Population and Development Review, Hoover Institution Conference

SUSTAINABLE DEVELOPMENT : FROM CONCEPT AND THEORY TOWARDS OPERATIONAL PRINCIPLES[1]

by
Herman E. Daly
Alumni Professor of Economics
Louisiana State University
and
Senior Economist
Environmental Division
World Bank

I. Introduction. Three conceptual issues seem to me critical for clear thinking about economic development and the environment in the next decade. I will begin by stating the three issues briefly, and then discuss each one in turn, along with the relations among them.

(1) The first issue is whether the basic conceptual starting point of economic analysis is to be the circular flow of exchange value, as it presently is, or the one-way entropic throughput of matter-energy. The latter concept is virtually absent from economics today, yet without it it is impossible to relate the economy to the environment. It is as if biology tried to understand animals only in terms of their circulatory system, with no recognition of the

---

[1]The views expressed in this article are those of the author and should in no way be attributed to the World Bank or any other institution. I am indebted for helpful comments to S. Davis, S. El Serafy, P. Ehrlich, R. Goodland, P. Knight, R. Overby, and the editors both of this journal and this special issue. Earlier versions of parts of this paper were presented at a conference on development and environment, sponsored by the Italian Ministry of the Environment, Milan, March, 1988.

fact that they also have digestive tracts. The metabolic flow is not circular. The digestive tract firmly ties the animal to its environment at both ends. Without digestive tracts animals would be self-contained perpetual motion machines. Likewise for an economy without an entropic throughput.

(2) The throughput has two dimensions of importance: its scale and its allocation. The concept of optimal allocation among alternative uses of the total resource flow (throughput) must be clearly distinguished from the concept of an optimal scale of total resource flow relative to the environment. Under ideal conditions the market can find an optimal allocation in the sense of Pareto. But the market cannot find an optimal scale any more than it can find an optimal distribution. The latter requires the addition of ethical criteria; the former requires the further addition of ecological criteria. The independence of allocation from distribution is widely recognized; the independence of allocation from scale is not widely recognized, but is easily understood.[1] In theory we can double the population and the per capita resource use rate, or cut them in half, and the market will still grind out a Pareto optimal allocation for every scale. Yet the scale of the economy is certainly not a matter of indifference. A boat that tries to carry too much weight will still sink even if that weight is optimally

---

[1] An example of the consequences of non-recognition of the distinction between scale and allocation can be found in H. E. Daly, "Review of National Research Council, Population Growth and Economic Development : Policy Questions", in Population and Development Review, Vol. 12, No. 3, September, 1986 :pp. 582-585. A similar distinction is made in D.W. Pearce, "Foundations of an Ecological Economics", Ecological Modelling, 38, 1987, in which he develops the idea of an ecologically bounded economy and argues that sustainability cannot be derived from the market mechanism.

allocated. Allocation is one thing, scale is something else. We must deal with both, lest even the efficiently allocated weight of the economy sink the environment. We need something like a Plimsoll line to keep the economic scale within the ecological carrying capacity. Economics has tried to reduce scale issues to matters of allocation (just get the prices right), and has thereby greatly obscured the relation between the economy and the environment. While an optimal allocation can result from the individualistic marketplace, the attainment of an optimal scale will require collective action by the community.

(3) The third issue, sustainable development, is already under intense discussion. Following the Brundtland Commission Report (Our Common Future) the concept has been endorsed by the United Nations and all its many development agencies, and urged upon all member countries. While this discussion is greatly to be welcomed it nevertheless suffers from considerable confusion. Underlying this confusion is the unresolved, indeed unaddressed, issue of steady state versus growth as the normal, healthy condition of an economy. Our growth-bound way of thinking makes it hard for us to admit the concept of throughput of matter-energy, because it brings with it the first and second laws of thermodynamics, which have implications that are unfriendly to the continuous growth ideology. The circular flow raises no such problems. The growth ideology is extremely attractive politically because it offers a solution to poverty without requiring the moral disciplines of sharing and population control. Also the obvious implication of recognizing an optimal scale is that growth should stop once the optimum is reached--that growth beyond the optimum scale is "anti-economic growth"--i.e. growth that makes us poorer rather than richer. Optimal allocation has no such growth-limiting implications.

The three issues I am raising are not difficult , arcane, or esoteric--they are no more than common sense--but it is hard for us to think clearly about them because doing so threatens the absolute priority of growth as the North Star of economic policy. Although the three issues are separable they are also related. Once throughput is recognized as a fundamental and indispensible concept, then the question of its optimal scale within a finite ecosystem naturally arises, along with the recognition that the question is different from that of optimal allocation. Once we face up to the question of limiting scale we recognize the collective or social nature of the task and the futility of leaving it up to the individualism of the market which can only deal with allocation. We are also faced with the problem of criteria for optimal scale, and the most obvious one is sustainability. The discussion of sustainable development will not get far without the recognition of throughput and the problem of its scale.

Much confusion could be avoided if we would agree to use the word "growth" to refer only to the quantitative scale of the physical dimensions of the economy. Qualitative improvement could be labeled "development".[1] Then we could speak of a steady-state economy as one which develops without growing, just as the planet earth, of which the economy is an open subsystem, develops without growing. Growth of the economic organism

---

[1]This distinction is not the result of any idiosyncratic redefinition. It is explicit in the dictionary's first definition of each term. To grow means literally, "to increase naturally in size by the addition of material through assimilation or accretion". To develop means, "to expand or realize the potentialities of; bring gradually to a fuller, greater, or better state." (The American Heritage Dictionary of the English Language)

means larger jaws and a bigger digestive tract. Development means more complete digestion and wiser purposes. Limits to growth do not imply limits to development.

