are assumed to be (immaterial) flows, which embody technical progress. For additional uses of material, also for physical capital, see the discussion in section 5.2. A firm $i$ producing an intermediate good $x_i$ uses labour $L_{xi}$ and natural resources $R_i$ under a production technology represented by $\tilde{F}$, so that $x_i = \tilde{F}(L_{xi}, R_i)$. $\tilde{F}$ is normally assumed to have constant returns to scale, but the degree of substitutability between $L_{xi}$ and $R_i$ is under debate; the elasticity of substitution may be below, equal to, or bigger than unity. With $n$ different firms producing intermediates and with symmetric costs between the firms, the quantity of each intermediate good $x_i$ is equal to $x$, the whole output of the industry is $X = n \cdot x$, and aggregate intermediate goods production is described by:

$$X = \tilde{F}(L_X, R)$$  \hspace{1cm} (5)

where $L_X$ and $R$ are total input of labour and resources in the intermediate goods sector.

Intermediates are used by final goods firms to produce final output $Y$. Following the seminal approach of Dixit and Stiglitz (1977), a CES-production function can be postulated, according to:

$$Y = \left( \int x_i^\beta \, di \right)^\frac{1}{\beta} = (n \cdot x^\beta)^\frac{1}{\beta} = n^{\frac{1-\beta}{\beta}} X$$  \hspace{1cm} (6)

This equation states that the production and thus the consumption of $Y$ depend on an input effect, captured by $X$, and a productivity effect, given by $n$ raised to the power of $(1-\beta)/\beta$. As can be seen from (6), an increasing number of varieties raises final goods production with a given $X$. Put differently, the economy grows in this model because output is produced with increasing specialisation in inputs. A different but similar interpretation is to introduce $Y$ as a utility index and to assume that increasing varieties lead to rising individual welfare because of the taste for variety-effect.

With the help of this basic model, several central implications can be derived. Regarding the relationship between inputs in equation (5), poor input substitution has different implications compared to the one-sector approach from section 2. When the resource becomes scarcer over the course of time, two effects arise simultaneously. On the one hand, $x$-producers like to substitute labour for the increasingly expensive resource, which leads to a reallocation of labour from R&D to intermediate goods production (substitution effect). On the other hand, $x$-goods and, at the same time, $Y$-goods become more expensive for consumers, so that demand and production of final output decreases and labour is released from intermediates production to the R&D-sector (output effect). When the substitution effect is smaller than the output effect, additional labour is allocated to the research sector and the innovation growth rate of the economy increases, see Bretschger (1998). The poorer input
substitution is, the faster labour is reallocated to the dynamic sector, that is, the larger the structural change in the economy. One has thus identified an additional important channel to substitute natural resources: the decrease of resource-intensive sectors (intermediate goods) and the increase of resource-extensive sectors (research). Put in an even more tapered form: an inverse relationship between input substitution and sectoral substitution can be found in this approach.

A further point in this endogenous innovation model is the effect of population growth. Here, the combination of labour intensive research, see (3), and the public good character of “knowledge”, see (4), leads to effects of population growth on per capita income which differ from earlier theories assuming investments as foregone consumption and capital as a private good. When additional labour is allocated proportionally to the different sectors, it can be derived from (3) that a larger labour force leads to higher innovation growth. The effect on per capita consumption depends on further model assumptions; it might well be positive in this simple set-up.

The previous results were obtained for the case of horizontal innovation, but models of vertical innovation – depicting rising product qualities – behave in a very similar way, see Grossman and Helpman (1991, ch.4). Moreover, the strict link between input and output in the research sector as given by (3) can be interpreted in a different (probably more realistic) way. When research outcome is uncertain for the single research unit but is governed by a Poisson process for the aggregate economy, the same type of results can be derived, see Aghion and Howitt (1998).

