Chapter 10: Green Economy Measurements and Indicators

Carmen Herrero*
José Pineda†
Antonio Villar‡
Eduardo Zambrano§

March 31, 2018

* Universidad de Alicante.
† University of British Columbia.
‡ Universidad Pablo de Olavide.
§ California Polytechnic State University.

1 Introduction

When we speak of green economy measurements and indicators our main interest is in evaluating the extent to which an economy is *inclusive* and *green*. Put simply, an Inclusive Green Economy (IGE) is one that successfully provides a response to three sets of challenges currently facing humanity: (a) persistent poverty; (b) overstepped planetary boundaries; and (c) inequitable sharing of growing prosperity.[1]

The goals of this chapter are to identify some *basic desiderata* that any inclusive green economy measurement framework ought to satisfy; to survey the existing initiatives aimed at tracking whether an economy is green, inclusive, and sustainable; and to provide an example of what a green economy measurement framework that satisfies the *basic desiderata* can look like in practice.

IGE measurement is not done for its own sake. At the country level, proper IGE measures can support policymakers by proposing a range of purposes along the main stages of the policymaking process: objective setting; planning, design, and implementation; and monitoring and evaluation (UNEP, 2014).[2] In other words, as Section 6 in Chapter 2 has stressed, the goal is for us to be able to have a measurement framework that is truly useful for guiding countries in their formulation and evaluation of their social, economic and environmental policies.

The rest of the Chapter is structured as follows: In Section 2 of this chapter we briefly discuss the limitations of the GDP to measure progress and introduce the *basic desiderata*, that is, a list of overarching principles that we would want a green economy measurement framework to satisfy. In Section 3 we review the main kinds of differences that exist among current composite indices and dashboard of indicators, and consider the role of indicators in supporting green economy policymaking. Section 4 presents an empirical application of a Green Economy Progress measurement methodology that satisfies the *basic desiderata* to a sample of 105 countries, and compares it to other methods. Section 5 concludes.

2 Moving 'beyond GDP'

Measuring human progress and its sustainability is a challenging task, fraught with a myriad of statistical and real-world complexities.

The most prevalent way, even to this day, in which most individuals evaluate the extent to which a society is making economic progress is the Gross Domestic Product

[1] The list of green economy measures discussed in this chapter varies significantly in terms of their methodologies and the indicators used. One example of such diversity of indicators is that some measurement frameworks use indicators for social inclusion while others do not. This could be explained, at least in part, by the evolution of the concept of green economy, as already discussed in chapter 1, since inclusiveness was not initially given much attention in the conceptualization of green economy, but it has recently received more attention as the concept has evolved subsequently.

[2] See UNEP (2015) for an application of the methodology to the cases of Ghana, Mauritius and Uruguay.

(GDP). As Section 6 of Chapter 2 explains, it is well known that GDP paints an incomplete picture of well-being. This is so for several reasons, and let's review here a few of those to better understand why we need to complement GDP with other measures. First, GDP tracks aggregate economic activity and it is insensitive to the distribution of the gains and losses of that economic activity across the individuals in society. Second, it is not adjusted for the depletion of existing natural, physical and other assets. Third, it does not track those factors that matter for well-being that lie outside of the sphere of market transactions.

Another way to say this is the following: some of what GDP tracks does not matter for well-being and some of what matters for well-being is not tracked by GDP. Moreover, even those parts that matter are evaluated 'on average,' and without regards to their distributional characteristics, or their sustainability. Thus, one quickly lands at the conclusion that, rather than 'repair' GDP one needs entirely different measures altogether if one wishes to evaluate societal well-being.

The first global Human Development Report in 1990 introduced the human development index (HDI) as an alternative to GDP in which people are put at the center. The HDI has since become a widely-used measure of human progress more related to the lives of people than GDP alone.

The search for alternatives to GDP in measuring progress have significantly expanded through the availability of new data and methodologies, including subjective measures of human well-being. The Better Life Initiative, developed by the Organization for Economic Cooperation and Development (OECD), is among the efforts to better capture what is important to people's lives. They have been significantly influenced by the Stiglitz-Sen-Fitoussi Commission (2010), which concluded in 2009 that a broader range of indicators about well-being and social progress should be used alongside GDP. The Report of the United Nations Secretary-General's High-level Panel on Global Sustainability also highlights that the international community should measure development beyond GDP, and it recommends the creation of a new index or set of indices that incorporate sustainability considerations.

The task of developing and testing such measures belongs to the field of Welfare Economics, the subfield of economic theory that encompasses social choice theory, the theory of fair allocation and cost-benefit analysis. Of course, the exercises carried out in these are heavily value laden, by necessity. As Section 2 of Chapter 2 explains, reasonable people will differ regarding how to carry out those measurements, depending on their philosophical postures about what matters most for well-being. In principle, one can identify relatively simple and uncontroversial ethical postulates one wishes a well-being measure to satisfy and these principles pin down the desired measure, $W(t)$. This is the ideal. Think of $W(t)$ as a function $W\big(y(t)\big)$ that depends on a profile of consumption flows (or other sources of well-being), $y(t)$, for the individuals belonging to generation t.[3]

Alongside such measure one would want a secondary measure, dV^*, to track the sustainability of such well-being. To motivate the measure, imagine that, apart from there it being agreement about W being the proper indicator of generational well-being, there is also agreement that intergenerational well-being can be evaluated via the function

$$V(y) = \int_t^\infty W\big(y(\tau)\big)\, e^{-\mu(\tau-t)} d\tau \qquad (1)$$

where $\mu > 0$ reflects either the risk of extinction of mankind or a discount factor on future generations. This formula calculates a discounted sum of the measures of well-being of different generations. This certainly is not the only way to aggregate intergenerational well-being and to account for extinction risk but entering this debate is not crucial for what we are trying to do here.

The evolution of the consumption flows $y(\tau)$ for the members of each generation τ, starting with generation t, depends on what happens to *key* physical, natural and other stocks $K(\tau)$. Then, given an initial condition $K(t)$ for those stocks, one can compute a projected joint trajectory for these consumption flows and stocks. If we denote such trajectory for the consumption flows by $y^*\big(K(t)\big)$ we can evaluate intergenerational well-being along such trajectory by replacing the value of those consumption flows in expression (1) as follows: $V\left(y^*\big(K(t)\big)\right)$, which for simplicity we can call $V^*(t)$. Let's now compute the change, $dV^*(t)$, in such intergenerational evaluation as follows:

$$dV^*(t) = \sum_j \frac{\partial V^*}{\partial K_j(t)} \cdot dK_j(t) \qquad (2)$$

where $\frac{\partial V^*}{\partial K_k(t)}$ tracks how a change in stock k affects intergenerational well-being and $dK_j(t)$ tracks the change in stock *j* for generation t. Notice that expression (2) is a weighted sum of the changes in the set of relevant stocks, where the weights correspond to the importance of each of those stocks for intergenerational well-being.

Present well-being $W(t)$ can be said to be *sustainable* if future generations experience a well-being of that is as large that enjoyed by the current generation, that is, if $W(\tau)$ is at least as large as $W(t)$ for the members of each generation τ following generation t. We show in the appendix that, if present well-being $W(t)$ is sustainable then $dV^*(t)$ is positive.

[3] The rest of this section has been strongly influenced by Fleurbaey and Blanchet (2013, Ch. 2).

In other words, we have established tha if $dV^*(t)$ is negative then present well-being $W(t)$ is not sustainable in the sense that there will be by necessity a future generation with a level of well-being that will be below the well-being of the present generation, $W(t)$.

Thus, the ideal set of indicators for a comprehensive evaluation of a country's current economic situation and its sustainability is perhaps the profile

$$\langle GDP, W(t), dV^* \rangle$$

To see how this profile could be used in practice let's look at their expected behavior in the case of an unexpected negative shock to physical capital. If the economy increases its work intensity to try and reconstitute lost capital this effort will register as more GDP, but the W index will drop, "sending the correct message that the initial catastrophe was definitely not a blessing. However, under this kind of reaction the dV^* index can correctly tell us that, thanks to this temporary effort, sustainability is not threatened: the current generation pays its share of the bill generated by the catastrophe."[4]

If, on the other hand, the economy does not try to reconstitute its lost capital and aims instead to maintain its pre-catastrophe standard of living, as measured by W, then the message can be a decrease in economic activity as measured by GDP due to the fact that there is less capital available for production. Then whether the level of well-being W is sustainable depends on whether or not the economy was, before the shock, on a more than sustainable path. If, for example, the economy was close to strict sustainability and the shock is large, then the dV^* index would tell us that the level W is no longer sustainable. "This is the correct expression of the fact that the bill will have to be paid, sooner or later." [5]

As attractive as all this sounds, we face serious practical difficulties in the computation of the panel $\langle GDP, W(t), dV^* \rangle$. Even if we were to include on the GDP measure all that we ought to include, and if we were to agree on a methodology for the calculation of the well-being index W, we face multiple layers of uncertainty that make it very difficult to accurately calculate dV^*. We cannot escape the fact that informing about sustainability is informing about the future. In other words, as Section 7 of Chapter 2 has already identified, we don't merely face a measurement problem, we also face a forecasting problem.

We learn, however, from the theoretical exercise that at the very least we ought to be:

(*D1*) identifying as many of the ingredients that matter for present well-being in order to estimate the extent to which countries are making progress in their levels of well-being,

[4] Fleurbaey and Blanchet, 2013, p. 60. See also Sections 3 and 5 in Chapter 2.
[5] Fleurbaey and Blanchet, 2013, p. 60.

($D2$) identifying as many of those assets that matter for future well-being, their current stocks, and the measurement of how they evolve over time,

($D3$) completed where possible with relevant information about what we think to be critical thresholds for the stocks of those assets,

($D4$) with the understanding that we believe it to be near impossible to combine all these into a synthetic indicator of sustainability, or sustainable development, in a manner that will satisfy all constituents.[6]

These four ingredients, D1 to D4, which we call the *basic desiderata* in what follows, become the starting point of what we would want any Green Economy measurement framework to have.

To summarize, the logic behind the desiderata is the following: D1 is about identifying the profile of consumptions flows, $y(t)$, that matter in the evaluation of present well-being. D2 is about the identification of the changes in the relevant stocks, $dK_j(t)$, critical for the computation of any indicator of sustainability dV^*. Ideally, we would have, for each such stock that would enter into the computation of dV^*, a measure of how the stock affects intergenerational well-being but, since this is bound to be difficult, a minimal requirement would be to know the thresholds for those stocks below which the marginal intergenerational value of those stocks would be very high or very low. This is the rationale behind D3. To see the rationale behind D4, we notice that, as we have seen above, the technological, ethical and environmental uncertainties we face make a precise calculation of dV^* very difficult. "Doubts about our ability to build an all-purpose scalar index of sustainability are too strong (…) This suggests concentrating efforts on a well-defined set of warning indicators covering separately the various dimensions of sustainability" (Fleurbaey and Blanchet, 2013, p. 249).