II. Entropic Throughput of Matter-Energy Versus the Circular Flow of Exchange Value. Georgescu-Roegen[1] has pointed to *"the standard textbook representation of the economic process by a circular diagram, a pendulum movement between production and consumption within a completely closed system"* as proof of the mechanistic epistemology of modern economics. There is only reversible motion, a circular flow, and no recognition of irreversible entropic change. There is only mechanical time, no historical time. This does not mean that economists deny historical time or the entropy law--but it does mean that they cannot deal with them at the most basic conceptual level of economics, and have to introduce them in ad hoc and unintegrated ways outside the structure of formal models.--i.e., externalities

In addition to the inability of the mechanistic epistemology to embrace irreversible phenomena, there was a practical reason for ignoring the entropic throughput. Economists are interested in scarcity, and during the formative years of economic theory the environment was considered an infinite source of raw materials and an infinite sink for waste materials. Therefore the throughput was not considered scarce and was naturally abstracted from. Only scarce items entered into exchange. Free goods were appropriated without need of a transaction. Since exchange value flowed in a

---

[1]Nicholas Georgescu-Roegen, The Entropy Law and the Economic Process, Harvard University Press, Cambridge, 1971.

circle, the circular flow became the paradigm within which we sought to understand the economic process. Once the economy reached the scale at which throughput itself became scarce, then the circular flow vision became economically , as well as physically, misleading. It totally obscured the emerging scarcity of environmental services. The circular flow has no beginning and no end, no points of contact with anything outside itself. Therefore it cannot possibly register the costs of depletion and pollution, nor the irreversible historical effects induced by the entropic nature of the throughput.

The concept of throughput was introduced into economics by Kenneth Boulding[1] , and more fully elaborated and integrated into economic theory by Georgescu-Roegen (op. cit.), who called it the "metabolic flow" and emphasized the manifold consequences of its entropic nature. Others (Kneese, Ayres and d'Arge[2]) have paid respect to the concept by way of emphasizing the importance of "material balances", thus recognizing the constraint on the economic process of the first law of thermodynamics, but neglecting that of the second law. The first law is consistent with the circular flow vision--the same indestructible building blocks of matter-energy could simply cycle faster and faster around the production-consumption loop. Nothing gets used up. But the second law says that something does indeed get used up--not matter-energy itself, but its capacity for rearrangement.

---

[1]Kenneth Boulding, "The Economics of the Coming Spaceship Earth" in Henry Jarrett, ed., Environmental Quality in a Growing Economy, Baltimore: Johns Hopkins University Press, 1966.

[2]A. V. Kneese, R.V. Ayres, and R.C. d'Arge, Economics and the Environment: A Materials Balance Approach, Washington,D.C., 1970.

Energy is conserved, but its capacity to do work is used up. To my knowledge no economics textbook has paid attention to any of these important contributions. Instead they continue to perpetuate the circular flow vision without so much as a reference to the concept of throughput.[1,2] Naturally, if the very concept of throughput is not admitted it will be impossible to consider the issue of its optimal scale, a theme to which we now turn.[3]

III. Optimal Allocation Versus Optimal Scale. Standard economics is about the optimal allocation of resources, which in this broad sense includes labor and capital as well as natural resources. But natural resources are not viewed as the components of an entropic metabolic flow from and back to the environment. Rather they are seen as building blocks which are indestructible elements in the circular flow. Allocation of these elements among competing uses is the only question raised for standard economics by its partial recognition of throughput. As mentioned earlier a Pareto optimal allocation can be achieved for any scale of population and per capita resource use. The concept of economic efficiency is indifferent to the scale of the economy's physical dimensions, just as it is indifferent to the

---

[1]Some explicitly announce in bold print that" the flow of output is circular, self-renewing, and self-feeding."See Robert Heilbroner and Lester Thurow, The Economic Problem, New York, 1981.

[2] A recent treatise by Charles Perrings, (Economy and Environment, Cambridge University Press, 1987) is an important theoretical contribution toward integrating the concept of throughput and the laws of thermodynamics with standard economics.

[3]H.E. Daly,"The Circular Flow of Exchange Value and the Linear Throughput of Matter-Energy : A Case of Misplaced Concreteness", Review of Social Economy, December,1985.

distribution of income. Equity of income distribution and sustainability of scale are outside the concept of market efficiency. Yet the environment is sensitive to the physical scale of the economy, and human welfare is sensitive to how well the environment functions. To optimally allocate resources at a non-optimal scale is simply to make the best of a bad situation. If the economy continues to grow beyond optimal scale, then optimal allocation means simply to keep making the best of an ever worsening situation. This anomaly is absent from the circular flow vision: if the economy is an isolated system with no dependence on its environment then naturally it can never exceed the capacity of the environment. Its scale relative to the environment is a matter of complete indifference. But once we recognize the central importance of the throughput, then we must concern ourselves with its optimal scale as well as its optimal allocation.

Optimal scale of a single activity is not a strange concept to economists. Indeed microeconomics is about little else. An activity is identified, be it producing shoes or consuming ice cream. A cost function and a benefit function for the activity in question are defined. Good reasons are given for believing that marginal costs increase and marginal benefits decline as the scale of the activity grows. The message of microeconomics is to expand the scale of the activity in question up to the point where marginal costs equal marginal benefits, a condition which defines the optimum scale. All of microeconomics is an extended variation on this theme.

