

# Incorporating Sustainability Concerns in the Better Life Index: Application of Corrected Convex Non-parametric Least Squares Method

Hideyuki Mizobuchi<sup>1</sup>

Accepted: 26 February 2016 / Published online: 5 March 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** The OECD recently released a comprehensive set of 11 well-being indicators, the so-called Better Life Index (BLI), for 36 countries. The BLI covers a wide range of socio-economic aspects of life, which are essential to well-being. This well-being dataset allows us to compare countries' overall well-being. However, in spite of the BLI's wider coverage of variables, it fails to consider sustainability concerns. This study provides a practical proposal for comparing overall well-being by incorporating sustainability concerns. Using the World Bank's adjusted net savings data as a sustainability indicator, we add an extra dimension to the BLI. Then, we apply a composite indicator and aggregate these 12 indicators for each country into a single number. Moreover, we improve the current method for constructing composite indicators by adopting corrected convex non-parametric least squares ( $C^2NLS$ ). It is a typical problem in a non-parametric approach based on linear programming for countries' scores of composite indicators to become equal and their performance cannot be distinguished. This becomes even more severe if the number of sample countries is small or the number of aggregated indicators is large, which is the case of the present study dealing with 12 indicators for 36 countries. The use of  $C^2NLS$ , which is based on quadratic programming, overcomes this problem and allows us to order all countries in the sample completely. The empirical results show that the introduction of a sustainability indicator for comparisons does not change countries' overall rankings significantly. However, it certainly changes the ranking of some countries in both directions.

**Keywords** Sustainability · Composite indicators · Better Life Index · Data envelopment analysis · Benefit of the doubt approach

---

✉ Hideyuki Mizobuchi  
mizobuchi@econ.ryukoku.ac.jp

<sup>1</sup> Faculty of Economics, Ryukoku University, 67 Fukakusa Tsukamoto-cho, Fushimi-ku, Kyoto 612-8577, Japan

## 1 Introduction

Per capita GDP has long been used as a proxy measure of well-being. However, it is now widely recognized that income data provide only a partial perspective on the array of factors that affect people's lives. Given the problems with using GDP per capita as a measure of well-being, many researchers have been searching for alternative measures. In particular, the importance of incorporating a wider range of socio-economic conditions rather than income alone is now widely recognized. Drawing upon the recommendations for research on economic measurement problems by Stiglitz et al. (2009), the OECD identified 11 dimensions as being essential to well-being. The dimensions cover material living conditions, such as income and wealth, as well as quality of life (QOL), such as community, environment, and work–life balance. These dimensions are explored and analysed in detail by the OECD (2011). The OECD released 11 types of well-being indicators, known as the OECD Better Life Index (BLI), which covers the 34 OECD member countries and 2 non-member countries.<sup>1</sup> However, evaluation of overall well-being by summarizing the 11 individual indicators is left to users of the statistics.<sup>2</sup>

The 11 well-being indicators allow us to compare countries by the comprehensive well-being of their populations. However, there is an important component missing from these indicators, namely, sustainability. While the 11 well-being indicators capture the well-being of the current population, it is also a critical issue whether current well-being can be sustained in the future.

There is much in common between, on the one hand, the literature and debates on measures of well-being and, on the other hand, those of sustainability.<sup>3</sup> Levels of well-being are essentially what sustainability advocates would like to sustain. Thus, it is necessary to measure well-being before discussing its sustainability. On the other hand, without sustainability concerns, a country that guarantees the current generation better life circumstances by depleting natural resources at the cost of future generations is evaluated similarly to another country that sustains current well-being in the future, as long as the well-being of the current generation is the same in both countries. This, however, is entirely unconvincing. While the OECD concedes it is necessary to introduce sustainability concerns into the BLI, this has been left as a future issue. The present study attempts to provide a practical proposal on how to measure well-being by incorporating sustainability concerns. First, we add an extra indicator of sustainability concerns to the 11 well-being indicators of the BLI. Second, we aggregate these 12 indicators by the composite indicators.

As Dasgupta (2001) and Arrow et al. (2004) advocate, the productive base of economies, which consists of produced and natural capital and intangible assets, determines the well-being of people. Thus, a smaller productive base predicts lower well-being of future generations. The World Bank's adjusted net savings (World Bank 2011), which are considered a good measure of sustainability, capture the change in the productive base. Thus, we define the sustainability indicator by the adjusted net savings.

---

<sup>1</sup> There were 34 countries covered in 2011. A revised dataset released in 2012 includes 36 countries, incorporating Brazil and Russia.

<sup>2</sup> *Your Better Life Index* (<http://www.oecdbetterlifeindex.org/>) was designed as an interactive tool that allows users to assign the importance of each of the 11 topics and track the performance of countries.

<sup>3</sup> However, the two strands of research, such as measures of well-being and sustainability, tend to have been separated.

Composite indicators are used in order to measure multidimensional concepts, which are characterized by multiple individual indicators. Since individual indicators may trend in different directions to each other, the set of multiple indicators itself is not enough to provide an overall picture of multidimensional concepts across countries. Among a number of techniques to construct the composite indicator, the 'benefit of the doubt' (BOD) approach, which has received increasing attention from researchers, avoids subjectivity in the determination of weights (Mahlberg and Obersteiner 2001; Cherchye et al. 2004, 2007; Despotis 2005; OECD 2008). Under BOD, the weights are country-specific and endogenously determined such that they maximize the value of each country's resulting composite indicator. Thus, larger weights are given to the individual indicators (topics of well-being) on which each country performs well. The core idea is that a good relative score of a country on an individual indicator shows that it considers the individual indicator as relatively important. Therefore, for international comparisons based on BOD, a country cannot attribute the lower score of its composite indicator to a harmful or unfair weighting scheme.

