Ecological Economics 198 (2022) 107470

Contents lists available at ScienceDirect

Ecological Economics

journal homepage: www.elsevier.com/locate/ecolecon

Sustainability and Human Development Indicators: A Poset Analysis

Flavio Comim [a,*], Tadashi Hirai [b]

[a] IQS School of Management, Ramon Llull University and Land Economy, University of Cambridge, Via Augusta, 390, Barcelona 08017, Spain
[b] Centre Development Studies, University of Cambridge, 7 West Road, Cambridge CB3 9DT, UK

ARTICLE INFO

Keywords:
Sustainable Human Development Indicators
Sustainability
Human Development Index
Capability Approach
Poset Analysis

ABSTRACT

The paper introduces the methodology of partially ordered sets to evaluate sustainable human development indicators. It shows its importance to the integration of human and environmental indicators, clarifying the limitations of composite indicators and emphasising the need to move towards higher levels of conceptual consistency and solid informational pluralism, incorporating richer social and environmental spaces towards an analytical structure that avoids subjective, arbitrary and potential obscure choices of variables in the analysis of sustainable human development. The paper uses Hasse diagrams and sensitivity analyses to evaluate the evidence offered by the new pressure-adjusted human development index (PHDI). The results indicate that the introduction of CO_2 and material pressure indicators produce an excessive number of incomparabilities and affect the role of GNIpc in the determination of countries' partial rankings.

1. Introduction

A new chapter has been written in the history of integrating environmental dimensions into the Human Development Index (HDI): UNDP (2020) has officially proposed a new pressure-adjusted human development index (PHDI). The index combines indicators of greenhouse gas emissions and material footprint that are used to adjust human development achievements in relation to the environmental pressures that they produce. By doing so, UNDP distils a 20-year academic dispute on how 'to green' the HDI (Morse, 2003; Hirai, 2017; Kalimeris et al., 2020) and consolidates a reorientation of the political objectives of the human development agenda towards sustainability goals.

During this time much of this debate has concentrated on the choice of suitable environmental variables used to complement the human development picture, focusing mostly on pollution indicators (e.g. air, soil and organic water pollutants) and on natural resource consumption indicators (e.g. energy and minerals, forest coverage, arable land, water withdrawals). On the side of pollution, CO_2 emissions has often been used either as a unique environmental factor alongside HDI dimensions (De La Vega and Urrutia, 2001; Togtokh, 2011; Bravo, 2014; Biggeri and Mauro, 2018) or as part of wider environmental accounts (Tarabusi and Palazzi, 2004; Costantini and Monni, 2005; Hermele, 2010; Pineda, 2012; Jain and Jain, 2013; Türe, 2013; Ray, 2014; Xiao et al., 2018; Liang et al., 2019; Hickel, 2020; Jin et al., 2020). Similarly, on the side of natural resource consumption, material footprint and forest coverage have been the most used factors (Costantini and Monni, 2005; Hermele, 2010; Jain and Jain, 2013; Türe, 2013; Ray, 2014; Xiao et al., 2018; Liang et al., 2019; Jin et al., 2020). Thus, it seems that in the formulation of the new PHDI, UNDP has just considered two sets of factors already established in the literature.

In addition, distinct academic contributions have put forward different formulas to tackle environmental variables in the formulation of a 'green' HDI. Whereas the majority suggested to include environmental variables as independent dimensions (Tarabusi and Palazzi, 2004; Hermele, 2010; Jain and Jain, 2013; Ray, 2014; Bravo, 2014; Biggeri and Mauro, 2018; Jin et al., 2020), there is a well-established procedure to adjust the HDI by the use of a discounting formula (e.g. De la Vega and Urrutia (2001) and Costantini and Monni (2005) discount gross domestic income and Pineda (2012), Türe (2013) and Hickel (2020) discount the overall HDI). Some others such as Distaso (2007) and Liang et al. (2019) added environmental variables to correct specific HDI variables, whereas Xiao et al. (2018) opted for adding the environmental dimension on top of the overall HDI. On the same line, Neumayer (2013) conceptually suggests complementing the HDI with genuine savings and the ecological footprint as measures, respectively, of weak and strong sustainability. Chhibber and Laajaj (2008) also put forward a proposal for including two kinds of environmental costs, namely, those inflicted to countries' own nature and people through air pollution, soil erosion and poor water quality and those imposed on other countries through CO_2 emissions, ocean acidification and

* Corresponding author.
E-mail addresses: fvc1001@cam.ac.uk (F. Comim), th299@cam.ac.uk (T. Hirai).

https://doi.org/10.1016/j.ecolecon.2022.107470
Received 21 September 2021; Received in revised form 5 May 2022; Accepted 7 May 2022
Available online 17 May 2022

biodiversity loss. Within this context, UNDP's proposal of using a planetary boundaries framework (Rockström et al., 2009a, 2009b; Steffen et al., 2015) allows the PHDI to dialogue with all these issues.

Another line of research defines sustainability as the relative efficiency of converting resources in terms of human development outcomes (Assa, 2021; Mahlberg and Obersteiner, 2001; Despotis, 2005). More concretely, the use of data envelopment analysis (DEA) and the triple Index approach has from its inception been concerned with the issue of converting income, wealth and human capital into quality-of-life indicators (Mariano and Rebelatto, 2014; Mariano et al., 2015; Morais et al., 2013). More recently, Ferraz et al. (2020) expanded this framework by introducing the Capability Index Adjusted by Social Efficiency (CIASE) and Panagiotis and Giannis (2021) developed the Value Efficiency Analysis (VEA) to examine the effectiveness of countries in translating income into key human capabilities. Whilst this is a much valuable and welcome field, it leaves some loose (conceptual and empirical) ends, in particular in regard to the lack of reference of environmental elements and the need for a broader and more analytically rigorous approach to confront the incomparabilities and tensions between different dimensions of the HDI.

Given this context, the objective of this paper is to introduce a poset analysis, namely, a tool for working with partially ordered sets, that is, with rankings that are not complete (Fattore and Bruggemann, 2017). As much as we engage with conceptual and empirical issues, our main research question is essentially methodological. We apply an empirical assessment tool advocated by Sen, 1970a, 2017 to examine the newly proposed PHDI. This can be justified on four grounds. First, it offers a more solid conceptual use of the capability approach in the analysis of sustainable development indexes; secondly, it faces the key empirical issue of incomparabilities among different dimensions that is part of the debate between weak versus strong sustainability; thirdly, it empirically scrutinises the newly proposed PHDI, examining how it changes previous HDI rankings and demonstrating, fourthly, it shows how differences in rankings can also happen at all levels of human development. To our knowledge, this is the first investigation to explore the use of poset analysis to the key and recurrent issue of integrating sustainability issues into the HDI.

This paper critically addresses key vulnerabilities in the elaboration and analysis of sustainability and development indicators, namely: the addition of variables without assessing their informational contribution to the composite indicator, the elaboration of composite indicators without a solid conceptual structure and the usual dismissal of incomparabilities in merging environmental and social variables in understanding the main messages behind these composite indicators. We make a case for poset analysis on the grounds that it best reflects Sen's capability perspective and that it translates a conceptual weak vs strong sustainability debate, based on a priori arguments, into an evidence-based general approach informed by Hasse diagrams.

