Biophysical Constraints to Economic Growth

Cutler J. Cleveland

Department of Geography
and
Center for Energy and Environmental Studies
Boston University
675 Commonwealth Ave.
Boston, MA 02215
phone: (617) 353-3083
fax: (617) 353-5986
email: cutler@bu.edu

TABLE OF CONTENTS

Introduction

The Standard Model of Economic Growth

The Ecological-Economic View of the Economy

Limits of the Market and Technology

The Role of Energy in Technical Change
Do Rising Incomes Improve Environmental Quality?
Countervailing Forces: Rising Affluence and the Rebound Effect
Thermodynamics Limits Substitution
Complementarity Limits Substitution
Physical Interdependence and Scale Limits Substitution
Irreversibility Limits Substitution
Market Signals Aren’t Always a Reliable Compass
Uncertainty, Ignorance, and the Unintended Side Effects of Technology

Is There a Carrying Capacity of the Earth For Humans?

Indicators of Scale and Carrying Capacity

Alternative Models of Production, Wealth and Utility

Will resource depletion limit growth?
Will the environment’s ability to process wastes limit economic growth?
To what degree can human-made capital substitute for natural capital?
To what degree can an educated work force substitute for natural capital?

The Search for Prometheus III

Conclusions

References
Introduction

Stable consumer prices, full employment, and increasing per capita wealth are economic and political goals in nearly every nation. Aggregate economic growth has been the principal means for realizing these goals. Yet comprehensive and independent scientific investigations provide compelling evidence that the growth of the global economy is not sustainable because it consumes many of the environmental services that underpin the production of goods and services (e.g., Houghton et al., 1996; Heywood, 1995; Postel et al., 1996; Vitousek et al., 1997). There also is a growing realization that economic growth does not necessarily go hand-in-hand with growth in the well-being of people. Standard measures of economic output such as Gross National Product do not reflect the growing disparity between rich and poor in most nations (UNDP, 1996), or the environmental degradation which diminishes the health of people, communities, ecosystems, and the economy (Daly and Cobb, 1989).

Underlying the universal prescription for economic growth are theoretical models that describe the process of growth itself. These models (and their derivatives) reflect the conventional wisdom about the driving forces behind the historic growth in living standards, the role of the environment in the economic process, and the ability of substitution and technical change to overcome resource scarcity and environmental degradation.

These models fundamentally misrepresent these important relations, and therefore contribute to the expectation that the type of economic growth we have experienced in this half-century is sustainable. There has been much discussion over the last quarter century about the role of resources in economic development and the compatibility of growth with environmental conservation. As indicated by a recent exchange between mainstream and ecological economists (Daly, 1997a, b; Stiglitz, 1997; Solow, 1997) this debate has not been settled.

The Standard Model of Economic Growth

The inclusion of environmental concerns in standard growth models is an active area of research in environmental economics (Tahvonen and Kuuluvainen, 1991, 1993; Bovenberg and Smulders, 1995; Baranzini and Bourguignon, 1995; Beltratti, 1996). Many applications of the neoclassical theory of economic growth to environmental problems downplay the likelihood that resource depletion and environmental degradation can significantly constrain economic growth (e.g., Nordhaus, 1994). However, there are a number of reasons to question this conclusion. The basic growth model in the Nobel-prize winning work by Solow (1956) does not include resources at all. This model subsequently was extended with nonrenewable resources, renewable resources, and some waste assimilation services (for surveys see Kamien and Schwartz, 1982; Toman et al., 1994). A common interpretation of standard growth theory is that substitution and technical change can effectively de-couple economic growth from resources and environmental services. Depleted resources or degraded environmental services can be replaced by more abundant substitutes, or by “equivalent” forms of human-made capital (people, machines, factories, etc.).
The neoclassical literature on growth and resources centers on what conditions permit continuing growth, or at least non-declining consumption or utility. I use the short-hand "sustainability" to refer to either continuing growth or non-declining consumption. Technical and institutional conditions determine whether or not sustainability is possible. Technical conditions refer to things such as the mix of renewable and nonrenewable resources, the initial endowments of capital and natural resources, and the ease of substitution among inputs. The institutional setting includes things such as market structure (competition versus central planning), the system of property rights (private versus common property), and the system of values towards future generations.

The elasticity of substitution (\(s\)) between what economists call capital (factories, machines, etc.) and inputs from the environment (natural resources, waste assimilation, ecosystem services) is a critical technical term that indicates by how much one of the inputs must be increased to maintain the same level of production when the use of the other input is reduced. A large \(s\) implies that the cost impact due to the rising price of one input, say natural resources, can easily be escaped by switching to a different technique of production that favors the use of another input, say capital.

As neoclassical economists are primarily interested in what institutional arrangements, and not what technical arrangements, will lead to sustainability, they typically assume *a priori* that sustainability is technically feasible. A unitary elasticity of substitution (\(s=1\)), referred to as "perfect substitutability," means that as the ratio of the two inputs is changed by a given percentage holding output constant, the ratio of their marginal products changes by the same percentage (in the opposite direction). Perfect substitutability does not mean that resources and capital are equivalently useful - in fact as resource availability declines its marginal productivity rises *ad infinitum*. Even so, as we discuss below, perfect substitutability is an unrealistic assumption from a biophysical perspective. Economists such as Solow (1974) explicitly dispose of cases where \(s\) for non-renewable resources and capital is greater or less than unity. In the former case substitution possibilities are large and therefore the possibility of non-sustainability is not an issue. In the latter case, sustainability is not feasible if an economy uses only non-renewable resources. Of course, where there are renewable resources sustainability is technically feasible, at least in the absence of population growth.