When we move to macroeconomics, however, we never again hear about optimal scale. There is no optimal scale for the macro economy. There are no cost and benefit functions defined for growth in scale of the economy as a whole. It just doesn't matter how many people there are, or how much they each consume. But if every micro activity has an optimal scale then why does not the aggregate of all micro activities have an optimal scale? If I am told in reply that the reason is that the constraint on any one activity is the fixity of all the others and that when all economic activities increase proportionally the restraints cancel out, then I will invite the economist to increase the scale of the carbon cycle and the hydrologic cycle in proportion to the growth of industry and agriculture. I will admit that if the ecosystem can grow indefinitely then so can the aggregate economy. But until the the surface of the earth begins to grow at a rate equal to the rate of interest one should not take this answer too seriously. The total absence in macroeconomics of the most basic concept of microeconomics is a glittering anomaly, and it is not resolved by appeals to the fallacy of composition. What is true of a part is not necessarily true for the whole, but it can be and usually is unless there is some aggregate identity or self-cancelling feedback at work. (As in the classic examples of all spectators standing on tiptoe to get a better view and each cancelling out the better view of the other; or in the observation that while any single country's exports can be greater than its imports, nevertheless the aggregate of all exports cannot be different than the aggregate of all imports). But what analogous feedback or identity is there that allows every economic activity to have an optimal scale while the aggregate economy remains indifferent to scale?

In the circular flow vision there is an aggregate identity--total expenditures equal total receipts, one person's expenditure is another person's income. Costs and benefits are conflated in transactions. In circular flow accounting we add up transactions rather than compare costs and benefits at the margin, so the question of an optimal scale of the circular flow never arises. It is the throughput that has an optimal scale. When growth pushes scale beyond the optimum we begin to experience generalized pervasive externalities, such as the greenhouse effect, ozone layer depletion, and acid rain, which are not correctable by internalization of localized external costs into a specific price.

Probably the best index of the scale of the human economy as a part of the biosphere is the percentage of human appropriation of the the total world products of photosynthesis. Net primary production (NPP) is the amount of solar energy captured in photosynthesis by primary producers, less the energy used in their own growth and reproduction. NPP is thus the basic food resource for everything on earth not capable of photosynthesis. Vitousek, et al.[1] calculate that 25% of potential global (terrestrial and aquatic) NPP is now appropriated by human beings. If only terrestrial NPP is considered the fraction rises to 40%. Taking the 25% figure for the entire world it is apparent that two more doublings of the human scale will give 100%. Since this would mean zero energy left for all nonhuman and nondomesticated species, and since humans cannot survive without the services of ecosystems, which are made up of other species, it is clear that two more doublings of the human scale is an ecological impossibility, although arithmetically possible. More than two doublings is even arithmetically impossible! Furthermore the terrestrial figure of 40% is

---

[1]Peter M. Vitousek, Paul R. Ehrlich, Anne H. Ehrlich, and Pamela A. Matson, "Human Appropriation of the Products of Photosynthesis" BioScience, Vol. 34, No. 6, May, 1986, pp. 368-373. The definition of human appropriation underlying the figures quoted includes direct use by human beings (food, fuel, fiber, timber), plus the reduction from the potential due to degradation of ecosystems caused by humans. The latter reflects deforestation, desertification, paving over, and human conversion to less productive systems (such as agriculture).

probably more relevant since we are unlikely to increase our take from the oceans very much. Total appropriation of the terrestrial NPP is only a bit over one doubling time in the future. Perhaps it is theoretically possible to increase the earth's total photosynthetic capacity somewhat, but the actual trend of past economic growth is decidedly in the opposite direction.

Assuming a constant level of per capita resource consumption the doubling time of the human scale would be equal to the doubling time of population, which is on the order of 40 years. Of course economic growth currently aims to increase the average per capita resource consumption and consequently to reduce the doubling time of the scale of the human presence below that implicit in the demographic rate of growth. Unless we awaken to the existence and nearness of scale limits then the greenhouse effect, ozone layer depletion, and acid rain will be just a preview of disasters to come, not in the vague distant future, but in the next generation.[1]

---

[1] In view of the seriousness of the situation it is distressing to find economists of the caliber of W. J. Baumol still writing complacent articles, "On the Possibility of the Continued Expansion of Finite Resources" (Kyklos, , May, 1986). Baumol flatly states that "Technological progress makes it feasible both absolutely and in economic terms to obtain ever increasing amounts of a usable resource from a given source (such as an oil well)." (p. 170) One suspects that such nonsense resulted from careless editing, because the point seems to be only partially reaffirmed in Baumol's conclusion that ".....measured in terms of their prospective contributions to human welfare , the available quantity of the world's exhaustible resources, may rise forever, year after year. However, even though they may never approach disappearance, the consumption of their services will eventually have to decline and, ultimately, approach zero asymptotically."(p. 178) To the extent that (contrary to the first quotation, but implicit

As growth increasingly turns previously free goods into scarce goods the standard solution is to put positive prices on the newly scarce goods. Once a good has become scarce it is important that it have a positive price in order to be properly allocated. But there is a prior question: How do we know that we were not better off at the previous scale when the good was free and its proper price was zero? In both instances the prices were right. But that does not mean that the scale was right. Furthermore, the new exchange value created when previously free goods become scarce reflects a cost, not a benefit as currently reckoned. The classical economist Lauderdale recognized that private riches could expand while public wealth declined. This perversity will occur whenever formerly abundant objects with great use value, but no exchange value, become scarce and thus acquire exchange value. Although scarcity is necessary for value in the sense of measurable

---

in the second) Baumol's argument is based on the possibility of increasing the efficiency of resource use rather than the amount of resources used (development instead of growth) then the issue is simply one of expectations about how far efficiency can increase. If Baumol's optimism on this score is correct that will make it less painful to limit scale, but will in no way make it possible for scale to continue growing. Indeed, in Baumol's model scale is absent: there is no consideration of population, and resource extraction approaches zero (in contradiction to the first quotation) even though the contribution of resources to human welfare rises forever! While the article makes many good points about efficiency and substitution, I confess that the vision of infinitesimal rates of depletion of resources that have become infinitely productive strikes me as mathematical fun and games with infinity, rather than serious economics.

exchange value, " the common sense of mankind would revolt at a proposal for augmenting wealth by creating a scarcity of any good generally useful and necessary to man."[1] The revolt has been slow in coming, but let us hope that Lauderdale was right!