The paper is organised into five parts. The first part provides a conceptual analysis of the use of the CA in examining sustainability issues. The second part introduces the poset analysis and the terminology used to navigate within diagrams known as Hasse diagrams. They are very useful to display the trade-offs among different informational spaces and can supplement aggregate indicators, showing in the case of the HDI how comparabilities vary among different groups of countries. The third part presents the data and software used in the paper. The fourth part examines the key poset results of combining social and environmental indicators. The fifth part discusses the main implications for the elaboration of sustainable human development indicators and offers some concluding remarks. In particular, we show how neither the efficiency approach nor the PHDI are sufficient answers to solve the problem of integrating sustainability indicators within the capability approach.

2. Sustainability and the Capability Approach

The capability approach is one of the perspectives informing the human development paradigm. It has been developed by Amartya Sen, Martha Nussbaum and a rich and diverse community of capability scholars during the last decades (Robeyns, 2017; Comim, 2018). Starting from his famous "Equality of What" (1979) lecture till the recent revisions of his social choice work (2017, 2020), Sen has linked the approach, first and foremost, with his defence of informational pluralism in normative evaluations. More specifically, he sees the CA as a key strategy for broadening the informational basis of social evaluations within the context of social choice theory (Qizilbash, 2007; Comim, 2018). Sen argues that the CA (2009: 232) "does not, on its own, propose any specific formula about how that information [on capabilities or other informational spaces] may be used" nor any blueprint about how to deal with conflicts. In this regard, he emphasises that we should not expect resulting indicators to produce complete rankings (Sen, 2017). The main reason for that, and his consequent defence of partial rankings, is that different people may not agree on the importance of distinct outcomes and as a result, the best we can do is to look for some agreed common basis among them (Sen, 2002). His argument clearly reflects his advocacy for democratic consensus. It is not an exaggeration to claim that Sen has been arguing for the use of partial rankings for the last 50 years (Sen, 1970a, 1970b, 2017). What is remarkable is that this key aspect of CA remains fully ignored in the debates about the HDI.

Whereas in complete rankings we can always compare any pair of alternatives (indicating whether one is superior to another or whether we are indifferent between them), in partial rankings this total comparability is not secured, but it is possible to find grounds for some comparability between distinct options. Thus, Sen (1970b) introduces the term partial comparability as a generic term to be used for every case of comparability lying between full comparability and non-comparability. In this sense, partial rankings might reflect a minimum standard, a lower limit of comparability among alternatives that can be asserted without contradiction of any other rankings. Rankings can then be different in terms of their specific ordering of options and as how they assemble different criteria. From this perspective, the analytical possibility of incompleteness in ranking alternatives is not an embarrassment but rather freedom from what Sen has called "the tyranny of required completeness" (Sen, 1985: 20–21).

Alternatives or options are often represented by different informational spaces. An informational space is a class or category of information that refers to a common conceptual basis, such as utilities, resources, rights, primary goods and capabilities, among others. The capability as an approach is pluralist and as such it may include all relevant informational spaces. This is different from capability as an informational space that represents specific vectors of functionings (different beings and doings of an individual, a country, etc). In the case of the HDI, the index is composed of two kinds of informational spaces, namely functionings (health and education) and resources (standard of living). The main advantage of using functionings or even capabilities as informational spaces is that functioning indicators are objectively defined, avoiding biases and distortion metrics problems associated with conversion of resources into valuable doings and beings and the use of subjective evidence in welfare economics (Sen, 1985, 1996). We can argue that there are three essential tenets to Sen's approach (Sen, 1993, 2000, 2009, 2017):

1. informational pluralism: it is about the admissibility of different informational spaces (such as resources, subjective well-being, rights, primary goods, functionings and capabilities) in normative assessments;

2. comparative analysis: it concentrates on the practical reason behind choices and avoids using ideal objectives as reference points; rather, it focuses on alternatives that are possible to be contrasted;
3. reasoned scrutiny: it makes room for further scrutiny and reassessment; it takes into account that people's views depend on their positionality and as such can reflect their different interests. This puts public reason and public scrutiny at the centre of the process of normative evaluations. Without informed debates and interactive discussions, it is difficult to achieve impartial results and reflected evaluations.

The road of informational pluralism is paved with partial rankings because the sheer complexity and practical difficulty of defining what is individually and socially best imposes limitations that should be acknowledged. Partial rankings are at the centre of Sen's CA because they articulate informational pluralism required for comparative analysis and reasoned scrutiny. They offer an operational tool to face the difficulties and challenges in trying to achieve consensus about normative evaluations. When full agreement cannot be achieved for all options, the best we can do it to achieve partial resolutions in comparing alternatives. This is essential for the discussion about bringing together social and environmental aspects of sustainable human development. Quite often the processes of integration of social and environmental indicators in the context of the HDI entails a view of weak sustainability[1] in the sense that they allow trade-offs between distinct variables, being "inconsistent with the principles of ecology, and inconsistent with the planetary boundary framework", as argued by Hickel (2020).

At the core of this issue is the concept of incomparability. By incomparability we understand a situation that happens when "the order of objects due to one indicator contradicts the order of another indicator" (Bruggemann and Patil, 2011: 5). Indeed, incomparabilities between human and environmental aims might turn geometrical means and discount formulas insufficient to tackle issues about what should be relevant for an operational definition of value (Biggeri and Mauro, 2018). It is important to distinguish between social and technical incomparabilities because one thing is the irreducibility among different visions about what is important in life and another about technical difficulties in integrating a broad set of various and conflicting points of view (Munda, 2004). The problem is that the concept of sustainable development, in its different formulations, tends to amalgamate different viewpoints without making explicit their tensions and conflicts. This is not the case of indicators such as the Genuine Progress Indicator (GPI) more focused on consumption of resources, monetary outcomes and sustainable incomes, but it becomes the case whenever environmental factors have not only an instrumental but a constitutive (intrinsic) role (Ramos et al., 2018; Kalimeris et al., 2020; Fox and Erickson, 2020), as it happens with the PHDI.

The demands of sustainability are also closely associated with the claims for intergenerational justice and impartiality towards future generations that can be affected by overexploitation of natural and environmental resources (Anand and Sen, 2000). Although from a capability perspective, all resources should be seen as imperfect indicators of human well-being (Sen, 2009), in the context of sustainability, environmental resources can be understood as proxies for a general concern with intra and intergenerational equity, given that what really matters are the freedoms that current and future generations will have to lead worthwhile and meaningful lives. We should also remark that freedoms from pollution and other forms of environmental degradation (such as natural hazards) are also important to be considered. The demands of sustainability call attention to possible incomparabilities between human and environmental factors for the current and future generations.

The use of partial rankings, together with informational pluralism, comparative analysis and reasoned scrutiny are necessary elements in grounding the creation of human and environmental indicators on Sen's CA. In this paper we emphasise the use of partial rankings, but this does not mean that the other elements should not be kept in mind. In what follows we move from this conceptual background into operational tools such as the use of posets and Hasse diagrams as part of a capability perspective.