Substitution that is technically possible will not occur unless society invests in sufficient capital over time to replace the depleted natural resources and ecosystem services. How much investment does take place depends on the institutional setting of the economy. For example, in an economy where sustainability is just technically feasible (\(s=1\)) and there are only non-renewable resources, sustainability will not occur in either a competitive or centrally-planned economy where the decision rule is the maximization of the discounted flow of utility of future generations using a constant and positive discount rate. Consumption per capita will eventually decline to zero after an initial period of economic growth because resources and ecosystem services are depleted faster than capital can be accumulated to replace them (Stiglitz, 1974; Dasgupta and Heal, 1979). Sustainability is achieved under certain institutional settings (Solow, 1974). If the utility of individuals is given equal weight without regard to when they happen to live and the aim is to maximize the sum of utilities over time, then growth in consumption can
occur indefinitely. This is equivalent to maximizing net present value with a zero discount rate. Obviously, therefore, a constant level of consumption over time also is feasible. An important result in this context is the Hartwick rule (Hartwick, 1977) which shows that if sustainability is technically feasible, a constant level of consumption can be achieved by reinvesting resource rents in other forms of capital, which in turn can substitute for resources. Dixit et al. (1980) extended the rule to multiple capital stocks while Hartwick (1995) extended the rule to open economies.

How well do economic models reflect the material basis of the economy? Neoclassical economists argue that the class of growth models that include resources can account for mass balance and thermodynamic constraints with the “essentiality condition.” If $s$ is greater than one, then resources are "non-essential." If $s$ is less than or equal to one, than resources are "essential." Essential in this case means that given positive non-resource inputs, output is only zero when the resource input is zero, and strictly positive otherwise. The Cobb-Douglas production function, a frequent form used in growth models, has the essentiality condition. Economists argue that this at least accounts for the fact that some amount of energy and materials are required to produce goods and services. But when the elasticity of substitution is unity this “essential” amount can be infinitesimal if sufficient manufactured capital is applied. Economists also note that resources and capital are interdependent in the neoclassical models in that some positive quantity of resources is required to produce capital assets. Thus, the capital stock cannot be increased without depleting the resource stock. Some economists acknowledge that an assumed value for $s$ of one or greater between energy and other inputs violates the laws of thermodynamics (Dasgupta and Heal, 1979, p.211). But, in general, this important constraint—and its implications for substitution—have not been integrated into the main body of work on sustainability.

Modern growth theory has sought to improve the standard theory by “endogenizing” technical change through more explicit modeling of investments in human capital (education, health care) and new technology (research and development). These may prove to be important advances. In Romer's (1990) model there are decreasing returns in the acquisition of knowledge, which is surely more physically realistic, but this school still assumes that human-made capital is perfectly substitutable for resources and environmental services (e.g. Gradus and Smulders, 1993).

In summary, environmental economists have paid increasing attention to the environment, extending the standard tools of micro- and macro-economics to problems of resource depletion and waste assimilation, and in doing so have provided insight into some of the costs and benefits of alternative plans to ameliorate environmental problems. Some environmental economists have engaged natural scientists and policy-makers in constructive debate about what if anything should be done about environmental problems. Yet, despite the increased emphasis by environmental economists on accounting for the role of the environment in economic production, their treatment of the topic remains incomplete. While some of the relevant mechanisms have been incorporated into individual models, models incorporating all of the important feedbacks have not been developed, and some models used in applied work continue to ignore resources and the environment.
Perhaps more importantly, Common (1997) argues that it is difficult to avoid the conclusion that economics as opposed to individual environmental economists does not take the material basis of the economy seriously. The majority of degree programs in economics at the undergraduate and graduate level do not require students to take courses in resource and/or environmental economics. The majority of standard texts pay little attention to resource and environmental issues; the indexes of some popular texts do not even include entries on energy, natural resources, pollution, or the environment (Common, 1997).

**The Ecological-Economic View of the Economy**

Ecological economists have a fundamentally different “pre-analytic vision” of the economic process than neoclassical economics. The economic process is sustained by a flow of low entropy (high quality) energy, materials, and ecological services from the environment. Collectively, these resources and services are called natural capital. Ecological economists distinguish between natural capital, which generates natural resources and ecological services, and the more familiar form of capital manufactured by or residing in humans and their economies, cultures, and institutions. The latter form of capital takes two broad forms. Human-made or manufactured capital refers to factories, buildings, tools, and other physical artifacts. Human capital refers to the stock of education, skills, culture, and knowledge stored in human beings themselves.