BOD is rooted in data envelopment analysis (DEA), which is designed to compute efficiency. DEA is an established technique to measure the relative efficiency of decision-making units based on inputs and outputs of observations in a sample. It measures the efficiency of each unit by its distance from the production frontier, which is represented by the best-practice units. Formally, BOD is tantamount to the input-oriented DEA model in multiplier form, with all individual indicators as outputs and a 'dummy input' equal to one for all countries.

A well-known problem associated with DEA (thus, BOD) is that it often fails to differentiate the performance of all decision-making units completely, with the result that some units are ranked equally. Poor discriminatory performance of DEA is found when the sample size is small relative to the number of inputs and outputs. This arises from the DEA procedure of constructing the production frontier based on a linear-programming technique. Kuosmanen and Johnson (2010) introduce an alternative method, namely, corrected convex non-parametric least squares ( $C^2NLS$ ) for computing the efficiency measure.<sup>4</sup>  $C^2NLS$  constructs the production frontier based on quadratic programming. This new method offers certain advantages to the existing DEA. Kuosmanen and Johnson (2010) show that the estimates based on  $C^2NLS$  are consistent, and asymptotically unbiased, and yield smaller mean-squared error than the corresponding DEA efficiency estimators.

In addition to these advantages,  $C^2NLS$  has better discriminatory power than DEA, which allows for the complete ordering of the efficiency scores of all the units in a sample. Several other methods for improving the discriminatory power of DEA have been proposed already. Weight restriction is one of the most widely used method among them. The weights assigned to two well-being indicators should be made according to their relative importance in our application. There are two problems. First, setting common restrictions on weights removes the advantage of flexible weighting of BOD. Second, subjective judgement needs to be involved in deriving the weight restriction, since no widely accepted tool for setting a restriction is available. It is particularly difficult to reach consensus on the relative importance of different socio-economic conditions. Therefore, we believe that

---

<sup>4</sup> The efficiency score based on  $C^2NLS$  is constructed from residuals in non-parametric least squares subject to continuity, monotonicity, and concavity constraints, and this is referred to as convex non-parametric least squares (CNLS). Constructing the production frontier based on CNLS is part of the entire process of estimating the efficiency measure known as stochastic non-parametric envelopment of data (StoNED). See Kuosmanen (2008) and Kuosmanen and Kortelainen (2012).

C<sup>2</sup>NLS is more appropriate especially in the context of constructing a composite well-being indicator.<sup>5</sup> As suggested by Kuosmanen and Johnson (2010) and reported by Kuosmanen et al. (2015), the C<sup>2</sup>NLS method can be used for estimating shadow prices, setting performance targets, and identifying benchmarks in a similar fashion to the standard DEA. To the best of our knowledge, this is the first study that applies C<sup>2</sup>NLS to construct composite indicators.<sup>6</sup>

Mizobuchi (2014) applies the BOD method to construct a composite indicator which aggregates the 11 well-being indicators of the BLI. However, the problem of equal rankings among many countries associated with the BOD is established but left unresolved. It is known that the BOD method has a property that an additional indicator increases the score of the composite indicator. Thus, the more indicators the composite indicator aggregates, the more countries are likely to be ranked the highest, leading to weaker discrimination of country performance in terms of well-being. Since the procedure proposed by the present study is not subject to such a limitation, it is clearly a more appropriate tool for analysing the effect of introducing a new indicator into countries' performances.

Other than GDP per capita, the United Nations' Human Development Index (HDI) is the most popular measure of well-being. In addition, it is a composite indicator which aggregates fewer aspects than ours, such as income, education, and health. In the last 2 decades, a series of papers has introduced sustainability concerns into composite indicators (Desai 1995; Neumayer 2001; Costantini and Monni 2005; Ray 2014). Adjusted net savings or ecological footprints are used as sustainability indicators. While these indicators are, like ours, motivated by integrating sustainability concerns into measures of well-being or human development, their procedures of constructing composite indicators involve a simple geometric mean with ad hoc constant weight over countries, which has been adopted for the HDI. On the other hand, the C<sup>2</sup>NLS method allows for a more general and flexible weighting scheme, which assigns different and favourable weights to each country, like BOD.

The rest of this paper unfolds as follows. Section 2 discusses two approaches to construct a composite indicator. Section 3 explains the data of well-being indicators and sustainability. Section 4 computes composite indicators under different cases and compares them across countries. Section 5 examines the robustness of our findings. Section 6 concludes.

## 2 Methodology

The present study aggregates each of 36 countries' 11 well-being indicators and a single sustainability indicator into composite indicators. This is to compare countries' performance in terms of well-being, along with accounting for sustainability concerns. We adopt two approaches, namely, the BOD and C<sup>2</sup>NLS methods, to construct composite indicators.

---

<sup>5</sup> Principal component analysis has been applied to reducing the number of inputs and outputs in DEA. While each principal component is a linear combination of underlying variables, negative weight is often assigned to each variable. Since underlying well-being indicators are constructed so that larger values indicate better socio-economic conditions, it is not appropriate to apply this to the current problem of constructing the composite well-being indicator.