4. Refining innovation conditions

4.1 Demand conditions

One of the first interesting extensions of the previous model is to extend the idea of input substitution to the demand side of R&D. If we assume that final output $Y$ is produced with both intermediate inputs and natural resources, the demand for intermediate inputs (and, therefore, research) depends on the relative intensity in the use of the resource in the final goods sector. Quite naturally, it turns out that here – regarding the demand for innovations – a high substitutability between inputs is favourable for dynamics, as is the case in the one-sector resource model. The reason is that a high elasticity of input substitution guarantees a high compensation for innovative activities, just as it guarantees high returns to capital in the traditional one-sector approach with capital accumulation. Put differently, when natural resources become more scarce in the course of time, the income share to compensate successful innovators rises, which favours innovative activities. By including the demand side, it thus becomes possible to demonstrate how endogenous innovation models move closer to the spirit of earlier results of resource economics. However, the opposite result regarding input substitution, as presented in section 3, still applies for the supply conditions of innovations. Again, poor input substitution guarantees that costs in the research sector remain moderate or decrease over time. Taken together, one can show that innovation growth is
positive in the long run as long as the elasticity of substitution on the demand side is larger than the elasticity of substitution on the supply side, where both elasticities may lie below unity, see Bretschger and Smulders (2003).

Another issue is to reconsider the effects of specific gains from specialisation (and of the degree of market power) in (6), assuming that they are not necessarily tied to the parameter $\beta$, as is the case in the simple models we presented here. A further extension is to explicitly specify resource supply and to distinguish between non-renewable and renewable resources. To do so, one can either introduce the Hotelling rule for prices of non-renewable resources, or a natural regeneration rate for renewable resources, which extends the conventional Hotelling rule by a term referring to natural resource growth.

4.2 Supply conditions
Regarding the supply conditions in the research sector, an important issue is the possible use of natural resources as an input. More generally, the question arises whether and how resources are used for the “growth engine” of an economy. Provided that $R$ is an inessential input into R&D, the problem is solved in the long run, because new varieties can be invented without resources as inputs. But Groth and Schou (2002, p. 386) suggest that “non-renewable resources are clearly an important element in the technologies of present-day economies.”

In this case, we have to reformulate R&D technology in (3), according to:

$$\dot{n} = b \cdot L_n^\alpha \cdot R_n^{1-\alpha} \cdot \kappa$$

$(0 < \alpha < 1)$

(3'')

When $R$ is an essential input into R&D as in (3''), the problem of obtaining constant innovations in the long run becomes more difficult, of course. This applies because the model sketched in section 3 generates a reward to research which is bounded from above. If $R$ is a non-renewable resource, its price increases without bound according to the Hotelling rule. *Ceteris paribus*, this decreases the direct return on innovation. To maintain a minimum level of profitability in R&D, increasing returns to scale in $x$-production (equation 5) can be assumed, but this is not unambiguously plausible, see Groth and Schou (2002). Another possibility is that structural change leads to a steady labour-inflow into research, which could keep research at a constant level. Provided that input substitution in $x$-production is poor, this is the expected outcome, see Bretschger (2003).

A further aspect is the impact of innovation on knowledge. According to (4), the spillover intensity $\eta$ has a big effect on the value that total knowledge converges to in the long run. Indeed, what a “realistic” intensity of knowledge spillovers - that is the return on knowledge - actually is, has been under debate. Many endogenous growth models assume proportional spillovers, i.e. $\eta = 1$, leading to constant returns in R&D, see Romer (1990) and Grossman and Helpman (1991). The well-known critique of Jones (1995) concerns the empirical experience in developed countries showing that the relationship between input and output in research was not linear in the past. In addition, $\eta = 1$ is a knife-edge assumption. A
value of $\eta$ slightly exceeding unity leads to an explosion of the growth rate in finite time, while a $\eta$ slightly smaller than unity does not lead to endogenous growth in general. Only additional assumptions like sectoral change can lead to constant growth in the latter case.

