A sequential benefit-of-the-doubt composite indicator*

Barnabé Walheer†

November 24, 2023

Abstract

In many contexts, performances are measured by aggregating indicators. To do so, practitioners have to choose how to normalize and weight the selected indicators. A popular method is the benefit-of-the-doubt (BoD) which constructs composite indicators based on relative weights and avoids normalization. When dealing with panel data, the BoD computes composite indicators using contemporaneous data only. A consequence is that composite indicators are over-estimated because the accumulation of best practices is ignored. Inspired by the production economics literature, we suggest new sequential composite indicators keeping the BoD spirit. These indicators are not based on contemporaneous data but include current and past information. By comparing the two approaches, we define the new concept of knowledge accumulation ratio. We use the sequential composite indicators to evaluate the vulnerability, readiness, and resilience to climate change of more than 180 countries over the 1995–2020 period. Our results highlight two main groups of countries: those with great need of new investments and an urgency for adaptation, and those well positioned but with some adaptation challenges.

Keywords: Data envelopment analysis; benefit-of-the doubt; composite indicator; sequential; Notre Dame Environmental Change Initiative.

---

*We thank the Editor Michael Yearworth and the three referees for their comments that have improved the paper substantially.

†HEC Liège, Université de Liège, Belgium. email: barnabe.walheer@uliege.be. This work is supported by the Fonds de la Recherche Scientifique (FNRS) under Grant No. J.0193.23.

1 Introduction

Composite indicators are used to measure and monitor the performances of entities, such as countries, regions, provinces, firms, plants, and schools. Such indicators have faced a growing popularity as they are easy to construct and interpret. The starting point of the definition of a composite indicator is the observation of a set of indicators for several entities. Next, it is required to normalize, weight, and aggregate the selected indicators. For each of these three steps, several approaches have been suggested (OECD, 2008; Greco et al., 2019).[1] To normalize the indicators, practitioners can use the minimum-maximum, the distance to a reference, or the percentage change approach. Next, to define the weights, an exogenous approach such as expert weighting or social survey, or an endogenous approach such as principal component analysis, factor analysis, regressions, and multicriteria analysis can be chosen. Finally, aggregation available techniques include the arithmetic and geometric averages and non-compensatory approaches.

Amongst all methods available to construct composite indicators, the benefit-of-the-doubt (BoD) approach (Cherchye et al., 2007b) has grown in popularity. The BoD uses linear programming to endogenously compute the weights. At this point, it is fair to note that the most popular linear method to construct endogenously composite indicators is the principal component analysis. The basic idea is to extract the maximum variance from a data set with a few orthogonal components. Indicators on the first component and the associated weights are used to define the composite indicators. For the endogeneous non-linear method, the multicriteria analysis is probably the most suitable as it can combine criteria of different natures and it provides enough flexibility in the decision-making process to practitioners.

For the BoD, the weights are selected such that every entity is evaluated under the best possible light, i.e. it gives the "benefit-of-the-doubt" to the entities. Some entities are therefore found as the best performers while others are lagging behind. The set of best performers defines the best practice possibility. Also, another advantage of the BoD is that it is not required to select a normalization procedure. A less desirable feature is that the endogenous weights can be too extreme when full flexibility is given. Such an aspect can be mitigated by selecting adequate weight restrictions

[1] We refer to Freudenberg (2003), Munda and Nardo (2005), Saltelli (2007), Walheer (2018b), and Barclay et al. (2019) for an overview of all available approaches to construct composite indicators and a critical assessment.

(Cherchye et al., 2008; Decancq and Lugo, 2013).

The BoD has been applied to a large set of different entities and topics revealing its practical attractiveness: the Europe 2020 objectives for countries (Walheer, 2019), waste management of regions (Rouge et al., 2017), teachers' effectiveness (Rogge, 2011), tourism vulnerability of provinces (Duro et al., 2021), quality of water companies' services (Sala-Garrido et al., 2021), ranking of world tourist destinations (Gómez-Vega et al., 2019), human development (Mariano et al., 2015), countries' financial inclusion (Lui and Walheer, 2022), ranking universities (Johnes, 2018), climate change vulnerability of nations (Edmonds et al., 2020), farmers' agricultural sustainability (Talukder, 2017), competitiveness of firms (Lafuente et al., 2020), transport sustainability of cities (Reisi et al., 2014), sustainable development goals of countries (Karuaihe et al., 2018), and arming sustainability (Chopin et al., 2021).

When entities are observed for several time periods, it is natural to explore how the performances have changed over time. The BoD offers an extra advantage in that case: it allows us to compute and decompose composite indicator changes. In particular, two dimensions can be distinguished (Cherchye et al., 2007a): a first index capturing pure performance changes and a second one measuring best performer changes. The former captures the extent to which the evaluated entities get closer to their best practice benchmark over time, i.e. the catching-up effect. The latter measures shifts in the best practice performances, i.e. the environmental change.

A potential drawback of the BoD when dealing with several time periods is that composite indicators of each time period are computed using contemporaneous data only. In other words, entities are evaluated using the information of a specific time period ignoring what has happened in the past. A consequence is that the set of best performers of a specific time period, defining the best practice possibility, ignores the best performers of previous periods. It turns out that a best practice possibility of previous periods may become unfeasible in the following periods. Also, best-practice possibility intersections are possible. Putting this differently, the performance evaluation exercise may be incomplete because the set of best performers does not include all relevant entities.

Moreover, both the pure performance and environmental changes can be positive or negative with the BoD. While this makes sense for pure performance changes, an entity can over- or under-perform the best practice over time, we believe that a negative environmental change has little sense in some contexts. In fact, a negative environmental change would imply that a degradation of the best practice is possible. Also, this might lead to confusion between pure and best practice performance regressions as best practice possibility intersections are possible.

Inspired by the sequential production possibility in production economics developed by Diewert (1980), Färe et al. (1985) and Tulkens and Vanden Eeckaut (1995) and used in several works (e.g. Shestalova, 2003; Henderson and Russell, 2005; Walheer 2016, 2021; Perelman and Walheer, 2020; Chambers and Pieralli, 2020), we suggest a sequential composite indicator. Such an indicator, while keeping the BoD spirit, is based on the fact that the best practice possibility can only improve over time. The resulting indicator is complete and performance slowdowns are only due to negative pure performance changes as environmental change can only be positive.

To define the sequential composite indicators, we introduce the new notion of peer set. Such a set allows us to naturally generalize the BoD linear programming as initially defined. Moreover, by comparing the contemporaneous and sequential composite indicators, we define the new concept of the knowledge accumulation ratio. Such a new ratio quantifies the impact of the accumulated knowledge on the entity's performance. In panel data settings, the knowledge accumulation ratio can be used to define new components for the catching-up and environmental change effects.

We use our new sequential BoD composite indicator to measure countries' vulnerability, i.e. capacity to adapt to the negative effects, and readiness, i.e. ability to leverage investments, to climate change. From a practical perspective, the sequential approach has to be used when it is difficult to assume that what has happened in the past has no impact on today's performances. This is especially the case at the aggregate level (e.g. sectors, industries, regions, countries). In these contexts, temporary slowdowns are more likely to be due to negative pure performance changes rather than environmental degradation. This is the case for the vulnerability and the readiness of countries to climate change. We may argue that, over time, countries have learned how to adapt to climate change and leverage investments resulting in an improvement of the best practices. Of course, worse performances are always possible but they cannot be attributed to a negative environmental change. They are more likely explained by country-specific reasons.

Using a unique database compiled by the *Notre Dame Environmental Change Initiative*, we evaluate the catching-up and environmental effects of more than 180 countries over the 1995–2020 period. Our findings reveal that countries are less vulnerable over time while they have fewer adaptative abilities. Also, the difference between the contemporaneous and sequential approaches is more pronounced for the vulnerability dimension, and the contemporaneous approach overestimates the catching-up effects while the environmental one is underestimated. Next, although knowledge accumulation has an average comparable effect on the vulnerability and the readiness of countries, disparities are more important for readiness. Finally, we cross countries' vulnerability and readiness to capture their resilience. A minority of countries are well-positioned to adapt to climate change. Most of the countries have an improvement in one dimension only. For vulnerability, countries with larger initial composite indicator values are those pushing the best practice possibility, and those with smaller initial composite indicator values have more benefited from the catching-up effect. For readiness, there is a path dependence for the catching-up effect as better performers have, on average, larger catching-up effects.

The rest of this paper unfolds as follows. In Section 2, we define the composite indicators and the relevant indexes. We present our application in Section 3 and conclude in Section 4.

