SPATIAL SUSTAINABILITY, TRADE AND INDICATORS: AN EVALUATION OF THE "ECOLOGICAL FOOTPRINT"

Jeroen C.J.M. van den Bergh[1]
Department of Spatial Economics
Vrije Universiteit
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
jbergh@econ.vu.nl

and

Harmen Verbruggen
Institute for Environmental Studies
Vrije Universiteit
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
harmen.verbruggen@ivm.vu.nl

[1] Corresponding author; fax.: +31-20-4446004.

Abstract

The search for frameworks and indicators of sustainable development has taken a prominent place in this journal. However, some specific aspects have received little or no attention, notably the spatial dimension and the role of international trade in indicator development. Moreover, many sustainable development indicators comprise implicit valuations, weighting schemes and policy objectives, which are insufficiently recognised as such. This contribution tries to highlight these issues by means of a review of a recently proposed indicator for ecological-economic analysis, namely the Ecological Footprint, that has been developed by Wackernagel and Rees. Its concept and calculation procedure will be criticised on a number of points, and it is concluded that the Ecological Footprint is not the comprehensive and transparent planning tool as has been assumed or suggested. In explaining our position we will address the notions of spatial or regional sustainability and sustainable development, and argue that they are critical in this context, but have not been precisely discussed so far, neither in the literature on trade and environment, nor in that on sustainable development, nor in the Ecological Footprint literature. We will defend the view that, as it presently stands, trade can be both good and bad for the environment, but in the long run trade may be the only way to combine economic welfare and global sustainability.

1. Introduction

The search for indicators of sustainability or sustainable development is a recurrent theme in this journal. Recently, the Ecological Footprint (abbreviated as EF hereafter) is suggested to offer a concept and method that can generate one of the most objective, non-biased, aggregate, single-dimension indicators for evaluating sustainability:

"In summary, by putting sustainability in simple but concrete terms, the Ecological Footprint concept provides an intuitive framework for understanding the ecological bottom-line of sustainability. This in turn stimulates public debate, builds common understanding and suggests a framework for action. The Ecological Footprint makes the sustainability challenge more transparent - decision makers have a physical criterion for ranking policy, project or technological options according to their ecological impacts." (Wackernagel and Rees, 1996, p. 57).

These claims are substantiated by an indicator expressing the ecological impact of human activities in terms of (hypothetically) required land areas to sustain these activities. The question that will be addressed in this contribution is to what extent the EF fulfils these pretensions. This is needed, because the EF concept and indicator seems to be accepted almost without any critique by many scientists and policy makers, and especially by environmental organisations. The rapid increase of its popularity and influence over a short period of time - e.g., witness the amount of internet sites mentioning it - provide the motivation to systematically examine its pros and cons.

Our evaluation bears in mind that the search for operational indicators for sustainable development should be guided by a number of specific criteria. For instance, the calculation procedure should be scientifically sound; the indicators should relate to clear policy objectives; the indicators should have a clear interpretation and be understandable to non-scientists; the indicators should cover the functioning of a system as a whole; the indicator should be based on parameter values that are stable over a long period of time (see Kuik and Verbruggen, 1991). It will subsequently be argued that the EF suffers from serious shortcomings regarding all these criteria. As a result, it may provide a wrong direction for our intuition, i.e. give rise to unsustainable, inefficient or even immoral policy options.

The organization of this article is as follows. Section 2 provides a short explanation of the EF concept and calculation procedure. Section 3 offers a critique, focusing on methodological issues, the notion of land use, the transformation of energy use, the spatial dimension and the policy relevance. Section 4 discusses some alternative views on spatial sustainability and trade. Section 5 looks at the implications of these for EF analysis. Section 6 concludes.

2. A short explanation of the Ecological Footprint

Wackernagel and Rees (1996) have introduced the concept of, and method for calculating the EF (see for a short account Wackernagel and Rees, 1997). It is presented as a simple operational indicator to aid in monitoring progress towards (un)sustainability, i.e. maintenance (loss) of natural capital. It accounts for the flows of energy and matter to and from a specific economy or activity, converted into corresponding land and water area needed to support these flows. Six land categories are included in the procedure, namely consumed/degradaded land (built environment), gardens, crop land, pasture land and grasslands, productive forest, and energy land. EF analysis is suggested to be useful in determining the human appropriation of ecological production, measured in area units. The power of the method is the fact that all human exploitation of resources and environment is reduced to a single dimension, namely land and water area needed for its support.

An EF can be assessed for persons, activities or regions. How is it calculated? First, consumption is determined in a particular spatial domain for each relevant category. This includes food, housing, transportation, consumer goods and services. Next, the land area appropriated by each consumption category is estimated for different land categories. This includes land appropriated by fossil energy use, built environment, gardens, crop land, pasture/grassland and managed forest. This is based on both resource and waste flows, and leads to a consumption/land-use matrix. Summing all the area figures in this matrix gives an estimate of the EF of the region considered. Application of the method shows that many developed countries have an EF in the range of 3 to 5 ha/person, with the USA ranking highest. The world average is estimated at 1.8, while the EF of a developing country like India is 0.4 ha/person.

Wackernagel and Rees also calculate the EF/actual-productive-area ratio for a region, as an indication of its (un)sustainability. Especially small developed countries and densely populated cities score high on this ratio measure: for instance, Belgium and the Netherlands between 10 and 20; and London 120. In addition, one can compare per capita EFs with the per capita available ecological space on Earth. It is estimated that the latter has decreased from approximately 6 to 1.5 hectares since the beginning of the century. Note that an EF larger than a region's actual land area is possible due to the EF representing hypothetical instead of concrete land use. This feature will be considered in detail later on. In addition, trade may of course cause the EF to exceed land availability on a regional level.