3 An overview of measurement frameworks

Using the understanding we just gained about what we would ideally want to have, and what we are likely to be able to achieve in practice, in this section we review the most important initiatives aimed at monitoring well-being and its sustainability that have been developed in recent years.

There are four types of measurement frameworks represented in this literature: adjusted economic measures, dashboard of indicators, composite indices, and index-dashboard combos. Below we summarize 18 of the numerous initiatives aimed at measuring development "beyond GDP" that formally incorporate sustainability considerations as an important part of their methodological approach.

[6] "Doubts about our ability to build an all-purpose scalar index of sustainability are too strong," (Fleurbaey and Blanchet, 2013, p. 249).

3.1 Adjusted economic measures

The System of Environmental-Economic Accounting (SEEA)

The SEEA 2012 Central Framework is a statistical framework consisting of a comprehensive set of tables and accounts, which guides the compilation of consistent and comparable statistics and indicators for policymaking, analysis and research. It is a conceptual framework for understanding the interactions between the economy and the environment, and for describing stocks and the trend of environmental assets. Its main purpose is to put statistics on the environment and its relationship to the economy at the core of official statistics. This framework has been produced as part of a joint effort between the United Nations, the European Commission, the Food and Agriculture Organization of the United Nations, the Organisation for Economic Co-operation and Development, the International Monetary Fund and the World Bank Group.[7]

The SEEA Central Framework is based on agreed concepts, definitions, classifications and accounting rules. As an accounting system, it enables the organization of information into tables and accounts in an integrated and conceptually coherent manner to inform decision-making. The SEEA Central Framework allows for its implementation in parts or as a whole, adjusting to the different needs, priorities and resources of its users.

The SEEA Central Framework allows for the measurement of physical flows and stocks while providing guidance on the valuation of renewable and non-renewable natural resources and land within the System of National Accounts (SNA) asset boundary. While not having an inclusivity emphasis *per se*, this framework is the base of indicators such as Inclusive Wealth Index (IWI) or Adjusted Net Savings (ANS), which will be discussed below. However, some of the critics of such measures[8] have highlighted important limitations in their weighting systems, because within the SEEA Central Framework full valuation of assets and flows related to natural resources and land beyond the valuation included in the SNA is still an outstanding challenge (see also Section 3.1.1 below).

The Genuine Progress Indicator (GPI)

The GPI is designed by the Centre of Sustainable Economy and the Institute for Policy studies "genuine progress project". This adjusted measure is applied in some U.S. states and it is one of the several attempts made to substitute Gross Domestic Product and provide a better measure of the economic welfare. The GPI is related to an earlier measurement initiative that led to the creation of the Index of sustainable economic

[7] The SEEA Central Framework (2012) was adopted as an international standard by the United Nations Statistical Commission at its forty-third session in March 20121. It is the first international statistical standard for environmental-economic accounting. See United Nations (2014)

[8] See Thiry and Roman (2014).

welfare (ISEW). This index measures the economic welfare of a country by using the personal consumption and making deductions to account the costs of crime, environmental degradation, and loss of leisure, as well as additions to account for the flow of services from consumer durables, public infrastructure, and the benefits of volunteering and housework. Inclusivity enters the calculation by also allowing for deductions to account for income inequality.

Inclusive Wealth Index (IWI)

The IWI is designed by the United Nations University International Human Dimensions Programme on Global Environmental Change (UNU-IHDP) and UNEP in collaboration with the UN-Water Decade Programme on Capacity Development (UNW-DPC) and the Natural Capital Project. The inclusive wealth index measures the wealth of nations by carrying out a comprehensive analysis of a country's productive base. This adjusted measure covers 140 countries over the span of 20 years from 1990 to 2010. The report is produced every two years and each edition is focused on a specific topic: the 2012 report was focused on natural capital while the 2014 report focused on Human Capital. The index tracks the amount of capital in a nation by adding the "*social worth*" of three forms of capital: Manufactured, Human and Natural Capital. Produced capital and human capital are directly calculated by use of formulas derived from the theory behind the measurement exercise, whereas natural capital is an aggregate of natural stocks.

Adjusted Net Savings (ANS)

The ANS concept was first introduced by Pearce and Atkinson (1993) and the data is reported by the World Bank. It aims to assess an economy's sustainability based on the concepts of extended national accounts. This adjusted measure comprehends 213 Countries and it is calculated by adding fixed capital, human capital and environmental capital to the standard savings. There is no explicit accounting for inclusivity in this methodology.

3.1.1 Comments about the adjusted measures

The purpose of any dashboard of sustainability indicators is to track the evolution of key stocks of built, human, intellectual, natural, cultural and institutional capital at the country and at the planetary level that are priorities to sustain life on the planet. We explained in Section 2 that in the ideal world such dashboard could contain only a single number: the indicator $dV^*(t)$ since, as we saw in that Section, if $dV^*(t)$ is negative, then present well-being is not sustainable. *It is this theoretical ideal, $dV^*(t)$, what the adjusted measures are trying to proxy*. But we also discussed in that section the serious difficulties of computing $dV^*(t)$ with any degree of accuracy. If the economy was perfectly competitive and had complete contingent claims markets the weights $\frac{\partial V^*}{\partial K_j(t)}$ could be recovered from the behavior of the market prices of the different stocks. But it is a "fiction that all the relevant information about future trends is adequately reflected in current observed prices."[9] It is thus difficult to take the adjusted measures as currently computed as good estimates of what we would ideally like to measure to assess sustainability, namely, indicator $dV^*(t)$, and for that reason these measures are in a conflict of sorts with desideratum D4.

3.2 Dashboard of indicators

Green Growth Indicators

The OECD Green Growth Indicators are developed by the Organization for Economic Cooperation and Development. This dashboard is composed of more than 50 indicators and is intended as a guideline for Countries that want to assess themselves in terms of green growth. The OECD collects data for the 34 OECD countries, as well as Brazil, China, India, Indonesia, the Russian Federation, and South Africa. The indicators are available on the OECD website. Actually, OECD extends the set to countries in Latin America and the Caribbean, Central and East Asia, and the Caucasus. Green growth indicators are also being integrated into OECD work, including country reviews and policy analysis.

The indicators are grouped into four main categories: environmental and resource productivity, natural asset base, environmental quality of life, economic opportunities and policy responses. Indicators describing the socio-economic context and the characteristics of growth complete the picture. There is no explicit accounting for inclusivity in this methodology. Together with these indicators, a small set of six representative indicators were chosen to facilitate the communication with policy makers, media and citizens.

Sustainable Development Indicators (SDIs)

The SDIs are produced by Eurostat. This dashboard of indicators covers 28 EU countries. It was first published in 1997, and the most recent changes to the indicator set were related to the adoption of the Europe 2020 strategy. Eurostat publishes a report every two years with key trends in each sector for the EU as a whole. The Sustainable Development Indicators dashboard is composed of more than 100 indices. Twelve of them have been selected as headline indicators. The indicators are divided into 10 categories: Socioeconomic development, Sustainable consumption and production, Social inclusion, Demographic changes, Public health, Climate change and energy, Sustainable transport, Natural resources, Global partnership, and Good governance. The set of EU SDIs is divided in three levels of indicators, complemented by contextual indicators, which provide useful information about the countries but do not directly monitor progress.

3.2.1 Comments about the dashboard of indicators

While all these dashboards contain information on variables that affect current well-being as well as variables that affect the sustainability of such well-being, it is difficult

[9] Fleurbaey and Blanchet, 2013, p. 63. See also Chapter 2, Section 4 in this textbook.

sometimes to get a clear sense of whether or not a country is making progress overall, especially when the dashboards contain a large number of variables. This task of overall country evaluation could be easier if we knew how close the country was to meeting its goals or targets, and whether or not it was perilously close to reaching thresholds it wishes to avoid in key capital stocks. None of those initiatives, however, identify in a systematic manner critical thresholds or targets for the stocks of the assets they monitor for each country, as would be required by desideratum D3.

3.3 Composite indices[10]

Yale Environmental Performance Index (EPI)

The EPI is designed by Yale (Center for Environmental Law & Policy) and Columbia University (Center for International Earth Science Information Network) and it is calculated for 178 countries. The first report published is the pilot in 2006, and then EPI was calculated every two years until the last edition of 2014. EPI is divided into two main environmental protection objectives: environmental health and ecosystem vitality. These two areas are further divided into nine core policies categories: Health impact, Air quality, Water and sanitation, Climate and energy, Biodiversity and habitat, Fisheries, Forests, Agriculture and Wastewater management. Those are further divided into 20 indicators. EPI provides an overall and a category score to all the 178 countries considered. The indicators were chosen according to some generic guideline: relevance, performance orientation and data quality. Moreover, the principles of time series availability and global and temporal completeness should be respected by this data. Normalization was done by converting the transformed and logged data into one indicator through the proximity-to-target method. This methodology measures each entity's performance on any given indicator based on its position within a range established by the lowest performing entity (equivalent to 0 on a 0-100 scale) and the target (equivalent to 100). The indices are aggregated through the arithmetic mean. There is no explicit accounting for inclusivity in this methodology.

Low Carbon Competitiveness Index (LCCI)

The Low Carbon Competitiveness index is projected by The Climate Institute with the support of Vivid Economics. This composite index, first released in 2009 and last published in 2013, is currently used in the G20 countries. Its objective is to measure the current capacity of each country to be competitive and generate material prosperity to its residents in a low carbon world, based upon each country's current policy settings and indicators. LCCI derives from the aggregation of 19 variables, divided into three categories: *Sectoral composition*: how well, or otherwise, the composition of the economy is currently structured towards less emissions intensive activities; *Early preparation*: the steps that countries have already taken to move towards a low carbon economy; *Future prosperity*: the impact on the level of

[10] The first three sub-sections of this Section are based on Pineda, J. and Galotto, L. (2015). The fourth sub-section is based on PAGE (2017a).

production of goods and services. The variables were selected according to their relationship with the carbon competitiveness (GDP per tonne of emissions). Among the 36 variables likely to be linked to a country's low carbon competitiveness, only the 19 with the strongest association were chosen, according to the results of a backwards stepwise ordinary least squares regression. The normalization of the variables is made through a min-max transformation, while aggregation of indices in the categories is made by arithmetic mean. There is no explicit accounting for inclusivity in this methodology.