2 Methodology

We start off by defining the benefit-of-the-doubt composite indicator and the new notions of peer set and knowledge accumulation ratio. Then, we explain how they can be used to measure performance changes over time using indexes. Finally, we establish several interesting relationships for the composite indicators and the indexes.

2.1 Benefit-of-the-doubt composite indicator

We assume that there are N entities (e.g. firms, plants, schools, countries, regions, sectors) and, that for every entity, we observe I indicators at time t, denoted by $\mathbf{x}_t = (x_{1t}, \ldots, x_{it}, \ldots, x_{It})$. The particularity of the benefit-of-the-doubt (BoD) approach is to evaluate the performance of each entity by means of a score in the $[0, 1]$ interval using linear programming where peers are used as the benchmark. As defined by Cherchye et al. (2007b), the linear programming to compute the composite indicator for a specific entity operating at $\mathbf{x}_t$ is given by:

$$\begin{aligned} CI(\mathbf{x}_t) = & \max_{\omega_{it}\ (i \in \{1, \ldots I\})} \sum_{i=1}^{I} \omega_{it} x_{it} \\ & \textbf{(C-1)}\ \sum_{i=1}^{I} \omega_{it} x_{ikt} \leq 1,\ \text{for } k = 1, \ldots, N; \\ & \textbf{(C-2)}\ \omega_{it} \geq 0,\ \text{for } i = 1, \ldots, I. \end{aligned} \tag{1}$$

$CI(\mathbf{x}_t)$ is by definition between 0 and 1. A value of 1 reflects a situation where the evaluated entity has the best performance at time t. Smaller values indicate worse performances. Also, note that the indicators are not required to be between 0 and 1, i.e. applying a normalization procedure to the indicators is not required. This is an advantage of the BoD.[2]

In light of our methodological contribution, we reformulate the previous linear programming using a set of peers at time s, labeled $D(s)$. The composite indicator for a specific entity operating at $\mathbf{x}_t$ with respect to the peer set $D(s)$ is given by:

$$\begin{aligned} CI(\mathbf{x}_t, D(s)) = & \max_{\omega_{it}\ (i \in \{1, \ldots I\})} \sum_{i=1}^{I} \omega_{it} x_{it} \\ & \textbf{(C-1)}\ \sum_{i=1}^{I} \omega_{it} x_{iks} \leq 1,\ \text{for } k \in D(s); \\ & \textbf{(C-2)}\ \omega_{it} \geq 0,\ \text{for } i = 1, \ldots, I. \end{aligned} \tag{2}$$

As $CI(\mathbf{x}_t, D(s))$ depends on the peer set $D(s)$, it is not necessarily bounded by unity. In fact, it is bounded by unity only when the evaluated entity is included in the peer set $D(s)$.[3] It turns out that the linear programming in (2) is a generalized version of (1) as it gives more flexibility to practitioners. For example, counterfactual peer sets, used when evaluating performances over time (see Section 2.2), can be considered. To get the linear programming in (1), it suffices to consider the following

[2] We point out that (1) gives full flexibility to the weights when computing the composite indicator. That is, the entities are evaluated in the best possible light. To avoid too extreme weights, additional constraints for the weights could be included in (1). See Cherchye et al (2007b, 2008) and Decancq and Lugo (2013).

[3] In that case, we have $\sum_{i=1}^{I} \omega_{it} x_{it} \leq 1$ as the evaluated entity belongs to $D(s)$.

contemporaneous peer set in (2):

$$C(t) = \{k \mid \text{entity } k \text{ is observed at time } t\}. \tag{3}$$

Such a notion of peer set is not used in the initial definition of the BoD composite indicator. In fact, at this point, it does not bring much to existing definitions. When using $C(t)$ in (2), we obtain $CI(\mathbf{x}_t, C(t))$ that is equal to $CI(\mathbf{x}_t)$ in (1). Intuitively, $C(t)$ contains the information at period t; ignoring what has happened before. We note that there is no particular relationship between contemporaneous sets of different time periods. A consequence is that the set of best performers, defining the best practice possibility, changes at each time period when relying on contemporaneous peer sets. It may thus be the case that there is an intersection between best practice possibilities of contemporaneous sets of different time periods. Also, a degradation of the best practice possibilities based on contemporaneous sets is possible.

As explained in the Introduction, composite indicators based on a contemporaneous peer set may present several drawbacks such as an incomplete performance analysis and a confusion between pure performance changes and best practice modifications. To overcome such potential issues, we suggest relying on a sequential peer set. It is defined at time s as follows:

$$S(s) = \{k \mid \text{entity } k \text{ is observed at time } s \text{ or before}\}. \tag{4}$$

The basic idea of sequential modeling is to include current and past observations. In a sense, by including past observations such an approach takes the accumulated best practices into account. $S(s)$ therefore simply accumulates all available observations until time s:

$$S(s') \subseteq S(s) \text{ if } s \geq s'. \tag{5}$$

This means that previous best practices are taken into account. That is, best practice possibilities can only explode over time. Graphically, it implies that intersections between best practice possibilities based on sequential peer sets are impossible. Moreover, the sequential peer set is linked to the contemporaneous peer sets defined before as follows:

$$S(s) = \bigcup_{r \leq s} C(r). \tag{6}$$

$S(s)$ is thus the union of all contemporaneous peer sets until time s.[4] This highlights, once more, that $S(s)$ contains the accumulation of best practices until t, while $C(s)$ only contains the knowledge at time s ignoring what has happened in the past. Note that both sets coincide only in the first year of observation and that, in the final year of observation, $S(s)$ contains all peers, i.e. all observed information.

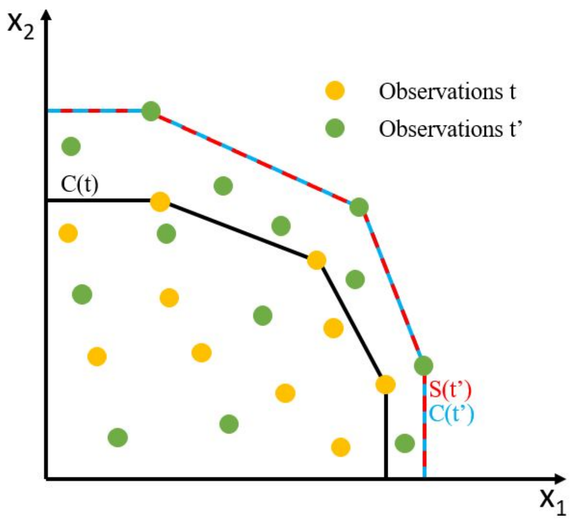
We illustrate the concept of sequential and contemporaneous peer sets in Figure 1. To simplify the graphical representation, we consider an illustrative example with two consecutive periods (t and t') and two indicators (x_1 and x_2), and assume convexity. We plot the observed points and the best practice frontiers defined by each peer set, i.e. the frontiers that envelop the observed points. In case 1, sequential modeling does not bring much as the best practice frontiers of both approaches coincide. The best practice frontier of the contemporaneous peer set moves up and right.

In cases 2 and 3, it is a very different story: there is an intersection between the best practice frontiers in case 2, and, in case 3, it is a degradation of the best practice possibility. In case 2, the best practice frontier based on the sequential peer set at t' is the union of the best practice frontiers of the contemporaneous sets at t and t'. In case 3, the best practice frontier of the sequential peer set at t' is equal to the one of the contemporaneous set at t. It means that, in both cases, the contemporaneous peer sets are not appropriate to conduct a fair performance evaluation exercise as there may be confusion between pure performance and best practice changes.

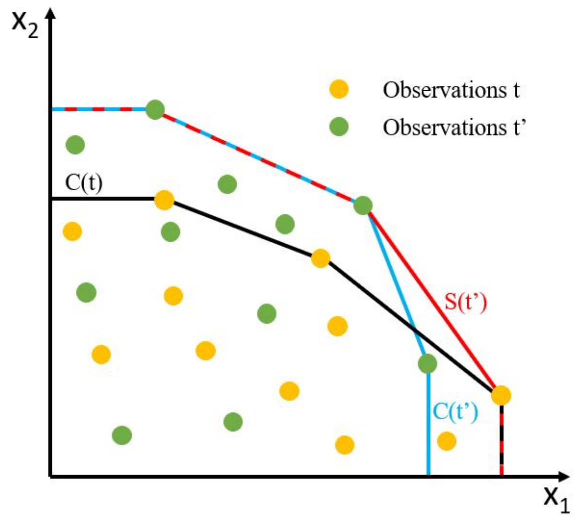
To compute the composite indicators based on the sequential peer set at time t, it suffices to use $S(t)$ as the peer set in (2). The resulting composite indicators, denoted $CI(\mathbf{x}_t, S(t))$, are also between 0 and 1 where greater values mean better performances over time. The distinguished feature with $CI(\mathbf{x}_t, C(t))$ is that all available information about the indicators until t has been used in the linear programming. It turns out that the sequential composite indicator is always smaller or equal to the composite

[4] Note that the union is non-convex but the linear programming in (2) is. There is no particular reason to do so except that it is how the initial BoD linear programming was defined. It is, however, possible to consider a non-convex version. For more discussion on the consequence of assuming convexity in linear programming technique refer, for example, to Walheer (2018b) and Jin et al. (2020).