3. Critique on the EF concept and method

Conversion, aggregation and weighting

In explaining our objections against the measurement and aggregation procedures underlying the EF, we depart from the “global EF indicator”, for the moment overlooking regional and spatial dimensions. A first objection has to do with the supposed attractiveness and strength of the EF, namely that it provides a one-dimensional indicator by summing up all consumption - of a region, activity or person - related direct and indirect ecological impacts in terms of land use. This requires that different consumption categories are translated into land area. Evidently, this conversion is necessarily incomplete, rough, based on sometimes arbitrary data, while no account is taken of regional and local features of land types and land use. But the main problem is that physical consumption-land conversion factors are used that function as implicit weights in the conversion as well as the aggregation. Physical weights do, however, not necessarily correspond to social weights. This becomes problematic when the EF is meant to serve as a criterion for ranking policy options. This problem is magnified by the choice of a fixed weighting scheme. This reflects neither relative scarcity changes over time, nor variation over space - due to, for instance, the level of economic development. Consequently, the EF procedure may produce odd results that are unwanted from a societal point of view.

Of course, this problem is a characteristic of many aggregate ecological indicators, which are often based on the choice of a fixed and physical aggregation scheme. Another such indicator that has been discussed in this journal is “materials inputs per service unit” (MIPS), in which kg of any type of material are added to arrive at an aggregate indicator of material intensity per service. As a result, heavy metals and sand thus receive implicitly equal weights. The objective of Factor 4, for instance, has been based on this indicator (see Von Weizsäcker e.a., 1997).

Finally, the EF indicator may be regarded too aggregate in the sense that it lumps together population size, environmental pressure per capita, and technological efficiency in physical/environmental terms. As a result, one cannot judge much on the basis of (regional) EF alone, neither what is the main problem nor what can be policy solutions. For this reason a decomposition type of approach is needed, which distinguishes between population density, materials use per capita, economic structure (production and consumption) unsustainable land use per capita, and technological efficiency. This implies a logical complete system of multiple, complementary indicators.

Land use and the EF

A second objection against the EF is that it does not distinguish between sustainable and unsustainable use of land, however defined. In order to measure the degree of unsustainability of an economy or activity, we would argue that indicators are needed that take explicitly into account processes contributing to unsustainability, such as unsustainable resource use and environmental degradation in a broad sense, rather than just an overall (and hypothetical) land area measure. Thus indicators need to reflect the quality and quantity of renewable resources, and be confronted with safe margins or threshold values. Soil degradation is an important issue to be addressed in this context.

Serrão *et al.* (1996) are able to indicate for a range of current land use systems in the Brazilian Amazon the degree of sustainability, distinguishing even between agronomic, ecological, economic and social components. Clearly, non-renewable resources cannot be used and maintained at the same time, so that more complicated sustainability rules have been devised, like maintaining the productivity of the non-renewable resource base through technological progress, or replacing reductions in non-renewable resource stocks by additions to alternative stocks of renewable resources. Neither of these subtle issues are addressed by the EF. A distinction between sustainable and unsustainable land use seems a minimum condition for any procedure aimed at determining to what extent an activity or region is contributing to (un)sustainable development. This is not to say that such a procedure can easily be implemented. For instance, one will have to address difficult questions like “what is sustainable land use?”. However, ignoring this question is even worse. Some authors have tried to come to grips with this issue by relating economic activities, demography and other variables to land use and land cover patterns (see Darwin *et al.*, 1996). Global modelling (“integrated assessment”) has also devoted much attention to land use (see, e.g., Rotmans and de Vries, 1997). These exercises conclude that the relationships are usually complex, so that simple aggregation procedures to arrive at single indicators, like the EF, are bound to loose information and be biased.

Another important issue completely neglected by the EF calculation procedure is that land use is regarded to be associated with single functions only. However, in many cases land use (and cover) provide multiple services or functions, and each of these cannot easily be separated and linked in a non-arbitrary way to specific land areas. For this reason the terms multifunctionality and multiple use have been employed in the study of ecosystems under stress (Bowes and Krutilla, 1985; van der Ploeg 1990; Braat, 1992). Neglecting multifunctionality associated with land use and cover will bias the EF upwards.[2]

---

[2] This point was noted by Helias Udo de Haes of Leiden University, The Netherlands.

Wackernagel and Rees admit that the way their indicator has been constructed neglects various issues related to unsustainability, notably acidification, disturbance of ecosystems via infrastructure and noise, emission of toxic substances to natural systems, and the whole in the ozone layer. Other indicators are needed to deal with some of these issues (see, e.g., Ayres, 1996). The EF focuses attention on CO_2 emissions (other greenhouse gases are omitted), implicitly giving a greater (infinite) weight to this problem than to the previous ones. Of course, this means that the EF value is an overestimate when the intensified greenhouse effect is not real after all, or an underestimate when both the included and omitted processes previously mentioned have a direct (and linear) link with sustainability. Wackernagel and Rees clearly recognize this deficiency (e.g., 1996, p. 62; they leave out the “human footprint” in the sea; see p.64-65), but they regard the resulting EF as a minimum estimate. Likewise, any comparison of the EF with actually available land in a region is considered to underestimate the degree of unsustainability of such a region. However, this neglects other biases: given the previous point on neglect of multifunctionality, among others, nothing can be said in general about the direction or sign of the bias of the EF, and it may differ significantly between different applications.

An implication of the foregoing critique is that the EF does neither take account of, nor allows for - given its aggregation level - a trade-off between environmental sustainability and intensive/extensive land use, notably in agriculture. Clearly, this is an important issue for policy and science to examine, but it is completely lost in the EF procedure. In agricultural production, for instance, intensive land use, translated into a small contribution to the EF, is usually associated with high environmental pressure in both space and time: i.e. in terms of use of pesticides and fertilisers, and groundwater control and irrigation to improve productivity at each moment in time, as well as to prolong the crop season.