5. Long-run effects and predictions

In principle, it is never unobjectionable to use theoretical models for predictions of the very long run. Yet, the sustainability debate unambiguously requires attributing a lot of weight to the long run, so that a discussion of long-term trends predicted by the model economy is unavoidable. In fact, the importance of the long run underlines the need for sophisticated models in this field. To point out the issues most clearly, we again assume $R$ to be a non-renewable resource (with an increasing price - and scarcity - according to the Hotelling rule). This brings about the result that the per-period use of the resource becomes lower and lower over the course of time. Following our basic model and, specifically, (6), consumption growth evolves according to:

$$\hat{Y} = [(1 - \beta)/\beta]g + \hat{X}$$

(7)

Following (7), we first regard $[(1 - \beta)/\beta]g$ and then the prediction for $\hat{X}$ in the very long run.

5.1 Long-run innovation growth

According to (4), the intensity of the spillovers determines the marginal return on R&D investments, as discussed under 4.2 above. Moreover, the productive value of innovations is determined by the effect of the gains from diversification as given by $(1 - \beta)/\beta$. The intuition behind this term goes at least back to Adam Smith’s pin factory parable. But is the assumption of a positive and constant $(1 - \beta)/\beta$ good enough to predict the effects of technology in the long run? Two points of critique have been put forward. First, one can argue that the gains from diversification are overstated if the costs of doing so are not fully captured. Indeed, costs of coordination, transportation, assembly of components etc., which reduce the overall gains, rarely appear in the expansion-in-varieties models. This is especially critical when natural resources are involved in these activities. Second, it has been questioned whether the process of diversification can be viewed as continuous in the very long run or is likely to stop at a certain point in time.

On the other hand, the international division of labour has increased dramatically over the last years and few economists would argue that there have been no gains from subsequent trade because of increased coordination activities (using more resources). Also, as any theory is open to critique, the expansion-in-varieties approach has to be compared with other concrete specifications of technology effects, in order to find out what represents current knowledge about the effects of innovations in the best (yet still incomplete) way. Finally, it is well known that there are different forms of technology with additional effects, see section 6.
Thus it is fair to say that predictions for the development in the very long run, which are based on the expansion-in-varieties approach, are not the only and ultimate truth. In fact, they were never meant to be. But nevertheless, they are a serious attempt to capture the complex effects of technology on income and should, therefore, be complemented or modified with more characteristics of the innovation-income-mechanism in the future.

5.2 Intermediate goods production
When labour and resources are poor substitutes in $x$-production, see (5), the sectoral income share of labour approaches zero and the income share of the resource goes to unity in the very long run. Thus $\dot{X}$ is negative because of the decreasing input of $R$ into intermediate goods production. Is it possible to produce (intermediate) goods with only little resource input? Again, we have to note that the values of the elasticities of substitution we estimate are based on the part of the isoquant we know, and are not necessarily valid when one extends the range of observations. In particular, with low resource input the elasticity could become low or even zero, as it seems difficult to imagine that the resource coefficients of all goods can asymptotically approach zero. In addition, physical capital is a product of materials and resources, and production with virtually no physical capital hardly seems feasible. This reasoning calls for a minimum requirement for $R$ in intermediates production.

However, a closer look at the nature of the resources depicted by $R$ can help to clarify matters in this case, that is for the (very) long run. When we think of fossil fuels and, in a somewhat broader sense, of energy supplies, so-called “backstop technologies” like solar, tidal or wind energy will be profitable after the price of the non-renewable resource has reached a certain level. If these energy forms are a perfect substitute for the considered resource, a renewable input will substitute for the non-renewable resource in the distant future so that the minimum requirement can probably be met. It all depends on whether the quantity of the resource demanded at the price of the backstop technology is sufficient to meet the minimum material requirement.