3. Poset Analysis: Key Concepts

Proposals for integration of environmental dimensions in the HDI have regularly ended up with the elaboration of composite indicators (CIs). Their use has evident potential merits. They provide a more comprehensive picture than the use of single indicators, offer a normative benchmark for handling multidimensionality and might facilitate political action based on evidence. Indeed, it might even sound strange prima facie to criticise CIs because they often deliver more comprehensive accounts of the reality. Within this context, bringing on board incomparabilities among different variables seems to be positively related to plural and richer informational spaces rather than to a limitation. However, there are different levels of incomparabilities within and between rankings that could result in CIs concealing important differences among its elements. For instance, in the case of the HDI, countries can be perceived as successful because of their high rates of economic growth rather than because they are able to translate their affluence in health and education achievements to their populations. More importantly perhaps, we might not even realise the presence of trade-offs derived from conflicting variables when comparing countries by their overall HDI. One example might be helpful here. The United States has a GNIpc of USD 83.826 and compared with the UK's GNIpc of USD 46.071 (both in 2017 PPP) is clearly in advantage. But if we include another attribute, the CO_2pc (in tonnes, 2018), the US produces a value of 16.6 t whereas the UK produces only 5.6 t. In this last case, the UK is in a much better situation than the US. We could also mention that in the UK people have higher life expectancy at birth than in the US (81.3 versus 78.9) but that the US has higher mean years of schooling (13.4 versus 13.2 years). As much as collapsing all variables into a single composite indicator can be politically helpful to call attention to the importance of the planetary pressures in shaping human development, looking at their incomparabilities might provide a richer picture.

One alternative, such as the one advocated by UNDP (2020), is the use of a dashboard approach as a way of complement the use of composite indicators. Nonetheless, this exercise is not without their own practical challenges because the interpretation of many indicators in a dashboard might be difficult without a clear notion of weights or relevance of the separate pieces of information. These difficulties might even be magnified in the particular case of the HDI and PHDI that are calculated for a large number of (189) countries. It seems that we are methodologically left to choose between two evils: the first represented by CIs that are criticised (Fattore, 2016; Barclay et al., 2019) for being extremely subjective, arbitrary, potentially misleading and prone to obscure essential information, hiding key qualitative differences among different indicators and the second epitomised by dashboards that leave all value judgments and complexity at the discretion of their users. But poset (partially ordered sets) analysis might provide a fruitful path for overcoming this concrete empirical dilemma and in so doing, it could give concreteness to CA's defence of partial rankings (Comim, 2021). This argument does not depend on whether HDI's informational spaces are actually characterised by capabilities, functionings or mixed dimensions, but rather on the role of the capability approach as a key

[1] The analytical categories of comparability and non-comparability are not the same as the ones for weak and strong sustainability. Whereas the first group refers to how we analyse indicators, the second is much broader, including actions and policies. For instance, society can move forward towards stronger notions of sustainability even under the circumstances that academics maintain full comparability in their composite indicators. We are grateful to one of the anonymous reviewers for alerting us to this important distinction.

element of Sen's social choice approach (Hirai, 2021; Sen, 2017).

The creation of CIs (by adding dimensions or by using discount factors) entails the transformation of different attributes into a common metric and the generation of complete orderings. In simple terms, a complete ordering is generated whenever every pair of alternatives can be ranked against each other. In the case of two alternatives x and y we could say that either x is better than y or x is equal to y or x is worse than y. Mathematically and empirically speaking, complete orderings can be seen as particular cases of partial orderings (Bruggemann and Patil, 2011; Fattore and Bruggemann, 2017). In the case of the HDI, a complete order means that we can always compare the situation between two countries. In other words, we could say that either a country A is better than a country B, or a country A is worse than a country B or simply that they are equal (or have the same performance). Countries can be defined as objects in a set X of countries and evaluated according to a set of attributes Q that would include variables q1, q2, q3, etc., such as those related to health, education, income or the state of the environment. Thus, the situation of a country can be defined by a vector of attributes that can be ordered as a binary relation to other countries. These comparisons are normally defined by a set of properties such as reflexivity (one object can always be compared with itself), antisymmetry (if a country A is better than a country B and a country B is better than a country A this means that the countries are identical) and transitivity (if a country A is better than a country B and a country B is better than a country C this implies that a country A has to be better than a country C).

Thus, whenever we can say that one country is better or worse or equal than another, we can claim that they are comparable, and when this does not happen, we say that they are incomparable (for a given attribute). The important point about CIs is that they can hide incomparabilities found among their different dimensions. How should we assess the development of a country, vis-à-vis another, with (for instance) a higher GDPpc but a higher pressure on ecosystem services? Shall we collapse the different dimensions into a common metric and pretend that these incomparabilities does not exist? The interesting point to explore here is that there is a partial[2] equivalence between the categories of comparability and incomparability and the concepts of weak and strong sustainability given that they all assume a range of substitutability options (Neumayer, 2001, 2013; Jain and Jain, 2013). To the extent that they overlap, both comparability and substitutability are grounded on notions of commensurability of different dimensions (Nussbaum, 1990). What is important about the poset analysis is that it can translate a conceptual debate justified by a priori arguments into an evidence-based general approach. In other words, in the case of a poset analysis both alternatives (comparability and incomparability) can be tackled simultaneously.

Following Bruggemann and Patil (2011) and Fattore et al. (2012) we define a chain or a linear order as a partial order P where two elements are comparable and an antichain when this is not the case. In an extreme situation where all objects could be arranged in a simple sequence, we would have a complete order. But this is not what normally happens. In particular, when we bring together environmental and social variables there are several conflicts among the different dimensions that are either hidden by CIs or left open with dashboards. Alternatively, the use of posets allows us to empirically scrutinise trade-offs between different variables and dimensions. More concretely, we can consider all different social and environmental attributes of different countries and represent them as sets of ordinal data. The profile of a country can then be represented by a collection of rankings for all the different attributes. The partial order of the profile of a country would indicate that this profile would be equal or worse than the profile of a different country when all its attributes are equal or lower than the attributes of the other country.

In the case of the HDI, countries that have better attributes than others can be organised in chains. We can then have a ranking of countries where countries that are above in the ranking are those that have better values for all attributes than the subsequent ones. These chains have heights that are defined by the number of objects (in the case of HDI they are countries) in their ranking. But when countries cannot be compared, for instance, because they perform better in one attribute and worse in another, we say that we have an antichain. If we consider the number of elements in the largest antichain we can then define the width of the poset. Sometimes it is useful to have a visual representation of these partial orders. For this reason, we use a Hasse diagram. We can define a Hasse diagram as a visual representation of all combinations of partial rankings, with their comparabilities and incomparabilities (Caperna and Boccuzzo, 2018). The main advantage of using this representation is that it provides an overall picture of all comparabilities and incomparabilities in sets of objects and their corresponding attributes. It can provide evidence about the level and density of incomparabilities found among different variables. Whenever objects (in the case of HDI and PHDI we refer to countries) can be compared they are connected by a line, which means that they are also transitive. We call successors of particular elements in a Hasse diagram downsets, as we call their predecessors upsets.