Energy and the EF

A third objection against the EF has to do with the measurement and aggregation procedure used to address environmental impacts associated with energy use. The land appropriated by fossil energy use makes up more than 50 % of the EF estimate for most developed countries. This component consists of estimating the land area needed to catch (assimilate) the CO_2 emissions from burning fossil fuels, i.e. "carbon sink" land. The idea behind this is that sustainability is realized if the carbon sink is not exceeded, thereby focusing only on the emission and not on the resource scarcity side of fossil energy use. This is questionable: CO_2 assimilation by forests is one of many options to compensate for CO_2 emissions, and indeed a very land-intensive option. Other sustainable solutions that are less land-intensive may result when the cost of CO_2 emissions is significantly increased to cover the external costs. These include shifts to other fuels, less fuel use, increasing energy efficiency of processes, or other ways to prevent CO_2 buildup in the atmosphere, such as carbon storage underground. The EF translation of energy use to land area is meant by Wackernagel and Rees to reflect the land required to support activities in an environmentally sustainable manner.

However, choosing a specific option based on the argument that it is presently the most cost-effective at the margin is incorrect. The cost-effective solution depends on the cost of land and the productivity of reforestation. Both are likely to differ between countries or regions, depending on the level of development, the technological expertise available, and geographical circumstances (including climate and soil type). Moreover, land scarcity and costs of land use differ significantly between countries - compare, for instance, Denmark with Australia. A mix of the above mentioned options would be adopted if policies are implemented to ameliorate the environmental impact of fossil fuel based energy use.

More importantly perhaps, the EF is not consistent with marginal cost thinking of economics as it does not consider marginal changes. Even if “carbon sink land” is the cheapest sustainable option at the margin, given the enormous distance energy use - in quantity and composition - is away from any sustainable scenario, large structural changes an all levels are to be expected under any such scenario. Given such “nonmarginal changes” it is extremely unlikely that the cheapest option chosen will be “carbon sink land”. Moreover, the more land is reforestated, the more expansive and unattractive this option becomes. Therefore, the EF procedure will significantly overestimate sustainable land use required for a sustainable economy with the present structure - even in the absence of any technological innovations in the area of energy conservation that moving to such an economy would surely entail.

A way to improve the EF procedure in this respect may be to make it dependent on an (static) energy scenario, rather than to fix the scenario. But if one scenario is adopted - as is done at present - then it should at least be realistic. Blueprinting on the basis of rigid models and scenarios will be insufficient; results of complex modelling are needed for this purpose. Given the nature of the changes needed to reach environmental sustainability - large and structural, and possibly over a longer period of time - such models can only generate rather realistic patterns if account is taken of costs, prices, behaviour, sectoral interactions, as well as dynamic effects such as technological change, including those in the area of energy use and conservation. The EF should take such economic mechanisms into account, and try to start from hypothetical land uses that are realistic, i.e. economically feasible and stable. If not, it is bound to overestimate the land “appropriation” for sustainable energy provision.

Concluding this critique point, the EF is too much dominated by energy use, which seems an indirect (and incomplete) approach to focus the attention on environmental problems related to energy use. Clearly, fossil fuel use or CO_2 emissions provide more direct and fair indicators to capture energy/environment effects. An evident conclusion in this light seems that cleaner (and likely more costly) energy will be the most essential step towards sustainability for a large part of the world.[3]

---

[3] Folke *et al.* (1997) separate clearly between ecosystem appropriation for natural resource production and for waste assimilation. For the latter they produce a low range and a high estimate, to address the uncertainties involved. Also here it is unclear whether land use scenarios are realistic.

Space, region and the EF

A fourth objective refers to the arbitrariness of the spatial scales at which the EF is calculated. Wackernagel and Rees calculate EFs at global, regional, national and local (cities) scales, and on a per capita basis. However, from an environmental point of view, sub-global EF estimates are rather arbitrary. To begin with, national boundaries are of a geo-political and cultural nature, and have no environmental meaning. On the contrary, countries show large discrepancies in environmental and resource endowments, soil characteristics, climate conditions and assimilative capacities. Exactly these differences have largely determined human settlements, the location of industrial activities and agriculture and the growth of urban areas. Over time, people have tended to concentrate in coastal zones, where presently the largest proportion of the world population resides, near harbours and (intersections of) waterways, where trade, commerce and other economic activities could flourish, largely driven by the existence of positive externalities. They include economies of scale and scope, agglomeration effects, and compact city advantages. Given the existence of positive effects one cannot be generally against regional concentration of people. This natural allocation pattern is mainly responsible for the separation between the places where people live and work, and where they derive their food and resources from. This pattern developed inter-regionally and among countries. Not surprisingly, there is a significant correlation between population density and resource endowment, in the sense that relatively densely populated countries and regions are often resource poor. By this natural fact alone, densely populated countries, regions and cities show large EFs. This is, however, not a sign of global unsustainability, but rather the outcome of allocation factors and specialisation patterns. International and interregional trade is then primarily a derivative. Trade and specialisation improves welfare for those countries and regions that take part in this trade. Apart from the fact that the EF concept completely overlooks the merits of international and interregional specialisation patterns, the calculations of EFs for countries, regions and cities adds an element of wrongful arbitrariness. In any case, this makes inter-country, inter-regional and inter-city comparisons of EFs meaningless, also on a per capita basis. Only an inter-temporal comparison for a particular country, region or city can be informative.