Some economists argue that futures markets and present value maximization automatically deals with the scale issue because the costs of excessive scale are merely the future costs of the present use of resources. But even in a single period analysis in which there is no future, there is still the possibility of having exceeded optimal scale in the sense of sacrificing current ecosystem services that are worth more than the current extra economic product whose production required the sacrifice of those services. It is true that many of the costs of increasing scale do fall on the future. But neither present value maximization nor an imaginary futures market is adequate for taking acount of these future costs.[2,3]

---

[1] Lauderdale, An Inquiry into the Nature and Origin of Public Wealth and into the Means and Causes of its Increase, second edition, Edinburgh: Archibald Constant and Co., 1819 p.44. I am indebted to Mr. George Foy for this reference.

[2] Talbot Page, Conservation and Economic Efficiency, (Baltimore: Johns Hopkins University Press, 1977).

[3] Present value maximization attempts to allocate resources efficiently over time. But once intergenerational time periods are encountered we escape the domain of allocation and must speak instead of distribution. Different generations are different people. Dividing the resource base among different people is distribution; dividing it among different uses for the same group of people is allocation. The former is a matter of justice, the latter of efficiency. Present value maximization (discounting) over intergenerational time conflates allocation and distribution.

Optimal allocation at least has a definition, however restrictive and limited in relevance it may be. But how do we define optimal scale? This is an enormous question that involves not only much greater knowledge of carrying capacity and ecological relations, but also much clarification and deeper understanding of our own purposes. Many economists keep the scale question out of sight by rejecting the concept of carrying capacity on the grounds that it is not clearly defined. But by that criterion they should also refuse to talk about "time", one of the most difficult concepts of all to define. Some say it is absolute, others say it is relative, still others insist that it is pure illusion. Even "money" should not be spoken of, since what is really money, M1 or M2? Or M1A? One of the temptations of debate is to demand an unreasonable standard of precision for concepts that have troublesome implications for one's position, while being more informal and relaxed in the company of concepts known not to raise impolite questions. But there is one thing we know about the optimum scale: it must at least be sustainable. So for the time being we can devote our practical policies toward sustainability, while we puzzle over the deeper philosophical issues of optimal scale.

One further criterion for optimal scale suggested indirectly by Charles Perrings[1] is that the economy be small enough to avoid generating feedbacks from the ecosystem that are so novel and surprising as to render economic calculation impossible. Perrings begins with the first law of thermodynamics, pointing out that an increasing throughput (extractions and insertions in his language) provokes ever greater feedbacks from the environment (external

[1] Op. cit.

costs) as the scale of extractions and insertions grows. Since we do not understand the ecosystem very well the feedbacks from it provoked by our actions come as surprises to us. These surprises are nearly always unpleasant ones, since random interferences in a complex system nearly always disrupt the functioning of the system, and since our welfare depends on the proper functioning of that system. Novelty and surprise begin to outweigh the calculated projections of the costs and benefits of the increasing scale of our activities. We can react to this loss of predictability in two ways: (1) increase the sphere of control so as to internalize the "surprises" (Boulding's image of the spaceman economy in which the entire life support system is planned and controled,--i.e. everything is economy and nothing is environment, leading to what we might call "full-world economics") ; or (2) decrease the scale of human interference to a level such that the ecosystem can function on "automatic pilot" (Boulding's image of the cowboy economy in which nearly everything is environment, and sinks are automatically recycled into sources without any planning by the cowboy-- "empty-world economics"). Our ability to centrally plan economies does not inspire optimism about the likelihood of our success in the vastly more difficult task of planning the ecosystem. One of the main criteria for an optimal scale, therefore, is that the economy be small enough to avoid unmanageable interference with the "ecological invisible hand" or automatic pilot. Ecological laissez faire requires social control of the scale of the economic subsystem. Refusal to limit scale because of an overextended belief in providentialistic individualism (Adam Smith's invisible hand) will lead to a situation in which even constrained individualism becomes impossible because we will, like the spaceman, be faced with the burden of planning and regulating our entire life support system. A full-world economy will not have enough slack between the carrying capacity of its supporting ecosystem and its actual load to permit the luxury of free-market trial and error. The world now is too full for the empty-world economics of laissez faire. The only way to retain some of the freedoms of the empty-world economy is to control scale.

We could of course simply accept the eventual destruction of life support capacity as the price we must pay for freedom from restriction of individual rights to grow. It is widely believed, however, by persons of diverse religious commitment, that there is something fundamentally wrong in treating the earth as if it were a business in liquidation. The value of sustainability is so basic that it is usually tacitly assumed in our economic thinking. It should come as no surprise (but often does) to learn that sustainability is built into the very concept of income. J. R. Hicks[1] defined income as the maximum amount that a person or a nation could consume over some time period and still be as well off at the end of the period as at the beginning. Hicks further argued that the whole practical reason for calculating income is to have a guide to how much we can consume year after year without eventually impoverishing ourselves. Income equals maximum sustainable consumption.