Resource-augmenting technical change and substitution between natural and human capital are the core of the debate about limits to growth and sustainable development (Turner, 1997), as evidenced by the long history of the debate. The most renowned exchanges in this debate are those between Herman Daly, Robert Solow, and Joseph Stiglitz. Daly (1979; 1997a) criticizes the growth models of Solow (1974) and Stiglitz (1979) because the production functions they use assume perfect substitutability of manufactured capital for natural capital. Daly argues that the two forms of capital largely are complements because human capital ultimately is derived from and sustained by energy, materials, and ecological services. Similar arguments are made by other ecological economists (Ayres and Nair, 1984; Costanza and Daly, 1992; Cleveland and Ruth, 1997; GutÄs 1996; Stern, 1997a; Victor et al., 1995). But as van den Bergh (1997) notes, the substitute-complement debate has shed less light on the issue than it could because it tends to characterize the human-natural capital relationship at one end of the spectrum or the other, when in fact it is very complex issue that defies a single, universal label.

**Limits of the Market and Technology**

To what extent can technical and substitution change de-couple the production of goods and services from energy and material inputs and waste outputs? To what extent can human ingenuity “substitute” for depleted resources and degraded ecosystem services? This is the core of the debate about limits to growth and sustainable development (Turner, 1997).

The idealized response of the market and technology to depletion and degradation is seductively simple. However, there are a number of factors that reduce or override the potential
of the market to ameliorate environmental problems. These are: 1) technical change has heavily relied on using energy of increasing quality; 2) rising affluence may not automatically lead to improvements in environmental quality; 3) improvements in the efficiency of energy and material use can produce a “rebound effect” that actually increases resource use or waste generation; 4) thermodynamic limits substitutions that decrease energy and material use; 5) human capital and natural capital are largely complements, which limits the degree to which the former can substitute for the latter; 6) human capital is made from and operates on flows of energy and materials, which limits substitution possibilities; 7) Ecosystems that provide critical life support services that have no human equivalent cannot be reduced below minimum threshold levels which once breached, produces an irreversible loss ecological service; 8) the market often does not provide the right signals for technical change; 9) technical change often has unanticipated side effects. I discuss these in turn.

**The Role of Energy in Technical Change**

Among the countless technologies humans have developed, only two have increased our power over the environment in an essential way. Georgescu-Roegen (1979) called these Promethean technologies. Promethean I was fire, unique because it was a qualitative conversion of energy (chemical to thermal) and because it generates a chain reaction that sustains so long as sufficient fuel is forthcoming. As Georgescu-Roegen (1982) described:

> The mastery of fire enabled man not only to keep warm and cook the food, but, above all to smelt and forge metals, and to bake bricks, ceramics, and lime. No wonder that the ancient Greeks attributed to Prometheus (a demigod, not a mortal) the bringing of fire to us (p. 30).

Promethean II was the heat engine. Like fire, heat engines achieve a qualitative conversion of energy (heat into mechanical work), and they sustain in a more complex way a chain reaction process by supplying surplus energy. Surplus energy or (net energy) is the gross energy extracted less the energy used in the extraction process itself. The Promethean nature of fossil fuels is due to the much larger surplus they deliver compared to animate energy converters such as draft animals and human labor (Cottrell, 1955; Odum, 1971; Cleveland et al., 1984; Hall et al., 1986; Cleveland, 1993).

The unparalleled ability of fossil fuels, and especially oil, to produce economic wealth is due to another attribute: energy quality. This refers to the fact that heat units of different fuels have different abilities to do work. The ability of one heat unit to do work varies among energy types because heat is the lowest common denominator among fuels. But humans use energy for tasks other than supplying heat, so the standard practice of aggregating fuel types by heat equivalents misses important difference in energy quality. For example, a kcal of electricity used to power an electric locomotive can move a train about three times further than a kcal of diesel fuel used to power a diesel locomotive (Adams and Miovic, 1968). Open hearth and electric arc furnaces require different quantities of coal and electricity respectively to produce a ton of steel.
Kaufmann (1994) demonstrates the profound economic importance of the physical and engineering aspects of energy quality. It is not possible to measure in physical units the economic work associated with many services provided by fuels, such as heating homes, driving cars, and dispelling darkness. Furthermore, people want a particular service delivered in a particular way, such as motive power with safety, style, comfort, etc. Thus, the economic significance of a fuel is its marginal product: the amount of economic value generated by a heat unit. The results of Kaufmann's (1994) econometric model indicates that the marginal product of fuels in the U.S. economy varies over time, but that there is a consistent ranking of fuel quality: primary electricity is the highest quality, followed by oil, gas, and coal.

Energy quality plays a dominant role in determining the quantity of energy a society requires to produce wealth. The decrease in the energy/real GDP ratio in most industrial nations often is attributed to energy-saving technical change and substitutions caused by the energy price shocks. But detailed empirical analyses indicate that much of the variation of the energy real/GDP ratio is due to shifts changes in the composition of fuel use, and hence changes in the quality of fuel use. Much of the variation in the energy/real GDP ratio for the five largest industrial nations in the post-war period is due to changes in energy quality (Kaumfann, 1992; Cleveland et al., 1984; Ko et al., 1998). Kaufmann (1992) finds no statistical evidence for autonomous energy saving technical change in this period. As he states:

[This] should not be interpreted as an argument that substitution or technical change cannot reduce the amount of energy used to produce a unit of output...Technical change has reduced the amount of energy (as measured in heat units) used to produce a unit of output. But characterizing that technical change as "energy saving" is misleading. Over the last forty years, technical change has reduced the amount of heat energy used to produce a unit of output by developing new techniques for using oil, natural gas, and primary electricity in place of coal. These technical innovations ... take advantage of the physical characteristics of these energies that allow oil, natural gas, and primary electricity to do more useful work per heat unit than coal. This interpretation implies that technical change is not something shaped solely by the mind of man... but rather technical change is shaped in part by the physical attributes of energies available from the environment (p. 53).