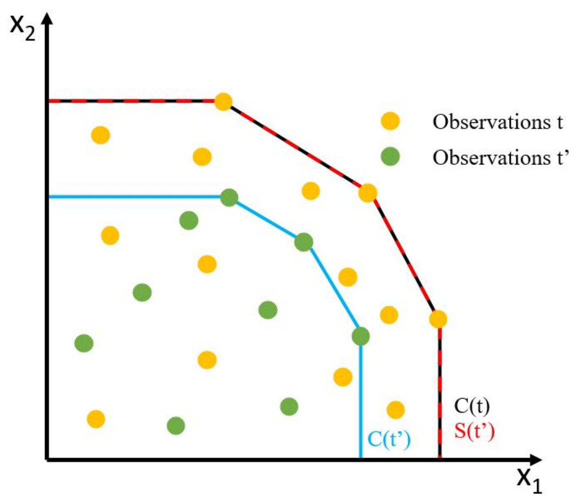
Figure 1: Illustrative examples



(a) case 1



(b) case 2



(c) case 3

indicator based on a contemporaneous peer set for a fixed time period. This follows directly from the definitions of the sequential and contemporaneous peer sets as the former is based on present and past information while the latter only uses current information.

Proposition 1. $\forall s, t : CI(\mathbf{x}_t, S(s)) \leq CI(\mathbf{x}_t, C(s))$.

Proof. It suffices to observe that $S(s)$ contains all information in $C(s)$. □

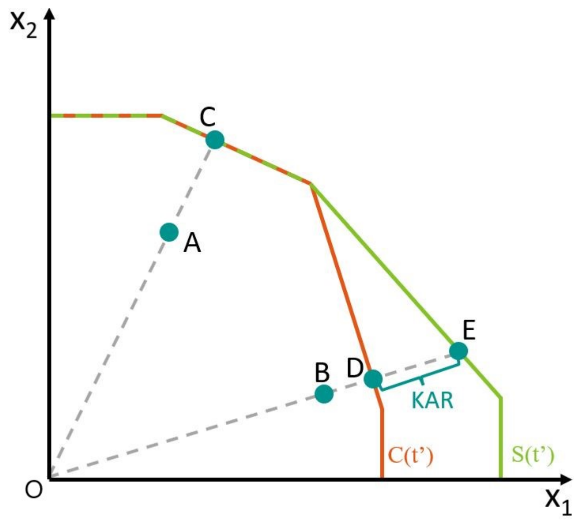
An important consequence of this result is that the composite indicator based on a contemporaneous peer set gives us an over-estimated performance measurement because the accumulated best practices are not taken well into account. To relate both versions of the composite indicator, we suggest the following knowledge accumulation ratio:

$$KAR(\mathbf{x}_t, C(s), S(s)) = \frac{CI(\mathbf{x}_t, C(s))}{CI(\mathbf{x}_t, S(s))}. \tag{7}$$

Such a ratio captures the distance between the best practice frontiers defined by the sequential and contemporaneous peer sets. In a sense, it captures the impact of the accumulated knowledge on the entity's performance: higher values mean that the accumulated knowledge has a larger impact. By construction, KAR has a benchmark value of unity. We illustrate the knowledge accumulation ratio in Figure 2. There, we represent the best practice frontiers for the contemporaneous and sequential peer sets at period t' of case 2 in Figure 1 (for better readability, we remove the observed points). We also show two fictitious points A and B that both present possible performance improvements.

In practice, the composite indicators are given by relating the distance from the origin to the fictitious point to the distance from the origin to the best practice frontier. For point A, we obtain $|OA|/|OC|$ for both the contemporaneous and sequential approaches making $KAR = 1$. For point B, $|OB|/|OE|$ and $|OB|/|OD|$ measure the composite indicator for the sequential and contemporaneous approaches, respectively. Using (7), we find that $KAR = |OE|/|OD| > 1$. That is, there is an impact of the accumulated knowledge for point B. Finally, we highlight that the accumulated knowledge may have an impact for the best performers. This is the case for point D that has a composite indicator of one based on the contemporaneous approach but a KAR larger than unity.

Figure 2: Knowledge accumulation ratio



We can rewrite (7) to obtain a useful decomposition of the composite indicator based on a contemporaneous peer set as follows:

$$CI(\mathbf{x}_t, C(s)) = CI(\mathbf{x}_t, S(s)) \times KAR(\mathbf{x}_t, C(s), S(s)). \tag{8}$$

This equation highlights the decomposition of the composite indicator based on a contemporaneous peer set into two parts: a first component capturing the pure performance, the composite indicator based on the sequential peer set, and a second measuring the accumulated knowledge impact, the knowledge accumulation ratio.

Finally, we point out that at the final time period, all knowledge has been accumulated. The corresponding sequential set therefore contains all peers, i.e. all available information. Such a set is conceptually similar to the global (Pastor and Lovell, 2005) or overall set (Afsharian and Ahn, 2015) developed in production economics contexts, and the corresponding knowledge ratio to the best practice gap.

2.2 Dynamic changes

When entities are observed over time, it is natural to measure the changes in their performance by means of indexes. Usually, such indexes are evaluated between an initial and a final time periods, denoted here by b and c.[5] It is, of course, possible

[5]For example, Shestalova (2003) takes $b = 1970$ and $c = 1990$, Henderson and Russell (2005) $b = 1965$ and $c = 1990$, Walheer (2016) $b = 1995$ and $c = 2008$, Walheer (2021) $b = 1965$ and

to do so for intermediate time periods but this raises several issues. First, it is often the case that indexes are not comparable across time periods (this is true for the majority of famous indexes used by practitioners). This is known as non-circular indexes (Diewert and Fox, 2017; Walheer, 2022). Intuitively, this means that there is no 'circle relationship' between indexes of different time periods, i.e. it is difficult to establish a stable ranking of the time periods. Second, when studying performance changes over time, it is required to have enough data variations to obtain interesting results. This would not be the case if time periods were too close. All in all, to avoid possible abnormal results and capture long-term patterns, the safest option is to compute indexes between b and c. Note that when data are observed over a very long time period, we may consider it as a collection of time intervals $[b, c]$.

Two main focuses of the performance change are, first, how entities perform over time with respect to the best practices, and, second, how the best practices have evolved over time. Both dimensions have to be used in a complementary fashion to understand the full picture. To evaluate entity performance change a simple and intuitive procedure is to take the ratio between the composite indicators in b and c:

$$CU(\mathbf{x}_b, \mathbf{x}_c, D(b), D(c)) = \frac{CI(\mathbf{x}_c, D(c))}{CI(\mathbf{x}_b, D(b))}. \tag{9}$$

$CU(\mathbf{x}_b, \mathbf{x}_c, D(b), D(c))$ captures the catching-up of an entity with the best practices between the base and current periods.[6] It has a benchmark value of unity and values greater (smaller) than one implies a progression (regression). In light of our results in Proposition 1, we immediately obtain the following relationship:

Proposition 2. $CU(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c)) \geq CU(\mathbf{x}_b, \mathbf{x}_c, S(b), S(c))$.

Proof. First, the composite indicators are equal for the base period: $CI(\mathbf{x}_b, C(b)) = CI(\mathbf{x}_b, S(b))$. Next, as sequential peer set at time c contains the contemporaneous peer set at time c, it implies that $CI(\mathbf{x}_c, S(c)) \leq CI(\mathbf{x}_c, C(c))$. □

Intuitively, such a result means that the performance changes are over-estimated when relying on $CU(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c))$ since past values are ignored. A part of this

$c = 2014$; Perelman and Walheer (2020) $b = 1965$ and $c = 2014$; Chambers and Pieralli (2020) $b = 1961$ and $c = 2004$.

[6] Another option is to use a difference. The ratio is preferred for several reasons such as it is unit-free, easy to interpret, and to measure over time. It is fairly easy to adapt the indexes developed here to a different version.

change is, in fact, due to the best practice change. The overestimation is due to the possible intersections between best practice possibility frontiers of contemporaneous peer sets. $CU(\mathbf{x}_b, \mathbf{x}_c, S(b), S(c))$ therefore reflects pure performance changes as it avoids the drawbacks of $CU(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c))$.