IV. Sustainable Development

Lack of a precise definition of the term "sustainable development" is not all bad. It has allowed a considerable consensus to evolve in support of the main idea that it is both morally and economically wrong to treat the world as a business in liquidation. If development is to be the major policy

---

[1] J. R. Hicks, Value and Capital, p.172.

goal of nations then it should mean something that is generalizable both to all members of the present generation and to many future generations. The popularity of the notion of sustainable development derives from the increasing recognition that present paterns of economic development are not generalizable. Present levels of per capita resource consumption underlying the economies of the US and Western Europe (which is what is generally understood by development) cannot be generalized to all currently living people, much less to future generations, without destroying the ecological sources and sinks on which economic activity depends.[1] The Brundtland

---

[1]This lack of generalizability can be seen from the following back-of-the-envelope calculation, based on the crude estimate that the US currently uses 1/3 of annual world resource flows (derived from Material Needs and the Environment Today and Tomorrow, National Commission on Materials Policy, 1973, Washington, D.C.: U.S. Government Printing Office). Let R be current world resource consumption. Then R/3 is current US resource consumption, and R/3 divided by 240 million is present per capita US resource consumption. Current world per capita resource consumption would be R divided by 5 billion. For future world per capita resource consumption to equal present US per capita consumption, assuming constant population, requires that R increase by some multiple, call it M. Then M times R divided by 5 billion must equal R/3 divided by 240 million. Solving for M gives 7. World resource flows must increase seven-fold if all people are to consume resources at the present US average. Current resource use rates are provoking ecological feedbacks such as the greenhouse efect and ozone depletion. Imagine the consequences of a seven-fold increase!

But even the seven-fold increase is a gross underestimate of the increase in environmental impact, for two reasons. First, because the calculation is in terms of current flows only with no allowance for the increase in accumulated stocks of capital goods necessary to process and transform the greater flow of resources into final products. Some notion of the magnitude of the extra stocks needed comes from Harrison Brown's estimate that the "standing crop" of industrial metals already embodied in the existing stock of artifacts in the ten richest nations would require more that 60 years' production of these metals at 1970 rates. Second, because the seven-fold increase of net, usable minerals and energy will require a much greater increase in gross resource flows, since we must mine ever less accessible deposits and lower grade ores. It is the gross flow that provokes environmental impact.

Commission Report (Our Common Future) was wise not to forclose the emergence of this vague but important consensus by insisting on a precise analytical definition from the outset.

But the term is now in danger of becoming an empty shibboleth. For example, many people in the development community who use the term cannot tell you what is being sustained in sustainable development--whether a level of economic activity or a rate of growth of economic activity! Some, therefore, have become impatient with the concept and want to abandon it. That would be a great mistake. After all we do not have a precise definition of money either (M1, M2....M4), but we certainly cannot abandon the concept. Nor is income a precise concept , yet in practical affairs we can hardly do without it. Even though we must not expect analytical precision in reasoning with dialectical concepts, it is nevertheless possible and very necessary to clarify the notion of sustainability, and to offer a few first principles of sustainable development.

Two terms are frequently used more or less synonymously: "sustainable growth" and "sustainable development". Earlier I suggested the following distinction: that "growth" refer to expansion in the scale of the physical dimensions of the economic system, while "development" refer to qualitative change (improvement or degradation) of a physically nongrowing economic system in a state of dynamic equilibrium maintained by its environment. By this definition the earth is not growing, but it is developing. Any physical subsystem of a finite and nongrowing earth must itself also eventually become nongrowing. Therefore the term sustainable growth implies an eventual impossibility, while the term sustainable development does not. It is development that can have the attribute of sustainability, not growth. What is being sustained in sustainable development is a level, not a rate of growth, of physical resource use. What is being developed is the qualitative capacity to convert that constant level of physical resource use into improved services for satisfying human wants.

The concept of sustainability is by no means new in economics, although the word is. As noted earlier, sustainability is implicit in J. R. Hicks' definition of income as the maximum amount that a person or community could consume over some time period and still be as well off at the end of the period as at the beginning. Remaining equally well off means maintaining capital intact, or maintaining the wealth and population of the community. Growth in Hicksian income is by definition sustainable. Any consumption that is not sustainable cannot be counted as income. Exploiting renewable resources at a profit-maximizing sustainable yield is an application of the Hicksian concept of income to resource management.

How, then, can there be a problem of lack of sustainability if that notion is implicit in the very concept of income? The problem is that the capital that we have endeavored to maintain intact is manmade capital only. There is also the important but relatively unappreciated category of natural capital--natural stocks that yield flows of natural resources and services without which there can be no production.[1] In practice we do not maintain

---

[1] Natural capital may be divided into marketed and nonmarketed natural capital. The former yields the flow of priced natural resources; the latter yields the flux of unpriced natural life-support services. The term "natural capital" is a bit awkward because capital has traditionally been defined as produced (manmade) means of production. The term "land" in earlier times meant something equivalent to natural capital, but has now lost that meaning. The term natural capital is used to call attention to the fact that there is a <u>stock</u> of natural assets that yields a <u>flow</u> of resources and services and that require

natural capital constant in the process of production, and consequently the NNP generated is not Hicksian income. The present System of National Accounts treats receipts from liquidating natural assets as income, thus giving countries the illusion that they are beter off than they really are.

Why has natural capital been left out of our accounts? There are two main reasons.

(1) The scale of the economy (population times per capita resource use) relative to the environment used to be negligible, and consequently natural capital regeneration was either automatic or perceived as unimportant because it was not a limiting factor. Between 1950 and 1986 the scale of the world population doubled (from 2.5 to 5.0 billion), while the scale of gross world product and fossil fuel consumption each quadrupled. The physical presence of the economy within the ecosystem was not negligible even in 1950 and is certainly not now. The humanly directed flows of matter and energy through the economy rival in magnitude the flow rates of many natural cycles and fluxes. As previously noted human beings now appropriate 40% of terrestrial net primary productivity. In the past the limitative factor in economic development was the accumulation of manmade capital. We are now entering an era in which the limitative factor will be the remaining natural capital. The notion of limitative factor implies less than perfect substitutibility between factors,--i.e,. that factors are to some extent complementary. This leads us to the second reason why natural capital has been neglected.

---

maintenance in the face of depreciation, and whose consumption cannot be counted as income.