Related to the foregoing point is population pressure. As indicated, it does not seem fair to compare large (in terms of economies or land area) and small economies (like cities or small countries/regions), as the latter will always be relatively more open. In other words, there is always a spatial economic scale, where one can find relatively many cross-boundary flows of materials. Our most serious worry is that regions and regional demarcations are arbitrary, which makes regional EFs and their comparison also arbitrary. For instance, the comparison of densely populated small countries (e.g., various European countries) with sparsely populated large countries (e.g., North-American), is a bit like comparing cities with continents. The implication of a high concentration of people in a region is clearly a large EF. But in itself, this is not good or bad. Interestingly, Wackeragel and Rees (1996) even themselves state that living in densely populated urban areas leads to smaller per capita EF. But how should this be evaluated in the light of the remarks on trade and ecological deficit? To address population pressure a more direct indicator is trivial, namely population density per unit of area. The trade-off between environmental advantages and disadvantages of alternative spatial compositions and concentrations of activities is not reflected in the EF concept or procedure.

Policy relevance, objectives and implicit value judgements

A final objection underplays the EF as a planning tool. Wackernagel and Rees refer to EF analysis as offering “... a planning tool that can help to translate sustainability concerns into public action.” (1996, p.3). Bearing above objections in mind, regarding fixed arbitrary weights in the aggregation procedure underlying the EF, and on the problems associated with spatial scales and trade, it is hard to see how such an indicator can be used as a planning device. For any planning there should be one or more objectives, one or more constraints and one or more instruments. All three categories are unclear in the

EF context, and hence, no alternative sustainable policy strategies are identified to reduce the EF. It should be stressed that no objectives are explicitly specified. For instance, neither minimizing land use nor maximizing regional land productivity are mentioned. Minimizing land use could of course be regarded as most relevant in the context of the EF concept, but the obvious question then is under what conditions with respect to income and welfare. The common use for regions seems to be to compare the EF - i.e. hypothetical and unrealistic land area - with the available land area in a region. This is clearly uninformative for global sustainability, and unclearly related to regional sustainability, since the latter is undefined in the EF context. Implicit in comparing the EF with the actually available land in a region seems to be the interpretation that autarky is the ideal situation. The ideal situation is then a level of population and economic development that "uses" more land than is available in the region. But if trade is allowed, then the arbitrariness becomes even more evident: trade is allowed up to the point where the sum of land use domestically and abroad equals according to the EF equals the land available in the region. Only expansion of land by war would allow for a larger footprint. However, as we shall argue later, trade based on sustainable land use is more desirable - and peaceful. In any case, the foregoing points pose the question as to the implicit value judgements and objectives underlying the EF concept.

The EF should be placed in the context of the pursuit of environmental sustainability and maintaining "natural capital" - interpreted in a very broad sense, including flows of goods, materials and services, as well as all ecosphere relationships. This pursuit comes very close to the notion of "strong sustainability", which requires that all components of the "natural capital" should be maintained (see Pearce *et al.*, 1998). This is clearly a normative choice, which should be open to debate, instead of being fixed in an indicator procedure. In particular given the regional or national focus of many EF exercises, one may ask whether substitution among different spatial scales is allowed for? If so, interregional and international trade is possible. Cross-region and cross-country differences in environmental endowments and natural conditions can create environmental gains through trade and specialisation. In that sense, spatial optimisation of the use of environmental goods through trade can contribute to sustainability. But how should we measure, value or compare sustainability at different spatial levels if trade is involved. The EF clearly does not address this important issue, and can therefore not serve as useful information for policy decisions relating to sustainable development of regions and nations. It seems to suggest regional and national autarky as the most desirable outcome from an environmental perspective, which is not only politically unrealistic and dangerous, but also untrue as it completely neglects comparative advantages of countries and regions related to endowments of environmental and ecological resources, or simply in terms of space and population density (e.g., The Netherlands versus Canada).

4. Spatial sustainability and trade

The issue of spatial sustainability and spatial sustainable development, i.e. the spatial dimension of environmental sustainability, has been largely neglected by environmental and ecological economists alike, even in very systematic approaches (see, e.g., Costanza and Patten, 1995). One reason for this may be that it involves an integration of insights from such diverse fields as international economics, regional economics, transport economics, economic development and growth theories, ecology and environmental science. A short and necessarily incomplete discussion is given here of economic approaches to spatial dimensions of environmental problems, and particularly the interaction between growth, trade and environment. The policy lessons for environmental policy makers from the advanced literature on externalities in environmental economics includes various corrections on the standard Pigouvian charges and taxation. Such corrections are motivated by a number of phenomena: endogenous locations, imperfectly operating markets; large firms, cartels and countries using their international trade power; policy competition between countries; international transboundary pollution flows (e.g., Anderson and Blackhurst, 1992; Beladi, and Frasca, 1996; Markusen *et al.*, 1993; Verhoef and van den Bergh, 1996). However, a mix of these insights with (endogenous) long-run growth, trade and environmental sustainability, has so far been lacking in the standard approach of environmental economics. To some extent this is due to the fact that the externality concept of welfare economics has not been clearly linked to the concept of sustainability. The latter is usually approached by economics from the perspective of economic growth theory. Trade theory is closer to welfare (or micro) economics (see van Beers and van den Bergh, 1996).

A logical starting point in the context of growth and environment is the existence of a finite natural carrying capacity (CC) which would act as a limiting factor to the scale of the economy. Some possible patterns of economic scale and CC over time are shown in Figures 1, for two situations: no trade or a closed, autarkic system (CCautarky), and trade (CCtrade). The “density dependent growth” pattern, restrained by some limiting factor (resource) is represented by curves 1 for the autarky situation, and by curve 2 for the trade situation. Two-way interactions between economic and ecological systems may cause variations in the carrying capacity over time. For instance, due to trade regional production and consumption can rise above the regional carrying capacity (“overshooting”), as shown by curve 2. Subsequently, environmental degradation of factors that compose the CCautarky may occur, due to the fact that higher production and consumption lead to excessive waste and toxic emissions to natural systems. This is illustrated by replacing the constant CCautarky by the variable carrying capacity represented by curve 3, which in turn may affect the CCtrade. The latter can be regarded as the sum of CCautarky and a part representing a “pure trade effect”. The variable carrying capacity CCtrade for the open system, as given by curve 4, can ultimately give rise to a negative feedback to economic change, in which case curve 2 is replace by curve 5. Of course, it should be realized here that curve 3 does not need to follow from trade, as long as environmental policy and management are adequate and aimed at long run sustainability, i.e. constant carrying capacities in the present abstract model. This simple picture provides a general conceptual framework for studying the interaction between growth and environment on global and regional scales. In van den Bergh (1993) a simple formal model is considered that represents the described processes.