To capture the change of the best practices over time there are two natural candidates: $\frac{CI(\mathbf{x}_c, D(b))}{CI(\mathbf{x}_c, D(c))}$ that is the environmental change when year c is chosen for the evaluated entity; and $\frac{CI(\mathbf{x}_b, D(b))}{CI(\mathbf{x}_b, D(c))}$ when year b is chosen. In both cases, it is the same spirit: fix a time period for the evaluated entity and vary the peer sets. The benchmark value for both candidates is one, meaning that a value greater than unity implies a best practice performance progress while the converse reflects a performance regress. To avoid choosing a particular peer set, i.e. a certain path, Cherchye et al (2007a) suggested defining the environmental change as the geometric average of the environmental changes where years b and c are taken as the reference year for the evaluated entity:

$$EC(\mathbf{x}_b, \mathbf{x}_c, D(b), D(c)) = \left[\frac{CI(\mathbf{x}_c, D(b))}{CI(\mathbf{x}_c, D(c))} \times \frac{CI(\mathbf{x}_b, D(b))}{CI(\mathbf{x}_b, D(c))} \right]^{1/2}. \tag{10}$$

Such a procedure has been popularized by Caves et al. (1982) in production economics. Environmental changes based on the contemporaneous and sequential peer sets are related as follows:

Proposition 3. *$EC(\mathbf{x}_b, \mathbf{x}_c, S(b), S(c)) \geq EC(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c))$.*

Proof. First, the composite indicators are equal for the base period regardless the time period of the evaluated entity: $CI(\mathbf{x}_b, C(b)) = CI(\mathbf{x}_b, S(b))$ and $CI(\mathbf{x}_c, C(b)) = CI(\mathbf{x}_c, S(b))$ making the numerators of the environmental changes equal. Next, as sequential composite indicators are always smaller or equal to composite indicators – here: $CI(\mathbf{x}_c, S(c)) \leq CI(\mathbf{x}_c, C(c))$ and $CI(\mathbf{x}_b, S(c)) \leq CI(\mathbf{x}_b, C(c))$ – the denominator of the environmental change of the sequential composite indicator is larger. □

Again by taking all available information, environmental change is correctly measured. Another interesting result for the environmental change based on sequential peer sets is the following:

Proposition 4. *$EC(\mathbf{x}_b, \mathbf{x}_c, S(b), S(c)) \geq 1$.*

Proof. $CI(\mathbf{x}_c, S(b)) \leq CI(\mathbf{x}_c, S(c))$ and $CI(\mathbf{x}_b, S(b)) \leq CI(\mathbf{x}_b, S(c))$ as the evaluated entity is fixed and because $S(b) \subseteq S(c)$. $\square$

This result means that environmental degradation is impossible. Worse performances of an entity can only be due to bad pure performance. This result also implies that intersections between best practice possibilities of sequential sets are impossible.

Finally, we can follow a similar procedure to obtain a catching-up and environmental change effects for the knowledge accumulation dimension:

$$KACU(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c), S(b), S(c)) = \frac{KAR(\mathbf{x}_c, C(c), S(c))}{KAR(\mathbf{x}_b, C(b), S(b))} = \frac{CU(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c))}{CU(\mathbf{x}_b, \mathbf{x}_c, S(b), S(c))}. \tag{11}$$

$$\begin{aligned} KAEE(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c), S(b), S(c)) &= \left[\frac{KAR(\mathbf{x}_c, C(c), S(c))}{KAR(\mathbf{x}_b, C(b), S(b))} \times \frac{KAR(\mathbf{x}_c, C(c), S(c))}{KAR(\mathbf{x}_b, C(b), S(b))}\right]^{1/2}, \\ &= \frac{EC(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c))}{EC(\mathbf{x}_b, \mathbf{x}_c, S(b), S(c))}. \end{aligned} \tag{12}$$

We highlight that the catching-up and environmental effects of the knowledge accumulation indexes are linked to the similar indexes defined before for the contemporaneous and sequential peer sets. Furthermore, the indexes have a benchmark value of unity. By construction $KACU(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c), S(b), S(c))$ is bounded from below by one, while $KAEE(\mathbf{x}_b, \mathbf{x}_c, C(b), C(c), S(b), S(c))$ is bounded from above by one. Values further from one indicate a larger effect of knowledge accumulation for a specific factor.

3 Application

Climate change has an increasing impact on countries that differ in their geographical location and their socio-economic conditions. An important feature is the ability of countries to adapt to such increasing impact by taking concrete actions at various levels such as leveraging investments, government policy, and community awareness. To measure the impact and the adaptation, *Notre Dame Environmental Change Initiative* has developed a unique database of 74 variables to measure two main dimensions: vulnerability and readiness. The former measures the exposure, sensitivity, and capacity of countries to adapt to the negative effects of climate change; and the latter gives the ability of countries to leverage investments and convert them to adaptation actions.

Such initiative has recently received important attention in the policy and academic worlds (Russo et al., 2019; Tellman et al., 2020; Beirne et al., 2021; King et al., 2021; Dechezleprêtre et al., 2022; and D'Orazio, 2022). We can partition current studies into two main categories. The first one uses the composite indicators as given by *Notre Dame Environmental Change Initiative*, while the second tries to construct its own composite indicators using a certain approach. Our empirical study belongs to the second category.

Using the sequential composite indicators, we recompute vulnerability and readiness composite indicators, evaluate the catching-up effect and environmental change for each dimension, and quantify the effects of the knowledge accumulation. We believe that the sequential approach is more accurate than the contemporaneous approach in this context as it is difficult to believe that a degradation of the best practices over time is possible. Indeed, we may argue that, over time, countries have learned how to adapt to climate change and leverage investments resulting in an improvement of the best practices. Therefore, worse performances are more likely explained by country-specific reasons.

3.1 Definitions and data

The database constructed by *Notre Dame Environmental Change Initiative* consists of 74 variables to measure vulnerability and readiness. Data are given from 1995 to 2020.[7] Due to data availability, the vulnerability of 182 countries can be computed, while this number reaches 184 countries for readiness.

Vulnerability is defined using six main indicators ($\mathbf{x}_t$): food, water, health, ecosystem services, human habitat, and infrastructure. Each of the main indicators is obtained as an arithmetic average of 12 normalized variables as defined in Tables 5 and 6 in the Appendix. Readiness is based on three indicators: social, economic, and governance. The social and governance indicators are computed as arithmetic averages of four variables, while the economic one is based on one variable; see Tables 7 and 8 in the Appendix.

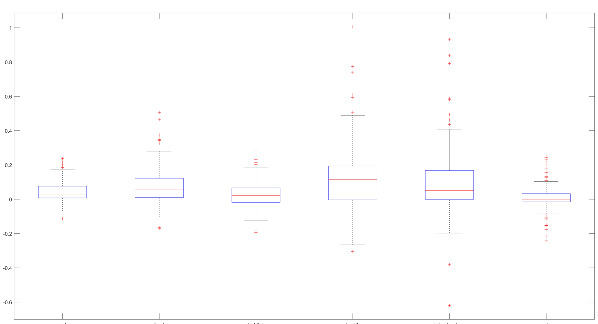
The variables are normalized and scaled using an adapted minimum-maximum

[7] Data can be freely retrieved from https://gain.nd.edu/our-work/country-index/download-data/

procedure. Instead of using the minimum point in the numerator, a reference point is adopted.[8] The reference point is the status of perfection, i.e. the best performance, as given by the data. Reference points are generally defined by common sense or by default. For example, the reference point for child malnutrition is 0%, and for reliable drinking water is 100%. For the quality of trade and transport-related infrastructure, it is 5 because the raw data ranged from 1 to 5. In all other cases, reference points correspond to the minimum or maximum of the variable. Finally, a direction is used to be sure that we have comparable composite indicators.[9] In our setting, we define the direction to be unity for both the vulnerability and the readiness to have comparable composite indicators.[10]

Descriptive statistics for the vulnerability and readiness indicators are given in Tables 1 and 2, and boxplots for the indicator (average) changes are in Figures 3 and 4. A positive change implies a performance improvement, i.e. less vulnerability or more readiness.

Figure 3: Vulnerability indicator change boxplot



[8]Applying the minimum-maximum procedure would give the following normalized variables: normalized variable $= \frac{\text{variable}-\text{minimum}}{\text{maximum}-\text{minimum}}$

[9]All in all, the normalized variable is given as follows: $\text{normalized variable} = \left|\text{direction} - \frac{\text{variable}-\text{reference point}}{\text{maximum}-\text{minimum}}\right|$.