(2) Neoclassical economic theory has taught that manmade capital is a near perfect substitute for natural resources, and consequently for the stock of natural capital that yields the flow of these natural resources. Even if this assumed near perfect substitutability were true, it would still be necessary to maintain total capital (manmade plus natural) intact in calculating Hicksian income--i.e., the running down of natural capital would still have to be offset by the accumulation of an equivalent value of manmade capital. Even this is not done. Moreover, substitution in economic theory is reversible, while the substitution of manmade for natural capital is frequently irreversible. Contrary to neoclassical assumptions, natural and manmade capital are more complements than substitutes, with natural capital increasingly replacing manmade as the limitative factor in development.

Maintaining total capital intact might be referred to as "weak sustainability", in that it is based on generous assumptions about substitutability of capital for natural resources in production. By contrast "strong sustainability" would require maintaing both manmade and natural capital intact separately, on the assumption that they are really not substitutes but complements in most production functions. For example the manmade capital represented by a sawmill is worthless without the existence of the complementary natural capital of a forest. In the strong sustainability case economic growth would require the increase of whichever type of capital is limitative at the margin. At the current margin in many countries natural capital is limitative for sustainable development, yet is routinely sacrificed for more manmade capital under the prevailing model of unsustainable development based on national accounts that treat consumption of natural capital as income.

We might distinguish a third concept of sustainability that I would label "very weak sustainability". Some authors define sustainability as the maintenance of a constant level of utility.[1] What is being sustained is now a psychic state rather than a physical state. This subjectivist definition incorporates psychological substitution in the utility function as well technological substitution in the production function. In this view we can learn to enjoy the services of manmade capital more relative to the services of natural capital and remain equally happy as the former is continually substituted for the latter (the"Disneyland effect"). The appeal to economists of the subjectivist view is that it allows sustainability so defined to fit directly into the discounted utility maximizing theoretical framework of neoclassical economics. The overwhelming operational disadvantage is that it defines one imprecise concept (sustainability) in terms of something even less definable --utility, nay, discounted, future, aggregated utility! [2] It is

---

[1]See for example John Pezzy, "Economic Analysis of Sustainable Growth and Sustainable Development",World Bank, Environmental Department Working Paper No. 15, March, 1989. Pezzy is well aware of the difficulties discussed above, but set for himself the task of analyzing sustainability from within the neoclassical paradigm. Although I think it is fair to say that Pezzy aimed to demonstrate the usefulness of the neoclassical aproach, the basic honesty of his scholarship resulted in what is to my mind a demonstration of its severe limits.

[2]Discounting is an operational concept when applied to money in the bank that is growing at a rate of interest. By extention, it can, within limits, be applied to trees in a forest or fish in a pond, as long as we remember that there is a limit to how many trees there can be in the forest and how many fish there can be in the pond--while there is no limit to how much money there can be in the bank. But to discount utility, a psychic

better to aim at something more operational by sticking with the physical approach of the ecologist. It is impossible for the present to bequeath happiness or utility to the future. The only thing that can be passed on is natural and manmade capital (also knowledge, although that has to be taught and learned, not just bequeathed). The physical approach can provide a definition of sustainability that can be imposed as a constraint on the maximization of utility in neoclassical models. Sustainability should not refer to a psychic state, but rather to a state of the biophysical world, namely a condition of dynamic equilibrium between the physical dimensions of the economy and the larger environment of which it is an open subsystem. In this view the major determinant of sustainability is likely to be the physical scale of the economic subsystem relative to the containing ecosystem.

An operational approach to sustainability that does not hinge on resolution of the substitutability question is to adjust national accounts so as to arrive at a closer approximation of Hicksian income than that given by NNP. One way to do this is to subtract from NNP two categories of expenditure that measure nonsustainable activities. First subtract an estimate of the value of natural capital depreciation. Second, subtract an

---

experience that cannot be accumulated or saved, and which has no natural tendency to grow in any case, is to commit Whitehead's fallacy of misplaced concreteness. Furthermore, to aggregate this future psychic experience across individuals before discounting it by a nonexistent natural growth rate is to lose touch completely with any possibility of a real world counterpart to the paper-and-pencil operation.

estimate of defensive or regrettably necessary expenditures made to protect ourselves against the unwanted side effects of other production.[1]

The main idea of Hicksian income is captured in the definition of sustainable development offered by the Brundtland Commission[2] as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs". Two questions will arise in any attempt to make this definition operational. First, the "needs of the present" requires some distinction between basic needs and extravagant wants. If needs of the present include an automobile for each of a billion Chinese then sustainable development is impossible. Sustainable development is about sufficiency as well as efficiency. Second, the "ability of the future generations to meet their own needs" may be interpreted as requiring either strong or weak sustainability--that pervasive issue of substitutability surfaces again.

At what level of community is sustainability to be sought? [3]International trade allows one country to draw on the ecological carrying capacity of another country, yet both together might be sustainable in their

---

[1]This and other ways of adjusting the national accounts are discussed in Environment and Resource Accounting and their Relevance to the Measurement of Sustainable Income, Ernst Lutz and Salah El Serafy, eds., Washington, D.C.: World Bank, 1989).

[2] World Commission on Environment and Development, Our Common Future. (Great Britain: Oxford University Press, 1987, p.8)

[3] For a discussion of community, trade, and ecological issues, see Chapter 12 in Herman E. Daly and John B. Cobb, Jr., For the Common Good: Restructuring the Economy Toward Community, the Environment, and a Sustainable Future, Boston: Beacon Press, 1989 (forthcoming).

symbiotic relationship. How does trade affect sustainabiliy defined at the national level? This brings us again to the question of complementarity versus substutability of natural and manmade capital. If we follow the path of strong sustainability (natural and manmade capital are more complements than substitutes) then this complementarity must be respected either at the national or at the international level. A single country may substitute manmade for natural capital to a very high degree if it can import the products of natural capital (i.e. the flows of natural resources and natural services) from other countries which have retained their natural capital to a greater degree. In other words, the demands of complementarity can be evaded at the national level, but only if they are respected at the international level. One country's ability to substitute manmade for natural capital depends on some other country's making the opposite (complementary) choice.