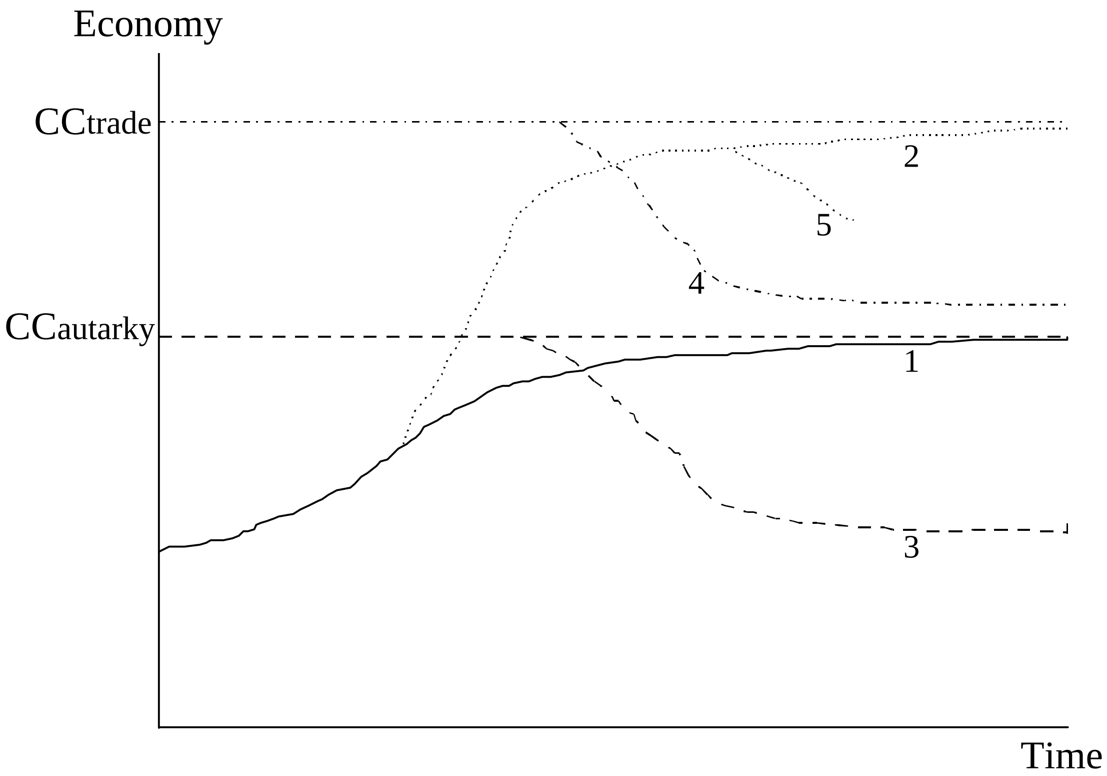

Figure 1. Development and carrying capacity under autarky and trade.

A balanced analysis of the relation between environment and trade requires that trade theory and environmental economics are merged. Welfare theory which takes due account of negative external effects provides a proper framework of integrated analysis. However, environmental economics has mainly focused on closed economies. Only since the second half of the 1980s much attention has been devoted to open economies, in which context the impact of environmental policy on international trade, and the international coordination of environmental regulation, are important topics. The traditional trade theories that have substantiated the plea for free trade are based on the assumption that there are no negative external effects. Integration of environmental economic insights in these theories implies that that free trade only optimizes social welfare when all external effects are "internalised" via property rights, integration of polluters and victims, price corrections/incentives or otherwise (for an overview of standard economic approaches, see van Beers and van den Bergh, 1996; and for some critical notes on the standard perspective Folke *et al.*, 1994).

An integration of standard neoclassical economics of international trade and environmental externalities is needed, but insufficient to address sustainability issues. Economic efficiency and sustainable development may be conflicting even (Pezzey, 1989; Pezzey and Withagen, 1998). More generally, we are concerned that externality-based policies are inadequate to help us realize safe and sustainable environmental quality. A fundamental reason is that externalities are between economic agents, usually framed in a single generation, with the environment playing only an abstract and hidden role, and dynamics mostly being mitigated. A more pragmatic argument is that externality valuation is always partial, and based on peoples' preferences that do not extend to long term impacts of our present decisions, i.e. dynamic externalities are represented incompletely. So it is not clear that socially optimal trade, base on optimal externality pricing, is consistent with environmental sustainability, neither on global nor on regional levels.

However, the most important interaction between trade and environment may not be physical but perceptional, namely that open economies are restricted in their regional policies, including environmental ones, by the international competition they face and the goal of remaining a sufficiently competitive international position.

5. Sustainable trade and the EF

The notion of "ecological deficit" is introduced by Wackernagel and Rees (1996, 1997) as an indicator of unsustainability. As it is compensated by either trade or depletion of regional natural assets, they draw the conclusion that trade should be minimized.

In our view, trade can in principle spatially distribute the environmental burden among the least sensitive natural systems. Since it is not realistic that either human societies as a whole, or immobile natural resources, change locations, it is obvious that trade of commodities and resource materials remains the only other way to spatially match consumption, production and resource use. This requires that correct national and international incentives or regulations are operative, preferably at sources of environmental pressure. Only then full ecologically comparative advantages can be enjoyed.