Do Rising Incomes Improve Environmental Quality?

A new line of argument actually has rising incomes reducing depletion and degradation. The hypothesis underlying environmental Kuznets curve (EKC) is that resource depletion and pollution tend to fall as incomes rise, producing an inverted U-shape function. The initial research on EKCs suggested that some pollutants follow an inverted-U curve with respect to income (Grossman and Krueger 1995; Shafik and Bandyopadhyay 1992; Panayatou 1993; Shafik 1994; Selden and Song 1994). These results have been extrapolated by some to be an omnipresent outcome of economic development. The theoretical EKC model consistently
appears in the World Development Report of the World Bank (World Bank 1992), and in statements such as "the strong correlation between incomes and the extent to which environmental protection measures are adopted demonstrates that, in the longer run, the surest way to improve your environment is to become rich" (Beckerman 1992, p. 491).

These conclusions are the subject of considerable scrutiny by many analysts (Arrow et al. 1995, Stern et al. 1996) and special journal issues (Ecological Economics 1998; Environment and Development Economics 1996; Ecological Applications 1996). This body of work indicates the EKC hypothesis is just that: a tentative hypothesis about the relation between income an environmental quality. A number of unknowns, uncertainties, and errors have been identified:

- Many of the regression models that find the existence of an inverted U function may be misspecified or suffer from omitted variable bias. In particular, they omit important variables such as the composition of production and consumption, international trade and the density of economic activity, to name a few (Kaufmann et al. 1998).
- Most of the improvements in environmental quality identified in EKC studies have been achieved in part due to specific environmental policies, which are indirectly related to income.
- The inverted U curve has been examined for only a few pollutants, usually those that have local health effects that can mitigated with existing technology at moderate economic expense.
- None of the EKC work has assessed the broad array of ecosystem services that underpin our biological and economic existence.
- Different studies with similar data have produced different results (Ekins et al., 1994).
- The existing work provides limited insight into the actual mechanisms that diminish pollution after particular income levels.

Thus, contrary to the sweeping claims of some analysts (Larson et al. 1986, Bernardini and Galli 1993) and potentially contrary to the seductive idea that growth itself is the antidote to depletion and degradation (Beckerman 1992), the quality and quantity of evidence does not yet support the hypothesis that the EKC is an ironclad, universal phenomenon.

Countervailing Forces: Rising Affluence and the Rebound Effect

There is no guarantee that technical change will reduce pressure on the environment faster than overall economic growth increases that pressure. Wernick (1994) and others note that population growth and rising affluence increase energy and material use, offsetting substitution, technical change, and other forces that promote de-linking. Industrial growth in Japan has more than offset the significant improvement in the efficiency of fuel, electricity, and water use in industry (Jänicke et al. 1997). Aggregate economic growth in the US helped drive the consumption of wood products by enhancing the effects of an increased intensity of use of paper and by offsetting a decrease in the intensity of use of lumber (Waggoner et al., 1996). The decline in the intensity of use of metals in telecommunications and broader electronics and computer markets brought about by miniaturization has been offset by the overall growth in these industries (Key and Schlabach, 1986). Of course, the opposite effect also is possible; slower or negative growth can reinforce declining intensity of use. Tilton (1990) and Roberts (1988) find
that slower growth of the global economy helped reduce the demand for many metals by reinforcing the decline in their intensity of use.

Another important force is the so-called “rebound effect.” Distilled to its essence, the rebound effect applied to energy is this: an energy efficiency gain looks to the consumer a lot like a price reduction. This spurs an increase in the demand for energy either directly through price elasticity effects (e.g., people buying more gasoline when its price drops), or indirectly through released purchasing power redirected to energy-using goods and services (Saunders 1992). The implication is that one cannot look at just an individual material or an individual sector to assess the net benefit to the economy from improved energy or material efficiency. The effects of change in efficiency in one sector or for one resource ripple through the economy, affecting energy and material use in other sectors and in future time periods.

There is theoretical and empirical evidence that supports the existence of a significant effect for energy. Saunders (1992) uses a macroeconomic Cobb-Douglas and CES (constant elasticity of substitution) production function to show that, in general, energy efficiency gains increase energy use by making energy appear effectively cheaper than other inputs, and by stimulating economic growth, which pulls up energy use. Several analysts have estimated the size of the rebound effect that is caused by gains in automobile efficiency in the US. The effect is measured by the percent increase in miles driven associated with a one percent increase in the energy efficiency of automobiles. Values range from 0.05 to 0.40, with most estimates between 0.1 and 0.2. This means that 10 to 20 percent of the motor gasoline saved due to increased energy efficiency is “lost” by increased driving. Khazoom (1980, 1987, 1989) claims that Lovins (1986) overstates the potential energy savings from more efficient appliances because he ignores the rebound effect. In a similar vein, Brookes (1990) argues that relying on energy efficiency to mitigate the greenhouse effect is fundamentally flawed because “reductions in energy intensity that are not damaging to the economy are associated with increases, not decreases, in energy demand.” Lovins (1988) and Grubb (1990) take issue with the arguments of Khazoom and Brookes.