In the example we discussed above between the US and the UK they would be clearly in a position of an antichain, given that their attributes are not fully comparable. But if we were to include Switzerland in this comparison with higher (or equal) values for all attributes (GNIpc of USD 69.394, life expectancy at birth of 83.8, mean years of schooling of 13.4 years and CO_2pc of 4.3 t), we could have a chain with Switzerland on top and the UK and the US on the second line. We could expand this ranking by including for instance Saudi Arabia that has a poorer performance on all variables mentioned here (GNIpc of USD 47.495, life expectancy at birth of 75.1, mean years of schooling of 10.2 years and CO_2pc of 18.4). The resulting Hasse diagram would have 3 levels (heights) (1.Switzerland, 2.US and UK and 3.Saudi Arabia) with a width of 2 because of the incomparabilities between the US and the UK.

Finally, it is important to acknowledge that not all attributes (social or environmental) play the same role in the characterisation of a given state-of-affairs. Some attributes can make conceptual sense but can provide little, if any, comparative or discriminatory power. For this reason, it is useful to consider attribute-related sensitivity measures that can show how an attribute influences (locally, around an object or globally, for the entire poset) the position of objects in a Hasse diagram. This involves a comparison between several posets obtained from different data matrices originated from distinct sets of attributes (Bruggemann and Patil, 2011). Measures of distance are conceptualised by counting ordinal changes between any two pairs. In the case of the HDI, we can count the ordinal differences that countries have in their GNIpcs and compare with the ordinal difference that they have in another variable, say, life expectancy. The main point here is about comparing ordinal rankings corresponding to distinct informational bases. Thus, measures of sensitivity can be obtained by counting the mismatches using downsets.

4. Data

The capability approach is quintessentially pluralist. This means that it invites the use of a wide variety of types of data. In principle, one could use the CA to combine spaces as diverse as resources, subjective well-being and rights, without any reference to capabilities as informational spaces. In the case of the HDI, two dimensions, namely, health and education, are used as proxies for capabilities whereas one dimension, represented by GNIpc (gross national income per capita) indicates levels of resources. In the elaboration of the new PHDI, additional measures of

[2] The equivalence between these concepts is only partial because the concepts of weak and strong sustainability are about technical possibilities that might not be fully reflected by country orderings, given that public choices are also influential at a point in time. We are grateful to one of the anonymous referees for raising this point.

	CO_2	matfoot	lexp	eys	mys	gni
CO_2	1					
matfoot	0,516	1				
lexp	-0,487	-0,616	1			
eys	-0,458	-0,605	0,799	1		
mys	-0,517	-0,607	0,771	0,813	1	
gni	-0,722	-0,792	0,726	0,705	0,698	1

Fig. 1. Correlation matrix among key variables.

CO_2pc (carbon dioxide emissions per capita) and material footprint per capita are used to denote the challenges of shifting away from fossil fuels and closing material cycles (UNDP, 2020). They are transformed in indices and their arithmetic mean represents an adjustment factor that multiplies the HDI. In the exceptional case that a country puts no pressure on the planet, there is equality between their HDI and PHDI; otherwise, the PHDI falls below the HDI (i.e. by discounting it) in the proportion of this pressure.

In this paper, disaggregated HDI data per dimensions and per country was obtained for 2019 from the table of the HDI and its components (UNDP, 2020). Data about life expectancy at birth (in years - *Lexp*), expected years of schooling (in years - *eys*), mean years of schooling (in years -*mys*) and GNIpc (2017 PPP, purchase power parity, in US dollars -*gni*) were compiled for 189 countries. Similarly, data for the dimensions of the PHDI was obtained for 2019 from the table of the PHDI for the same set of countries (UNDP, 2020). In this case, data for CO_2pc and material footprint per capita were not used in tonnes but in the format index that inverts their scales. By doing so they are harmonic with HDI variables and thus the interpretation of the results becomes more intuitive (i.e. the higher their figures, the better the situation that they describe). A quick look at the resulting correlation matrix (Fig. 1) discloses three main patterns: first, it displays a positive association between material footprint (*matfoot*) and carbon dioxide emissions per capita (CO_2); secondly, it highlights how these environmental elements show a strong negative association with all other economic and social variables and finally, it indicates how all these last sets of variables are also strongly positively associated among themselves.

The software PyHasse, developed by Professor Rainer Bruggemann, was used in its "pro" version, kindly shared by him. More specifically, the modules mHDC17_3.py and sensitivity24_5.py were used to produce the Hasse diagrams and the corresponding similarity analysis. When all attributes (variables) were merged we had complete data for 169 countries which were represented in the Hasse diagrams by three-digit ISO codes.

5. Results

Both the HDI and the PHDI are composite indicators and as such lump together incommensurable features of countries, hiding conflicts among their different dimensions. To a certain extent, the complete rankings that they provide are artificial once they hide these conflicts and incomparabilities. This problem is particularly salient considering the idiosyncratic nature of the integration between social and environmental variables and the unsolved controversies about weak vs strong sustainability. However, the use of a Hasse diagram can unveil with precision all chains and antichains involved in ranking the countries based on their multidimensional profiles and clarify where conflicts are more relevant for an evidence-based characterisation of sustainable human development.

Fig. 2 displays the Hasse diagram for the 2019 HDI. Several results are worth noticing. First, even before considering the integration of environmental variables, there are quite a few incomparabilities among the HDI variables. Secondly, this means that the official UNDP classification of countries within four groups (very high human development, high human development, medium human development and low human development) oversimplifies their differences, dismissing the complex picture that emerges from the consideration of incomparabilities between different attributes. Thirdly, the density of incomparabilities is not homogeneous along the diagram. We can see a clear picture of the distribution of these incomparabilities per level in Fig. 3. One could (correctly) object to the use of a measure of total number of incomparabilities on the grounds that it could be promoted by a higher density of number of countries placed at these intermediary levels (the higher the absolute number of countries in a given level, the higher the absolute number of incomparabilities). To solve this problem, we can consider the average number of comparabilities. A comparison between the total and the average number of incomparabilities confirms the evidence of a heterogeneous (almost in the format of an inverted U) density of incomparabilities along the poset.

This Hasse diagram shows that all those countries on its first line, from Austria to the UK (GRB), are the ones with the best attributes but that they all have incomparabilities among themselves and for this reason they are not fully comparable. This does not happen, however, when we look at the countries on the second line, linked by a line. This is, for example, the case of Belgium and Canada, but not of the US and the UK which are not linked by a line despite their ordering. Countries at the bottom of the chain in the Hasse diagram are those with the worst set of attributes. Whereas the best attributes situate the countries in the higher categories, the worst ones place them in the lower categories of the Hasse diagram. The attributes are also defined by their power of discrimination, as explained by the sensitivity analysis below.

Overall, this Hasse diagram includes 9526 relations of comparability and 4670 relations of incomparability. A simple spearman correlation matrix suggests that the associations between all variables are very high (for instance, the correlation between $r_{Lexp\&eys} = 0.839$; $r_{Lexp\&mys} = 0.761$; $r_{Lexp\&gni} = 0.87$). However, this does not imply that one variable can be used to represent others, least that these correlations imply a harmonious picture. Rather, it suggests that we should be more cautious in comparing countries that are in the middle of the distribution vis-à-vis those at the top or at the bottom.