The EF hides the favourable impact of specialization, not merely in terms of efficiency, but also in terms of environmental sustainability given the clustering of people in space. This is possible by taking advantage of cross-regional differences in environmental endowments an natural (climatic and geo-physical) conditions. On the basis of the traditional mechanism of Ricardian comparative advantages, environmental gains can be realized through trade and specialization. Essential in this respect is that natural resources are not uniformly distributed or heterogeneous over space and are immobile. Daly and Cobb (1989), and Daly and Goodland (1994), have argued that capital is mobile nowadays, and therefore trade theories do not apply anymore. However, it is straightforward that the lessons of trade theory, notable the existence of comparative advantages, still holds with regard to unique and immobile natural resources.[4]

According to Wackernagel and Rees, trade is allowed up to a level where the sum of regional and indirectly land uses equals regionally available productive land. Comparing regions or countries in this sense is intrinsically normative: why should not a country be allowed to use its space for producing for a country where land is more scarce. This is just an expression of comparative advantage from trade theory. We trade implicitly or explicitly all sorts of concrete and abstract features of our environment (publicly or privately "owned") - including climate, knowledge, and culture-, so why should an exception be made for land or space? If we do not accept such trade, we should also forbid the existence of urban areas (see Folke *et al.*, 1997). And we should accept that some countries have so much more land per capita which is unused while other countries are overusing their scarce land. Or, alternatively, the EF approach could be used to motivate a more uniform distribution of people over space, or in any case enforcing

---

[4] Of course comparative advantage also applies because capital is not completely and perfectly mobile (investments in plants), and knowledge and labour population characteristics differ between countries. Furthermore, mobility is certainly not perfect and complete over short or medium time periods.

migration of people to spatially match population density and ecosystem carrying capacity.

One can use a variation of the well-known decomposition formula for environmental pressure to write the EF as follows:

$$EF = P*C*(E_R+E_X),$$

with P for Population, C for (average) consumption per capita, and E_i for environmental pressure per unit of consumption. The latter is distinguished between regional (i=R) and external pressure (i=X). In this model, the EF changes via alterations in 4 factors. The interesting one is clearly E_x which may change via either less trade, less unsustainable land use outside the region. Alternatively, E_R and E_x may change due to substitution of sustainable land use outside the region for unsustainable land use inside the region.

Trade and regional carrying capacity factors

An important issue raised by Wackernagel and Rees (1997), as well as by van den Bergh and Nijkamp (1991, 1994, 1995), is that trade seems to increase regional carrying capacity, but may actually harm it. Wackernagel and Rees only discuss this issue shortly. They argue that global sustainability is reduced as all regions are encouraged to exceed local limits, mainly because risks attached to local natural capital depletion are undervalued. We tend to agree, although we do not see trade as the main problem or cause, but rather the perceptions of people and policy makers. This creates “policy failures”. In other words, if perceptions are corrected, and if in line with this regional policies are adequately adjusted to protect regional environmental carrying capacities, then trade can perfectly go along with regional sustainability. Furthermore, this is consistent with global sustainability, although not sufficient, even if all regions act in this way.

It remains true, however, that trade can alleviate specific limiting factors underneath regional carrying capacity, thus opening the way for other previously inoperative, potentially limiting factors. This point is also raised, in a somewhat different way perhaps, by Wackernagel and Rees (1997), when they discuss the “Law of the Minimum”, which refers to the idea that “...systems and processes are governed by that *single necessary factor* in least supply ...” (p. 15), or perhaps more precisely, the factor which is relatively most scarce. One can imagine that by trade a region starts to experience growth such that the regional environment becomes subjected to new types of environmental pressure - new resource extraction, new substances emitted, new uses of space, etc. - which are not well controlled as they were not experienced before, so that there is, for instance, no adequate environmental policy developed for such new problems. However, here again our conclusion is that a good environmental policy is the best solution and not restricting trade. Clearly, openness of regions, not merely physical trade, have led to spreading of all sorts of habits, ideas, knowledge, technique, goods, and so on, which has brought about various changes, some to the benefit and some to the harm of mankind and local people and cultures. We are of course not claiming that this is perfectly desirable, but a return to a world composed of isolated regions seems highly unrealistic, so that a better approach would be to strive for each region to be seriously concerned about sustainably managing its own ecosystems and resources, as well as to pursue international environmental governance and coordination. In that case, trade in itself is not bad for the environment, but instead may be good as it allows for spatial matching of environmental pressure with (in)sensitivity or carrying capacities of natural systems.

By the way, is it true that trade and globalization reduce the value of regional resources and ecosystems? In terms of use values, this is not always the case. Assume an isolated region with a non-unique resource (supply) and a given demand and imagine this region opens up trade in the resource (either import or export). Then for a larger demand the price will increase, reflecting additional use values, while a larger world supply, given adequate regional environmental policies, just means that the resource is “less scarce” on a world than regional scale, so that the regional pre-trade price will fall once trade is started. In other words, the answer to the question raised is no, while the drop in market value just says something about relative scarcity, providing a mechanism for allocation. Obviously, (coordinated) regional environmental policies can raise the resource price level relative to other prices/costs, and make sure that conservation of resources and ecosystems becomes a standard approach.

None of the above considerations supports the conclusion by Wackernagel and Rees (1997, p.21): “A shift [is needed] from the present emphasis on global economic integration and inter-regional dependence toward greater regional autonomy and self-reliance”. This is a normative judgement, which we respect as such. However, we would like to add that some of the considerations mentioned would lead to the opposite conclusion. Furthermore, one can imagine various negative consequences of minimizing trade between countries that are not directly related to environmental issues. So even if one is not convinced by our arguments, then it should be realized that a trade-off is required between taking environmental risks and other risks. The latter are obvious: worsening of international relationships between countries, trade wars and other conflicts, lack of diffusion of knowledge and information, and widening of the gap between rich and poor regions in the world.