<sup>6</sup> Kuosmanen and Johnson (2010) refers to Kuosmanen et al. (2006) as an attempt to measure efficiencies of all units simultaneously in a single large problem like C<sup>2</sup>NLS. It is worth noting that Kuosmanen et al. (2006) are motivated by the research on composite indicator of sustainable development of (Cherchye and Kuosmanen 2004). This point is credited to an anonymous referee.

Since they are sufficiently versatile to be applicable to a variety of problems and situations, we explain these methods below in a more general setting independent of the number of countries and underlying individual indicators.

We assume there are  $K$  countries and that the well-being of people in a country  $k$  is characterized by a set of  $M$  individual indicators,  $\mathbf{y}_k = (y_{1,k}, \dots, y_{M,k})'$ , with  $y_{m,k}$  representing the value of the  $m$ -th individual indicator of country  $k$ . Suppose that there are some sustainability indicators among  $M$  indicators, constituting  $\mathbf{y}_k$ .<sup>7</sup> BOD aggregates these individual indicators using their weighted average. We denote a set of weights for country  $k$  by  $\boldsymbol{\mu}_k = (\mu_{1,k}, \dots, \mu_{M,k})'$ , whose component  $\mu_{m,k}$  represents the weight of the  $m$ -th individual indicator. The composite indicator based on BOD for country  $c$ ,  $CI_{BOD,c}$ , is formulated as follows:

$$CI_{BOD,c} = \max_{\mu_{1,c}, \dots, \mu_{M,c}} \left\{ \sum_{m=1}^M \mu_{m,c} y_{m,c} \mid \sum_{m=1}^M \mu_{m,c} y_{m,k} \leq 1 \right. \\ \left. \text{for } k = 1, \dots, K; \mu_{m,c} \geq 0 \text{ for } m = 1, \dots, M \right\} \tag{1}$$

For the international comparison, the abovementioned procedure is repeated for every country in our sample. The weight  $(\mu_{1,c}, \dots, \mu_{M,c})$  is determined endogenously to maximize the value of the composite indicator for country  $c$ . Thus, a larger weight is assigned to an individual indicator on which the country performs well. In this procedure, a good performance of country  $c$  on an individual indicator is considered to indicate that the country prioritizes this indicator. Therefore, countries cannot excuse their poor performance by an unfair weighting scheme, because any weight other than that used for their evaluation would not improve their position. The first constraint in (1) is that every country in a sample has a resulting composite indicator smaller than one when applying the most favourable weights for the evaluated country  $c$ . Thus, the resulting composite indicator for country  $c$  will be less than or equal to one.

As Mahlberg and Obersteiner (2001) graphically illustrate, an alternative interpretation of  $CI_{BOD,c}$  is possible. Given individual indicators  $\mathbf{y}$  as outputs and a dummy input equal to one for all countries,  $CI_{BOD,c}$  is considered as evaluating the performance of country  $c$  in terms of its productive efficiency.<sup>8</sup> Strictly speaking,  $CI_{BOD,c}$  equals the distance between country  $c$ 's well-being indicator  $\mathbf{y}_c$  and the production frontier constructed over countries' input and output sample data by DEA. The production frontier represents the optimal practices to produce well-being. Countries whose well-being indicators  $\mathbf{y}$  are on the frontier are considered the most efficient and are ranked the highest under BOD. The farther from the frontier and the closer to the origin the individual indicators of a country are, the lower its performance is evaluated.

One of the problems associated with BOD is that multiple countries are located on the production frontier and they are evaluated the highest. Thus, we fail to distinguish their performance. Such weak discriminatory power of the composite indicator based on BOD would be more evident in a case in which the observations are small relative to the number

<sup>7</sup> Sustainability indicators and well-being indicators are treated alike. Thus, we do not differentiate them in our notation.

<sup>8</sup> The dummy input can be considered as a helmsman in each country, and is intended to provide people with a better life. This interpretation goes back to Lovell et al. (1995).

of underlying indicators. As shown in Sects. 3 and 4, we apply composite indicators for aggregating 12 or 11 indicators of 36 countries. BOD fails to fully discriminate countries' performances in this case. Weak discriminatory power is a well-known problem of DEA. C<sup>2</sup>NLS has a decisive advantage over DEA and BOD by improving discriminatory power significantly.

C<sup>2</sup>NLS is implemented in two stages. First, the production frontier is estimated by solving convex non-parametric least squares (CNLS). In the situation of a dummy input that is equal to one, the production frontier is formulated as follows:

$$\min_{\substack{\varepsilon_1, \dots, \varepsilon_K, \\ \mu_1, \dots, \mu_K}} \left\{ \begin{array}{l} \sum_{i=1}^K \varepsilon_i^2 \\ \sum_{m=1}^M \mu_{m,i} y_{m,i} + \varepsilon_i = 1; \\ \sum_{m=1}^M \mu_{m,i} y_{m,i} \geq \sum_{m=1}^M \mu_{m,j} y_{m,i} \\ \text{for all } i, j = 1, \dots, K; \\ \mu_{m,k} \geq 0 \text{ for all } m = 1, \dots, M \text{ and all } k = 1, \dots, K \end{array} \right\} \quad (2)$$

Let  $\mu_1^* = (\mu_{1,1}^*, \dots, \mu_{1,M}^*), \dots, \mu_K^* = (\mu_{K,1}^*, \dots, \mu_{K,M}^*)$  be a solution to optimization problem (2). The composite indicator  $CI_{BOD,c}$  is the efficiency measure for country  $c$  based on DEA.<sup>9</sup> The corresponding efficiency measure based on CNLS,  $CI_{CNLS,c}$ , is derived so that  $CI_{CNLS,c} = \sum_{m=1}^M \mu_{m,c}^* y_{m,c}$  for all  $c = 1, \dots, K$ . Second, the efficiency measures are adjusted so that the maximum value becomes one. Then, the composite indicator based on C<sup>2</sup>NLS for country  $c$  is defined as follows<sup>10</sup>:

$$CI_{C^2NLS,c} = CI_{CNLS,c} - \left( \max_{i \in \{1, \dots, K\}} CI_{CNLS,i} - 1 \right) \quad (3)$$