Considering raw materials like metals, it is often assumed that a certain amount of throughput is necessary to sustain economic activities in the long run. Here, recycling is the key to keep the quantity of available materials constant. But even with a high price for metals and a possibly high compensation for recycling activities, it is questionable whether it will be (physically) possible to recycle one hundred percent of material stock. Thus, once it is used for economic activities, this kind of input could decline over time, but possibly at a very slow rate. Lastly, when adopting the perspective of material input in food production, we turn to the field of renewable natural resources. Nevertheless, limited regeneration and complementary inputs like land or water present possible bottlenecks for production. More research efforts on these specific topics in the future seem especially rewarding.
6. Directions of innovations

6.1 Biased research

Empirical results and observations have led various authors to the conclusion that a significant but not predominant fraction of innovation in the energy and environment area is indeed induced and thus responds to market signals, see Popp (2001, 2002), and that environmental regulation is likely to stimulate innovation and technology adoption, see Porter and van der Linde (1995) and Jaffe and Stavins (1995). Grubb, Köhler and Anderson (2002, p. 282) emphasise that “the evidence ... clearly suggests that much technical change in the energy sector is induced, not autonomous; and models should embody this fact.” Popp (2002, p. 160) finds with the help of U.S. patent data that “both energy prices and the quality of existing knowledge have strongly significant positive effects on innovation.” In their simulation model, Gerlagh and van der Zwaan (2003) conclude that the effect of endogenising technological change is large, both in terms of the timing of carbon emission abatements and the cost of such policies. So even if firms are not fully rational and far from being fully informed about the return on innovations, on average they seem to react systematically to incentives, e.g. relative prices and market size. This was already the hypothesis of Hicks (1932), who suggested that technical progress is affected by changed relative input prices, which is crucial in the field of environment and resource saving technologies. It corresponds to the Environmental Kuznets Curve approach, where technical progress is, at least partially, induced. In addition, the assumption of endogenous innovations allows us to derive long-run “if/then”-relationships in more complex dynamic resource models, which can be tested empirically afterwards. This is what theory should provide to keep scientific progress going. The alternative is to use the agnostic “manna from heaven” assumption for technology which does not help in evaluating the policy alternatives in the environmental sector.

Once innovation is assumed to be determined by incentives, the labels “endogenous” or “induced” innovation are appropriate; to express an asymmetric impact of innovation on factor productivities, the attribute “biased” seems to fit best. Using endogenous technology in environmental and resource economics, the greatest challenge consists of showing that the history of economic development can be reversed in the future: while in the past low prices for the environment have led to extensive natural resource use and rising polluting activities, increasing prices of natural resource use should be able to change this general pattern. Corresponding to the Environmental Kuznets Curve in mature economies, income should rise in the future while natural resource use will decrease. With the help of the theory, the probability of such a development and the underlying mechanisms should be described in detail.

When we assume that two different inputs are combined with differentiated intermediate goods for production, biased technical progress can be analysed, see Smulders and de Nooij (2003). These authors postulate that labour and one kind of aggregate capital good are needed to produce one type of output, while natural resources (energy) and another kinds of aggregate capital goods are used to produce a second type of output. Innovative activities can go in both directions, that is they can increase the number of capital varieties in
both sectors. Depending on sector-specific reward, research can be more directed at one of the two sectors. When the output related to natural resources becomes relatively more expensive (e.g. according to the Hotelling rule as stated above), poor substitutability between the two types of outputs leads to an increase in compensation and demand for innovations in the resource-intensive sector. However, poor substitutability between produced outputs is empirically debatable, and supply reactions have to be supplemented to obtain the full effects in the model.

6.2 General purpose technologies

The analysis above has concentrated on incremental technical progress, but the introduction of new major technologies (general purpose technologies), see Bresnahan and Trajtenberg (1995), obviously has a large impact on the path of natural resource use as well. Undoubtedly, the model analysis is easier for gradual innovations than for general purpose technologies. But only when considering the latter can we explain relevant phenomena like path dependency and lock-in effects. All these issues are far from being broadly treated in the context of environment and growth. A possibility of combining the two types of technical progress is provided in Smulders and Bretschger (2000).