[10]Note that we take one as the direction for both dimensions while it is 0 for vulnerability and 1 for readiness in the *Notre Dame Environmental Change Initiative*. The main reason for their choice is that they define an overall composite indicator, named $ND-GAIN$, that aggregates both vulnerability and readiness. As greater readiness indicators imply better performances, while greater vulnerability indicators imply worse performances, it is natural to select such a direction. In our empirical study, we do not wish to aggregate vulnerability and readiness but rather to have comparable composite indicators. This is why, we use the same procedure for both vulnerability and readiness. Note that another option is to rely on directional vectors (Fusco et al., 2020; Lahouel et al. 2022).

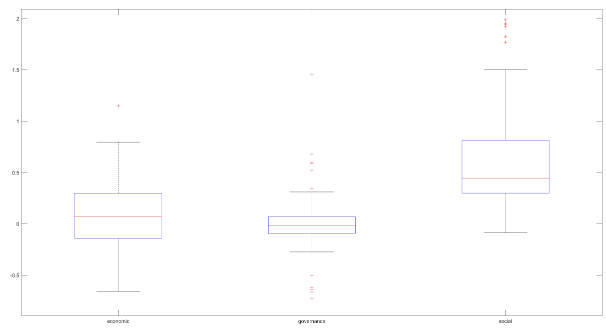
Table 1: Vulnerability indicator descriptive statistics

1995						
statistics	ecosystem	food	habitat	health	infrastructure	water
min	0.23	0.11	0.33	0.15	0.08	0.01
mean	0.45	0.43	0.51	0.43	0.34	0.36
median	0.44	0.42	0.52	0.41	0.33	0.34
max	0.66	0.78	0.79	0.84	0.73	0.72
std	0.09	0.15	0.09	0.17	0.10	0.11
2020						
min	0.22	0.14	0.35	0.21	0.08	0.01
mean	0.47	0.46	0.53	0.47	0.37	0.36
median	0.47	0.46	0.53	0.46	0.36	0.34
max	0.68	0.80	0.79	0.85	0.81	0.72
std	0.09	0.15	0.09	0.15	0.12	0.11

Table 2: Readiness indicator descriptive statistics

1995			
statistics	economic	governance	social
min	0.29	0.15	0.08
mean	0.41	0.51	0.25
median	0.41	0.46	0.20
max	0.53	0.87	0.64
std	0.06	0.18	0.13
2020			
statistics	economic	governance	social
min	0.11	0.10	0.13
mean	0.45	0.50	0.38
median	0.43	0.47	0.33
max	0.77	0.89	0.80
std	0.15	0.19	0.15

Figure 4: Readiness indicator change boxplot



All vulnerability aggregated indicators have overall averages and medians in the $[0.3 - 0.6]$ interval. Higher values are found for the habitat and lesser ones for the infrastructure and water. Also, they all present a positive average change but the number of countries with a positive change varies. It barely reaches 50% for water while it is almost 80% and 85% for ecosystem and food, respectively. More volatilities are found for the water and health core indicators. For readiness, we see that performances are overall worse for the social component while almost all countries have a positive change for that dimension. The governance and economic components have, on average, larger values but fewer positive changes. This is particularly true for the governance component.

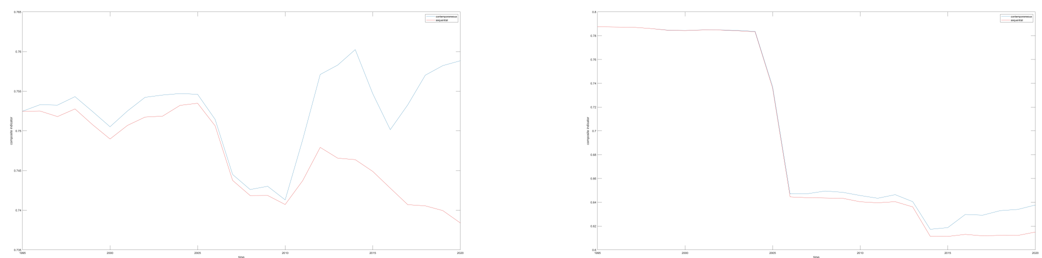
3.2 Results

Using the methodology explained in Section 2, we evaluate the vulnerability and readiness performances between 1995 and 2020. This means that we compute composite indicators in 1995 and 2020 for each country. We recall here that vulnerability is defined using six main indicators ($\mathbf{x}_t$) and readiness three (see Section 3.1). When following the contemporaneous approach, only data from one specific time period is used, while the sequential approach makes also use of all data before the evaluated time period. Once the composite indicators are computed, we can obtain the knowledge accumulation ratio and the catching-up and environmental effects. We note that the vulnerability and readiness composite indicators are computed as a simple arithmetic average based on contemporaneous data by *Notre Dame Environmental Change Initiative*.

3.2.1 Composite indicators

We start off our presentation of the results by showing the composite indicator averages per year obtained with the sequential approach in Figure 5. For comparison purposes, we also give the composite indicator averages when relying on the contemporaneous approach in the same figure.

Figure 5: Composite indicator averages



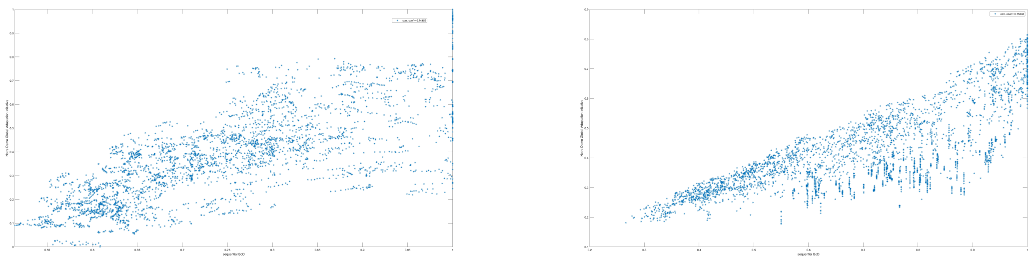
(a) Vulnerability (b) Readiness

First, we see that there is a decrease in the readiness performances over time while the performances are rather stable for vulnerability. This is true regardless of the approach used. Next, without surprise, the composite indicator averages for the sequential approach are always smaller than the ones for the contemporaneous approach. This is due to the theoretical relationship between both approaches as established in Proposition 1. Intuitively, by using both past and current information, the sequential approach better measures the performances. Finally, we also note that the gap between the sequential and contemporaneous composite indicators increases over time. They were similar until 2005 for readiness while they were always different for vulnerability (except, of course, for 1995). As explained in Section 2.1, the gap between the sequential and contemporaneous approaches can be understood as the knowledge accumulation impact; this aspect is studied in Section 3.2.2.

It is also important to compare our results to the ones obtained by the *Notre Dame Environmental Change Initiative*. After normalizing the variables as explained in Section 3.1, the indicators and composite indicators are obtained as simple arithmetic averages based on contemporaneous data. As the indicators are between 0 and 1, composite indicators lie also in the $[0, 1]$ interval. Nevertheless contrary to the BoD, absolute weights are used making a value of unity much more difficult to obtain.

Therefore it is better to use scatter plots and Pearson correlation coefficients to obtain a fair comparison. These are displayed in Figure 6.

Figure 6: Composite indicator scatter plots



(a) Vulnerability

(b) Readiness

The scatter plots highlight the positive connection between the composite indicators based on the sequential approach and the ones given by *Notre Dame Environmental Change Initiative*. This is true for both vulnerability and readiness. This claim is confirmed by the Pearson correlation coefficients which are both close to 0.75.

3.2.2 Catching-up and environmental effects

The catching-up effect tells how countries performed in 1995 and 2020 with respect to the best practices of that year. The environmental effect captures the shift in the best practice possibilities between 1995 and 2020. Descriptive statistics and distributions are given in Table 3 and Figures 7 and 8, respectively. We also give the composite indicators in 1995 in Table 3 to ease the interpretation and put the changes in perspective.