**Thermodynamics Limits Substitution**

Thermodynamics can tell us a lot about the limits to substitution and technology at the level of individual processes or industries. The limits to substitution are easily identified for individual processes by an energy-materials analysis that defines the fundamental limitations of transforming materials into different thermodynamic states and on the use of energy to achieve that transformation (Ruth 1993). These types of analyses have shown where technological improvements exhibit strong diminishing returns due to thermodynamic limits, and where there is substantial room for improvements in the efficiency of energy and material use. For example, the energy efficiency of power plants and the synthesis of ammonia are approaching their thermodynamic limits. In the area of energy technologies, thermodynamic analyses suggest good reasons for not pursuing research on thermal methods for generating hydrogen from water.


**Complementarity Limits Substitution**

Production is a transformation process in which two agents of transformation, human labor and manufactured capital transform a flow of materials, energy, and information. The flow of energy, materials and services from natural capital is what is being transformed (the material cause), while manufactured capital effects the transformation (the efficient cause). For example, all machines require energy for their operation and they function by acting on a flow of materials from natural capital (Victor, 1994). Thus, adding to the stock of pulp mills does not produce an increase in pulp unless there also is the wood fiber to feed them. Material and efficient causes clearly are complements. Historically, manufactured capital and natural capital have been developed as complements, not substitutes (Daly, 1991). The stock of manufactured capital such as tractors, oil rigs, and fishing vessels has been increased with the express intent of increasing the use of natural capital such as fertile soil, oil deposits and fish populations. For these reasons, many ecological economists argue that this complementarity limits the degree to which the agents (machines) can be substituted for the flows (materials) they transform (Costanza and Daly, 1992; Victor, 1991, 1994; van den Bergh, 1997).

**Physical Interdependence and Scale Limits Substitution**

There is a biophysical interdependence between manufactured and natural capital. The construction, operation, and maintenance of tools, machines, and factories require a flow of materials, energy from natural capital. Similarly, the humans that direct manufactured capital energy and materials (i.e., food and water). Thus, producing more of the “substitute,” i.e. manufactured capital, requires more of the thing that it is supposed to substitute for.

Conventional wisdom about economic growth does not account for this interdependence, and thus assumes a degree of substitutability that may not exist (Georgescu-Roegen, 1979; Cleveland et al., 1984; Ayres and Nair, 1984; Kaufmann, 1992; Daly, 1997). It is critical to distinguish between the micro- and macro- level. Substitution is fundamentally more constrained at the macro-level of analysis than at the micro-level (Stern, 1997). For example, home insulation directly substitutes for heating fuel, a clear substitution of manufactured capital for natural capital within the household sector. But interdependence means that insulation requires fuel to manufacture, so for the economy as a whole the net substitution of insulation for fuel is less than that indicated by an analysis of the household sector in isolation from the rest of the economy.

Empirical analyses of the substitution issue are few in numbers and varied in their results. Some suggest that manufactured capital is a good substitute for major metals (Brown and Field, 1979, while others find a wide range of substitutability between manufactured capital and aggregate material inputs that is highly dependent on a priori model specification (Moroney and Trapeni, 1981). Some studies find little or zero possibility for substitution between manufactured capital and specific strategic metals (Deadman and Turner, 1988). All of these are industry level elasticities, and most cover only major nonrenewable resources such as metals. There are no empirical estimates of the degree of substitution between manufactured capital and any major ecosystem service.
Models that account for scale and interdependency constraints find that over a broad range of plausible substitution possibilities, any reduction in environmental life support lowers the long-run growth path of the economy (Kaufmann, 1995). Degradation or depletion diverts more capital and labor to the extractive sector, reducing investment and/or consumption in the rest of the economy.

Inattention to scale also leads some analysts to ignore how most nations substitute or supplement domestic natural capital with trade. For example, Pearce and Atkinson (1993) suggest that the current paths of the US and Japanese economies are sustainable because they invest in human-made capital faster than they depreciate all forms of *domestic* human-made and natural capital. But all industrial nations use and in some instances degrade natural capital from foreign and global stocks, and thus live beyond the means sustainable by domestic sources. For example, the 30 largest cities in the Baltic Sea drainage basin use 200 km$^2$ of terrestrial and aquatic ecosystem for every 1 km$^2$ of urban area to produce their use of agricultural, forestry, and fishery products (Folke et al., 1996).

**Irreversibility Limits Substitution**

Ecosystems are multi-functional. A forest produces a range of energy and materials (wood, chemicals) and services (habitat for biodiversity, climate regulation, flood protection). Some functions may be substitutable by manufactured capital, e.g., wood as a raw material. But species and their environment are connected in a complex web of interrelations that fundamentally are non-linear and evolutionary, with lags, discontinuities, thresholds and limits (Perrings et al., 1995). Change in ecosystems, whether “natural” or human-induced, often is not continuous. Rather, it often is episodic and rapid (Holling *et al*., 1995). Degradation of ecosystems, and hence the services they provide, often is irreversible. This means that some ecosystem services (e.g. climate regulation) are non-substitutable. In such cases no amount or type of human-made capital can replace natural capital; the elasticity of substitution is zero.