We explain the characteristics of  $CI_{C^2NLS}$  in comparison with  $CI_{BOD}$ .  $CI_{C^2NLS}$  share flexible weighting with  $CI_{BOD}$  in the sense that every country is allowed to adopt a favourable weight. However, the determination of weights differs between the two measures. While the weights in  $C_{BOD}$  are solved independently for each country in (1), the weights in  $C_{CNLS}$  and  $C_{C^2NLS}$  are solved simultaneously for all countries in (2). Thus, the well-being indicators  $y$  of countries with lower  $CI_{BOD}$  have no impact on the  $CI_{BOD}$  of other countries. On the other hand, the  $C_{CNLS}$  and  $C_{C^2NLS}$  of all countries are affected by the well-being indicators  $y$  of other countries.

Optimisation problem (1) is formulated alternatively by the following Eq. (4),<sup>11</sup> which helps us to explain the greater power of discrimination of  $CI_{C^2NLS}$ , compared with  $CI_{BOD}$ .

<sup>9</sup> Formally,  $CI_{BOD,c}$  is known as the input-oriented Farrell efficiency.

<sup>10</sup> Following the formulation of Kuosmanen and Johnson (2010), the C<sup>2</sup>NLS efficiency estimator for unit  $c$ ,  $\varepsilon_{C^2NLS,c}$  is constructed by using the CNLS efficiency estimator  $\varepsilon_{CNLS,c}$  such as  $\varepsilon_{C^2NLS,c} = \varepsilon_{CNLS,c} - \min_{i \in \{1, \dots, K\}} \varepsilon_{CNLS,i}$ . Since  $CI_{CNLS,i} + \varepsilon_{CNLS,i} = 1$  and  $CI_{C^2NLS,i} + \varepsilon_{C^2NLS,i} = 1$  for all  $i = 1, \dots, K$ , it leads to Eq. (3).

<sup>11</sup> Kuosmanen and Johnson (2010) introduced this alternative formulation. It has the advantage of including all countries in a single large problem. Thus, the weights in  $C_{BOD}$ , which need to be solved repeatedly for each country in (1), are solved simultaneously for all countries in (4).

$$\min_{\varepsilon_1, \dots, \varepsilon_K, \mu_1, \dots, \mu_K} \left\{ \begin{array}{l} \sum_{i=1}^K \varepsilon_i^2 \\ \varepsilon_i \geq 0; \\ \sum_{m=1}^M \mu_{m,i} y_{m,i} + \varepsilon_i = 1; \\ \sum_{m=1}^M \mu_{m,i} y_{m,i} \geq \sum_{m=1}^M \mu_{m,j} y_{m,i} \\ \text{for all } i, j = 1, \dots, K; \\ \mu_{m,k} \geq 0 \text{ for all } m = 1, \dots, M \text{ and all } k = 1, \dots, K \end{array} \right\} \quad (4)$$

Equation (4) is simply a sign-constrained variant of the CNLS problem of Eq. (2). We can interpret these equations as follows: both BOD and CNLS maximize the value of each country’s composite indicator by adopting its most favourable weight. While BOD faces a constraint that the resulting composite indicator is below one, CNLS is free from such a constraint. Thus, there are countries whose composite indicators become larger than one under the CNLS approach. In the case in which multiple countries are ranked the highest with the value of one for their composite indicators under the BOD, the application of the CNLS approach allows us to differentiate the performances of these countries.

### 3 Data

#### 3.1 OECD Better Life Index

Amid growing concerns about identifying an alternative approach to measuring well-being, in 2011, the OECD launched the Better Life Initiative and released a set of 11 well-being indicators covering the 34 OECD member countries, comprising advanced and emerging economies. The data were updated in 2012 and more dimensions were added to calculate indicators. Moreover, the country coverage was expanded beyond the OECD to include Brazil and Russia. We use the most recent data covering individual indicators, which were released in 2014. The data are cross-sectional for a single year around 2011, as explained later in this subsection.

The 11 individual well-being indicators evaluate topics that the OECD considers essential to people’s well-being. Each individual indicator corresponding to each topic is based on between one and four underlying secondary indicators, which are expressed in different units, such as dollars, years, or numbers of people. To compare and aggregate values expressed in different units, the values are normalized. This normalization is performed according to a standard formula which converts the original values of the individual indicators into numbers between 0 and 10, as follows:<sup>12</sup>

$$\frac{\text{value to convert} - \text{minimum value}}{\text{maximum value} - \text{minimum value}} \times 10 \quad (5)$$

Within each topic, the secondary indicators are averaged with equal weight. For example, while the topic of the environment is constructed using two secondary indicators, *water quality* and *air pollution*, first, their scores are normalized in a range between 0 and 10. Then, they are aggregated as follows:  $\frac{\text{waterqualityscore} + \text{airpollutionscore}}{2}$ . The 11 individual indicators and their corresponding 24 secondary indicators are shown below.

<sup>12</sup> This is the normalization adopted by the OECD.