Global Sustainable Competitiveness Index (GSCI)

The GSCI is developed by the World Economic Forum (WEF) and it has been published since 2014. The last report (2016) covers 180 countries. The GSCI is based on 109 quantitative performance indicators grouped into the 5 pillars of sustainable competitiveness. Data sets have been scored both for the current levels as well as the recent development of the indicator in order to not only reflect current standing, but also development potential. The GSCI aims to evaluate the ability of countries to create and sustain wealth that does not negatively affect the underlying fundament of wealth creation, based on the definition of Sustainable Development. The data, taken from international organizations and an internal survey, was aggregated in countries through a sector-weighted country average procedure. The normalization of the variables is made through a min-max transformation, while aggregation of indices in the categories is made by arithmetic mean. There is no explicit accounting for inclusivity in this methodology

Global Green Economy index (GGEI)

The GGEI is projected by Dual Citizen LCC. It measures both the green economic performance of 80 countries and how experts assess that performance. This aggregate index of 32 indicators studies 60 countries and their largest metropolitan areas. It was first published in 2010. Together with the Performance Indicator, Dual Citizen publishes a Perception Index. The performance index of the 2014 GGEI is defined by 32 underlying indicators and datasets, each contained within one of the four main dimensions of leadership & climate change, efficiency sectors, markets & investment and environment & natural capital. Data are selected according to a "top down" method: first, the dimensions and sub-categories to include in the GGEI were defined, and then were identified those third-party datasets able to provide a value measure or generate a system for calculating a qualitative scoring for each category. There is no explicit accounting for inclusivity in this methodology. The normalization approach uses GDP (PPP) to expressed values with inherent imbalances based on the size of the country economy while aggregation is made by calculating a z-score and the associated percentile of the standardized distribution. Then, these percentile values are aggregated in a uniform manner, generating a country score that is expressed on the spectrum of 0-100.

Human Green Development Index (HGDI)

Human green development is captured by this index through many types of indicators such as welfare indicators, green economy indicators, and environment, resources and ecology indicators, involving more than 20 indicators. Many green economy indicators were employed for comparison, including the wealth accounting and the valuation of ecosystem services (WAVES) initiated by the World Bank, UNEP's manual on the use of green economy indicators, OECD's green growth indicators, EEA's green economy indicators, and the global green economy index (GGEI), etc. Its design took great inspirations from the HDI initiated by the United Nations Development Programme (UNDP). Notably, the goals and targets of MDGs and SDGs gave more direct guidance on how to choose indicators. The indicators used are separated into two dimensions, a social and economic sustainable development: 1) Proportion of population below minimum level of dietary energy consumption, 2) Inequality- adjusted income index, 3) Inequality- adjusted life expectancy index, 4) Inequality- adjusted education index, 5) Proportion of population using an improved sanitation facility , 6) Proportion of population using an improved drinking-water source, and Sustainable development of resources and environment with 7) Primary energy intensity, 8) CO2 Emissions per capita , 9) an air pollution indicator measure with PM , 10) Terrestrial protected areas of total land area , 11) Forest area (% of land area) , 12) Percentage of threatened animal species in total number of species.

The weight system adopted for the HGDI was based on an approach similar to the Delphi method, after seeking adequate references from relevant studies at home and abroad and organizing Chinese and foreign experts to hold three seminars successively where they discussed about the importance of each indicator. Considering there may be different correlation among these indicators, it could be more scientific to assign them different weights accordingly. The indicators with high correlation such as forest, ecology and land may be assigned a slightly lower weight. Conversely, those with low correlation may be attached with a little higher weight. In consideration of this, the research group calculated the correlation coefficients among the six indicators of each dimension separately. The final decision was to give a 50 % weight to "society and economy" and "resources and environment" respectively, given that they were considered to be equally important to the sustainable development. Similarly, an equal weight of 8.33% was given to each of the 12 indicators. The final value of HGDI is obtained by the geometric mean of the two dimensions.

3.3.1 Comments about the composite indices

While they can communicate complex ideas more quickly, composite indicators can be difficult to interpret. Ravallion (2012) presents several criticisms of composite indices, including the substitutability between indicators that they imply, as well as the sensitivity of their resulting ranks to factors set by their producers (such as indicator weights and aggregation methods). A perhaps more problematic aspect of the single composite index approach in this arena is that it attempts to combine the variables behind desiderata D1 and D2 into a single measure. This is problematic because any composite index intended to be a measure of 'sustainability adjusted well-being' that rewards the growth in variables that matter for present well-being but that penalizes the growth in variables that threaten the sustainability of that well-being may end up classifying countries having, say, low life expectancy and low greenhouse gas emissions and a high value for both in a similar way, while their positions clearly need to be differentiated.[11] This follows as long as we adopt the principle that "we consider it our moral duty not to impose on future generations any form of sacrifice that we do not accept for ourselves."[12]

3.4 Index-Dashboard Combos

The Green Economy Progress Measurement Framework (GEP)

The Green Economy Progress measurement framework is composed of a GEP index, a companion dashboard of sustainability indicators, and a country ranking that is based on the index and the dashboard. It was designed with the specific aim of meeting the *basic desiderata* identified at the end of Section 2, and for this reason we explain this methodology in more detail.

<u>The Index</u>

As in expression (4) above the GEP index is a weighted sum of how much progress a country makes in each of the variables, y_i, that matter for present wellbeing ($i = 1, \dots, I$). We measure progress in each variable by the extent to which a country is meeting its **targets**. If dy_i^* denotes the desired level of change in the variable and $dy_i \coloneqq y_i^1 - y_i^0$ denotes the actual change in the variable we then measure progress in the variable by the computation of the ratio $Progress(i) \coloneqq \frac{dy_i}{dy_i^*}$. This ratio has a straightforward interpretation: when $Progress(i)$ for variable i equals one this means the country met its target for that variable, when it is greater than one this means the country exceeded its target, when it is positive but less than one it means that the country made progress but did not meet its target, whereas when it is negative it means the country regressed, or got worse, in that variable.

The GEP for a country is calculated as $GEP{:}= \kappa \cdot \sum_{i=1}^{I} Importance(i) \cdot Progress(i)$ for some (normatively determined) country-specific importance levels, $Importance(i)$ for each $i = 1, \dots, I$ and some country-specific proportionality constant κ.[13]

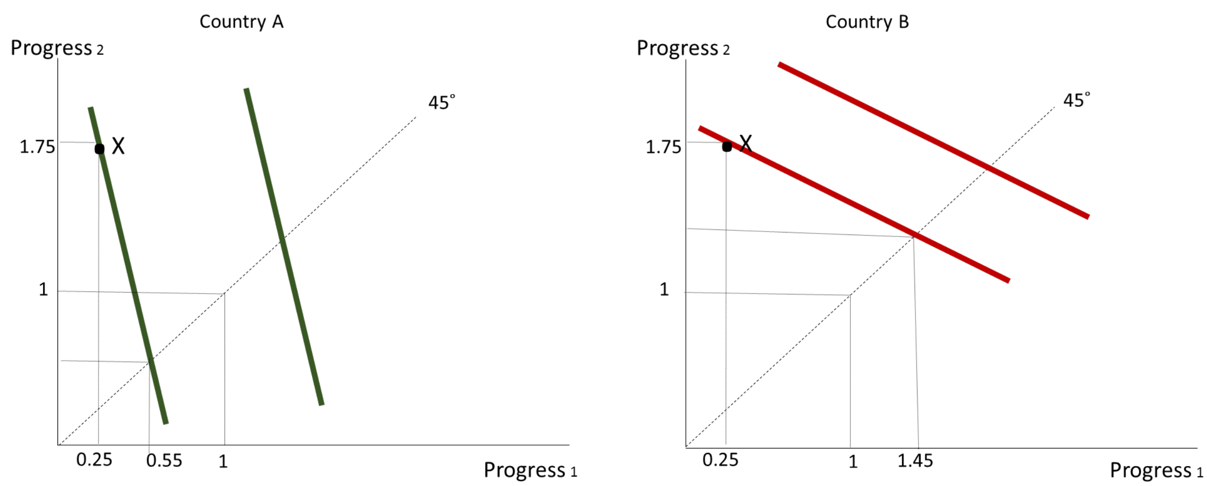
Figure 1 below illustrates the level curves for the GEP of two countries for the case where there are only two variables of interest. Variable 1 is one that we want more of, so progress is positive when the level of variable 1 grows over time. Variable 2 is one that we want less of, so progress is positive when the level of variable 1 shrinks over time. Notice that the slopes of these level curves differ across the two countries

[11] Example adapted from FB, p. 21.
[12] FB, p. 50.
[13] For more on the foundations of indicators of this sort see Villar (2011).

in the example. This illustrates the general principle that different countries may value certain kinds of progress more than others, depending on their initial conditions and other country characteristics.

As a way to put this general principle into practice we introduce the idea of **thresholds**: levels for the variables that the countries should critically try to avoid. If it is desirable that a country's level for a variable to be above a certain threshold assign a greater importance to progress in that variable if the country's current level is below the threshold than if the country's current level is above the threshold. On the other hand, if it is desirable that a country's level for another variable to be below a certain threshold we then assign a greater importance to progress in that variable if the country's current level is above the threshold than if the country's current level is below the threshold.

Figure 1. Level curves for the GEP for 2 countries



Appendix B.1 contains more details about the determination of the importance levels, which determine the slopes of the level curves in Figure 1 above, as well as the proportionality constant, which determines the units in which GEP is measured.

<u>The Dashboard</u>

The GEP measurement framework keeps track of the changes in relevant capital stocks, dK_j, and present those changes in a dashboard of progress for each country. Progress for each capital stock is calculated for each indicator K_j in the dashboard in the same way it is done for the variables in the index, that is, as $Progress(j) := \frac{dK_j}{dK_j^*}$, for all relevant indicators in the dashboard $j = 1, \dots, J$. The thresholds for these stocks are calculated with respect to planetary boundaries. This approach, of not combining all of the progress measures in the dashboard into a single measure intended to approximate $dV^*(t)$, is compatible with: (a) an outright acceptance of the intrinsically limited substitutability between the different forms of capital under consideration or, even if it wasn't limited; (b) the extraordinary difficulty, both ethical and technical, in identifying the proper "trade-offs" between forms of capital.

Despite the fact that the GEP measurement framework does not combine the variables in dashboard into a single scalar measure, determining the importance levels associated with each variable in the dashboard is of course relevant. We explain how the Green Economy Progress measurement framework determines those importance levels in Appendix B.2.

The Ranking

Even though the GEP index is not combined with the dashboard indicators into a composite measure of sustainable development, the information from both instruments can nevertheless inform which countries are in a comparably more favorable position than others.

To do this the GEP measurement framework defines the *achievement profile for each country* as follows: (i) The country's achievement for each stock j in the dashboard is given by the expression $Progress(j) \cdot \overline{Importance(j)}$. (ii) The country's GEP achievement is given by the expression $GEP \cdot \overline{Importance}$, where $\overline{Importance}$ is the average across all importance levels $Importance(i)$ for each $i = 1, \ldots, I$. The *achievement profile for each country* is then the (J+1)-vector given by the GEP achievement and the achievement for each of the stocks in the dashboard, as calculated above.

To determine how to use the information in the achievement profile of two countries to see which country is in a comparably more favorable position than another, the GEP measurement framework appeals to three normative principles: the *Pareto* principle, the principle of *Priority to the Worst Achievement,* and the principle of *Independence of Identical Achievements*. We explain why it makes sense to employ these principles in this setting in Appendix B.3.