Ecosystems that provide critical life support services that have no human equivalent cannot be reduced below minimum threshold levels which once breached, produces an irreversible loss ecological service. Fuelwood collection and arable production in Vietnam has converted about one-third of the area to barren land (Perrings et al., 1995). In theory, the land, soil and vegetative cover could be recovered. In practice, the cost of rehabilitation and restoration is enormous and far exceeds societal resources. The degradation is, for all intents and purposes, irreversible. The essential nature of many ecosystem services and their potential for irreversible change are the basis for policies such as the precautionary principal and safe minimum standards.

**Market Signals Aren’t Always a Reliable Compass**

Another question is whether technology will follow the “right” direction (Gutés, 1996). Technological optimists make a critical assumption: the relative price of inputs accurately reflects their relative scarcities. If this is the case, the depletion or degradation of natural capital
will lead to an increase in its price, and market forces will induce technological change towards saving or improving the productivity of natural capital. If prices fail to signal scarcity, there is no guarantee that technology will be biased in the “right” direction. There is ample evidence that in most nations market failures such as government regulation, subsidies, monopolies, and externalities significantly distort the price of energy and material resources. For most ecosystem services, markets, and hence prices, do not even exist.

**Uncertainty, Ignorance, and the Unintended Side Effects of Technology**

Environmental problems such as global climate change have larger uncertainties that are not easily reduced. Understanding global warming requires information on atmospheric physics, ocean circulation, and global photosynthesis. Most importantly, it requires an understanding of the interconnections among those factors. There is enormous scientific uncertainty about those interconnections. For example, there is a lot we do not know about deep ocean circulation and how it affects climate. The uncertainties multiply when you link the atmosphere, the oceans, and vegetation together in a complex system that determines the planet’s climate. Added to this is uncertainty about the human response to climate change. How will our economic well being be affected by a change in climate? What is the appropriate response of government to a change in climate? These are difficult questions that add to the complexity and uncertainty about climate change.

There are a multitude of examples of mechanical, medical, energy, chemical, and biological technologies that had the opposite of their intended effect, or which had unanticipated side or “revenge” effects (Tenner, 1996). These effects often are displaced in time and/or space, making their presence more difficult to detect. The net effect has been a systematic underestimation of the nature-consuming aspects of technology and systematic overestimation of its nature-saving aspects (Doeleman, 1992).

Nuclear power, once projected to be "too cheap to meter," now has been abandoned by many nations due to higher than expected costs, reliability questions, and waste disposal problems. Tall stacks on power plants alleviated local air pollution problems, but caused longer term, more severe problems hundreds of kilometers away. Pesticides kill not only target organisms but also natural predators, thereby exacerbating crop damage. The deliberate or inadvertent introduction of alien plant and animal species has had devastating effects on native plant and animal species.

No one denies the immensely positive effects that technology has had on human existence. But it is prudent to fully account for the costs as well as the benefits of new technologies. History teaches us that unquestioning faith in the market or another human institution to self-generate technical antidote to environmental is myopic. Given the global dimensions of the current generation of environmental problems, it also is dangerous.

**Is There a Carrying Capacity of the Earth For Humans?**
Carrying capacity is the maximum number of individuals of a population that can be maintained indefinitely by the life support services of a given area of the environment without degrading those life support services. In principle, the carrying capacity of non-human populations is straightforward. Each individual has roughly the same demand on its environment: food, habitat space for refuge, reproduction and waste assimilation, and so on. A given area of the environment has a relatively fixed amount of each resource that sets the carrying capacity for the population. The dynamics population growth is controlled by negative feedback loops that are characterized by density independent factors (e.g., weather) and density dependent factors (e.g., predation, disease, reproduction). A population can overshoot its carrying capacity due to a lag in the effect of density dependent factors. Overshoot is temporary because the negative feedback loop between population and density dependent factor tends to shrink the population when it exceeds its carrying capacity.

In reality, it is difficult to calculate the carrying capacity of non-human populations because populations often vary widely over time, often in response to external factors such as climate.

Applying the concept of carrying capacity to human populations is even more problematic because the relationship between human society and the environment is far more complex than that for non-human populations. There are several ways in which humans are fundamentally different from plants and animals. First, people use the environment to do much more than feed and clothe themselves. Most of the natural resources consumed in Finland or Japan are used to produce goods and services that are not necessary for biological survival. Thus, one has to define an average standard of living to even begin an assessment of carrying capacity. Second, there are rapid changes in the types and quantities of resources used by the human population. This flexibility implies that a critical resource now may be unimportant in the future, and that an unimportant resource now may be important in the future. Third, humans purposefully change their environment in ways that increase the amount of life support. Agriculture is an obvious example of this capability. Fourth, humans can increase carrying capacity by expanding the geographic extent of the environment from which they obtain life support. The Japanese economy often is described as an ‘economic miracle.’ The island of Japan is densely populated and has very few natural resources nonetheless, the people of Japan enjoy a rich lifestyle. Drawing life support from environments scattered across the entire planet produces that lifestyle. Almost all the paper and wood products used by the Japanese come from trees grown in Southeast Asia, and oil from the Middle East.