In innovation theory, the intensity of spillovers of R&D is decisive for the (social) marginal return on innovative investments. The models in sections 3 and 4 simply used proportional spillovers yielding constant returns on R&D. But even when it is argued that spillovers are less than proportional, see Jones (1995), income does not need to stagnate over time. When a constant fraction of resources is used in R&D, the result is arithmetic knowledge growth instead of geometric growth. Also, in a model with natural resources this can lead to rising incomes on the way to the steady state (which is never entirely reached). Finally, it has been noticed that the use of different energies, such as wind and biomass energy, produces non-linear learning curves; thereby, different points of the learning curve have been reached for the various energies. A possible extension is thus to introduce more than one natural resource into the model and to differentiate the intensity of spillovers between resources.

6.3 More complex R&D

In theory, the different types of R&D remain in a well-defined framework. Yet in reality, new technologies have a variety of impacts in very different fields. For instance, progress may affect several but not all sectors, which is more than incremental technical progress as in section 3, but less than general purpose technologies. Future research should try to capture these additional kinds of R&D and derive additional predictions regarding the effects of these innovations in the very long run.

An additional topic is the relationship between education of researchers and their productivity in research. Moreover, the institutional framework for performing research is an important factor for the success rate in R&D. Accordingly, the government has to design good research policies in the sense that government-financed education and institutions support the
general goals of resource saving and productivity increase. At the same time, the acceleration of knowledge diffusion has an impact on the productivity in research; public policy could also take a role in this respect.

Finally, all the different variants of innovations and the conditions which determine their economic success have to be evaluated according to their relevance for practical purposes. If, for example, only certain types of research activities depend on natural resources as essential inputs, resource-independent research can possibly be substituted for these activities.

7. More topics

Regarding the question of whether innovations are effective enough as a remedy for resource scarcity, there are, quite evidently, more topics which have to be put on the research agenda. Let us start with issues that are related to the class of models discussed so far. When emphasising structural change as a mechanism to achieve sustainable development, possible resource reallocation costs have to be considered. For instance, in the model of section 3, the assumptions of a single homogeneous labour factor and full employment at every moment in time may obscure real phenomena. In particular, the requirements of special skills in the research sector are an important topic. Accordingly, the efficiency of the education sector and adjustment costs of intersectoral labour mobility have to be carefully studied. In addition, necessary wage adjustments caused by structural change cannot be expected to materialise without a considerable time delay in the real world.

When inputs are not perfectly mobile between sectors due to resource reallocation costs, the dynamics of the economy are affected. In the class of models studied here, this happens because the restructuring of labour has not only level but also direct growth effects. In particular, the output of the research sector determines the innovation and, as a consequence, the consumption growth rates of the economy. Therefore, in this sector, labour in- and outflows have a direct impact on economic dynamics. Regarding the time horizon of sectoral change, restructuring may be modelled as a process which is never fully completed. For example, it can be assumed that a certain sector shrinks at a specific percentage rate in every period of time, such that it never completely disappears from the economy. Other sectors can be assumed to increase their share of total output and to approach a certain share but never entirely reach it. These are the assumptions used in the basic model of section 3.

On the other hand, sectoral change might stop at a certain point in the future, e.g. due to rising resource reallocation costs. Then, innovations are no longer supported by additional inflows of labour into the R&D sector. Provided that the prices of the natural resources are increasing, a negative impact on research can arise through fading compensation for innovations or increasing research costs. In the logic of the model of section 3, sectoral change also comes to an end when the resource price becomes constant. In the case of renewable resources, this happens as soon as the harvest rate is optimal and sustainable, meaning that the return on the resource equals the interest rate and the resource used each
period is regenerated by nature. Then, assuming $\eta = 1$ in (4), the system arrives at a dynamic equilibrium with constant innovation growth.

Direction, speed and optimal rate of sectoral change also depend on the utility functions of the households, that is on the demand for sectoral outputs. As long as resource-intensive goods are not superior goods, the scenario of welfare increasing sectoral change can be expected to hold. If, however, the demand for output of the expanding sector in consumption drops with the ongoing restructuring of the economy, sectoral change could come to a halt because of the demand side. In addition, environmental quality appearing in the utility function of the households may affect resource demand, see Smulders (2000) and Heal (2004), and thus sectoral change.