These principles allow us to order countries in terms of their worst achievement but only considering the dimensions on which they differ. This order is known as *the protective criterion*. This methodology allows us to rank all index-dashboard achievement profiles but not to combine the index and dashboard information into a synthetic index.

To summarize, when comparing progress based on the GEP index and the dashboard, countries will be ranked according to their comparatively least-performing type of achievements. This approach sends the policy message that a country that is only making substantive achievements on a few aspects of an inclusive green economy, at the cost of others, will not necessarily be doing better than one that is making small achievements in all areas. Ranking countries based on the area in which they are making comparatively least progress provides maximal incentives to countries to not to dismiss any area, and to develop a more balanced and integrated policy approach aimed at making progress in a large number of the dimensions that characterize an inclusive green economy.

3.4.1 Comments about the Index-Dashboard Combos

This comment is about the GEP measurement framework since it is the only index dashboard combo under consideration. Because the GEP measurement framework was structured, by design, to satisfy the *basic desiderata* identified in Section 2, the framework has a disciplined methodology for the identification of the kinds of variables that go in the index and in the dashboard; it uses country level targets and thresholds as key inputs for the determination of the importance that each variable of interest has; and is able to use the information from both the index and the dashboard to give guidance to a country as to its multidimensional performance without the need to combine all the relevant information into a single scalar measure. On the other hand, it achieves all this at the cost of added conceptual complexity, and this was perhaps unavoidable. Below we illustrate the payoff behind this conceptual effort by way of showing the kinds of insights one can derive about both country and planetary performance from its implementation.

4 Measuring progress in practice

From the multitude of approaches, we have surveyed, we see that there is no unanimous agreement on what analytical framework we ought to use to measure green economy. However, an important step towards developing such unifying framework has been done in a joint effort by the OECD, UN Environment, the World Bank, and GGGI as part of their collaboration on the Green Growth Knowledge Platform (GGKP, 2013). One potential useful framework that came out of this initiative is the use of the concept of production function, viewed in the context of government policies, economic opportunities and the underlying socio-economic background.[14]

The framework proposes considering the environment as natural capital which, together with other forms of capital, are essential inputs to many production processes. This is illustrated in Figure 2, which presents the transformation process from inputs into outputs passing through a production function, as in OECD, 2011. The main ingredients in this formalism are the following:

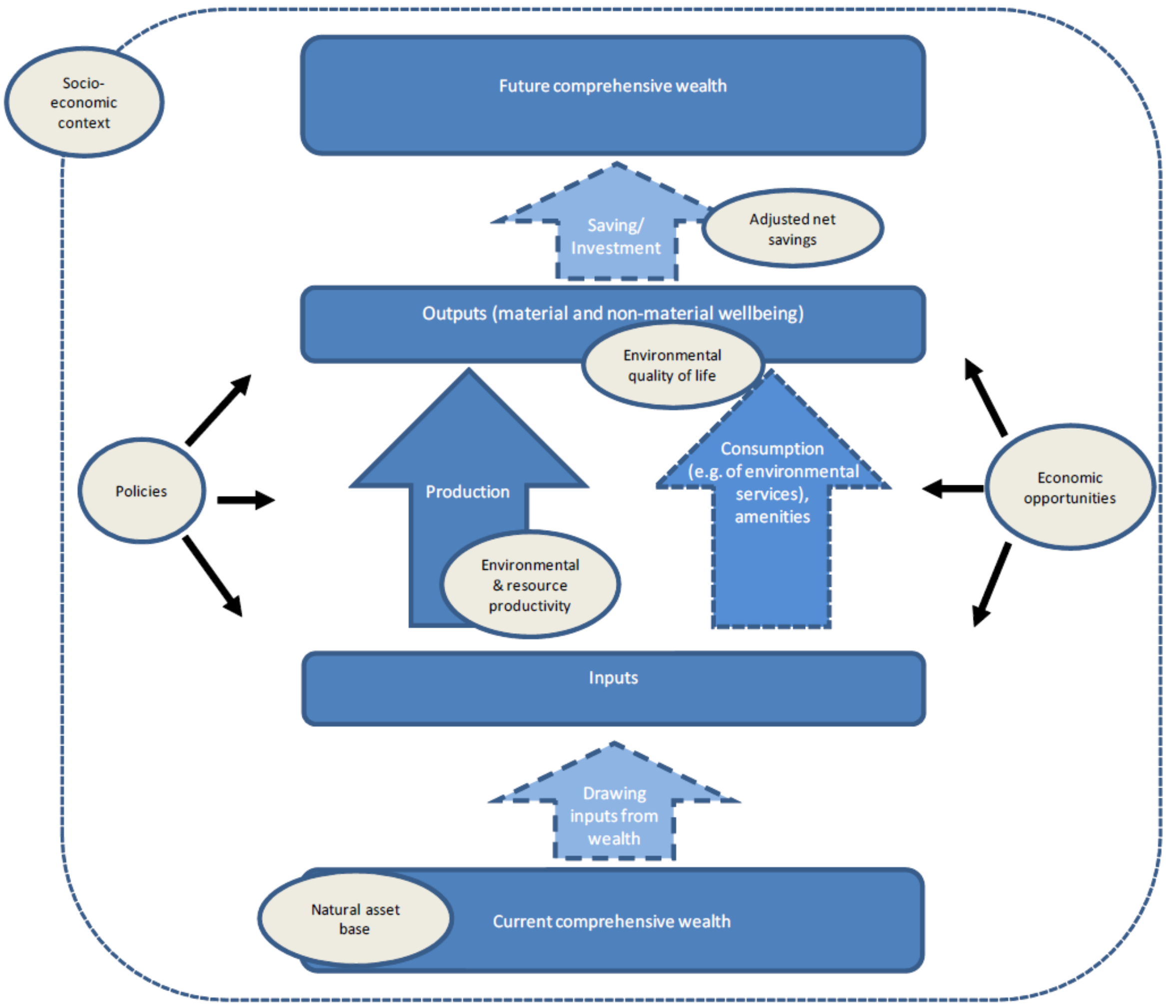
Inputs: the natural asset base. Natural capital provides services as well as natural resources *per se*, which provide crucial inputs into production or direct consumption. Indicators capturing the state of the natural asset base are crucial for identifying risks of overuse end/or depletion that may threaten future green growth.

[14] This section has been drawn from Green Growth Knowledge Platform (2013).

Production: intensity/productivity. These indicators focus on environment-related "productivity," or its inverse, "intensity." Progress can also be captured by measures of product-life environmental footprints or various proxy measures of innovation—which are important drivers of a green economy.

Outputs: material and non-material well-being. Outputs refers to broad notions of well-being that capture aspects that are not reported by conventional macroeconomic measures. This type of indicators attempts to capture the environment-related aspects of the quality of life and their impact economic processes (for example, the effect of good air quality on health that affects labor productivity).

Figure 2. The production framework for green economy indicators and wealth accounting

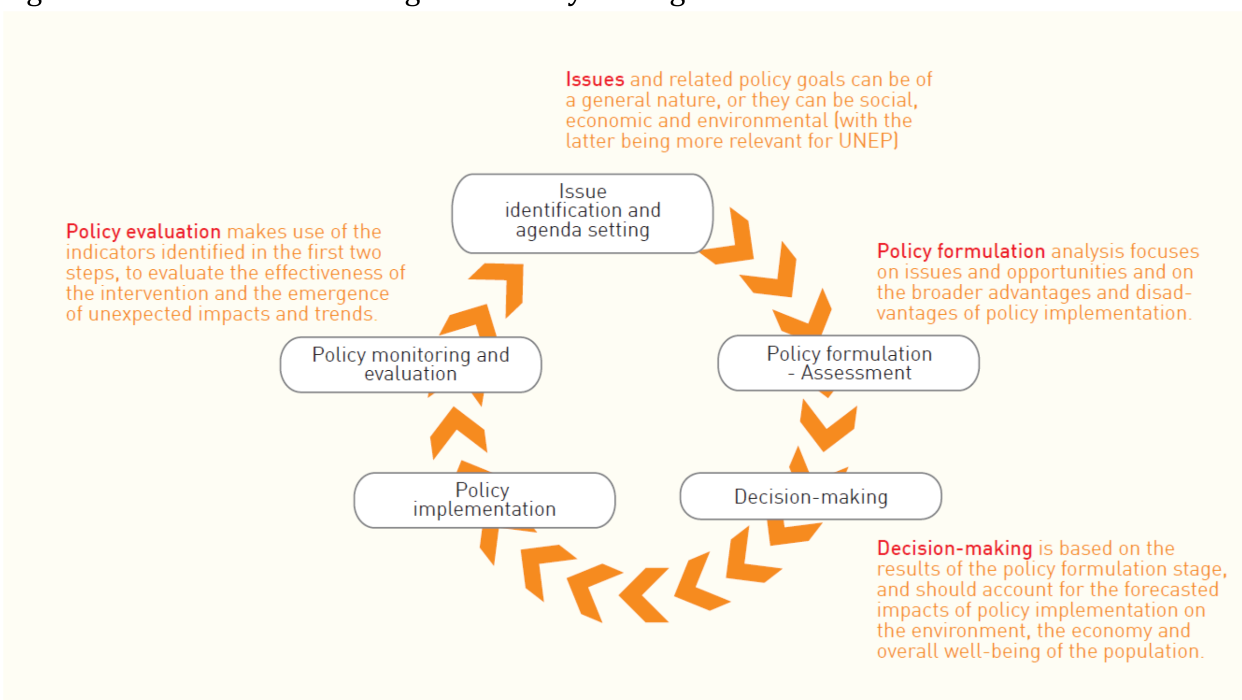


Source: GGKP (2013)
Note: White ovals represent indicator categories.

Measuring green economy can support policymakers along the main stages of the policymaking process: objective setting; planning, design, and implementation; and monitoring and evaluation (UNEP, 2014), as shown in figure 3.[15]

Regarding the objective setting, it is necessary to conduct diagnostics based on approaches and indicators that measure the present state, changes over time and future trends. Outcomes, such as expected climate change or health consequences of air pollution, and their drivers, such as emissions are the only elements taken into account in these diagnostics. These diagnostics identify challenges and opportunities that can lead to the formulation of respective policy priorities and goals. For example, it may involve establishing a long-term vision for green economy policy, developing baselines against which to compare developments over time and defining long-term targets aligned with domestic priorities (Mediavilla-Sahagun and Segafredo, 2014).

Figure 3. Overview of the Integrated Policymaking Process and the role of indicators



Source: UNEP (2015)

Finally, when desirable policies have been identified and measurable actions are implemented, indicators can be employed to monitor progress and assess the impact of policy actions (such as electricity production from renewables or the improvement in energy efficiency over time). The need to implement other policy interventions or mitigating actions can be determined through these indicators, thus supporting the achievement of the desired policy objectives.