**Indicators of Scale and Carrying Capacity**

The problems associated with estimating human capacity have not deterred people from making estimates. In 1679, Antoni van Leeuwenhoek, the Dutch inventor of the microscope, published the first quantitative estimate of the Earth’s carrying capacity: 13.4 billion people. Since then, there have been many attempts to estimate global carrying capacity (see Cohen, 1995 for a review). The estimates range from less than one billion to more than 1 trillion people. The enormous range in estimates is due in part to the range of methods employed. These range from assumed constraints from a single resource, usually food, assumed constraints from multiple
resources (food and water), mathematical curve fitting of population growth rates, generalizations from observed population densities, and categorical assertion (“It’s so because I say it’s so.”).

Despite all the pitfalls, there are some indicators of the scale of human existence relative to the global environment. These indicators measure human consumption or appropriation of key global resources or environmental services. Examples include:

- Humans are now the preeminent forces in many of the planet’s material cycles. The release of carbon stored by fossil fuel combustion, deforestation, and other biomass burning contributes to the increasing CO$_2$ concentration in the atmosphere. We fix more nitrogen annually than natural processes through the production of fertilizers and the combustion of fossil fuels (Vitousek, 1994). Our mobilization of trace metals such as lead and cadmium is enormous relative to natural sources (Nriagu, 1990), resulting in their accumulation in soils to levels that cause serious human and ecosystem health problems in many nations (Thomas and Sprio, 1994).
- Humans consume or control about 40 percent of global terrestrial net primary production, the total food resource on the planet, (Vitousek, 1994).
- Humans use about 26 percent of total terrestrial evapotranspiration (water taken up and eventually released by plants) and 54 percent of water runoff that is geographically and temporally accessible (Postel et al., 1996). Regional water scarcities are a growing constraint on economic development and are the source of a growing number of conflicts.
- Humans have fully exploited, overexploited or depleted two-thirds of the planet’s marine fisheries (Food and Agriculture Organisation, 1994). Chronic overfishing has wreaked havoc with the ecology of marine ecosystems and the economies of local communities whose livelihood is based on fishing.
- About one-third of the non-ice land surface of the planet has been converted to human-dominated landscapes, predominantly cropland and pasture (Buringh and Dudal, 1987). The resulting loss of original habitat is a major driving force behind the loss of biological diversity, and with it ecosystem functions that support human existence in countless ways. The Global Biodiversity Assessment (Heywood, 1995) sponsored by the United Nations Environmental Programme, the product of 1,500 scientists working on this issue, documents the magnitude of the biodiversity problem. Based on predicted future rates of tropical forest loss, the corresponding loss of biodiversity is 1 to 10% of all species in the next quarter century. These rates of extinction would be approximately 1,000 to 10,000 times the average expected “background” extinction rate.

These numbers describe a population whose demand for environmental life support is large relative to its environment. With population and affluence on the rise, these demands will rise regardless of any efficiency gains won by technical improvements.
Alternative Models of Production, Wealth and Utility

To evaluate the potential for sustainability and the viability of policies for moving economic activity in that direction, growth models must account for the relation between human-made and natural capital. A number of models have been developed that begin to address some of these issues (van den Bergh and Nijkamp, 1994; Gross and Veendorp, 1990). Revised growth models should be able to address the following questions.

Will resource depletion limit growth?

Much of the debate about the potential for resource depletion to limit economic growth focuses on the potential to substitute for scarce natural resources. The scale issue is critical here. Microeconomic analysis indicates that substitution can allow economic activity to continue, but they ignore the macroeconomic and global effects of substitution. Resource depletion often increases the quantity of human-made capital required to extract a unit of natural resource. These costs have been generally ignored in growth models (see Dorfman's (1982) critique of Dasgupta (1982)). Dorfman (1982) also criticizes the absence of capital deterioration in these models.

To represent the effect of resource depletion on economic activity, we suggest that the quantity of manufactured capital, labor, and natural resources required to extract a unit of natural resource should be modeled separately from production by other sectors. In addition, the feedback loop from resource depletion to the quantity of factor inputs required should be modeled explicitly. Technology has the potential to offset the effects of depletion, but the effects of technology should be represented consistent with the historical record of diminishing returns to the ability of technology to reduce the quantity of human-made and natural capital required to extract resources (Cleveland and Ruth, 1996).

Will the environment’s ability to process wastes limit economic growth?

Current models use damage functions to represent the reduction in economic output relative to levels that are possible in an unpolluted environment. For example, the DICE model specifies a 1.67 percent reduction in GDP when temperature rises 3$^\circ$ F relative to pre-industrial levels (Nordhaus, 1994). This damage function arbitrarily assumes that all types of environmental damage reduce output at a specific rate in the future.