An important question to answer is how do the different attributes (variables of the HDI) influence the position of countries in the HDI Hasse diagram? In other words, what would be the impact of the removal of a specific variable from a data matrix? Fig. 4 displays the results of the global sensitivity analysis of the HDI indicators. It shows that mean years of schooling (mys) and life expectancy at birth (Lexp) are the key influential attributes in defining the ranking of countries in the HDI Hasse diagram. This corresponds to the original purpose of the index, meant to convey more people's capabilities than their available resources (Hirai, 2017; Sen, 2009, 2017; Anand and Sen, 2000).

The inclusion of the environmental factors, namely, CO_2pc and material footprint per capita, as part of an adjustment factor in the PHDI,

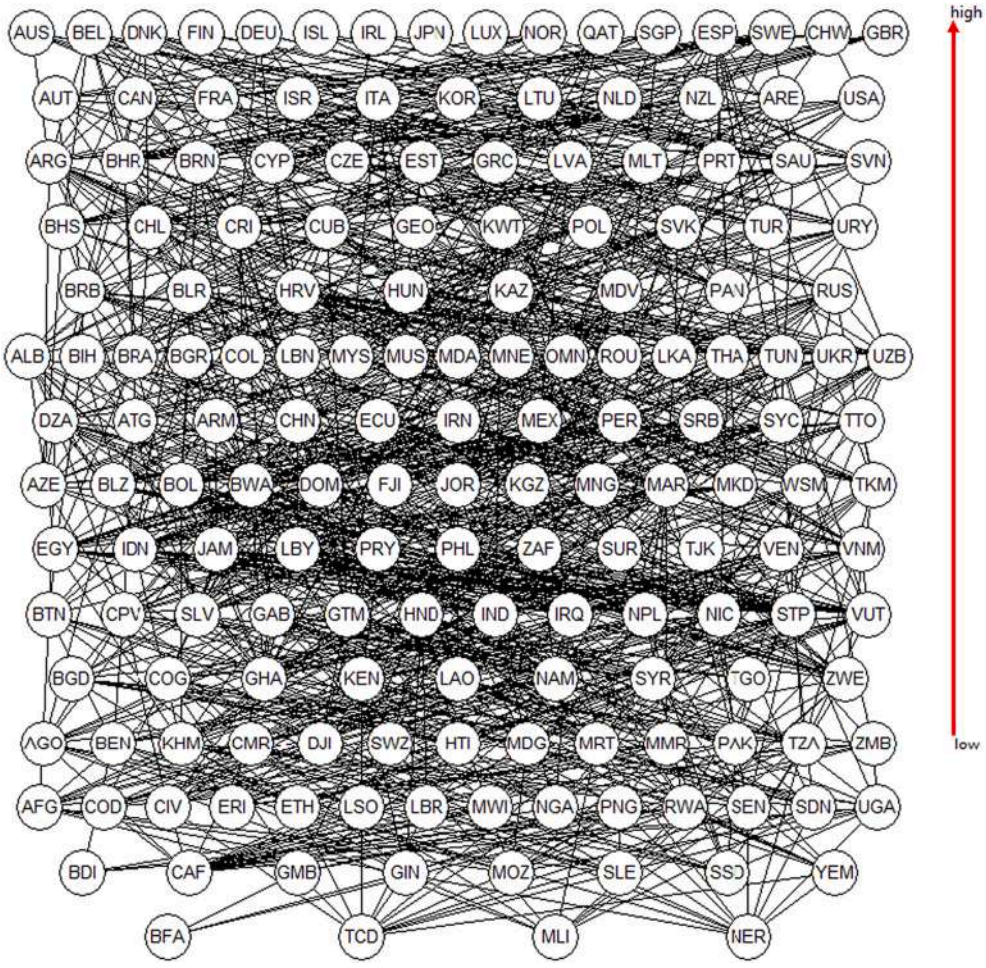

Fig. 2. Hasse diagram of 2019 HDI variables.

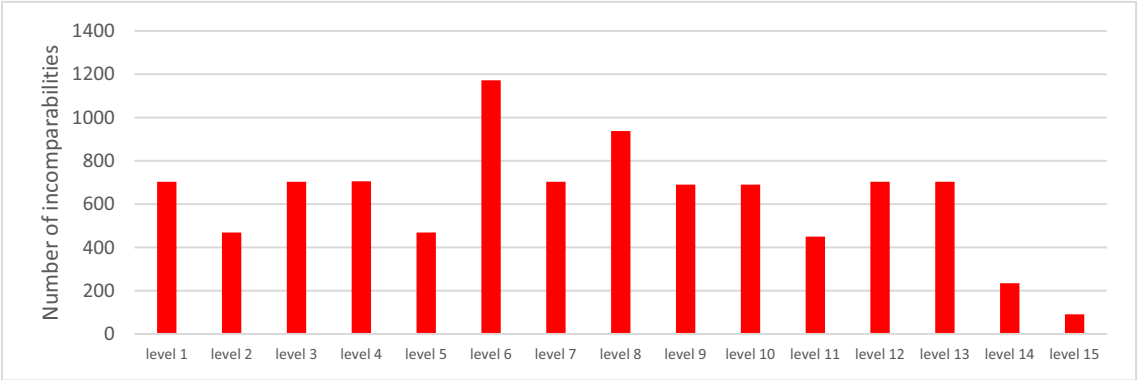

Fig. 3. HDI: Incomparabilities per level.

changes radically the ranking among countries. The number of comparability relations among the different attributes and countries is reduced to mere 138 (from 9526) and the number of incomparabilities jumps to 14,058 (from 4670). As a result, the Hasse diagram loses its power of discrimination; the antichains are dominant and the height of the posets is reduced from 15 to 3 levels (Fig. 5). The associations between human and environmental variables are negative and very high, ranging from $r = -0.753$ between CO_2 and Lexp to $r = -0.903$ between CO_2 and GNIpc. The correlation between CO_2 and material pressure (matf) is $r = 0.799$. But the overall degree of incomparability goes to 0.99.

Fig. 6 shows how the inclusion of the environmental variables provides an asymmetry in the incomparabilities, making the situation even more complex at the upper level. This differs from the almost inverted U shape (Fig. 3) in the case of the HDI without environmental variables.

Regarding the influence of each variable in the PHDI by means of the global sensitivity analysis, it is remarkable how the environmental

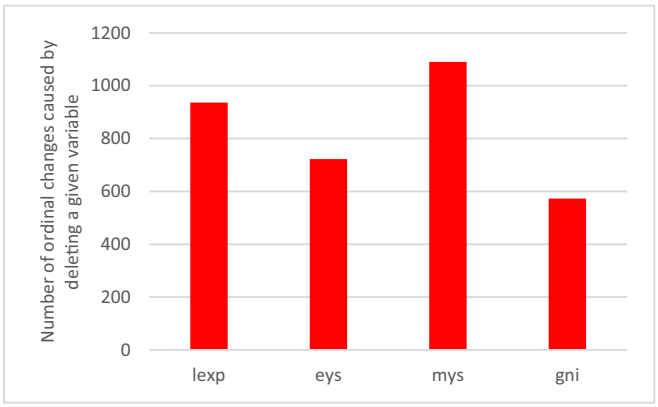

Fig. 4. HDI sensitivity analysis.