There are strong theoretical and commonsense reasons for believing that natural and manmade capital are complements. Natural resource stocks yield a flow of natural resource inputs that is physically transformed by stocks of manmade capital and labor into a flow of product outputs. There may be a great deal of substitutability between labor and manmade capital (the two agents of transformation), or between the various resource flows (that which is being transformed). But the main relation between that which is being transformed and the agent of transformation must be one of complementarity, not substitutability. Otherwise we could build the same wooden house with, say, half the lumber and twice as many saws and carpenters. Of course one could substitute brick or fiberglass for lumber, but that is the substitution of one resource flow for another, not the substitution of manmade capital stock for a natural resource flow. The agent of transformation (efficient cause) and the substance being transformed by it (material cause) must be complements.[1] Also we should not forget the obvious fact that production of capital itself requires natural resources--the production of the "substitute" requires the very thing being substituted for! For these reasons strong sustainability is the fundamental concept. Weak sustainability is an option for a single country only in the context of a set of trading countries which taken together meet the conditions of strong sustainability. Consequently we must distinguish closed from open economy concepts of carrying capacity. In the former the nation must draw only on its own ecosystem for everything. In the latter drawing on other ecosystems and economies is permitted as long as imports are paid for by current exports. Subsidies and continuing unpaid debts are excluded, though international trade is permitted.

Sustainable development ultimately implies a stationary population. Penultimately, however, there remain possibilities of substitution between population size and resource use per capita, since it is really the product of these two factors that is limited by biophysical constraints. Sustainability is compatible with a large population living at a low level of per capita resource use, or with a small population and high levels of resource use per capita. For many countries resource consumption levels are below sufficiency, yet ecological carrying capacity has already been exceeded (e. g. Haiti, El Salvador). In such cases population control is a precondition rather than an ultimate consequence of sustainable development.

---

[1]See N. Georgescu-Roegen, The Entropy Law and the Economic Process, Cambridge: Harvard University Press, 1971, pp.224-244.

Sustainable development does not imply constant technology, nor is the concept rendered unnessary by technological progress. New technology can have positive or negative effects. Technologies that increase resource productivity will reduce the pressure on natural capital stocks to yield increasing flows of natural resources. New technology which increases the productivity of manmade capital and labor frequently requires processing a greater flow of resources, and thus a tendency to reduce resource productivity in the interests of raising capital and labor productivity. Historically technological progress has favored capital and labor productivity at the expense of resource productivty (e.g., declining energy productivity in agriculture resulting from greater use of energy per unit of labor and capital, with consequent increase in labor and capital productivities). Sustainable development implies a different direction of technical progress, one that squeezes more service per unit of resource, rather than one that just runs more resources through the system--- one that is efficiency-increasing rather than throughput-increasing---one that does not sacrifice natural resource productivity, and if necessary will sacrifice labor or capital productivity instead. This will be politically difficult due to the tie between marginal productivity and income for all factors. In our society labor and capital are much stronger social classes than are landlords (resource owners). Naturally each class prefers those technologies that increase its own marginal productivity and income. In earlier times the landlord was dominant and ~~and~~ preferred labor-intensive technologies (and large populations) that increased the marginal product, and rent, of land. Nowadays the political demise of the landlord has left land (resources) without a social class to champion its higher price and productivity.

Resources tend to be used lavishly in the interests of labor and capital productivity. This works against sustainability.

The most obvious principle of sustainable development is that renewable resources should be exploited on a sustained yield basis. The choice among many sustained yield levels can be made on the criterion of profit maximization. In general the profit-maximizing stock level at which to maintain the exploited population will not be that corresponding to the biologically maximum sustainable yield. For wild populations it will be greater (assuming rising costs of capture); for cultivated populations it will be less (assuming rising costs of maintenance). Only with constant costs will the biological yield maximum coincide with the economic profit maximum.

A major problem for sustainable development is how to treat nonrenewable resources which by definition have no sustainable yield, at least on time scales relevant to human experience. A way of handling this problem is suggested in an ingenious paper by Salah El Serafy[1]. He shows how to divide net receipts from a nonrenewable resource into an income component that can be consumed each year, and a capital component that must be invested each year in a renewable asset that yields a rate of return such that, at the end of the lifetime of the nonrenewable resource (reserves divided by rate of depletion), a new renewable asset will have been built up to the point at which it can yield a perpetual stream equal to the income

---

[1]El Serafy, Salah (1989) "The Proper Calculation of income from Depletable Natural Resources," in Environment and Resource Accounting and their Relevance to the Measurement of Sustainable Income, Ernst Lutz and Salah El Serafy, eds., Washington, D.C. : World Bank, 1989).

component of the now depleted nonrenewable resource. A somewhat similar principle was suggested by economist John Ise back in the 1920's, namely to use up the nonrenewable resource at a rate such that its price will be equal to the price of its nearest renewable substitute. In other words, resources should be priced according to their long run replacement costs. El Serafy's rule is more oriented to the operational problems of proper income accounting rather than pricing, and does not require the identification of a specific long run renewable substitute. It does implicity assume, however, that there is something useful in the real world that is capable of growing at a rate equal to the rate of discount used in calculating the income component. It would seem that that something must be some renewable resource service, that is, the biological growth rate of renewable resources plus the technological rate of growth of the productivity of all resources (not the productivity of labor or capital, but of resources). Also the analysis assumes a chosen or given rate of depletion, which is often taken by economists as that which is to be determined. El Serafy's method does not answer the traditional question of what is the optimal rate of depletion, but rather tells us how much we can sustainably consume and how much we must invest of receipts from a nonrenewable resource under different discount rates, depletion rates, and reserves. It sets the guidelines for exploiting nonrenewable resources under a regime of sustainable development.