These types of damage functions do little to advance our understanding of how energy and materials flow between the economy and the environment and within economic systems feed back to the production of output. We must represent the flows of matter and energy explicitly in three ways. First, energy and material flows in the environment create the natural capital which humans extract and transform to raw materials. Second, the use and depreciation of human-made capital creates material and energy wastes. Third, those wastes are assimilated by natural capital. Growth models must represent the accumulation of wastes in the environment should economic activity generate wastes faster than natural capital can process them.
Growth models should account for the fact that the accumulation of wastes has negative feedbacks on economic output that often are lagged in time and space. The accumulation of waste materials can reduce the productivity of natural capital, as when ozone accumulates in the troposphere and depresses crop yields. This increases the quantity of human-made capital required to produce a given quantity of food. The accumulation of wastes also slows the rate at which natural capital can process waste material, as when sewage reduces the ability of aquatic ecosystems ability to process organic material. Finally, the accumulation of waste materials reduces the productivity of human-made capital in intermediate sectors, as when pollution increases workdays lost to illness.

To what degree can human-made capital substitute for natural capital?

The notion that the economy can grow ad infinitum depends in large part on the assumption that technical advances can reduce the quantity of natural capital required to produce a unit of economic output.

Growth models must account for the forces that produced the broad patterns of energy and material intensity of GDP. Much of the reduction in energy use per dollar of GDP in industrial nations is associated with the substitution of high quality for low quality fuels (Cleveland, et al., 1984; Kaufmann, 1992). Further reductions in the energy intensity of economic activity should be assessed relative to the availability of high quality fuels. Similarly, a significant portion of the reduction in the use of mineral and metals per unit output may be associated with changes in trade patterns. Many developed nations now import large quantities of metal and minerals embodied in finished goods such as automobiles. These embodied metal and mineral flows rarely are accounted for in growth models. Finally, models must represent the factors that cause resource use and waste production to change with income. In fact, income often is not the causal factor, but rather income is correlated with causal factors such as the spatial intensity of economic activity and imports (Kaufmann, et al., 1998).

To what degree can an educated work force substitute for natural capital?

Investments in education increase labor productivity, which we recognize as economic growth. Thus, it may be possible to sustain economic growth with little or no increase in the flow of services from natural and manufactured capital by producing a more educated work force. To assess this potential, growth models must represent the costs of increasing the level of education. Education consumes considerable amounts of natural and human-made capital. These costs may be small now, but they could become quite large over time. Some models assume that the skill level of the work force grows steadily over time. Over a long enough time horizon, the costs associated with such an increase could become quite large. For example, what is the educational effort required to increase the skill level a modest 2 percent per year over the next century? Does this increase imply that every person goes to college or get an advanced degree? The costs of such an effort would be large, both in terms of the educational work force and the
opportunity cost of extended schooling. The longer someone stays in the workforce the shorter time they will be in the workforce (Pezzey, 1992).

The Search for Prometheus III

The Promethean nature of fossil fuels aside, they have distinct drawbacks. Their use has produced a major alteration of the planet’s carbon cycle, which may be contributing to climate change. Fossil fuels are finite. World oil production will peak in 2020±10 years, and coal will follow suit a few decades later. More stringent restrictions on carbon emissions than those currently contemplated would shorten these lifetimes. It seems unlikely that we will tap the vast store of energy in tar sands and oil shales due to their high carbon intensities.

Thus, humanity faces a fundamental challenge: the need to replace fossil fuels with solar technologies that have Promethean qualities. Two qualities are critical. First, the new energy technologies must eliminate or substantially reduce carbon emissions. Second, they must approach the abilities of fossil fuels to generate economic wealth per heat unit. Solar technologies clearly meet the first criteria since they are free of carbon.

The extent to which they meet the second criteria is an open question because solar energy inherently is a lower quality than fossil fuels. The diffuse nature of incoming solar radiation requires a significant investment of energy and materials to capture, collect, and concentrate sunlight. This means that many solar technologies deliver a lower energy surplus than fossil fuels (Cleveland et. al, 1984; Hall et al., 1986; Gever et al., 1986). Equally important, the substantial “material scaffold” required to collect solar energy is made from fossil fuels.

Solar energy, therefore, currently is a “parasite” on fossil fuel systems because they cannot “reproduce” themselves (Georgescu-Roegen, 1979). Many biomass-based fuels such as ethanol are feasible recipes but fail the test of viable technologies because of low net energy yields and high environmental costs (Pimentel, 1991). On the other hand, some technologies such as solar parabolic collectors have become viable through innovations that improve their net energy yields (Cleveland and Herendeen, 1989). Photovoltaics and wind turbines also exhibit significant technical improvements. The issue is whether a sufficient number of solar technologies can move from “feasible” to “viable” status in terms of their net energy return, and whether they can be scaled-up in time to offset the economic effects of fossil fuel depletion.

Conclusions

Although national economies are in very different stages of development, in the aggregate we must view them as a global economy that relies on natural capital which transcends political boundaries. To guide the development of sustainable economies, we need economic models that properly reflect the physical and ecological basis of economic activity, and for the important feedbacks between the economy and the environment. Specifically, they must embody realistic limits to the degree to which we can substitute human-made capital for natural capital, and
account for critical role of ecosystem services as well as marketed natural resources. This is not a pessimistic assessment because I assume considerable potential for substitution between human and manufactured capital and among various forms of natural capital. However, some ecosystem services are irreplaceable, and complementarity among the factors of production effectively limits substitutions. Furthermore, the potential for nations to use foreign natural capital to sustain growth diminishes as rising populations and incomes deplete more global natural capital.
References