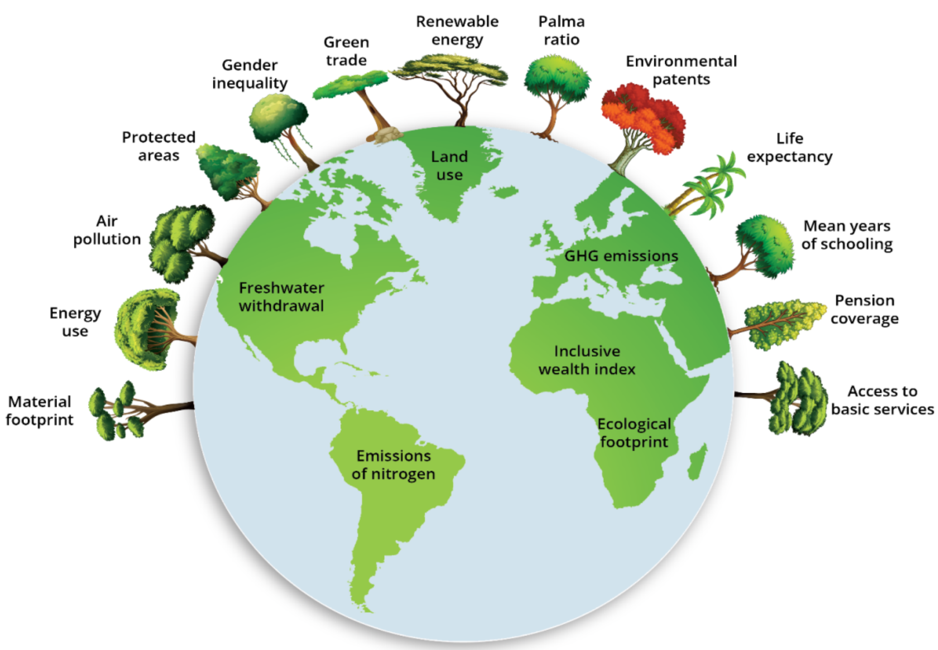
[15] See UNEP (2015) for an application of the methodology to the cases of Ghana, Mauritius and Uruguay.

Despite the significant efforts that have been done in order have a common framework (e.g. GGKP (2013)), the reality is that there is a variety of different approaches to work with indicators on green economy, as we already saw on Section 3. In this Section, we will present a practical application of the GEP measurement framework and how it compares with alternative methods for measuring green economy progress. In many cases, the indicators being used in the evaluations convey the same message, but in other cases this will not be so. This is, in essence, a reflection of the methodological differences behind these indicators.

4.1 Indicators in the GEP measurement framework

The GEP measurement framework is composed of a GEP index, a companion dashboard of sustainability indicators, and a country ranking that is based both on the index and on the dashboard (this country ranking is called the "GEP+" –more on this below). The GEP index is used to track the changes in GE indicators, relative to desired changes, which directly or indirectly impact current human well-being. It captures particular characteristics of the IGE concept by including a set of multidimensional indicators (e.g. indicators that capture the link between health and the environment). The dashboard of sustainability indicators aims to monitor the sustainability of well-being (i.e. the well-being of future generations). It tracks some of the main forms of natural capital (e.g. freshwater and land), as well as other key stocks of capital (e.g. human, health), as reflected in the Inclusive Wealth Index (UNU-IHDP/UNEP, 2014) and which affect long-term sustainability. Figure 4 gives an overview of the indicators included in the GEP measurement framework.

Figure 4 Indicators used in the GEP Measurement Framework



Source: PAGE (2017b)
Note: The 13 GEP Index Indicators are represented as trees while the elements of the dashboard of sustainability indicators are represented as the continents

Notice that in terms of the production function analogy, discussed on Section 3, it seems to be the case that indicators going to the GEP index tend to be more related to the production and outputs; while indicators of natural capital primarily belong in the dashboard of sustainability indicators.

4.2 Practical examples: Comparing the cases of China, Colombia and Ireland

To illustrate some of the main properties of the GEP measurement methodology, Tables 1a-c present the weights, the slope of relative weights, and the value of progress for the 13 indicators used in the GEP index for the case of China, Colombia and Ireland. The best way to understand how the weights are useful to set national priorities is the construction of the slope (ratio) of weights across indicators.

The weighting system of the GEP methodology allows us to understand the complexity of the multidimensionality of an IGE in a way that is useful for setting priorities both at the national and global level, and to understand the interplay between these two. For example, the three countries in Table 1 are exceeding the global threshold for material footprint per capita. For this indicator, the initial situation was worst for Ireland, followed by China and then Colombia.[16] However, China experiences other significant challenges, for example in air pollution, so the slopes of weighting relative to material footprint indicate that progress on material footprint will be important but not as important as progress on air pollution (relative slope of 4.535). So, even though the first weight indicates that it will be more important for China makes progress on material footprint than Colombia, the analysis of the slopes incorporates information on the other indicators for each country, making it clear that progress on material footprint is of higher relative priority for Colombia and progress on air pollution is of higher relative priority to China (relative slope of 4.535 vs 0.4317 for China and Colombia, respectively). In the case of protected areas, another indicator for which all three countries are exceeding the global threshold, the comparison of the relative slopes indicates that, in the case of China and Colombia, protected areas are higher priorities than material footprint while for the case of Ireland material footprint is the highest priority. In this way, the flexibility of the weighting system allows us to see when the global and national priorities coincide and when they differ, so that we can better articulate and act on national and global priorities.

[16] The initial situation was 33.16, 7.06 and 6.10 for Ireland, China and Colombia, respectively. Notice that the threshold for this indicator is 5 tons/person.

TABLE 1a. Progress, slopes and weights by indicator, China

China	material footprint	air pollution	protected areas	energy use	green trade	environmental patents	renewable energy	Palma ratio	gender inequality	access to basic services	mean years of schooling	Pension coverage	life expectancy
Weights π	0.0728	0.3303	0.2639	0.0704	0.0085	0.0355	0.0154	0.0360	0.0348	0.0400	0.0365	0.0116	0.0442
Slope with respect to mfp	1.0000	4.5350	3.6241	0.9660	0.1170	0.4881	0.2116	0.4942	0.4779	0.5487	0.5015	0.1597	0.6074
Progress	-3.8662	-0.1587	0.0387	0.5196	0.2306	0.4183	-0.5050	-0.1856	1.4626	0.6365	0.3624	0.6614	0.4121

TABLE 1b. Progress, slopes and weights by indicator, Colombia

Colombia	material footprint	air pollution	protected areas	energy use	green trade	environmental patents	renewable energy	Palma ratio	gender inequality	access to basic services	mean years of schooling	Pension coverage	life expectancy
Weights π	0.1222	0.0527	0.1334	0.0406	0.0520	0.0820	0.0218	0.1416	0.0878	0.0675	0.0709	0.0399	0.0874
Slope with respect to mfp	1.0000	0.4317	1.0921	0.3325	0.4256	0.6710	0.1784	1.1592	0.7188	0.5524	0.5805	0.3269	0.7156
Progress	-1.6227	-0.1889	0.1015	0.4568	-0.1547	0.4070	-0.1194	0.2291	0.3021	0.1972	0.2914	0.1621	0.3635

TABLE 1c. Progress, slopes and weights by indicator, Ireland

Ireland	material footprint	air pollution	protected areas	energy use	green trade	environmental patents	renewable energy	Palma ratio	gender inequality	access to basic services	mean years of schooling	Pension coverage	life expectancy
Weights π	0.3896	0.0414	0.1619	0.0269	0.0198	0.0397	0.1814	0.0470	0.0200	.	0.0250	.	0.0473
Slope with respect to mfp	1.0000	0.1063	0.4156	0.0690	0.0509	0.1020	0.4657	0.1205	0.0513	.	0.0642	.	0.1214
Progress	0.2588	-1.3650	0.2872	0.5965	0.0965	0.5188	1.1076	0.4897	0.6252	.	0.4115	.	0.5137

Source: Authors' calculations based on PAGE (2017b).[17]

Another important advantage of the GEP weighting system is that it helps to understand if progress is happening where it is most needed. One way to witness this is by comparing the differences between the simple average of progress across indicators (which is assumes equal weights across indicators), with the value of the GEP index, which is the weighted average of progress across indicators by using as weights. The values of the simple average of progress across indicators are 0.002044, 0.032693, and 0.325819 for China, Colombia and Ireland respectively. Now, the values of GEP index are -0.169356, -0.022191, and 0.400391 for China, Colombia and Ireland respectively. Notice that in the case of Ireland the GEP index is greater than

[17] The weights in the table are defined as $\pi_i = \kappa \cdot Importance(i)$ and the slope of indicator i with respect to mfp is $Importance(i)/Importance(mfp)$. See Appendix B.1 for details.

the simple average, reflecting that progress was important were it was of higher priority (Ireland made progress on the three indicators that were of the highest global and national priority).[18] In addition, notice that the GEP index is negative for Colombia and China as opposed to the positive simple average of progress values. This reflects the fact that these countries experienced regress on more indicators (four indicators each), but more importantly that regress was made in areas of high priority (material footprint for Colombia and material footprint and air pollution for China). These examples illustrate the richness of change that can be captured by the weighting system used in the construction of the GEP index.

Let's now compare the relative performance of these countries with two well-known measures, the EPI and the GGEI, which were discussed earlier.[19] Do they paint the same picture? Table 2 presents, for these three countries, the values of each of these indices. We can see that the relative order between these three countries is similar for the GEP index and the EPI, while there is a change in relative order between Colombia and Ireland, but China is still the worst performing country among the three. So, it seems to be the case, at least for these countries, that the different indices may provide some similar information. However, when we extend the analysis to all countries in the sample, it seems to be the case that the EPI and the GGEI have more in common among themselves than with the GEP index, as shown in Table 3. This lack of correlation of the ranking produced by these different measures could be a reflection of the differences in their methodologies and indicators covered. To illustrate this point, we have added as a measure the simple average of progress across indicators (a simplified version GEP index without its weighting system). As the rank correlation table shows, both EPI and GGEI have positive and statistically significant rank correlations with the simple average of progress but not with the GEP index, which could be partially explained by the differences in weighting across indices.[20]

TABLE 2. Rank and index value for a selected group of countries

	Rank GEP index	GEP index	Rank EPI	EPI	Rank GGEI
Ireland	5	0.4004	17	86.60	29
Colombia	81	-0.0222	49	75.93	16
China	93	-0.1694	77	65.10	52

Source: Authors' calculations based on PAGE (2017b), GGEI (2016) and Hsu, A. et al. (2016).