If we take sustainable development as our guiding principle, then the projects we finance should, ideally, each be sustainable. Whenever that is not possible (e.g., nonrenewable resource extraction) there should be a complementary project that would insure sustainability for the two taken together. A portion of the receipts from nonrenewable extraction should be invested in a renewable asset in an annual amount such that, given the renewable asset growth rate and the life expectancy of the nonrenewable asset, the former will provide a permanent income stream equal to the part consumed annually of the receipts from the latter. This is the basic principle underlying the El Serafy method, just discussed, only here it is applied at the project or micro level rather than at the macro level of national income accounting.

Also, if projects must be sustainable, then it is inappropriate to calculate the benefits of a sustainable project or policy alternative by comparing it with an unsustainable option--i.e., by using a discount rate that reflects rates of return on alternative uses of capital that are themselves in the majority of cases unsustainable. For example, if a sustainably managed forest can yield 4 percent, and is judged uneconomic in comparison with an discount rate of 6 percent, but on closer inspection the 6 percent discount rate turns out to be based on alternative uses of capital that are unsustainable (including perhaps the unsustainable use of that same forest), then clearly the decision boils down to sustainable versus unsustainable use. If we have a policy of sustainable development then we choose the sustainable alternative and the fact that it has a negative present value at an unsustainable discount rate is irrelevant. The discount rate must reflect the rate of return on alternative sustainable uses of capital, if we are to have a policy of sustainable development. The efficiency allocation rule (maximize present value) cannot be allowed to subvert the very goal of sustainable development by application of an unsustainable discount rate (i.e., a discount based on alternative uses of capital which are unsustainable).

Sustainability of an investment project is a benefit. In general an extra benefit usually requires an extra cost. A policy of sustainable development means that we are willing to pay that cost, at least within reason. The above discussion suggests two alternative ways of evaluating projects in a regime of sustainable developent. The first is a halfway measure, the second is more complete. (1) Use a discount rate that excludes nonsustainable projects from the alternative uses of capital when evaluating sustainable projects. Likewise investments in nonsustainable projects should be evaluated on the basis of a discount rate reflecting only alternative nonsustainable projects. Allocation between these two broad categories is not addressed by this splitting of the discount rate, and remains undetermined. (2) A better way is to pair unsustainable projects with sustainable ones and count only the income component of receipts in calculating rate of return on all projects. The single discount rate would then measure the rate of return on alternative projects, all of which (paired) are sustainable.[1] Perhaps the "pairing" of projects need not be explicit. Counting only the income component in calculating the rate of return on unsustainable projects may be sufficient, on the assumption that the capital component is invested in a sustainable project with a growth rate equal to the discount rate used in separating the income and capital components.

V. Summary and Conclusions

---

[1]This approach is in agreement with Markandya and Pearce's general principle that "where possible it is better to adjust the cost and benefit values than to adjust the discount rate"(p.58), "Environmental Considerations and the Choice of Dioscount Rate in Developing Countries" World Bank, May, 1988, Environmental Department Working Paper No. 3.

The major conceptual issue we must resolve in thinking about economic development and the environment in the next decade is to integrate the one-way throughput as the basic starting point of economic analysis, even more fundamental than the circular flow. Next is to distinguish clearly the problem of the optimal allocation of the throughput from that of its optimal scale. Our attention will then naturally become focused on how to limit the scale to an optimal, or at least sustainable level, thereby giving the sustainable development discussion a bit more of a theoretical foundation than it has had to date. From there we can begin to investigate operational principles of sustainability, such as those discussed here and summarized below.

(1) The main principle is to limit the human scale to a level which, if not optimal, is at least within carrying capacity and therefore sustainable. Once carrying capacity has been reached the simultaneous choice of a population level and an average "standard of living" (level of per capita resource consumption) becomes necessary. Sustainable development must deal with sufficiency as well as efficiency, and cannot avoid limiting scale. An optimal scale would be one at which the long run marginal costs of expansion are equal to the long run marginal benefits of expansion. Until we develop operational measures of cost and benefit of scale expansion the idea of an optimum scale remains a theoretical formalism, but a very important one. The following principles aim at translating this general macro level constraint to the micro level.

(2) Technological progress for sustainable development should be efficiency-increasing rather than throughput-increasing. Limiting the scale of resource throughput would induce this technological shift.

(3) Renewable resources, in both their source and sink functions, should be exploited on a profit-maximizing sustained yield basis and in general not driven to extinction, since they will become ever more important as nonrenewables run out. Specifically this means that:

(a) Harvesting rates should not exceed regeneration rates; and

(b) Waste emissions should not exceed the renewable assimilative capacity of the environment.

(4) Nonrenewable resources should be exploited, but at a rate equal to the creation of renewable substitutes. Nonrenewable investments should be paired with renewable investments and their rates of return should be calculated on the basis of their income component only, since that is what is perpetually available for consumption in each future year. If occassionally a renewable resource is to be depleted in a nonrenewable fashion (driven to extinction), then the same pairing rule should apply to it as for a nonrenewable resource. Thus the mix of renewable resources would not be static, but there would be a compensating renewable investment for every divestment.

Perhaps there are other principles of sustainable development as well, and certainly those listed above need to be refined, clarified, and made more consistent between the micro and macro levels. But these four are both an operational starting point and a sufficient political challenge to the present order. Will the nations seeking sustainable development be able to operationalize a concept from which such "radical" principles follow so logically? Or will they, rather than face up to population control, wealth redistribution, and living on income, revert to the cornucopian myth of unlimited growth, rechristened as "sustainable growth"? It is easier to invent bad oxymorons than to resolve real contradictions.