1. Income  
(1.1 Household income; 1.2 Household financial wealth)
2. Jobs  
(2.1 Employment rate; 2.2 Personal earnings; 2.3 Job security; 2.4 Long-term unemployment rate)
3. Housing  
(3.1 Rooms per person; 3.2 Housing expenditure; 3.3 Dwellings with basic facilities)
4. Work–life balance  
(4.1 Employees working very long hours; 4.2 Time devoted to leisure and personal care)
5. Health  
(5.1 Life expectancy; 5.2 Self-reported health)
6. Education  
(6.1 Educational attainment; 6.2 Years in education; 6.3 Students' skills)
7. Community  
(7.1 Social network)
8. Civic engagement  
(8.1 Consultation on rule-making; 8.2 Voter turnout)
9. Environment  
(9.1 Water quality; 9.2 Air pollution)
10. Safety  
(10.1 Homicide rate; 10.2 Assault rate)
11. Life satisfaction  
(11.1 Life satisfaction)

Among the 11 individual well-being indicators, the first 3 are categorized under material living conditions and the remaining 8 are categorized as QOL. According to the dataset released by the OECD Better Life Initiative, the data years of the underlying detailed indicators range from 2008 to 2013. Averaging them with each topic equally weighted suggests a year close to 2011. Thus, we consider that the 11 indicators of each country measure the socioeconomic situation of people around 2011.

Table 1 summarizes the statistics of the 11 well-being indicators; the complete data is provided in Table 9 of the Appendix. As the OECD (2011, 2013) finds, these tables show that while life is good in many dimensions in some countries, such as Australia, Canada, Denmark, New Zealand, Norway, and Sweden, it is significantly less so in other countries, such as Chile, Mexico, Portugal, Russia, and Turkey. While the latter group of countries is characterized by lower per capita income, except for Portugal, the former group does not necessarily comprise the richest countries.

Hereafter, we group countries based on per capita GDP to consider the link between well-being and economic development, which is well reflected in per capita GDP. There are three groups, as follows: four high-income countries with per capita GDP more than USD 45,000; 13 middle-income countries with per capita GDP between USD 30,000 and 40,000; and 19 low-income countries with per capita GDP less than USD 30,000.<sup>13</sup> Table 1 suggests that people's well-being improves in many aspects as income grows. However, this is not always true, especially in some of the topics categorized under QOL, such as community, education, civic engagement, and work–life balance. In these respects, the

<sup>13</sup> The 4 high-income countries are Luxembourg, Norway, Switzerland, and the US; the 13 low-income countries are Brazil, Chile, Czech Republic, Estonia, Greece, Hungary, Mexico, Poland, Portugal, Russia, Slovak Republic, Slovenia, and Turkey; and the other 19 countries are middle-income countries.



**Table 1** Descriptive statistics, OECD well-being indicators

	Housing	Income	Jobs	Community	Education	Environment	Civic engagement	Health	Life satisfaction	Safety	Work-life balance
<i>Mean</i>											
Overall	5.6	3.5	6.7	7.6	6.5	7.2	5.2	7.0	6.2	8.3	6.7
High income	7.0	7.0	8.8	8.4	6.6	8.3	5.5	8.5	8.8	8.7	7.3
Middle income	6.3	4.3	7.3	8.6	7.2	7.9	5.8	8.0	7.4	8.9	7.0
Low income	4.1	1.3	5.2	5.9	5.5	5.9	4.3	5.2	3.7	7.2	6.0
<i>Median</i>	6.0	3.6	7.1	8.2	7.2	7.7	5.3	7.8	7.3	8.8	7.2
<i>SD</i>	1.5	2.2	1.8	2.3	1.8	1.8	1.7	2.0	3.0	1.9	1.9
<i>Max</i>	7.8	10.0	9.6	10.0	9.2	9.8	9.5	9.4	10.0	9.9	9.8
<i>Min</i>	2.2	0.1	2.2	0.0	1.2	2.9	2.1	0.6	0.0	0.4	0.0

average person in middle-income countries enjoys a better life than the average person in high-income countries. It is also noteworthy that the life satisfaction indicator, which has the largest standard deviation, differs significantly across countries.

### 3.2 The World Bank's Adjusted Net Savings Dataset

Adjusted net savings, also known as genuine savings, are designated a sustainability indicator provided by the World Bank. Its theoretical grounding is the notion that sustainability requires the maintenance of a constant stock of the 'productive base'. This captures the extended wealth, which is not limited to natural resources but also includes physical, produced, and intangible capital, such as human capital and the rule of law. Adjusted net savings are considered as the change in this total wealth over a given time period. As Dasgupta (2001) and Arrow et al. (2004) advocate, the productive base is the source of well-being of future generations. Thus, negative adjusted net savings indicate future generations fail to be given an opportunity set which is at least as large as that available to current generations.

The World Bank computes adjusted net savings as follows:

$$\begin{aligned} \text{Adjusted net savings} = & \text{net national savings} + \text{education expenditure} \\ & - \text{natural resource depletion} - \text{carbon dioxide damage.} \end{aligned}$$

Net national savings is gross fixed capital formation minus the consumption of fixed capital, which indicates the amount of added produced capital. Education expenditure indicates the amount of added human capital, which makes up the larger share of intangible capital. Natural resource depletion is the sum of net forest depletion, energy depletion, and mineral depletion. Natural resource depletion with carbon dioxide damage captures the loss of natural capital. As the productive base consists of produced, natural, and intangible capital, adjusted net savings consists of changes in produced, natural, and intangible capital.