Another important issue is that new technologies are not neutral with regard to risk and uncertainty in the production process. One can reasonably argue that innovative techniques used for renewable resources, like wind and water power, yield little uncertainty because pollution is low and plants are decentralised. On the other hand, burning fossil fuels is cheaper than the use of these energies, but involves a higher risk because of the greenhouse effect. Accordingly, cheaper technologies with higher risk and lower costs have to be valued against technologies with opposite characteristics. Yet, new technologies could also be seen as increasing risks. Various aspects of nuclear power have been discussed in this way. Also, it could be that specific negative effects of new technologies do not become known until a certain time after they have been introduced to the market, leading to a bias in the original investment decisions. One (or maybe the best) of the possible consequences would then be to search for even better technologies.

Uncertainty has also to be taken into account when predicting the asymptotic properties of the production function used to evaluate the long-term behaviour of the economic system. Relating to individual or sectoral risks, expectations about payoffs and risks of new technologies and production processes have to be analysed from the perspective of financial markets. Decisions on the allocations of financial funds can have long-lasting impacts on the economic development through the choice and funding of certain technologies.

Finally, an important issue is to open the modelled economies with endogenous innovation to allow for foreign trade, factor movements and international spillovers, see e.g. Elbasha and Roe (1996) and the survey of Bretschger and Egli (2001). Interesting implications can also be derived in North-South models. All these topics are far from being fully explored and can be combined with the findings summarised in this paper to yield new insights.

8. Conclusions

We started our survey by asking whether new technologies are effective and efficient enough to act as a remedy for resource scarcity in the long run. It has become evident that the answer based on recent literature cannot simply be a clear yes or no, but has to be qualified in several respects. We have argued that a better understanding of the mechanisms driving innovations
contributes to forming more accurate predictions of long-run economic and environmental development. A comprehensive view of technology and the environment includes not only the relationship between these two issues, but also policies related to them. As far as political measures are able to change supply and demand incentives for investors, technology depends on environmental policy, which is an important aspect of induced and biased technical progress, see Jaffe, Newell and Stavins (2003, p. 463). On the other hand, optimal policy is determined by technological possibilities. Regarding the optimal use of instruments, it has to be observed that induced technical progress decreases costs of environmental policy, see Buonanno, Carraro and Galeotti (2003) for the Kyoto example. Moreover, in order to improve resource efficiency without negatively affecting economic growth, a combination of R&D and environmental policy can be indicated, see Van Zoon and Yetkiner (2003) and Carraro and Siniscalco (1994). Finally, the timing of optimal political measures is influenced by technology; if progress was largely autonomous, later measures would be preferred because they are more effective; but when technical progress exhibits large learning effects, earlier measures might be advantageous because of the larger knowledge accumulation, see Goulder and Mattai (2000). The combination of these ideas with the supply and demand determinants of innovations as in sections 3 and 4 promises to yield further insights in the field of environment and technology.

When modelling dynamic economic-ecological systems, the determinants of supply and demand in the innovative sector have a large impact on the long-run behaviour of endogenous variables like knowledge and income, as well as on optimal policy interventions. As a special characteristic, supply conditions of natural resources can entail immanent non-linearities of development paths, which leads to a methodological emphasis on out-of-steady-state behaviour. As policy makers demand an integrated assessment of environmental and resource policies, it is indispensable to evaluate empirical results valid for certain sectors from a general perspective on the aggregated level. In this respect, dynamic economic theory including natural resources has to play an important role.

This paper has shown that slight modifications in the model structure can, but need not, have large consequences for the interpretation of technical restrictions such as low elasticities of substitution. From this basic finding, it is argued that, in order to significantly increase our understanding of the relevant mechanisms, we have to extend the theoretical approaches in various directions and to explore the missing links in the theory. Increasing the variety of theoretical model variants can be viewed as a form of sensitivity analysis for the theory. Only the results that go through several model types can be seen as generally robust for a comprehensive theory, while other results can be clearly assigned to specific conditions.
References