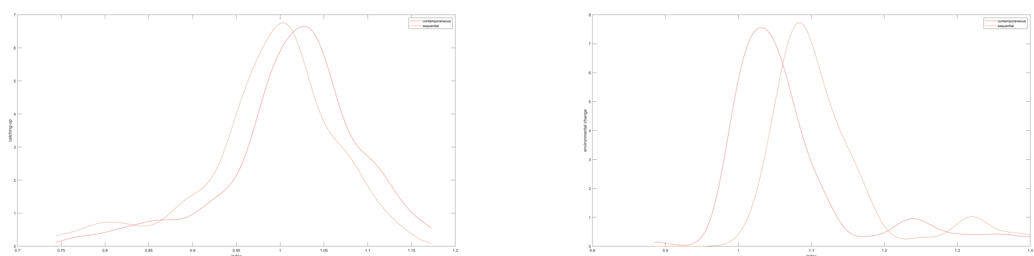
A first observation is that the initial average composite indicators for vulnerability and readiness are 0.75 and 0.78, respectively. Next, based on the catching-up effects, we see that there is a more important improvement for the vulnerability than for the readiness components regardless of the approach. To be fair, we note that the average value of the catching-up effect for vulnerability is 1.01 and the median is 1 indicating rather small improvement, if any. This means that, on average, countries are less vulnerable over time while they have less adaptative abilities. Finally, all countries present a positive environmental change for readiness, while a little bit more than 10% of the countries have a negative environmental effect for the vulnerability com-

Table 3: Descriptive statistics for dynamic changes

statistics	CI 1995	CU 1995–2020		EE 1995–2020	
		cont	seq	cont	seq
Vulnerability					
average	0.75	1.01	0.99	1.09	1.16
min	0.55	0.77	0.74	0.88	1.01
max	1	1.17	1.14	1.66	1.86
median	0.74	1.02	1.00	1.05	1.10
std	0.13	0.08	0.08	0.13	0.15
positive	–	65.47	50.36	89.21	100
Readiness					
average	0.78	0.80	0.77	1.77	1.95
min	0.55	0.41	0.37	1.10	1.13
max	1	1.46	1.28	2.21	2.82
median	0.79	0.82	0.79	1.78	1.92
std	0.12	0.18	0.17	0.35	0.50
positive	–	8.63	5.76	100	100

ponent. This implies that the difference between the contemporaneous and sequential approaches is more pronounced for the vulnerability dimension.

Figure 7: Vulnerability

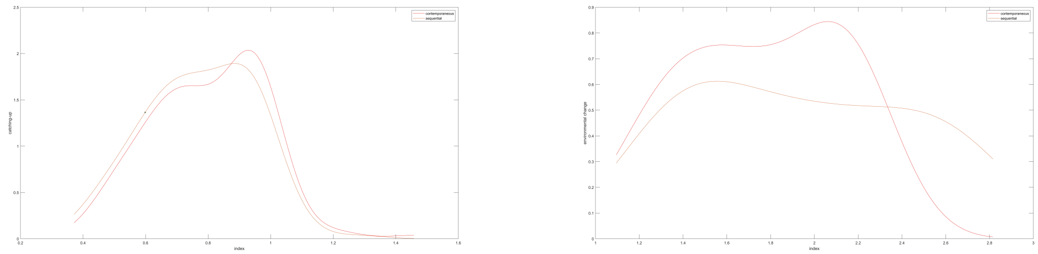


(a) Catching-up

(b) Environmental change

Next, thanks to the distributions in Figures 7 and 8, we see that the contemporaneous approach overestimates the catching-up effects, and the environmental one is underestimated. To verify whether the distributions are indeed different, we make use of the nonparametric Kolmogorov–Smirnov test (H_0: distributions are equal; H_1: distributions are different). Almost all $p-$values are smaller than 1% confirming our initial observations. The only exception is the catching-up effect of the readiness

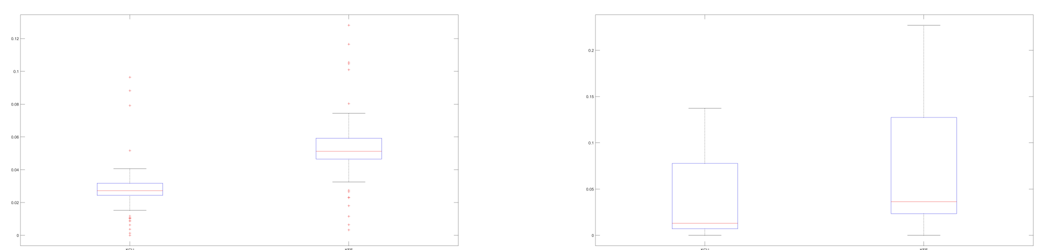
Figure 8: Readiness



(a) Catching-up (b) Environmental change

with a $p-$value of 0.29. The over- and under-estimation of the indexes based on the contemporaneous approach mean that a part of the environmental effect is put into the catching-up effect. The gap between both distributions is therefore interpreted as the knowledge accumulation effect. We plot this effect in Figure 9.[11]

Figure 9: Knowledge accumulation effects



(a) Vulnerability (b) Readiness

Although medians are comparable for the catching-up effect (0.0272 for vulnerability and 0.0130 for readiness) and the environmental change (0.0513 for vulnerability and 0.0364 for readiness), disparities are more important for readiness. Indeed, the interquartile ranges are 0.0074 for vulnerability and 0.0706 for readiness for the catching-up effect. For the environmental change, we find 0.0127 for vulnerability and 0.1039 for readiness. This is in line with the results in Table 3: the gaps between the indexes based on the contemporaneous and sequential composite indicators are greater for readiness even if negative environmental changes are only found for the vulnerability. Note that the median differences are supported by the $p-$values of

[11]Note that we subtract 1 and take the absolute values of both indexes to make them easily comparable. See Section 2.2 for more details about these knowledge accumulation indexes.

nonparametric Wilcoxon rank sum tests (H_0: medians are equal; H_1: medians are different) as they are all close to zero.

We end this part by crossing the catching-up and environmental effects for the vulnerability in Figure 10 and for readiness in Figure 11.

Figure 10: Vulnerability scatter plot

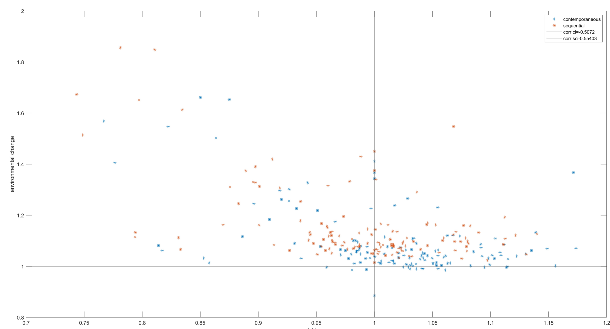
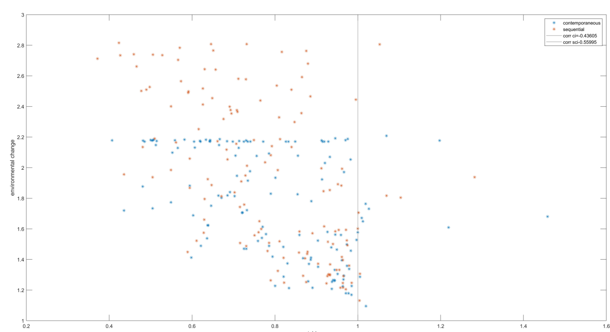


Figure 11: Readiness scatter plot



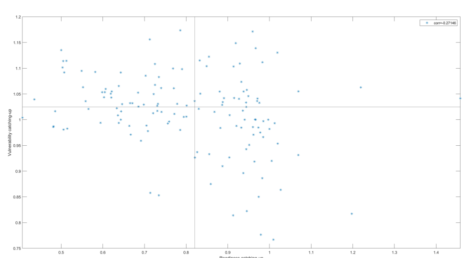
As pointed out before negative environmental changes are only found for the vulnerability. Figure 10 shows us that these negative environmental changes are rather close to unity (the minimum is 0.88) and only a short proportion of the countries are concerned (10.79 %). This supports our initial claim that negative environmental changes are not intuitive in our empirical context but are rather an undesirable side effect of the contemporaneous composite indicator methodology.

The scatter plots look similar for vulnerability and readiness even if more extreme points are found when using the sequential composite indicator approach. We compute Pearson correlation coefficients for vulnerability and readiness and for each approach. In all cases, negative coefficients are found significant at the 1% level, but they are larger under the sequential approach. This means that countries with larger environmental changes, i.e. those pushing the best practice possibilities, have, on average, smaller catching-up effects. For vulnerability, this is mainly due to a group of countries, on the upper right, with very large environmental changes and very low catching-up effects. For readiness, this is rather a tendency as very few countries have a positive catching-up effect. Finally, we highlight that there seems to be a fictitious limit for the environmental change for readiness in Figure 11 (around 2.2), this is probably another undesirable side effect of the contemporaneous composite indicator methodology.