Instead of using the variable of adjusted net savings released by the World Bank, we recompute the adjusted net savings, this time without including education expenditure, as follows:

$$\begin{aligned} \text{Adjusted net savings} = & \text{net national savings} - \text{natural resource depletion} \\ & - \text{carbon dioxide damage.} \end{aligned}$$

There are two reasons we exclude education expenditure from the construction of the sustainability indicator in the present study. First, education expenditure is not a good measure of the changes in intangible capital.<sup>14</sup> Education expenditure captures changes in human capital but lacks significant parts of other intangible capital, such as the rule of law and social capital. Second, the inclusion of education expenditure leads to double counting. An increase in government expenditure on education usually improves people's life conditions in terms of education. This might arise from smaller class sizes or more motivated teachers. The returns from educational investment are considered more immediate than changes in produced and natural capital. Since the *education* well-being indicator of the BLI already captures the impact of education expenditure, we exclude it from the sustainability indicator to avoid double counting. Finally, we normalize the value of adjusted

<sup>14</sup> Moreover, the depreciation of human capital is dismissed. Thus, strictly speaking, education expenditure is a dubious measure, even for changes in human capital.

net savings into the range between 0 and 10 based on Eq. (5) and this defines the sustainability indicator.

Table 2 summarizes the statistics of adjusted net savings and the ingredients thereof, along with the sustainability indicator. Net national savings, which indicate the net investment of produced capital, are much larger than the depletion of natural resources and carbon dioxide damage. The gap seems to be expanding as economies grow. While high-income countries seem to have larger natural resource depletion, once we exclude Norway as an exception, their average level of natural resource depletion is smaller than that of middle-income countries. Thus, the results show that as economies grow, natural resource depletion declines in general.

## 4 Results

We compute composite indicators based on BOD,  $CI_{BOD}$  and  $C^2NLS$ ,  $CI_{C^2NLS}$  in two cases: first, when 11 well-being indicators are aggregated, and second, when 12 indicators are aggregated (11 well-being indicators and 1 sustainability indicator). While the values of  $CI_{BOD}$  and  $CI_{C^2NLS}$  are originally set to be between 0 and 1 in Eqs. (1) and (3), we multiply them by 10 in this section so that their values are between 0 and 10, which follows the 11 well-being indicators of the BLI. The purpose of this section is to empirically show how the change in the methodology and the inclusion of a sustainability indicator changes the score and ranking of the composite indicators.

Table 3 presents the empirical results, containing the score and ranking of composite indicators along with existing HDI and GDP per capita. To ensure comparability with composite indicators, we rescale the HDI score so that its maximum value is 10, which is the same as the BLI. We compare the distribution of  $CI_{BOD}$ ,  $CI_{C^2NLS}$ , HDI, and GDP per capita among countries. According to Table 4, the mean, the variation characterized by the standard deviation, and the range of the distribution characterized by the difference between the maximum and minimum scores are roughly similar and comparable to each other for  $CI_{BOD}$ ,  $CI_{C^2NLS}$ , and HDI. While  $CI_{C^2NLS}$  and HDI each have a similar mean,  $CI_{BOD}$  has a higher mean than these two indicators. No matter which composite indicators we adopt, their scores are shown to grow as per capita income grows. However, the difference in the score of composite indicators  $CI_{BOD}$  and  $CI_{C^2NLS}$  between high-income and middle-income countries is much smaller than the difference in GDP per capita.

Table 3 shows that the lower discriminatory power of BOD becomes more evident in this study. More than 20 countries among 36 countries are assigned the highest value of one in both cases of aggregating 11 and 12 indicators. These are countries that have higher scores of HDI and GDP per capita among the sample. It is obvious that BOD fails to differentiate the performance of these countries and show its overall picture. Moving from BOD to  $C^2NLS$ , the comparison dramatically improves and we can completely distinguish countries' performances. Figures 1 and 2 compare two composite indicators,  $CI_{C^2NLS}$  and  $CI_{BOD}$ , along with the measure based on CNLS,  $CI_{CNLS}$ . Since the difference between  $CI_{C^2NLS}$  and  $CI_{CNLS}$  is constant for all countries, comparing  $CI_{C^2NLS}$  and  $CI_{BOD}$  illustrates how  $CI_{C^2NLS}$  improves  $CI_{BOD}$  in terms of discrimination power. It is shown that  $CI_{C^2NLS}$  differentiates the performance of the countries that are ranked equally under  $CI_{BOD}$  by holding the ranking of other countries almost constant. Thus, while the international comparison of well-being based on  $CI_{C^2NLS}$  is similar to that based on  $CI_{BOD}$ ,  $CI_{C^2NLS}$  enables us to undertake a more detailed comparison than  $CI_{BOD}$ .

**Table 2** Descriptive statistics, World Bank adjusted net savings, constant 2011 PPP dollars

	Sustainability index	Adjusted net savings	Net national savings	Natural resources depletion	Energy	Mineral	Net forest	Carbon dioxide damage
<i>Mean</i>								
Overall	4.7	1632.7 (100 %)	2333.5 (143 %)	614.5 (38 %)	434.9	164.0	14.6	87.3 (5 %)
High income	7.8	5182.8 (100 %)	7096.3 (137 %)	1810.6 (35 %)	1770.8	30.4	2.4	109.9 (2 %)
Middle income	4.8	1760.3 (100 %)	2182.2 (124 %)	345.4 (20 %)	202.3	142.0	1.8	75.7 (4 %)
Low income	3.5	353.9 (100 %)	1089.3 (309 %)	639.7 (182 %)	363.9	237.1	37.0	97.3 (28 %)
<i>Median</i>	4.3	1265.2	1599.3	101.3	46.9	12.1	0.0	75.2
<i>SD</i>	2.2	2481.4	3104.3	1286.3	1184.9	483.2	34.9	44.0
<i>Max</i>	10.0	7750.3	14439.5	6644.1	6608.8	2285.7	140.8	207.2
<i>Min</i>	0.0	-3704.6	-3583.6	0.0	0.0	0.0	0.0	24.0