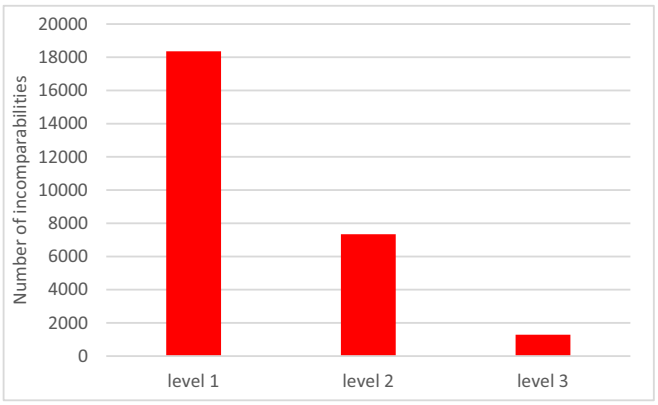

Fig. 6. PHDI incomparabilities per level.

variables exhibit a very high discriminating and comparative power in the reordering of the previous HDI ranking. Even the scale of the sensitivity values changes, moving from a maximum of 1100 for the HDI (Fig. 4) to a maximum of 5950 for the PHDI (Fig. 7). The variables CO_2 and material pressure influence much more the position of countries in the Hasse diagram than the other human variables previously did. They also promote a much higher number of internal conflicts in the data. Moreover, there is a key undesirable result that takes place in the PHDI poset: the gni becomes the third most influential factor, perverting the original purpose of the HDI of emphasising not the means (as resources) but the ends (as health and education) of development. As a consequence, income has more comparative power in the PHDI than in the HDI.

If incomparabilities are caused by conflicting indicators, it is likely that they will change (in density or levels and distribution) as we consider alternative sets of indicators. We illustrate this possibility for the 6 variables (attributes) that compose the PHDI. Fig. 8 presents the results of producing Hasse diagrams based on different combinations between socio-economic and environmental variables. If we read these different Hasse diagrams in a 3×3 matrix format, we can see that at 3×1 (the element at the third row and first column of the Hasse diagram) there is a reproduction of the HDI diagram that we can use as a benchmark (Fig. 2). All the other Hasse diagrams include either CO_2 emissions or material pressures in order to show the impact on the rankings that would be caused by incorporating an environmental variable into different readings of the HDI. It is interesting to note that when we combine both environmental variables (as displayed at the third column), the height (level) of the poset decreases up to 3 (as we can see at 1×3 and 2×3) with the maximum number of incomparabilities at 14000. Nevertheless, when we consider only one of them (as in the cases of 1×1 and 1×2) the height of the poset increases to 5 and 4, respectively. These results reinforce the early conclusions that environmental variables enhance the discriminatory powers of the rankings.

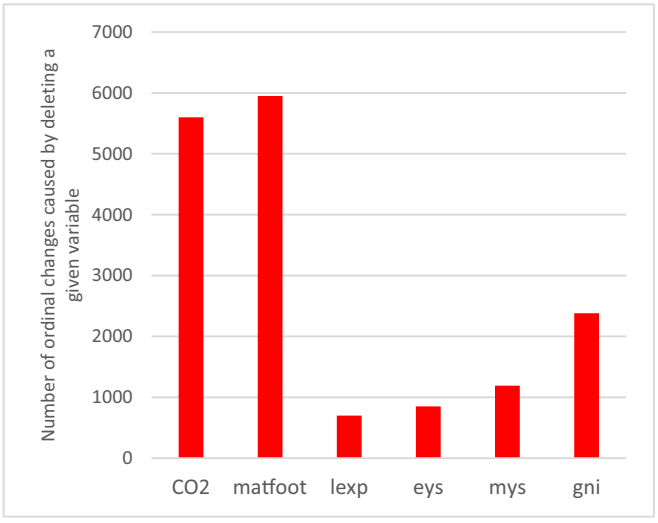

Fig. 7. PHDI sensitivity analysis.

However, should we take seriously the commitment of grounding the HDI and the PHDI on the CA, there is a potential problem in increasing the magnitude and distribution of incomparabilities within the inclusion of CO_2 and material pressures indicators.

Such information is usually hidden when composite indicators present their evidence in the format of complete rankings, as it is the case with the PHDI. However, this methodological decision in handling the interaction between human and environmental indicators is not without consequences.

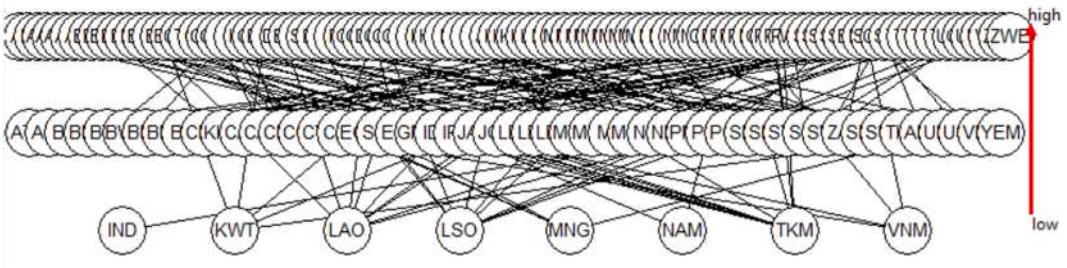

Fig. 5. Hasse diagram of 2019 PHDI variables.