Other indicators for sustainability and sustainable trade

Actual space or land used in and outside a country or region are more concrete and less confusing indicators than the EF, and they can be compared with actual land area available in the region and outside to provide for indicators of environmental pressure and risk, and environmental unsustainability. Hypothetical and unrealistic land use, as represented by the EF is tricky, as it is bound to be interpreted as realistic or even actual land use. Note that "hypothetical" refers mainly to the translation of energy use to land area.

An EF for one region does not say anything about sustainability. Only the global EF - compared with the global land availability - is useful to learn about sustainability or unsustainability. However, there are more informative indicators than such an aggregate indicator. Regional EFs do not make a difference between land use in one region or another (apart from the transport effect contribution, but that is a different matter) as they focus on the consumer perspective. However, regional carrying capacities are so different, and hence, actual land use in a region should be compared to available land and the quality/capacity of it, an approach many ecologists would probably feel more comfortable with. Much more could be said about alternative indicators, but this is not the place. For surveys of such indicators, see Kuik and Verbruggen (1991), Gilbert and Kuik (1999), Pearce *et al.* (1998), and Rennings and Wiggering (1997). However, the regional and spatial dimensions still need to be fully integrated in the process of indicator development.

6. Concluding remarks and recommendations

As much as we are in sympathy with Wackernagel and Rees in their concern about our impact on natural systems, and our "appropriation" of natural capital or carrying capacities of natural systems, we cannot subscribe to their view that EF analysis informs us sufficiently about this impact and "appropriation", and we completely disagree that the EF as it is presently constructed can serve as an indicator of (un)sustainability or provides useful information for assessing (un)sustainability. We would even go one step further, and argue that EFs are confusing, arbitrary, incomplete, normative and too aggregate. In the case of global sustainability, a hypothetical EF value for the world population can exceed the world's total available land, while its value is unbounded from above. Evidently, actual land use is restricted from above. Hence, the EF is not a concrete and transparent measure so that its interpretation is surrounded by arbitrariness.

Regional EFs, or better, comparisons of regional (or national) EFs with regionally available land (per capita) are very confusing. First, regional demarcations are abitrary (e.g., administrative, historical), so that whether regional EFs are within the regional land area is also arbitrary. Urban areas will never be sustainable, by definition almost, we would tend to say, but what is the lesson of this? There is no obvious realistic threshold below which urban EFs are acceptable. Trade between regions is a way, and likely the only way given the present state of the world, to arrive at sustainability, namely by matching spatial environmental pressure with spatial environmental capacities for “neutralizing” that pressure. This could be done without entering a “danger zone”, namely by staying well within “safe minimum standards” instead of reaching “determinist-like carrying capacity limits”. We repeat that this requires that regional environmental policies are “adequate”, i.e. aimed at protecting regional resources, and not extra-regional resources, which should remain the pure responsibility of the respective other regions. Of course, during a transition period or a process of uncoordinated policies things may be different (Kox and van der Tak, 1996). Furthermore, any restriction of trade for environmental reasons, given such “adequate” regional policies, would be purely arbitrary and normative, and restricting opportunities.

A first improvement would be to calculate actual instead of hypothetical “footprints” of two types, namely sustainable and unsustainable actual land use per capita. Especially the latter is relevant for environmental policy, while the sum would be relevant for equity evaluation among persons, or if in per capita (average) terms, among countries. In general, more disaggregation allows to adequately reflect the three central dimensions of ecological economics’ evaluation, i.e. efficiency, equity and sustainability. One aggregate indicator can impossibly do this.

In addition, more flexibility needs to be allowed for in the EF calculations. This should certainly apply to the aggregation and weighting of different physical and environmental dimensions. These should reflect social values, and require more motivation, and possibly a cases-to-case specification. Also energy transformation into hypothetical land needs to be done on the basis of more adequate schemes, reflecting regional differences as well as minimum costs. It should be realized that marginal costs estimated at present are not very informative when it comes to decide about nonmarginal changes in energy use and composition of energy sources. Lastly, one should be careful in trying to find a single, absolute value for the EF, and instead follow a scenario approach, which allows to deal with complex processes in the case of large, nonmarginal changes.

After so many critical points, we would like to provide a few suggestions for how to use the EF. When care is taken of the shortcomings due to its calculation noted above - which is not an easy task, and may even be impossible - it may be used in the following ways:

- For a comparison of people with different consumption patterns a corrected EF may be helpful. This is then based on calculating a per capita EF for different groups, regions or countries. These can be used as an indicator of distribution or equity. Also a fair earth-share can be determined and used as a benchmark, but rather in actual land use terms, complemented by other fair-share indicators (e.g., for CO_2 emissions). However, these footprints should not be compared with available per capita land in a region or country.
- For a comparison of an individual (or groups) over time a corrected EF could be useful. A positive change may then be interpreted as a increase of environmental pressure.
- For a comparison of techniques that do the same job, i.e. generate the same output in terms of a good or service, a corrected EF could also be useful. This approach is somewhat analogous to a life-cycle analysis of products.
- For a comparison of the EF for the world as a whole with the available land of the world as a whole a corrected EF could also be useful. This then approximates a global sustainability indicator.

As a last remark, it may be noted that our scrutiny of the EF indicator should be interpreted neither as a disconcern for ecological or environmental sustainability - quite the opposite, as will hopefully be clear by now -, nor as resulting from a narrow “free trade” dogma. We argue, however, that trade can in principle spatially distribute the environmental burden among the least sensitive natural systems, a point which does not seem to have attracted much attention in the literature so far. But this requires that correct incentives or regulations are operative, preferably at sources of environmental pressure, and international policy coordination for transboundary environmental issues. It does not require any trade controls or barriers. We agree with Wackernagel and Rees

(1997, p. 16) that trade may contribute to harming regional carrying capacities, however only due to a lack of regional environmental policy or a region’s concern for its regional natural systems and resources.