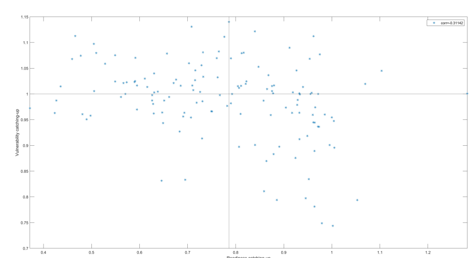
3.2.3 Resilience

As done by *Notre Dame Environmental Change Initiative*, we cross the vulnerability and readiness composite indicators to measure the resilience of countries in Figures 12 and 13. A particularity of our approach is to measure resilience from two points of view: catching up and environmental change. On each scatter plot, we add extra relevant pieces of information. On the one hand, we add the medians of the indexes by plotting horizontal and vertical lines (this allows us to partition countries into four groups). On the other hand, we give Pearson correlation coefficients (this allows us to measure the relationship).

Figure 12: Catching-up effect scatter plots



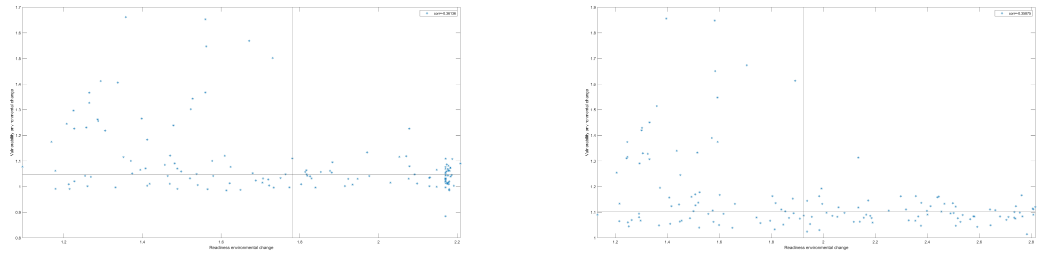
(a) Contemporaneous



(b) Sequential

In Figure 12, the coefficient of correlation is negative for both the contemporaneous and sequential approaches meaning that catching-up effects go usually in opposite directions. Note that the coefficient of correlation is slightly larger under the sequential approach. The best quadrant is the upper-right: countries present an

Figure 13: Environmental change scatter plots



(a) Contemporaneous (b) Sequential

improvement in both vulnerability and readiness. That is, they are well-positioned to adapt to climate change. The worst quadrant, with negative changes in both dimensions, lies on the lower left. Hopefully, few countries are present there. Between these two extreme cases, we have countries with an improvement in one dimension but a regression in the other one. These countries are in the upper-left and in the lower-right quadrants. In the upper-left, countries face less impact from climate change but have difficulties leveraging investments and converting them to adaptation actions. In the lower right, countries are more impacted by the climate change consequences but they have enough investments to adapt.

We again find negative significant correlation coefficients for the environmental changes in Figure 13. Indeed, very few countries present a very high environmental change for vulnerability and readiness at the same time (upper-left quadrant). We find two main groups. First, those with larger environmental changes have more often a smaller readiness for environmental change. Next, those with smaller vulnerability environmental change can present a low or high readiness environmental change.

Next, we verify whether larger catching-up and environmental effects occur more often for countries with smaller initial composite indicator values, i.e. in 1995. To do so, we regress the value of the catching-up or environmental index against the value of the composite indicator in 1995. GLS slope coefficients and associated $t-$statistics are displayed in Table 4.

All slope coefficients are significant for vulnerability. It is negative for the catching-up effect and positive for the environmental effect. This means that countries with larger initial composite indicator values are those pushing the best practice possibilities, and those with smaller initial composite indicator values have more benefited

Table 4: Regression results

statistics	CU		EE	
	cont	seq	cont	seq
Vulnerability				
slope coeff.	-0.26	-0.28	0.54	0.61
t-stat.	34.99	37.21	42.40	42.47
Readiness				
slope coeff.	0.37	0.38	-0.47	-0.51
t-stat.	19.93	21.16	1.64	1.55

from the catching-up effect. For readiness, only the slope coefficients of the caching-up effect are significant. They are positive implying a path-dependent for this effect as better performers have, on average, larger catching-up effects. No conclusions are found for the environmental change for that dimension as coefficients are insignificant.

4 Conclusion

The benefit-of-the-doubt (BoD) has faced increasing popularity among practitioners when computing composite indicators in various fields. This technique uses linear programming to endogenously compute the weights. Such weights are selected such that every entity is evaluated under the best possible light. In panel data contexts, the BoD computes composite indicators using contemporaneous data only. A consequence is that composite indicators are over-estimated because the accumulation of the best practices is ignored. Moreover, degradations and the intersections of the best practices are possible and this might lead to confusion between pure and best practice performance regressions.

Inspired by the production economics literature, we suggest new sequential composite indicators keeping the BoD spirit. These indicators include current and past information in the performance evaluation exercise. The resulting composite indicators are more informative and performance slowdowns are only due to negative pure performance changes. Finally, by comparing the contemporaneous and sequential composite indicators, we define the new concept of the knowledge accumulation ratio. Such a new ratio quantifies the impact of the accumulated knowledge on the entity's performance.

We use our new BoD sequential composite indicators to measure countries' vulnerability, i.e. capacity to adapt to the negative effects, and readiness, i.e. ability to leverage investments, to climate change. Using a unique database, we evaluate the performances of more than 180 countries over the 1995–2020 period. In such a context, it is difficult to believe that the best practices are worse over time and that the state of knowledge has decreased. It therefore represents a perfect fit to apply our new sequential composite indicators. Finally, we cross vulnerability and readiness to capture the resilience of countries.

Our findings reveal that the difference between the contemporaneous and sequential approaches is more pronounced for the vulnerability dimension and that the contemporaneous approach overestimates the catching-up effects and the environmental one is underestimated. Next, countries are less vulnerable over time and have less adaptative abilities. Also, in terms of resilience, some countries are well positioned to adapt to climate change, but this is not the rule: most of the countries have an improvement in one dimension but a regression in the other one.

References

[1] Afsharian, M., Ahn, H., 2015, "The overall Malmquist index: a new approach for measuring productivity changes over time", *Annals of Operations Research* 226, 1-27.

[2] Barclay, M., Dixon-Woods, M., Lyratzopoulos, G., 2019, "The problem with composite indicators", *BMJ quality & safety* 28(4), 338-344.

[3] Beirne, J., Renzhi, N., Volz, U., 2021, "Feeling the heat: Climate risks and the cost of sovereign borrowing", *International Review of Economics and Finance* 76, 920-936.

[4] Chambers, R. G., Pieralli, S., 2020, "The Sources of Measured US Agricultural Productivity Growth: Weather, Technological Change, and Adaptation", *American Journal of Agricultural Economics* 102(4), 1198-1226.

[5] Cherchye, L., Lovell, C. A., Moesen, W., Van Puyenbroeck, T., 2007a, "One market, one number? A composite indicator assessment of EU internal market dynamics", *European Economic Review* 51, 749-779.

[6] Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., 2007b, "An Introduction to benefit of the doubt composite indicators", *Social Indicators Research* 82, 111-145.

[7] Cherchye, L., Moesen W., Rogge N., Van Puyenbroeck T., Saisana M., Saltelli A., Liska R., Tarantola S., 2008, "Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index", *Journal of the Operational Research Society* 59, 239-251.

[8] Chopin, P., Mubaya, C. P., Descheemaeker, K., Öborn, I., Bergkvist, G., 2021, "Avenues for improving farming sustainability assessment with upgraded tools, sustainability framing and indicators. A review", *Agronomy for Sustainable Development* 41(2), 1-20.

[9] D'Orazio, P., 2022, "Mapping the emergence and diffusion of climate-related financial policies: Evidence from a cluster analysis on G20 countries", *International Economics* 169, 135-147.

[10] Decancq, K., Lugo, M. A., 2013, “Weights in multidimensional indices of well-being: An overview”, *Econometric Reviews* 32(1), 7-34.

[11] Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Chico, A. S., Stantcheva, S., 2022, “Fighting climate change: International attitudes toward climate policies (No. w30265)”, *National Bureau of Economic Research.*

[12] Diewert, W., 1980, “Capital and the theory of productivity measurement”, *American Economic Review* 70(2), 260-267.

[13] Diewert W. E., Fox K. J., 2017, “Decomposing productivity indexes into explanatory factors”, *European Journal of Operational Research* 256(1), 275-291.

[14] Duro, J. A., Perez-Laborda, A., Turrion-Prats, J., Fernández-Fernández, M., 2021, “Covid-19 and tourism vulnerability”, *Tourism Management Perspectives* 38, 100819.

