A summary of

A Critique of the Environmental Sustainability Index

Raghbendra Jha and K. V. Bhanu Murthy
(corresponding author email: r.jha@anu.edu.au)

The 2002 Environmental Sustainability Index (ESI) is a significant effort in the direction of developing an overall measure of environmental sustainability. It is based on the data of 142 countries. It has been developed by collaboration of the World Economic Forum, Geneva, Centre for International Earth Science Information Network, Columbia University, and Yale Centre for Environmental Law and Policy, New Haven. In the future, such a measure has the potential of seriously impacting domestic and international policy analysis. Hence, it is important that there be a widespread acceptance of the structure and methodology of the ESI.

Structure of the Index

The ESI has 5 broad indicators based on 20 core indicators. The core indicators are built-up from 68 variables. The broad indicators are classified as: -

1. Reducing Human Vulnerability
2. Environmental System
3. Reducing Environmental Stress
4. Social and Institutional Capacity Component
5. Global Stewardship

Clubbing of indices

There are at least three discernable standpoints. The focus in each standpoint is different.

(1) Environmental Degradation

(2) Effects of Degradation

(3) Environmental Management.

It is obvious that the first dimension affects the second. The third must be designed so as to affect the former two. How can a causal and an impacted indicator be clubbed into one grand index? Especially when the individual variables are causally linked and correlated. This is precisely what ESI does.

Problems

Classification of variables:

The classification raises several questions. For instance, it is not clear why variables like urban NO_2 and SO_2 concentration should be a part of the 'Environmental System' whereas NO_2 or SO_2 emissions per populated land area should be a part of 'Reducing Environmental Stress'. Had these variables been defined, as percentage fall there could be some reason to believe so.

Specification (type) of the variables:

Amongst other things, most surprisingly, 'percentage change in projected population between 2001 and 2050' has been specified as a variable for 'Reducing Environmental Stress'. How can a future (projected) rate of growth reduce present environmental stress? The variable 'proportion of under-nourished in total population' cannot be related directly to environmental degradation.

If the intention of the ESI is to be a 'near-complete' and an 'almost-correct' index, then a lot needs to be said about its coverage and correctness. For instance, under Environmental Governance certain variables that have been included are either antiquated

or politically incorrect. The variables that emphasize protected areas under law do ignore recent understanding on the subject of forest management. The approach that lays stress on enforcing protected forest is known as the 'fence and forget' approach. Not only are there real inroads, there is a fundamental change in thinking that has not been reflected. In real terms (though small), a significant part of forest management is coming under co-operative management of and by the local populations (tribal mainly). Not only is it more politically correct, it is also more appropriate because in large parts of the world the thorniest problem in environmental management is to reconcile the interest of people and nature, and such an approach tries to do so. In fact, the dynamics of management has gone much further. The concept management has already evolved to 'joint-management', that is, private, co-operative and government.

The broad category called 'Social and Institutional Capacity' is incomplete and, at least in parts, politically incorrect. The latter is because it ignores gender issues. The greatest inequity in forest management is in respect of gender. There is no variable that reflects this dimension. The indicator called 'Private Sector Responsiveness' has a corporate bias. It is biased towards industry against agriculture/forest. It is biased towards corporate governance against people's governance.

The code of variables:

But the serious lapse relates to the 'code' of the variables. For instance, in the core indicator 'Basic Human Sustenance' two variables are included:

(i) proportion of under-nourished in total population

(ii) percentage of population with access to improved drinking water

The problem with the 'code' is that if the index is low it favours the former variable and if it is high it favours the latter.