[18] It also reflects the fact that Ireland only had regress on one indicator (air pollution).
[19] Notice that this comparison is meant to be illustrative and by no means exhaustive. However, we consider that comparing these indicators could illustrate not only the complementarities across approaches but also their methodological differences and how they may explain differences in their results.
[20] For the EPI weights were selected according to the quality of the underlying dataset, as well as the relevance or fit of the indicator to assess the policy issue. The weightings given to Environmental health and Ecosystem vitality were chosen to balance the contribution of these indicators to the overall EPI; these explicit 60-40 weights provide an implicit 50-50 weighting because of the differences in variability of the two policy objectives (less variability across countries is observed for Ecosystem Vitality). Moreover, lower weights are used to lessen the impact of an indicator with a lower quality of data or for proxy indicators. For the GGEI the weightings for the four dimensions and sub-categories are applied equally, with the exception of the Leadership & Climate Change dimension, where the weighting for the head of state and media coverage sub-categories was lessened. See Appendix C for a more detailed description of the indicators used and their weights.

TABLE 3. Rank correlations for GEP index, Progress, EPI and GGEI

	Rank GEP index	Rank Progress	Rank EPI	Rank GGEI
Rank GEP index	1			
	100			
Rank Progress	0.5405	1		
	100	100		
Rank EPI	0.0368	0.382	1	
	100	100	100	
Rank GGEI	0.1689	0.2574	0.4255	1
	62	62	62	65

Source: Authors' calculations based on PAGE (2017b), GGEI (2016) and Hsu, A. et al. (2016).

Note: Progress is the simple average of the progress made in all indicators of the GEP index (Progress is a version of the GEP index with equal weighting).

4.3 Can composite indicators be combined with a dashboard of indicator to better capture progress in all aspects of green economy and sustainability?

Table 4 shows the results for a selected group of countries of the combination of the GEP index and the progress on the dashboard of sustainability indicators to make a final assessment of green economy progress.

Imagine that we base the progress on an IGE of a country by only using the GEP index, as we can see from Table 4 Bulgaria has a higher GEP index value than Ireland. However, Ireland is a country with progress in all indicators (the GEP index as well as all of the dashboard indicators). So, if we will assess progress just based on the GEP index, we could miss the fact that the progress achieved by Ireland seems to be more sustainable than the progress by Bulgaria, where there is regress on GHG and nitrogen emissions. Similarly, Germany has a higher GEP index value than Jamaica, but although both countries experienced progress in all indicators, the progress experienced by Germany on land use was the lowest of progress in all indicators considered. However, if we compare Germany and Singapore, we can see that Singapore exhibits more progress in all of the dashboard indicators. In addition, Germany has also a positive GEP index and therefore Germany will have a higher overall progress than Singapore.

What these examples illustrate is that using only a portion of the information could lead to misleading assessments of progress across countries. Some countries may have similar GEP indices, but different performances on the dashboard of indicators, others may have a high GEP index but low progress on its dashboard, compared to another country that may have progress on the dashboard of indicators but a negative GEP index. Given that we aim for a comprehensive assessment of progress, we want to be able to combine the information in all indicators, not just the GEP index. The GEP+ combines the information from the GEP index and the dashboard of sustainability indicators into a measure of progress that can give us an overall assessment, and that is conceptually sound (as it was explained in Section 3). It also produces very intuitive results, as illustrated in the examples discussed below.

Table 4: GEP+ (Rank among selected countries)

Country	Progress GHG (CO2e/cap)	Progress Nitrogen emissions	Progress Land use	GEP index	GEP+
Bulgaria	-0.5471	**-1.7996**	0.1097	0.5328	5
Ireland	2.3998	7.8447	**0.0012**	0.6197	3
Germany	0.5734	0.2181	**0.0039**	0.1664	2
Jamaica	1.1022	0.4906	0.1682	**0.1256**	1
Singapore	0.6208	0.4228	0.0211	**-0.1218**	4

Source: Authors' calculations based on PAGE (2017b).

Note: Observations in bold indicate the minimum value among all categories. The ranking presented in this table is based on the following four categories: (a) the GEP Index; (b) greenhouse gas emissions; (c) nitrogen emissions; and (d) the share of land used as permanent crops.

In Table 3, we showed that part of the lack of rank correlation between the GEP index and the EPI and the GGEI was due to the weighting system. In addition, the fact that both EPI and GGEI have indicators related to sustainability like CO2 emissions could also be explaining such differences. In fact, Table 5 shows that there is a positive and statistically significant correlation of the ranking from the GEP+ (which also takes into account the sustainability indicators from the dashboard) and those produced by the EPI and the GGEI.[21] Now we can see that there is a positive and statistically significant rank correlation between the GEP+ and both EPI and GGEI, as shown in the first column of Table 5. The correlations on Tables 4 and 5 illustrate that, although many indicators may be aiming to capture progress on green economy, their differences in terms of what indicators they include and how they're calculated could matter a great deal in terms of the assessment done. Knowing such differences is key in the interpretation and comparison of results among different methodologies.

TABLE 5. Rank correlations for GEP+, Progress, EPI and GGEI

	Rank GEP+	Rank GEP index	Rank EPI	Rank GGEI
Rank GEP+	1			
	100			
Rank GEP index	0.2338	1		
	100	100		
Rank EPI	0.3122	0.0368	1	
	100	100	100	
Rank GGEI	0.4003	0.1689	0.4255	1
	62	62	62	65

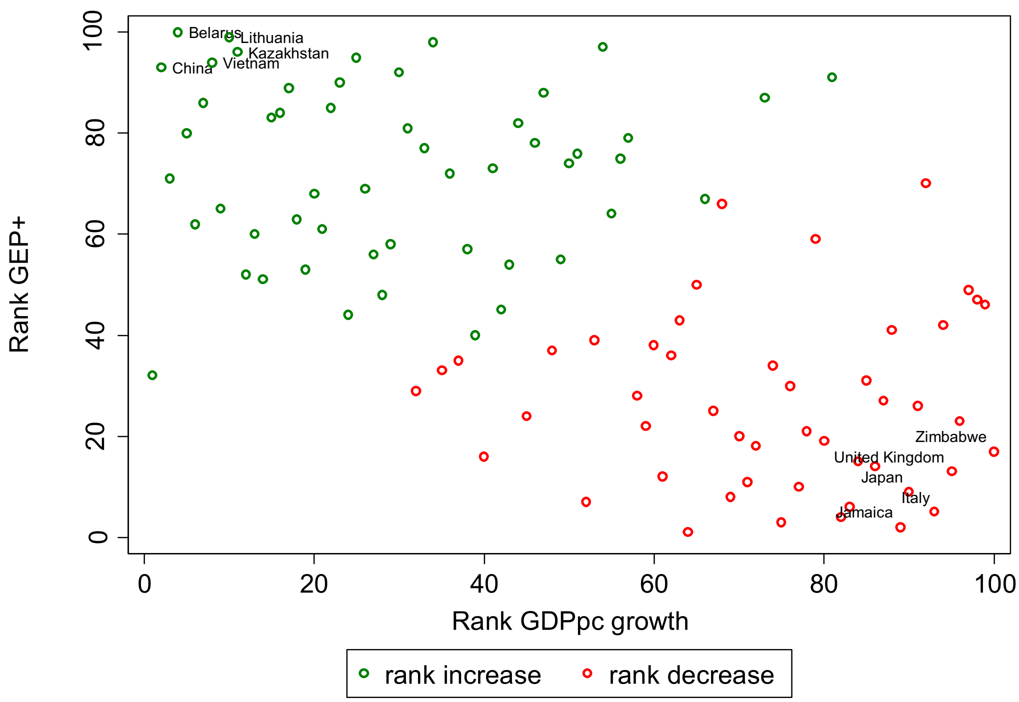
Source: Authors' calculations based on PAGE (2017b), GGEI (2016) and Hsu, A. et al. (2016).

[21] See Appendix B for a more detailed description of the indicators used in the EPI and GGEI.

4.4 Are measures like the GEP+, EPI or GGEI adding information relative to what we can learn from per capita GDP growth?

We started our conceptual discussion by arguing the need to go beyond GDP, which is particularly important for promoting a transition towards an IGE. In this section, we would like to check if our results from the GEP+ add relevant information relative to what can be learned from analyzing per capita GDP growth.

Figure 5: Rank differential of GEP+ vs GDP growth



Source: Authors' calculations based on PAGE (2017b).

Figure 5 presents the rank differences of GEP+ vs per capita GDP growth rank. Notice that an observation in red (green) means that the rank of GEP+ is higher (lower) that the rank of per capita GDP growth. As we can see from the figure, the information presented in the GEP+ is very different from the ranking information from per capita GDP growth. The rank correlation is -0.6051 and statistically significant at 1 percent. The countries highlighted in the figure represent the top and bottom 5 percent of the sample in terms of positions gained or when comparing rankings from the GEP+ with those from the per capita GDP growth. For example, if we only use per capita GDP growth, countries like China and Belarus will appear in the top positions of the ranking, but when we focus on progress towards an IGE (measured by the GEP+) these countries appear at the bottom of the distribution. The opposite is true for countries like Italy, Japan and United Kingdom. These countries are among the top 20 in the GEP+ but performs among the bottom 20 countries (out of 100) in terms of per capita GDP growth.

Table 6 shows the rank correlation between GEP+, EPI, GGEI and GDP growth. There is a statistically significant negative correlation between the ranking from GDP growth and the ranking from the GEP+ and GGEI, while the correlation is not statistically significant for EPI. This result illustrates that green economy indicators present a very different assessment of progress than GDP growth, which reinforce the need for pushing forward the work on green economy indicators as part of the broader beyond GDP agenda.[22]

TABLE 6. Rank correlations for GEP+, Progress, EPI, GGEI and GDP growth

	Rank GEP+	Rank EPI	Rank GGEI	Rank GDP growth
Rank GEP+	1			
	100			
Rank EPI	0.3122	1		
	100	100		
Rank GGEI	0.4003	0.4255	1	
	62	62	65	
Rank GDP growth	-0.6051	-0.0789	-0.4026	1
	100	100	62	100

Source: Authors' calculations based on PAGE (2017b), GGEI (2016) and Hsu, A. et al. (2016).

5 Conclusions

In this Chapter, we have presented some of the main methodological and practical aspects to consider when working with green economy indicators. We built on the discussion in Sections 6 and 7 from Chapter 2, presented an extensive list of the main approaches for the measurement of green economy, and discussed some of the challenges we face to achieve a unified framework. We have paid particular focus on the green economy measurement framework, recently developed by PAGE, since it is an example of a comprehensive methodology for the monitoring of progress towards key social, economic and environmental goals, integrating key theoretical aspects to its methodology and application. The GEP measurement framework follows an extraordinarily pragmatic approach, guided by theory.