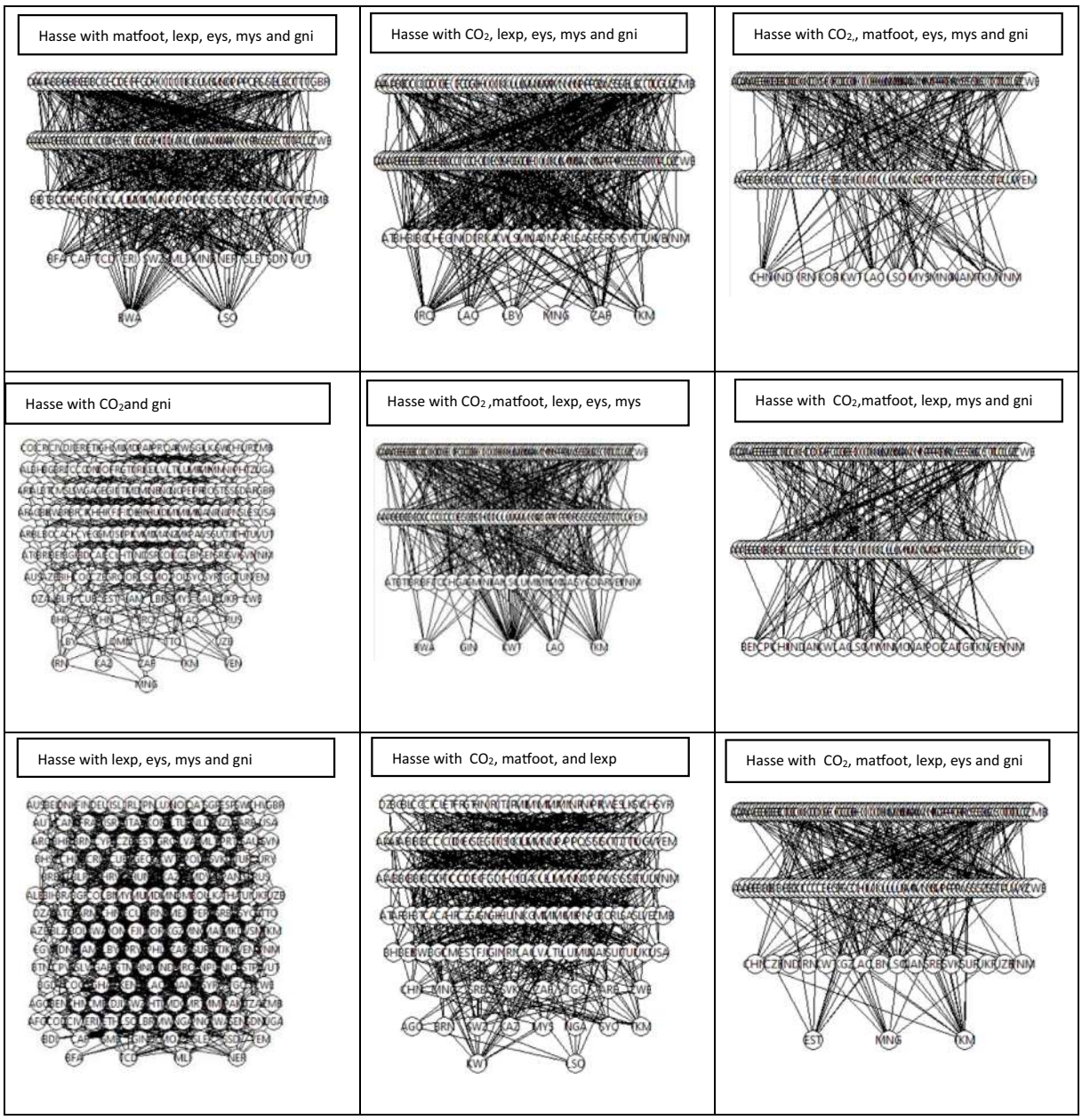

Fig. 8. Alternative Hasse diagrams based on PHDI variables.

6. Discussion and Concluding Remarks: Implications for the Elaboration of Sustainable Human Development Indicators

CIs have an important role of mediation between social values and desired policies. They have the power to shape mechanisms for change, such as incentives, regulations, etc. and can provide useful guidance and coordination among different agents. The human development approach, with its corresponding HDI, has been remarkably instrumental in the promotion of valuable capabilities in the last 30 years (UNDP, 2020; Hirai, 2017) and its recent move towards a sustainability evaluative framework deserves to be celebrated. Indeed, the importance of a joint assessment of human and planetary pressure indicators within a single framework cannot be overstated.

But this might come at a cost of making complete rankings based on composite indicators more of a challenge. And this happens because these new rankings based on the PHDI may hide much higher levels of incomparabilities than those already present in the HDI. This is a neuralgic point for sustainability indicators where the debate between the degree of substitutability and incommensurability of their different dimensions and variables has been constant from its inception (Bell and Morse, 2008; Neumayer, 2013; Davies, 2013).

Poset analysis shows that the debates between weak vs strong sustainability can benefit from empirical evidence from Hasse diagrams and sensitivity analysis where the conflicts can be identified for different groups of countries (or the most convenient unit of measurement). In the case of the HDI we were able to assess the overall level of incomparabilities and how they were distributed in different levels (groups). We were also able to verify how the discriminatory power of the original HDI variables was affected when environmental variables were included. Moreover, we saw how the new variables increased the level of incomparabilities in the poset and how they affected the impact of previous variables in the definition of the new poset.

Similarly, it indicated where the problem resides, namely, on the top level of countries where most of the incomparabilities can be found. By doing so it provides more systematic evidence about conflicts among the variables than the case-by-case exercise of listing differences in PHDI-

HDI rankings (UNDP, 2020, chapter 7). Composite sustainable human development indicators should pay particular attention to the issue of incommensurabilities because otherwise they will invariably become indicators of weak sustainability, missing part of the essence of the concern with sustainability that it is about the intrinsic value of nature.

The elaboration of sustainable human development indicators should take into account not only the theoretical nexus that may conceptually justify the relevance and role of the choice of dimensions and variables but also key empirically-based tenets such as:

- the structure of the rankings that they generate;
- the discriminatory power of the different variables (attributes) that they include;
- the density, level and distribution of incomparabilities (within their respective chains and antichains);
- the empirical significance of groups or levels of countries (objects) in terms of the public policies;
- the effects of new variables on old ones in comparative rankings.

Our results show that Sen's partial rankings become even more important when considering composite indicators with environmental variables. Posets can advance the debate on sustainable human development indicators by providing an evidence-based framework for tackling incomparability issues that are so dear to sustainability debates. They can also provide more conceptual consistency for the elaboration of sustainable human development indicators. In the particular case of the new PHDI, the poset evidence suggests that the introduction of environmental variables caused a dramatic change in the structure of the rankings in terms of the incomparabilities that they generate. They also changed the distribution of incomparabilities in the different levels and the role of variables, such as the one of GNIpc.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anand, S., Sen, A., 2000. Human development and economic sustainability. World Dev. 28 (12), 2029–2049.

Assa, J., 2021. Less is more: the implicit sustainability concept of the human development index. Ecol. Econ. 185, 1–13.

Barclay, M., Dixon-Woods, M., Lyratzopoulos, G., 2019. The problem with composite indicators. BMJ Qual. Safety 28 (4), 338–344.

Bell, S., Morse, S., 2008. Sustainability Indicators. Measuring the Immeasurable?, second ed. Earthscan, London.

Biggeri, M., Mauro, V., 2018. Towards a more 'sustainable' human development index: integrating the environment and freedom. Ecol. Indic. 91, 220–231.

Bravo, G., 2014. The human sustainable development index: new calculations and a first critical analysis. Ecol. Indic. 37, 145–150.

Bruggemann, R., Patil, G., 2011. Ranking and Prioritization for Multi-indicator Systems. Springer.

Caperna, G., Boccuzzo, G., 2018. Use of poset theory with big datasets: a new proposal applied to the analysis of life satisfaction in Italy. Soc. Indic. Res. 136, 1071–1088.

Chhibber, A., Laajaj, R., 2008. A Global Development Index: Extending the Human Development Index with Environment and Social Structures. http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=1523945C4A75654A3CF587C177155A47?doi=10.1.1.520.5089&rep=rep1&type=pdf.

Comim, F., 2018. Sen's capability approach, social choice theory and the use of rankings. In: Comim, F., Fennell, S., Anand, P.B. (Eds.), New Frontiers of the Capability Approach. Cambridge University Press, Cambridge, pp. 179–197.

Comim, F., 2021. A poset-generalizability method for human development indicators. Soc. Indic. Res. 158 (3), 1179–1198.

Costantini, V., Monni, S., 2005. Sustainable human development for European countries. J. Hum. Dev. 6 (3), 329–351.

Davies, G., 2013. Appraising weak and strong sustainability: searching for a middle ground. Consilience 10, 111–124.

De la Vega, M.C.L., Urrutia, A.M., 2001. HDPI: a framework for pollution-sensitive human development indicators. Environ. Dev. Sustain. 3 (3), 199–215.