[15] Edmonds, H. K., Lovell, J. E., Lovell, C. A. K., 2020, “A new composite climate change vulnerability index”, *Ecological Indicators* 117, 106529.

[16] Färe, R., Grabowski R., Grosskopf S., 1985, “Technical Efficiency of Philippine Agriculture”, *Journal of Productivity Analysis* 17, 205-214.

[17] Freudenberg, M., 2003, “Composite indicators of country performance: A critical assessment”, *OECD Science, Technology and Industry Working Papers* 16.

[18] Fusco, E., Vidoli, F., Rogge, N., 2020, “Spatial directional robust Benefit of the Doubt approach in presence of undesirable output: An application to Italian waste sector”, *Omega* 94, 102053.

[19] Gómez-Vega, M., Picazo-Tadeo, A. J., 2019, “Ranking world tourist destinations with a composite indicator of competitiveness: To weigh or not to weigh?”, *Tourism Management* 72, 281-291.

[20] Greco, S., Ishizaka, A., Tasiou, M., Torrisi, G., 2019, “On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness”, *Social indicators research* 141(1), 61-94.

[21] Henderson, D. J., Russell, R. R., 2005, “Human capital and convergence: a production-frontier approach”, *International Economic Review* 46(4),1167-1205.

[22] Jin, Q., Kerstens, K., Van de Woestyne, I., 2020, "Metafrontier productivity indices: Questioning the common convexification strategy", *European Journal of Operational Research* 283, 737-747.

[23] Johnes, J., 2018, "University rankings: What do they really show?", *Scientometrics* 115(1), 585-606.

[24] Kling, G., Volz, U., Murinde, V., Ayas, S., 2021, "The impact of climate vulnerability on firms' cost of capital and access to finance", , *World Development* 137, 105131.

[25] Lafuente, E., Leiva, J. C., Moreno-Gómez, J., Szerb, L., 2020, "A nonparametric analysis of competitiveness efficiency: The relevance of firm size and the configuration of competitive pillars ", *BRQ Business Research Quarterly* 23(3), 203-216.

[26] Lahouel, B. B., Zaied, Y. B., Taleb, L., Kočišová, K., 2022, "The assessment of socio-environmental performance change: A Benefit of the Doubt indicator based on Directional Distance Function and Malmquist productivity index", *Finance Research Letters* 49, 103164.

[27] Liu, F., Walheer, B., 2022, "Financial inclusion, financial technology, and economic development: A composite index approach", *Empirical Economics* 63, 1457-1487.

[28] Mariano, E. B., Sobreiro, V. A., do Nascimento Rebelatto, D. A., 2015, "Human development and data envelopment analysis: A structured literature review", *Omega* 54, 33-49.

[29] Munda, G., Nardo, M., 2005, "Constructing consistent composite indicators: the issue of weights".

[30] Nhemachena, C., Matchaya, G., Nhemachena, C. R., Karuaihe, S., Muchara, B., Nhlengethwa, S., 2018, "Measuring baseline agriculture-related sustainable development goals index for Southern Africa", *Sustainability* 10(3), 849.

[31] OECD, 2008, "Handbook on constructing composite indicators – methodology and user guide", *Paris: OECD Publications*.

[32] Perelman S., Walheer B., 2020, "Economic growth and under-investment: A nonparametric approach", *Economics Letters* 186, 108824.

[33] Reisi, M., Aye, L., Rajabifard, A., Ngo, T., 2014, "Transport sustainability index: Melbourne case study", *Ecological Indicators* 43, 288-296.

[34] Rogge, N., 2011, "Granting teachers the "benefit of the doubt" in performance evaluations", *International Journal of Educational Management* 25(6), 590-614.

[35] Rogge, N., De Jaeger, S., Lavigne, C., 2017, "Waste performance of NUTS 2-regions in the EU: A conditional directional distance Benefit-of-the-Doubt model", *Ecological Economics* 139, 19-32.

[36] Russo, S., Sillmann, J., Sippel, S., Barcikowska, M. J., Ghisetti, C., Smid, M., O'Neill, B., 2019, "Half a degree and rapid socioeconomic development matter for heatwave risk", , *Nature communications* 10(1), 136.

[37] Sala-Garrido, R., Mocholí-Arce, M., Molinos-Senante, M., 2021, "Assessing the quality of service of water companies: a 'benefit of the doubt'composite indicator", *Social Indicators Research* 155(1), 371-387.

[38] Saltelli, A., 2007, "Composite indicators between analysis and advocacy", *Social Indicators Research* 81, 65-77.

[39] Saltelli, A., D'Hombres, B., Jesinghaus, J., Manca, A., Mascherini, M., Nardo, M. Saisana, M., 2011, "Indicators for EU policies. Business as usual?", *Social Indicators Research* 102, 197-207.

[40] Shestalova, V., 2003, "Sequential Malmquist indices of productivity growth: an application to OECD industrial activities", *Journal of productivity Analysis* 19, 211-226.

[41] Talukder, B., W. Hipel, K., W. vanLoon, G., 2017, "Developing composite indicators for agricultural sustainability assessment: Effect of normalization and aggregation techniques ", *Resources* 6(4), 66.

[42] Tellman, B., Schank, C., Schwarz, B., Howe, P. D., de Sherbinin, A., 2020, "Using disaster outcomes to validate components of social vulnerability to floods: flood deaths and property damage across the USA", *Sustainability* 12(15), 6006.

[43] Tulkens, H., Vanden Eeckaut, P., 1995, "Non-parametric efficiency, progress and regress measures for panel data: methodological aspects", *European Journal of Operational Research* 80 (3), 474-499.

[44] Walheer, B., 2016, "Growth and convergence of the OECD countries: A multi-sector production-frontier approach", *European Journal of Operational Research* 252(2), 665-675.

[45] Walheer, B., 2018a, "Aggregation of metafrontier technology gap ratios: the case of European sectors in 1995-2015", *European Journal of Operational Research* 269, 1013-1026.

[46] Walheer, B., 2018b, "Decomposing the Europe 2020 index", *Social Indicators Research* 140, 875-905.

[47] Walheer, B., 2019, "Disentangling heterogeneity gaps and pure performance differences in composite indexes over time: the case of the Europe 2020 strategy", *Social Indicators Research* 143, 25-45.

[48] Walheer, B., 2021, "Labor productivity and technology heterogeneity", *Journal of Macroeconomics* 68, 103290.

[49] Walheer, B., 2022, "Global Malmquist and cost Malmquist indexes for group comparison", *Journal of productivity Analysis* 58, 75-93.

Appendix

Table 5: Vulnerability variables

Sector	Exposure variable	Sensitivity variable	Adaptive Capacity variable
Food	Projected change of cereal yields	Food import dependency	Agriculture capacity
	Projected population change	Rural Population	Child malnutrition
Water	Projected change of annual runoff	Freshwater withdrawal rate	Access to reliable drinking water
	Projected change of annual groundwater recharge	Water dependency ratio	Dam capacity
Health	Projected change of deaths from climate change induced diseases	Slum population	Medical staffs (physicians, nurses and midwives)
	Projected change of length of transmission season of vector-borne diseases	Dependency on external resource for health services	Access to improved sanitation facilities
Ecosystem services	Projected change of biome distribution	Dependency on natural capital	Protected biomes
	Projected change of marine biodiversity	Ecological footprint	Engagement in International environmental conventions
Human Habitat	Projected change of warm period	Urban concentration	Quality of trade and transport-related infrastructure
	Projected change of flood hazard	Age dependency ratio	Paved roads
Infrastructure	Projected change of hydropower generation capacity	Dependency on imported energy	Electricity access
	Projection of Sea Level Rise impacts	Population living under 5m above sea level	Disaster preparedness

Table 6: Vulnerability sectors

Sector	Definition
Exposure	The extent to which human society and its supporting sectors are stressed by the future changing climate conditions.
Sensitivity	The degree to which people and the sectors they depend upon are affected by climate related perturbations.
Adaptive Capacity	The ability of society and its supporting sectors to adjust to reduce potential damage and to respond to the negative consequences of climate events.

Table 7: Readiness variables

Dimension	Variables			
Economic	Doing business			
Governance	Political stability and non-violence	Control of corruption	Rule of law	Regulatory quality
Social	Social inequality	ICT infrastructure	Education	Innovation

Table 8: Readiness dimensions

Dimension	Definition
Economic	The investment climate that facilitates mobilizing capital from private sector.
Governance	The stability of the society and institutional arrangements that contribute to the investment risks
Social	Social conditions that help society to make efficient and equitable use of investment and yield more benefit from the investment.