References

Anderson, K. and R. Blackhurst (eds.), 1992. *The Greening of World Trade Issues*. Harvester Wheatsheaf, New York.

Ayres, R.U., 1996. Statistical measures of unsustainability. *Ecological Economics* 16: 239-255.

Beers, C. van, and J.C.J.M. van den Bergh, 1996, International trade and environment: An overview and comparison of methodological approaches. *Journal of World Trade* 30: 143-167.

Beladi, H., and R. Frasca, 1996. Regional pollution and multinational firms. *Ecological Economics* 17: 117-125.

Bergh, J.C.J.M. van den, 1993. A framework for modelling economy-environment-development relationships based on dynamic carrying capacity and sustainable development feedback. *Environmental and Resource Economics* 3: 395-412.

Bergh, J.C.J.M. van den and P. Nijkamp, 1991. A general dynamic economic-ecological model for regional sustainable development. *Journal of Environmental Systems* 20: 89-214.

Bergh, J.C.J.M. van den and P. Nijkamp (eds.), 1994. Sustainability, resources and region. *The Annals of Regional Science*, Vol. 28 (1), special issue.

Bergh, J.C.J.M. van den and P. Nijkamp, 1995. Growth, trade and sustainability in the spatial economy. *Studies in Regional Science* 25: 67-87.

Bowes, M.D. and J.V. Krutilla, 1985. Multiple use management of public forestlands. In: A.V. Kneese and J.L. Sweeney (eds.), *Handbook of Natural Resource and Energy Economics*, Vol. 2. North-Holland, Amsterdam.

Braat, L.C., 1992. *Sustainable Multiple Use of Forest Ecosystems: An Economic-Ecological Analysis for Forest Management in the Netherlands.* Ph.D. dissertation, Free University, Amsterdam.

Costanza, R. and B.C. Patten, 1995. Defining and predicting sustainability. *Ecological Economics* 15: 193-196.

Daly, H.E. and W. Cobb, 1989. *For the Common Good: Redirecting the Economy Toward Community, the Environment and a Sustainable Future.* Beacon Press, Boston.

Daly, H.E., and R. Goodland, 1994. An ecological-economic assessment of deregulation of international commerce under GATT. *Ecological Economics* 9: 73-92.

Darwin, R., M. Tsigas, J. Lewandrowski, and A. Raneses, 1996. Land use and cover in ecological economics. *Ecological Economics* 17: 157-181.

Folke, C., P. Ekins and R. Costanza, 1994. Trade and the environment. *Ecological*

Economics 9: 1-92 (special issue).

Folke, C., A. Jansson, J. Larsson and R. Costanza, 1997. Ecosystem appropriation by cities. *Ambio* 26: 167-172.

Kox, H.L.M. and C.M. van der Tak, 1996. Non-transboundary pollution and the efficiency of international environmental cooperation. *Ecological Economics* 19: 247-259.

Kuik, O.J., and A.J. Gilbert, 1999. Indicators of sustainable development. In: J.C.J.M. van den Bergh (ed.). *Handbook of Environmental and Resource Economics*. Edward Elgar, forthcoming.

Kuik, O. and H. Verbruggen (eds.), 1991. *In Search of Indicators of Sustainable Development*. Kluwer Academic Publishers, Dordrecht.

Markusen, J.R., E.R. Morey and N.D. Olewiler, 1993. Environmental policy when market structure and plant locations are endogenous. *Journal of Environmental Economics and Management,* Vol. 24: 69-86.

Pearce, D.W., G. Atkinson and K. Hamilton, 1998. The Measurement of Sustainable Development. In: J.C.J.M. van den Bergh and M.W. Hofkes (eds.). *Theory versus Implementation of Sustainable Development Modelling.* Kluwer Academic Publishers.

Pezzey, J., 1989. *Economic Analysis of Sustainable Growth and Sustainable Development*. Environmental Department Working paper no. 15, Environmental Department, The World Bank.

Pezzey, J., and C. Withagen, 1998. The rise, fall and sustainability of capital-resource economies. *Scandinavian Journal of Economics* 100: 513-527.

Ploeg, S.W.F. van der, 1990. *Multiple Use of Natural Resources*. Ph.D. dissertation, Free University, Amsterdam.

Rennings and Wiggering, 1997. Steps towards indicators of sustainable development: linking economic and ecological concepts. *Ecological Economics* 20: 25-36.

Rotmans, J., and B. de Vries (1997*). Perspectives on Global Change: The Targets Approach.* Cambridge University Press, Cambridge, UK.

Serrão, E.A.S., D. Nepstad and R. Walker, 1996. Upland agricultural and forestry development in the Amazon: sustainability, criticality and resilience. *Ecological Economics* 18: 3-13.

Verhoef, E.T., and J.C.J.M. van den Bergh, 1995. A Spatial Price Equilibrium Model for Environmental Policy Analysis of Mobile and Immobile Sources of Pollution. In: J.C.J.M. van den Bergh, P. Nijkamp and P. Rietveld (eds.). *Recent Advances in Spatial Equilibrium Modelling.* Springer-Verlag.

Weizsäcker, E. von, A.B. Lovins en L.H. Lovins, 1997. *Factor Four: Doubling Wealth -*

Halving Resource Use, A Report to the Club of Rome. Earthscan, London.

Wackernagel, M. and W. Rees, 1996. *Our Ecological Footprint: Reducing Human Impact on the Earth*. Illustrated by Phil Testemale. The new catalyst bioregional series vol. 9. Gabriola Island, BC and Philadelphia, PA: New Society Publishers.

Wackernagel, M. and W. Rees, 1997. Perceptual and structural barriers to investing in natural capital: Economics from an ecological footprint perspective. *Ecological Economics*, Vol. 20(1): 3-24.