The first and foremost principle in formation of an index is that the sum total of the variable must yield an interpretation that is unidirectional. That is, the ‘code’ must be the same. Once the index is aggregated such differences would not be known to the users but would have serious implications for analysis. An important principle in formation of an index is that the sum total of the variable must yield an interpretation that is unidirectional. That is, the ‘code’ must be the same. In a composite index somewhat disparate variables are clubbed together. The final index must have a magnitude and direction, so that the index can be uniquely interpreted. Once the index is aggregated such differences would not be known to the users but would have serious implications for analysis. For the purpose of yielding such an interpretation a greater magnitude, for each variable, should mean betterment. For instance, in the core indicator ‘Basic Human Sustenance’ two variables are clubbed, viz., 'proportion of under-nourished in total population’ and 'percentage of population with access to improved drinking water'. The problem with the ‘code’, in this case, is that if the magnitude of undernourishment were low it would go in favor of ‘Basic Human Sustenance’. On the other hand, with the second variable if the magnitude were high it would go in favor of ‘Basic Human Sustenance’. Now if these two variables are combined with the help of some weights then how would we interpret the index? If a country has high values for both undernourishment and drinking water could we unequivocally treat it as being favorable for ‘Basic Human Sustenance’?

Methodological problem:

There are serious problems in respect of methodology also. Some of these problems are being considered:

(1) On measurement it can be seen that there is inter-correlation amongst variables which may be indicative of cause and effect amongst variables.

(2) Although they have specified certain implicit weights for the variables while constructing the index they have used equal weights. In the process they have ignored more scientific methods for weighting.

(3) For convenience the extreme values have been ignored.

Data, Methodology and Results

We have used an alternative method for constructing an Environmental Degradation Index (EDI). The data is taken from the 2002 ESI report itself. Certain variables relating to environmental degradation were considered. Then a set procedure laid down below was used for constructing the Environmental Degradation Index (EDI). The methodology, in brief, consists of the following steps:

1. Check for correlations amongst variables. If they are significant proceed with following steps that are drawn from Principal Component Analysis literature.
2. Determine component scores of variables with respect to principal components.
3. Adopt procedure for discarding less important variables and retaining principal variables.
4. Verify 'explanatory power' by given procedure. If satisfactory.
5. Construct composite EDI by using determined weights with selected variables.
6. Rank countries in ascending order of index for obtaining ranks of countries from least to most degrading (1 - 142).

Rationale for procedure:

A basic criticism of the ESI methodology is that it neglects the use of Principal Component Analysis. Their own justification is supposedly due to the low correlation

amongst variables[1]. The very nature of many of the variables is such that, many of them are closely related, if not causally related. It is quite telling that, out of 45 correlation ((10x10)- (10 (own correlation))/ 2), only 16 are *not* significant, at the 5% level. All the rest of the variables are highly correlated and statistically significant. This means that around two thirds of the variables are correlated (See Table 2).

It is fairly well know that if variables are correlated then PCA is ideally suited for such a situation. This is further confirmed even in cases where there are a large number of variables. The ESI is based on 68 variables. However since

" Principal Component Analysis (PCA), is a statistical technique that linearly transforms an original (*large*) set of variables into a substantially smaller set of uncorrelated variables, that *represents most of the information* in the original set of variables"[2] (emphasis added) the use of PCA is justified. We report results on PCA here.

Environmental variables are usually closely related. By working with a large number of variables, the estimate (of the ESI) is prone to the vagaries of the accumulation or compounding of reporting or measurement errors. The data on environmental variables is highly prone to such errors. The ESI document is itself replete with such allusions. While many other controversial methods have been used, PCA has been consciously avoided.

The environmental variables at the global level are also prone to have non-normal or skewed distributions. Here again PCA has an advantage. It does not need the normality assumption. While observing worldwide data the variance is likely to be very high. Here again PCA has the advantage that it does not have to explain the correlation (or covariance) amongst the largest possible ("fully specified") set of variables. It is very economical because it uses the

[1] We have tested their raw variables to confirm this (Table 2).
[2] See Dunteman, G.H. " Principal Component Analysis (PCA)", in Lewis-Beck (1994). P.158. For technical details see Appendix in this paper.

least number of variables to explain the full contours (of a widely spread) phenomenon by accounting for the maximum possible variance.