We have made an effort to focus on measurement instruments that will be a practical guide to policy at the country level, and that allow us to track their performance in key stocks such as material footprint, rising emissions and increased freshwater withdrawal, while at the same time ensuring that further development is not put at

[22] Notice that in terms of the three countries discussed on table 2, the relative ranking in terms of GDP growth will be the opposite, with China being at the top with Colombia and Ireland many positions behind, which is not the case when we focus on indicators measuring green economy progress.

risk, and that economic opportunities are created, ecosystem services are preserved and social inclusiveness is promoted. Towards the end of the chapter, we have compared the results from the GEP framework with other well-known indices measuring green economy, the EPI and the GGEI, to illustrate the implications of using different indicators, discussing their main results and highlighting some of their important differences (in terms of methodology and indicators used). We have used in some cases specific countries examples to illustrate the main points and why these methodologies may be providing different, and many times complementary pictures of countries' progress on green economy. This kind of analysis is critical for assessing the policy implications of using any measurement framework, to help us select among competing frameworks, and to keep advancing the frontiers on this line of research.

The GEP measurement framework, in its current state, satisfies the *basic desiderata*, synthetizes the ideas of monitoring changes in key variables, taking into account: (a) global thresholds that should not be surpassed; and (b) ambitious but achievable targets that may help the countries to move in the right direction through policy interventions. These are critical to obtaining a useful measure of progress, which will be recognized as a valid instrument by practitioners, as well as by the wider community of researchers and academics working in the field.

The flexibility of the GEP measurement framework has allowed its application for policymaking at the country level. The methods used in the framework are flexible when it comes to selecting indicators, thereby making inter-country comparison possible as long as the underlying data is available, allowing for adjustments in the choice of indicators to specific country needs and priorities. A particular application of this framework was done by the Partnership for Action on Green Economy (PAGE) for China's Jiangsu Province in 2017. At the same time, the United Nations Industrial Development Organization (UNIDO) launched a project to construct a Green Industry Progress (GIP) index as a sectorial focused index, which would complement the ongoing PAGE development of a GEP index. The latest application of the methodology ranks the progress of 18 Chinese provinces based on seven indicators. Similar efforts are ongoing in other countries as South Africa, taking advantage of the methodology's ability to do benchmarking at the national and international level to inform policymaking.

Finally, we have shown that green economy indicators, some discussed in detailed in the final Section, portray a very different assessment of progress than just focusing on GDP growth. Although the challenges on green economy measurement are many, results like this illustrate the importance of going beyond GDP, in particular if we would like our measurement framework to support the transformative policy agenda implied by the green economy.

6 References

Barberà, S. and Jackson, M. (1988). Maximin, Leximin, and the Protective Criterion: Characterizations and Comparisons. *Journal of Economic Theory, 46:3444* (1988).

Dual Citizen (2016). "THE GLOBAL GREEN ECONOMY INDEX GGEI 2016 Measuring National Performance in the Green Economy". 5th Edition - September 2016.

Fleurbaey, M. and Blanchet D. (2013). *Beyond GDP Measuring Welfare and Assessing Sustainability*. Oxford: Oxford University Press.

Green Growth Knowledge Platform (GGKP). (2013). Moving towards a Common Approach on Green Growth Indicators. Seoul, Paris, Nairobi and Washington DC: Global Green Growth Institute, Organisation for Economic Co-operation and Development, United Nations Environment Programme and World Bank.

Hsu, A. et al. (2016). 2016 Environmental Performance Index. New Haven, CT: Yale University. Available: www.epi.yale.edu.

Mediavilla-Sahagun, A. and L. Segafredo. (2014). Establishing Vision, Baselines, and Targets. In Green Growth in Practice: Lessons from Country Experiences, Chapter: Establishing Vision, Baselines, and Targets, Publisher: Global Green Growth Institute, pp.59-78.

Moulin, H. (1998). *Axioms of Cooperative Decision Making*. Cambridge: Cambridge Univ. Press.

OECD (2011). "Towards Green Growth: Monitoring Progress".

PAGE (2017a). The Green Economy Progress Measurement Framework – Methodology.

PAGE (2017b). The Green Economy Progress Measurement Framework – Application.

Pearce, David and Giles D. Atkinson (1993). Capital theory and the measurement of sustainable development: an indicator of "weak" sustainability. Ecological Economics, 1993, vol. 8, issue 2, 103-108.

Pineda, J. and Galotto, L. (2015). Review of existing Green Economy Indicators. Mimeo.

Ravallion, M., "Troubling Tradeoffs in the Human Development Index, *Journal of Development Economics*, 99(2), 2012.

Stiglitz, J. E., Sen, A., and Fitoussi, J. P., (2010), "Report by the commission on the measurement of economic performance and social progress", Paris: Commission on the Measurement of Economic Performance and Social Progress.

Thiry, Géraldine and Philippe Roman (2014). The Inclusive Wealth Index. A Sustainability Indicator, Really?, FMSHWP-2014-71, juin 2014.

United Nations (2014). System of Environmental Economic Accounting 2012—Central Framework. New York.

United Nations Environment Programme (UNEP) (2014). Using Indicators for Green Economy Policymaking. Geneva: UNEP.

United Nations Environment Programme (UNEP) (2015). Indicators for Green Economy Policymaking – A Synthesis Report of Studies in Ghana, Mauritius and Uruguay. Geneva: UNEP.

United Nations Environment Programme (UNEP) (2017). The Green Economy Progress measurement framework. Geneva: UNEP.

University of the United Nations / International Human Dimensions Programme on Global Environmental Change, United Nations Environment Programme (UNU-IHDP/UNEP) (2014*). Inclusive Wealth Report 2014. Measuring progress toward sustainability*. Cambridge: Cambridge University Press.

Villar, A. (2011). Who Meets the Standards: A Multidimensional Approach, *Modern Economy*, 2(4): 614-624.

Appendix A: The sustainability of present well-being

Claim. If present well-being $W(t)$ is sustainable then $dV^*(t) \geq 0$.

Proof. Let $W(t)$ be sustainable. Then,

$$W(\tau) \geq W(t), for\ all\ \tau \geq t. \tag{A.1}$$

Therefore, along the trajectory $y^*(K(t))$ we obtain, integrating both sides of (A.1),

$$V^*(t) = \int_t^\infty W(y^*(\tau))e^{-\mu(\tau-t)} \geq \int_t^\infty W(y^*(t))e^{-\mu(\tau-t)} = \frac{W(y^*(t))}{\mu},$$

that is,

$$V^*(t) \geq \frac{W(y^*(t))}{\mu} \tag{A.2}$$

Now apply Leibnitz rule to $V^*(t) = V\left(y^*(K(t))\right)$ to obtain

$$dV^*(t) = -W(y(t)) + \mu V^*(t). \tag{A.3}$$

It follows from (A.2) that $\mu V^*(t) \geq W(y^*(t))$, and from (A.3) we obtain that $\mu V^*(t) - W(y^*(t)) = dV^*(t) \geq 0$. ■

Appendix B: The Green Economy Progress Measurement Framework. Further details

B.1 The determination of the importance levels of the components of GEP, and its units

In Section 3.4 we introduced a general principle we would want a green economy measurement framework to satisfy: that different countries can value certain kinds of progress more than others, depending on their initial conditions and other country characteristics. The GEP measurement framework uses the idea of thresholds to make this general principle operational. If it is desirable that a country's level for variable y_i be above a certain threshold t_i we assign a greater importance to progress in that variable if the country's current level is below the threshold than if the country's current level is above the threshold. On the other hand, if it is desirable that a country's level for another variable y_i to be below a certain threshold t_i we then assign a greater importance to progress in that variable if the country's current level is above the threshold than if the country's current level is below the threshold. The interpretation is that progress in a variable is more important the worse the country is in relationship to the relevant threshold for that variable. Specifically, we define the importance of progress in variable y_i, $Importance(i)$, to be equal to $\frac{t_i}{y_i^0}$ if we wish for variable y_i to grow (like with education, protected areas and other 'goods'), to be equal to $\frac{y_i^0}{t_i}$ if we wish for variable y_i to shrink (like with air pollution, inequality and other 'bads').

The implementation just described establishes the relevant tradeoffs between the different dimensions of progress (as represented the slopes of the level curves illustrated in Figure 2). It allows us to perform comparisons between different types of progress for each country and be able to make an overall determination of whether or not a country is making progress overall. Thus far, however, the implementation is silent about the units in which GEP is measured. This determination is important, for example, if we wish to know how to compare the GEP to its component parts (the progress measures for each variable of interest), and how to compare the GEP of different countries.

To make this determination the GEP measurement framework adopts the 'equivalence approach' from microeconomics. For a given progress profile $[Progress(1), Progress(2), \cdots, Progress(I)]$ of the indicators that matter for present well-being we can identify the GEP of this country with the level L such that, if the country were to progress in each of its dimensions by a magnitude equal to L, it would achieve the same level of green economy progress as when the country's progress profile is actually $[Progress(1), Progress(2), \cdots, Progress(I)]$. It follows, in particular, that if $Progress(i) = 1$ for all $i = 1, \dots, I$, then $GEP = 1$, which implies that $\kappa \cdot \sum_{i=1}^{I} Importance(i) = 1$.

As a consequence, we can interpret the magnitude and sign of the GEP in the same way we did for each individual measure of dimensional progress: when GEP is equal to one this means the country is equivalent to a position in which it met all its targets, when it is greater than one this means the country is equivalent to a position where it exceeded its targets, so on and so forth.

Figure B1 again illustrates how this works: In this Figure, countries A and B exhibit the same magnitude of progress in both variables in point x: they exceeded their target in variable 2 by 75 percent while they fell short of their target in variable 1 also by 75 percent. A simple average would put country A and B in the same level of progress overall, but that is not what the GEP does. In this example country A is below its threshold –a bad thing– in variable 1 and above its threshold –a good thing– in variable 2. It is therefore far more important for country A to improve in variable 1 than in variable 2, and the GEP consequently puts greater weight on variable 1 relative to variable 2. This is what makes the level curves for country A's GEP comparatively steep. The GEP for country A in situation x can be found by identifying where the level curve that passes through x crosses the 45-degree line. Because this level curve is steep, this intersection happens below the 'perfect progress' point (1,1), which means that the country gets a GEP less than one. In the example country B is above its threshold –a good thing– in variable 1 and below its threshold –a bad thing– in variable 2. It is therefore far more important for country B to improve in variable 2 than in variable 1, and the GEP consequently puts greater weight on variable 2 relative to variable 1. This is what makes the level curves for country B's GEP comparatively flat. The GEP for country B in situation x can be found by identifying where the level curve that passes through x crosses the 45-degree line. Because this level curve is flat, this intersection happens above the 'perfect progress' point (1,1), which means that the country gets a GEP greater than one. Put simply, country B's GEP is greater than country A's GEP because country B made comparatively more progress in the dimension where it needed it the most.