**Table 3** Composite indicators with HDI and GDP per capita

	Over 12 indicators		Over 11 indicators		HDI	GDP <sup>a</sup>
	BOD	C2NLS	BOD	C2NLS		
	Australia	10.00 (1)	9.66 (4)	10.00 (1)		
Austria	10.00 (1)	9.23 (18)	10.00 (1)	9.15 (18)	9.34 (20)	42887 (7)
Belgium	10.00 (1)	9.14 (22)	10.00 (1)	9.08 (22)	9.35 (19)	39839 (13)
Brazil	8.30 (34)	8.01 (34)	8.30 (33)	7.96 (34)	7.86 (36)	14300 (36)
Canada	10.00 (1)	9.63 (5)	10.00 (1)	9.60 (5)	9.56 (8)	41332 (11)
Chile	7.10 (36)	6.64 (36)	7.10 (36)	6.60 (36)	8.66 (32)	20154 (33)
Czech Republic	9.84 (22)	9.16 (20)	9.84 (22)	9.13 (19)	9.15 (25)	27047 (25)
Denmark	10.00 (1)	10.00 (1)	10.00 (1)	10.00 (1)	9.55 (10)	41830 (8)
Estonia	9.37 (28)	8.89 (29)	9.24 (29)	8.83 (29)	8.88 (27)	23310 (29)
Finland	10.00 (1)	9.55 (9)	10.00 (1)	9.55 (6)	9.34 (20)	38617 (14)
France	9.41 (27)	9.00 (26)	9.41 (26)	8.95 (25)	9.37 (17)	36263 (16)
Germany	10.00 (1)	9.39 (12)	10.00 (1)	9.37 (9)	9.65 (6)	40980 (12)
Greece	9.34 (29)	8.88 (30)	9.34 (28)	8.83 (28)	9.08 (26)	27045 (26)
Hungary	9.48 (25)	9.04 (25)	9.47 (25)	8.94 (26)	8.68 (31)	22413 (31)
Iceland	10.00 (1)	9.25 (16)	10.00 (1)	9.17 (17)	9.46 (13)	38215 (15)
Ireland	10.00 (1)	9.57 (7)	10.00 (1)	9.53 (8)	9.56 (8)	42945 (6)
Israel	9.62 (23)	9.08 (24)	9.57 (24)	8.99 (24)	9.40 (16)	30159 (23)
Italy	9.21 (30)	8.73 (31)	9.21 (30)	8.69 (31)	9.27 (23)	34626 (18)
Japan	10.00 (1)	9.24 (17)	10.00 (1)	9.22 (15)	9.43 (14)	34315 (19)
Korea	10.00 (1)	9.55 (8)	9.86 (21)	9.33 (12)	9.42 (15)	31327 (22)
Luxembourg	10.00 (1)	9.15 (21)	10.00 (1)	9.09 (21)	9.36 (18)	88848 (1)
Mexico	9.11 (31)	8.91 (28)	9.11 (31)	8.83 (30)	7.99 (34)	15886 (35)
Netherlands	10.00 (1)	9.40 (11)	10.00 (1)	9.37 (10)	9.71 (3)	43148 (5)
New Zealand	10.00 (1)	9.31 (14)	10.00 (1)	9.25 (14)	9.61 (7)	31711 (21)

Table 3 continued

	Over 12 indicators		Over 11 indicators		HDI	GDP <sup>a</sup>
	BOD	C2NLS	BOD	C2NLS		
	Norway	10.00 (1)	9.93 (2)	10.00 (1)		
Poland	10.00 (1)	9.36 (13)	10.00 (1)	9.31 (13)	8.82 (28)	21748 (32)
Portugal	8.78 (32)	8.47 (32)	8.78 (32)	8.41 (32)	8.70 (30)	25827 (27)
Russian Federation	8.30 (33)	8.14 (33)	8.25 (34)	7.99 (33)	8.24 (33)	22570 (30)
Slovak Republic	9.58 (24)	9.13 (23)	9.58 (23)	9.07 (23)	8.79 (29)	25128 (28)
Slovenia	9.42 (26)	8.93 (27)	9.38 (27)	8.86 (27)	9.29 (22)	28155 (24)
Spain	10.00 (1)	9.26 (15)	10.00 (1)	9.21 (16)	9.22 (24)	31731 (20)
Sweden	10.00 (1)	9.46 (10)	10.00 (1)	9.35 (11)	9.52 (11)	41762 (9)
Switzerland	10.00 (1)	9.70 (3)	10.00 (1)	9.64 (4)	9.71 (3)	51301 (3)
Turkey	8.09 (35)	7.56 (35)	8.01 (35)	7.50 (35)	7.99 (34)	17997 (34)
United Kingdom	10.00 (1)	9.61 (6)	10.00 (1)	9.54 (7)	9.47 (12)	34800 (17)
United States	10.00 (1)	9.18 (19)	10.00 (1)	9.12 (20)	9.68 (5)	49854 (4)