It is also economical because it minimizes the effort and time while achieving similar results. It reduces the cost of data collection. This is relevant especially, if the authors of ESI want it to be a model index for emulation. It can be sustained only if it is economical. Especially, if governments of poor countries are expected to collect bulky data from their own resources, the cost of collection becomes very relevant. These governments would either be dependent on donors (rich countries or Institutions) for funds for such purposes (which may have other implications) or would `cut corners` because of which data coverage, reliability and quality would suffer.

There are set procedures for scientifically selecting these variables from amongst many[3]. Such variables are known as "principal variables". There are certain measures that can be used for judging the utility ('explanatory power') of such variables. Define Total Variation Explained (TVE) = $n_r + \sum_{i \in d} R^2_{i,r}$ where the set d consists of all variables; n_r is the number of retained variables, and $R^2_{i,\,r}$ stands for the squared multiple correlation of the ith discarded variable with the *r* retained variables obtained by regressing each of the discarded variables on the four retained variables. The number of retained variables is added because each of the retained variables explains its own variation (variance = 1). The measure $R^2_{i,\,r}$ is summed over the discarded variables because it represents the variation in the discarded variables explained by the retained variables. Now, the (TVE/Total Variation) is a measure of the explanatory power. The measure of total variation is *d* x 1 = *d* (since the total variation can be 100% if all variables are included).

[3] See McAbe, G.P. (1984) "Principal Variables", Technometrics, 26: pp. 137-144.
And Jolliffe, T.I. (1986) *Principal Component Analysis*, New York, Springer-Verlag.

Another advantage with PCA is that, unlike Factor Analysis it does not have to assume any underlying hypothetical factors. Yet it is possible to have a meaningful interpretation with the help of select variables. *It only reduces the redundancy of data*. The method of rotation allows better interpretation while explaining the same amount of variance. For instance, it allows to pick up one air quality variable out of many, one bio-diversity variable amongst many, and so on. A spectrum of variables can be represented by a few.

We reduced the number of variables using three procedures: a) by eliminating some variables that have a `code' problem or are not very relevant; b) by sorting the variables into "Environmental Degradation Variables", "Degradation Effect Variables " and "Environmental Management Variables", which reduces the number in each category to around 15; and c) by using PCA to achieve data reduction.

Finally, the EDI was constructed on the basis of component scores. The ranks were established on the basis of ascending value of EDI. This was done to make the ESI and EDI comparable. The logic is, that a low value of EDI corresponds to a more sustainable environment, which can be represented by a higher value of the ESI. This makes the code of both comparable. Then the relative ranks of the 2002ESI and our EDI were compared (See Table 4).

Ideally, the rank correlation should have been (+) unity. This would have endorsed that there is no flaw in the estimation of 2002ESI. The computed EDI and the existing ESI would have been two different ways of saying the same thing. For verifying the correlation we measured the rank correlation. It was only 0.1067, that is rather low in value, and the Z value was only 1.2. Hence, the rank correlation was not significantly different from zero. When we test to see whether it is close to the expected value +1, the Z value is -10.60629. This also confirms that the rank correlation is well (statistically significantly) below +1.

There are wide differences in the ranks of many countries giving anomalous results. For instance, Australia has a difference of (-) 119 in rank (ESI minus EDI). This means that it is highly sustainable and extremely degrading (both simultaneously)!! Only very few countries retain the ranks. On the other hand Guinea-Bissau is hardly degrading but almost unsustainable!! This is true of most poor countries. Most of the rich countries have extremely high vehicular traffic and pollution and are by the EDI highly degrading but are fairly sustainable by the count of their ESI.

In conclusion, it may be said that there are only three logical categories of broad indicators. Any other division is not tenable. Finally, these three indicators cannot be collapsed into one index. There is a causal chain amongst the three indicators. The fallacy is obvious. If a country has a high index of sustainability, as things stand, it implies that it has a high degree of degradation, severe degrader effects and better management as well!!