Box 1. The details from the example illustrated in Figure B1

The parameters behind the example in Figure 2 are the following. Variable 1 is a good, variable 2 is a bad. $Progress(1) = .25$, $Progress(2) = 1.75$. These are the progress levels for these variables for both countries, which means that both countries improved their levels of variable 1 and reduced their levels of variable 2. The thresholds for all variables equal 1 for both countries. The rest of the variables are as follows: For country A we have $y_1 = y_2 = .5$, whereas for country B we have $y_1 = y_2 = 2$. It follows that the slope of the level curves of the countries' GEP are: -4 for country A and -.25 for country B, and that $Importance(1) = 2$, $Importance(2) = .5$ for country A, whereas $Importance(1).5$, $Importance(2) = 2$ for country B. The weights on progress for country A are then $\kappa_A \cdot Importance_A(1) = .8$ and $\kappa_A \cdot Importance_A(2) = .2$, while these same weights are $\kappa_B \cdot Importance_B(1) = .2$ and $\kappa_B \cdot Importance_B(2) = .8$ for country B. Finally, the GEP for country A is then 0.55 and the GEP for country B is 1.45.

B.2 The determination of the importance levels of the components of the dashboard

To determine the importance levels of the components of the dashboard the GEP measurement framework proceeds just as with the determination of the importance levels of the variables in the GEP index. In other words, $Importance(j)$ is determined to be equal to $\frac{t_j}{K_j^0}$, if we wish for stock K_j to grow, and to $\frac{K_j^0}{t_j}$, if we wish for stock K_j to shrink.

B.3 The Protective Criterion and its Properties

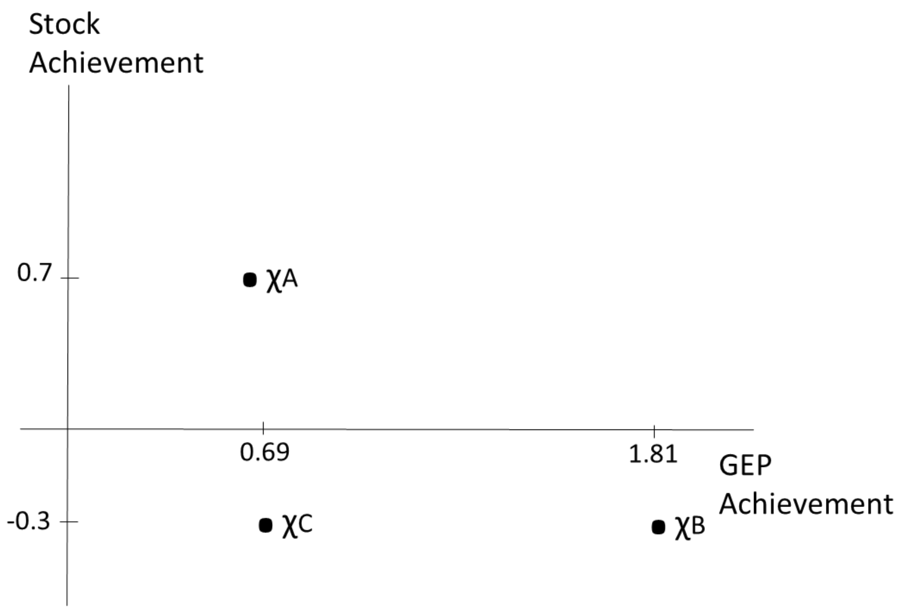
To understand how one could use the information in the achievement profile of two countries to see which country is a comparably more favorable position than another we consider first the case is one in which a country's achievement levels larger than the same respective achievement indicators for some other country. It is then perhaps suitable to assert that the first country is in a more favorable position than the second country. This is, after all, what the Pareto principle, when applied to this setting, would prescribe.

While less obvious, there are two additional cases in which one could argue that a comparison between countries is suitable.

Consider the case where x and z are the achievement profiles of two countries. Then it can be argued that if x is in a more favorable position than z it must be that the worst achievement in x is greater than the worst achievement in z. This is the principle of *Priority to the Worst Achievement*.

To illustrate the principle, go back to the example illustrated in Figure 3 and add a dashboard to each country, consisting of a single capital stock, with a measured level of achievement equal to .7 for country A and -.3 for country B. These index-dashboard achievement combos are illustrated in Figure A1 as points x_A and x_B.[23] The worst achievement for country A is therefore .69 whereas for country B is -.3. The principle says that we cannot rank country B over country A since .69>-.3. While in this example Country B obtains a high GEP achievement, equal to (1.81, it achieves it at the cost of depleting its capital stock, as represented by the -.3, and it cannot obtain a high rank as a consequence. Even if country B's GEP achievement level was much higher the conclusion would be the same.

Figure B1. The GEP Index-Dashboard space



Consider now a case where x and z be the achievement profiles of two countries such that they share the same achievement in some dimension. Then it is reasonable to assert that whether x is in a more favorable position than z is independent of how the countries do in the dimension in which they fare equally well. This is the principle of *Independence of Identical Achievements*.

[23] Because the average importance of the progress indicators is 1.25 in that example (see Box 1 above), the GEP achievements for the countries are, respectively, .55*125=.69 for country A and 1.45*1.25=1.81 for country B.

To illustrate the principle, go back to the example illustrated in Figure A1 and consider country C, with a GEP achievement of .69 (just like for country A) and a measured level of achievement for the capital stock of -.3 (just like for country B). Because achievement in the capital stock level is the same for B and C the principle says that we ought to rank countries according to the GEP achievement information alone, and we would rank country B over country C, since 1.81>.69. Similarly, the GEP achievements are the same for A and C and the principle says that we ought to rank the countries according to the achievements on the capital stock alone. We would thus rank A over C, since .7>-.3. Notice that this is a refinement over what the principle of Priority to the Worst Achievement would prescribe in this case, as that principle would allow countries B and C to be ranked equally, since the worst consequence in both cases is the same and equal to -.3.

These principles therefore allow us to order countries in terms of their worst achievement but only considering the dimensions on which they differ. This order is known as *the protective criterion*.[24] This methodology would allow us to rank all index-dashboard achievement profiles but not to combine the index and dashboard information into a synthetic index.[25]

Appendix C: Details on indicators and weights on selected composite indices

Table C1 Yale Environmental Performance Index (EPI)

Environmental health (40%)

Policy category	**Indicator**	**Indicator description**
Health impact (33%)	Child mortality (100%)	Probability of dying between a child's first and fifth birthdays (between age 1 and 5)
Air quality (33%)	Household air quality (33%)	Percentage of the population using solid fuels as primary cooking fuel.
	Air Pollution avg. Exp to PM2.5 (33%)	Population weighted exposure to PM2.5 (three- year average)
	Air Pollution avg. Exp to PM2.5 exceedance (33%)	Proportion of the population whose exposure is above WHO thresholds (10, 15, 25, 35 micrograms/m3)
Water and sanitation (33%)	Access to drinking water (50%)	Percentage of population with access to improved drinking water source
	Access to sanitation (50%)	Percentage of population with access to improved sanitation

[24] *Barberà and Jackson (1988).*

[25] This is so because the Protective Criterion (see Annex I.C in UNEP (2016)), like the leximin, does not admit a real-valued representation due to the lack of continuity of the preference ordering. For proof see the example of Moulin (1998, page 34).

Ecosystem vitality (60%)

Policy category	Indicator	Indicator description
Climate and Energy (25%)	Trends in CO2 emission per KwH	Change in CO2 emissions from electricity and heat production
	Change of trend in carbon intensity	Change in Trend of CO2 emissions per unit GDP from 1990 to 2000; 2000 to 2010
	Trend in Carbon Intensity (33%)	Change in CO2 emissions per unit GDP from 1990 to 2010
	Access to electricity (N/A)	Percent of population with access to electricity.
Biodiversity and habitat (25%)	Terrestrial Protected Areas (National Biome Weights) (25%)	Percentage of terrestrial biome area that is protected, weighted by domestic biome area
	Terrestrial Protected Areas (Global Biome Weights) (25%)	Percentage of terrestrial biome area that is protected, weighted by global biome area.
	Marine Protected Areas (25%)	Marine protected areas as a percent of EEZ
	Critical Habitat Protection (25%)	Percent of critical habitat sites as designed by the Alliance for Zero Extinction protected
Fisheries (10%)	Coastal shelf fishing pressure (50%)	Catch in metric tons from trawling and dredging gears (mostly bottom trawls) divided by EEZ area
	Fish Stocks (50%)	Percentage of fishing stocks overexploited and collapsed from EEZ
Forest (10%)	Change in forest cover (100%)	Forest loss - Forest gain in > 50% tree cover, as compared to 2000 levels.
Agriculture (5%)	Agricultural subsidies (50%)	Subsidies are expressed in price of their product in the domestic market (plus any direct output subsidy) less its price at the border, expressed as a percentage of the border price (adjusting for transport costs and quality differences).
	Pesticide regulation (50%)	Scoring of whether countries have signed on to the Stockholm Convention and allow, restrict, or ban the "dirty dozen" POPs that are common agricultural pesticides.
Water resources (25%)	Wastewater management (100%)	Wastewater treatment level weighted by connection to wastewater treatment rate.

Source: Hsu, A. et al. (2016). 2016 Environmental Performance Index. New Haven, CT: Yale University. Available: www.epi.yale.edu.

Table C.2 Global Green Economy index (GGEI)

Sectors	Sub-category	Source
Leadership & Climate Change (25%)	Climate Change Performance (50%)	International Energy Agency (IEA), Climate Change Performance Index (CCPI)
	International Climate Forums (20%)	Climate Action Network (ECO) reporting scored by Dual Citizen LLC on scale of 0-10
	Head of State (20%)	Google Analysis scored by Dual Citizen LLC on scale of 0-10
	Media Coverage (10%)	Google Analysis scored by Dual Citizen LLC on scale of 0-10
Efficiency Sectors (25%)	Buildings (25%)	LEED certification as reported by the U.S. Green Building Council (USGBC)
	Transport (25%)	International Energy Agency (IEA)
	Tourism (25%)	Scored by Dual Citizen LLC on scale of 0-10
	Energy (25%)	International Energy Agency (IEA)
Markets & Investment (25%)	Renewable Energy Investment (25%)	Renewable Energy Country Attractiveness Index (RECAI, Ernst & Young)
	Cleantech Innovation (30%)	Global Innovation Index (INSEAD), Cleantech Group, Heslin, Rothenberg, Farley & Mesiti p.c.
	Cleantech Commercialization (20%)	WWF Cleantech Group Global Cleantech Innovation Index 2014
	Green Investment Promotion & Facilitation (25%)	Scored by Dual Citizen LLC on scale of 0-10
Environment & Natural Capital (25%)	Agriculture (17%)	Environmental Performance Index 2014 (Yale University)
	Air quality (17%)	
	Water (17%)	
	Biodiversity & habitat (17%)	
	Fisheries (17%)	
	Forests (17%)	

Source: Dual Citizen (2016). "THE GLOBAL GREEN ECONOMY INDEX GGEI 2016 Measuring National Performance in the Green Economy". 5th Edition - September 2016.