Despotis, D.K., 2005. A reassessment of the human development index via data envelopment analysis. J. Oper. Res. Soc. 56 (8), 969–980.

Distaso, A., 2007. Well-being and/or quality of life in EU countries through a multidimensional index of sustainability. Ecol. Econ. 64 (1), 163–180.

Fattore, M., 2016. Partially ordered sets and the measurement of multidimensional ordinal deprivation. Soc. Indic. Res. 128 (2), 835–858.

Fattore, M., Bruggemann, R., 2017. Partial Order Concepts in Applied Sciences. Springer.

Fattore, M., Maggino, F., Colombo, E., 2012. From composite indicators to partial orders: Evaluating socio-economic phenomena through ordinal data. In: Maggino, F., Nuvolati, G. (Eds.), Quality of Life in Italy: Research and Reflections, Social Indicators Research Series, vol. 48, pp. 41–68.

Ferraz, D., Mariano, E.B., Rebelatto, D.A.N., Hartmann, D., 2020. Linking human development and the financial responsibility of regions: combined index proposals using methods from data envelopment analysis. Soc. Indic. Res. 150, 439–478.

Fox, M., Erickson, J., 2020. Design and meaning of the genuine progress indicator: a statistical analysis of the US fifty-state model. Ecol. Econ. 167, 1–11.

Hermele, L., 2010. Greening the human development index. In: Hornoborg, A., Jorgensen, A.K. (Eds.), International Trade and Environmental Justice: Toward a Global Political Ecology. Nova Science Publishers, New York, pp. 219–234.

Hickel, J., 2020. The sustainable development index: measuring the ecological efficiency of human development in the Anthropocene. Ecol. Econ. 167, 106331.

Hirai, T., 2017. The Creation of the Human Development Approach. Palgrave Macmillan, UK.

Hirai, T., 2021. Measuring capabilities: taking people's values seriously. World Dev. 148, 105657.

Jain, P., Jain, P., 2013. Sustainability assessment index: a strong sustainability approach to measure sustainable human development. Int. J. Sustain. Dev. World Ecol. 20 (2), 116–122.

Jin, H., Qian, X., Chin, T., Zhang, H., 2020. A global assessment of sustainable development based on modification of the human development index via the entropy method. Sustainability 12 (8), 3251.

Kalimeris, P., Bithas, K., Richardson, C., Nijkamp, P., 2020. Hidden linkages between resources and economy: a 'beyond GDP' approach using alternative welfare indicators. Ecol. Econ. 169, 106508.

Liang, M., Niu, S., Li, Z., Qiang, W., 2019. International comparison of human development index corrected by greenness and fairness indicators and policy implications for China. Soc. Indic. Res. 142, 1–24.

Mahlberg, B., Obersteiner, M., 2001. Remeasuring the HDI by data envelopment analysis. In: International Institute for Applied Systems Analysis (IIASA). Interim Report IR-01-069. Austria. https://doi.org/10.2139/ssrn.1999372 (accessed 5 January 2022).

Mariano, E.B., Rebelatto, D.A.N., 2014. Transformation of wealth produced into quality of life: analysis of the social efficiency of nation-states with the DEA's triple index approach. J. Oper. Res. Soc. 65 (11), 1664–1681.

Mariano, E.B., Sobreiro, V.A., Rebelatto, D.A.N., 2015. Human development and data envelopment analysis: a structured literature review. Omega 54, 33–49.

Morais, P., Miguéis, V.L., Camanho, A.S., 2013. Quality of life experienced by human capital: an assessment of European cities. Soc. Indic. Res. 110 (1), 187–206.

Morse, S., 2003. Greening the United Nations' human development index? Sustain. Dev. 11 (4), 183–198.

Munda, G., 2004. Social multi-criteria evaluation: methodological foundations and operational consequences. Eur. J. Oper. Res. 158, 662–677.

Neumayer, E., 2001. The human development index and sustainability – a constructive proposal. Ecol. Econ. 39 (1), 101–114.

Neumayer, E., 2013. Weak and Strong Sustainability: Exploring the Limits of Two Opposing Paradigms. Edward Elgar, Northampton.

Nussbaum, M., 1990. Love's Knowledge: Essays on Philosophy and Literature. Oxford University Press, New York.

Panagiotis, R., Giannis, K., 2021. Using VEA to assess effectiveness in the development of human capabilities. Econ. Chang. Restruct. 54 (1), 75–99.

Pineda, J., 2012. Sustainability and human development: a proposal for a sustainability adjusted human development index. Theoret. Pract. Res. Econ. Fields 2 (6), 73–100.

Qizilbash, M., 2007. Social choice and individual capabilities. Politics Philos. Econ. 6 (2), 169–192.

Ramos, S., Silva, J., Bolela, C., Andrade, M., 2018. Prediction of human development from environmental indicators. Soc. Indic. Res. 138, 467–477.

Ray, M., 2014. Redefining the human development index. Atl. Econ. J. 42, 305–316.

Robeyns, I., 2017. Well-being, Freedom and Social Justice: The capability approach re-examined. Open Book Publishers.

Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin III, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.A., 2009a. A safe operating space for humanity. Nature 461 (7263), 472–475.

Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin III, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.A., 2009b. Planetary boundaries: exploring the safe operating space for humanity. Ecol. Soc. 14 (2), 44.

Sen, A., 1970a. Collective Choice and Social Welfare. Holden Day.

Sen, A., 1970b. Interpersonal aggregation and partial comparability. Econometrica 38 (3), 393–409.

Sen, A., 1985. Commodities and Capabilities. Oxford University Press, Oxford.

Sen, A., 1993. Positional objectivity. Philos. Public Aff. 22 (2), 126–146.

Sen, A., 1996. On the foundations of welfare economics: utility, capability and practical reason. In: Farina, F, Hahn, F, Vannucci, S (Eds.), Ethics, Rationality and Economic Behaviour. Clarendon Press, Oxford, pp. 50–65.

Sen, A., 2000. Consequential evaluation and practical reason. J. Philos. 98 (9), 477–502.
Sen, A., 2002. Processes, liberty and rights. In: Sen, A. (Ed.), Rationality and Freedom. The Belknap Press Harvard, Cambridge.
Sen, A., 2009. The Idea of Justice. Harvard University Press, Cambridge.
Sen, A., 2017. Collective Choice and Social Welfare: An Expanded Edition. Harvard University Press, Cambridge.
Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., de Vries, W., de Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B., Sörlin, S., 2015. Planetary boundaries: guiding human development on a changing planet. Science 347 (6223), 1259855.
Tarabusi, E.C., Palazzi, P., 2004. An index for sustainable development. BNL Quart. Rev. 57 (229), 185–206.
Togtokh, C., 2011. Time to stop celebrating the polluters. Nature 479 (7373), 269.
Türe, C., 2013. A methodology to analyse the relations of ecological footprint corresponding with human development index: eco-sustainable human development index. Int. J. Sustain. Dev. World Ecol. 20 (1), 9–19.
UNDP, 2020. Human Development Report 2020: The Next Frontier, Human Development and the Anthropocene. New York.
Xiao, C., Wang, Q., van der Vaart, T., van Donk, D.P., 2018. When does corporate sustainability performance pay off? The impact of country-level sustainability performance. Ecol. Econ. 146, 325–333.