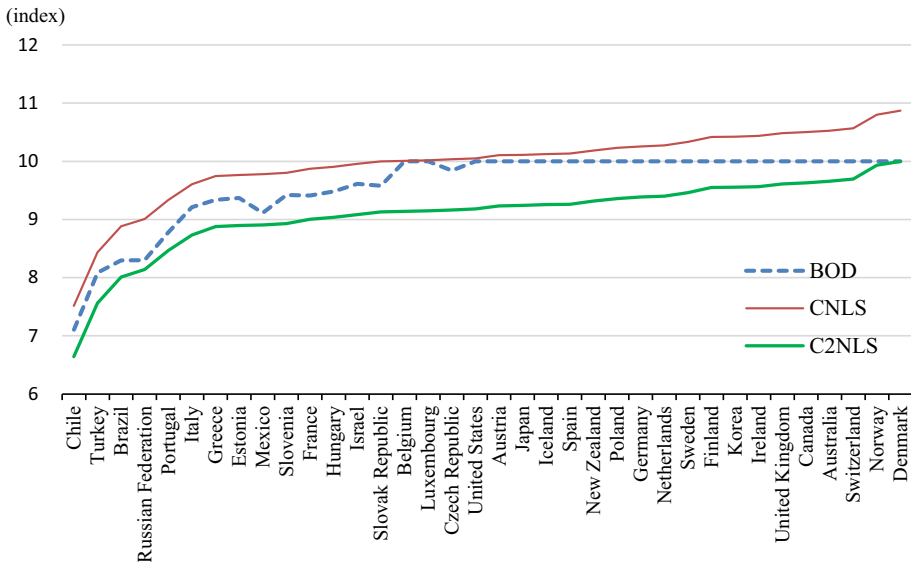
Parentheses indicate country ranking

<sup>a</sup> Unit: constant 2011 PPP US dollar

**Table 4** Descriptive statistics of composite indicators

	Over 12 indicators		Over 11 indicators		HDI	GDP <sup>a</sup>
	<i>BOD</i>	<i>C2NLS</i>	<i>BOD</i>	<i>C2NLS</i>		
<i>Mean</i>						
Overall	9.5820	9.0877	9.5684	9.0294	9.1944	35046
High income	10.0000	9.4893	10.0000	9.4297	9.6892	62975
Middle income	9.9073	9.3728	9.8978	9.3194	9.4793	37798
Low income	8.9780	8.5476	8.9543	8.4824	8.6258	22430
<i>Median</i>	10.0000	9.2069	10.0000	9.1399	9.3571	34471
<i>SD</i>	0.6882	0.6560	0.6955	0.6635	0.5339	14046
<i>Max</i>	10.0000	10.0000	10.0000	10.0000	10.0000	88848
<i>Min</i>	7.1041	6.6435	7.1041	6.5985	7.8640	14301
# of highest score	21	1	20	1	1	1

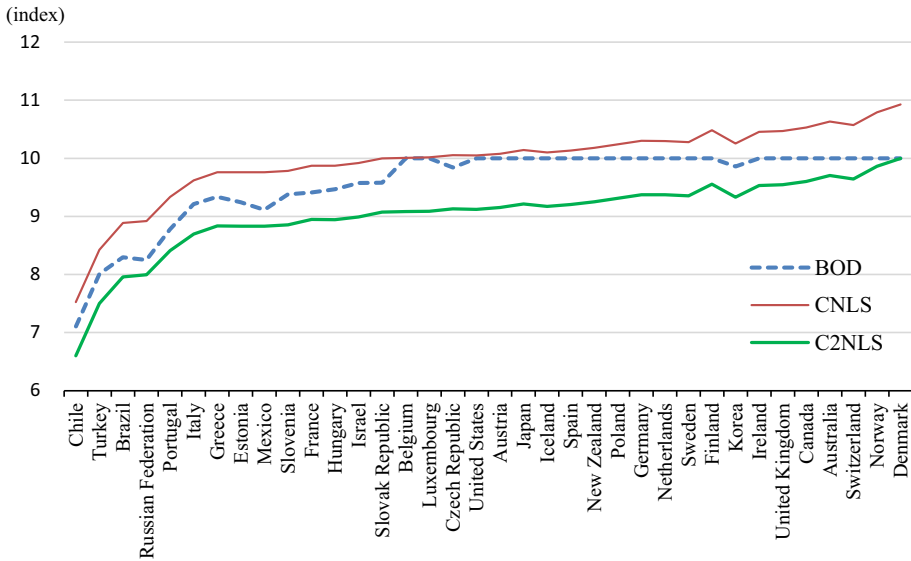
Unit: constant 2011 PPP US dollar



**Fig. 1** Comparison of composite indicators based on BOD and C<sup>2</sup>NLS (12 indicators)

Next, we consider how the inclusion of the sustainability indicator changes the composite well-being indicators by comparing the two cases. Since the present study deals with sustainability as just 1 among 12 well-being topics, the impact of the inclusion of the sustainability indicator is rather modest and it does not change the score and ranking of composite indicators dramatically. Table 4 shows that integrating the sustainability indicator slightly raises the values of the composite indicators and tightens their distribution on average.

As Figs. 1 and 2, and Table 3 show, Estonia, Israel, Korea, Russia, and Sweden are the five countries whose values or rankings of their composite indicators rise the most



**Fig. 2** Comparison of composite indicators based on BOD and C<sup>2</sup>NLS (11 indicators)

**Table 5** Correlation among composite indicators

	Over 12 indicators		Over 11 indicators		HDI	GDP
	BOD	C2NLS	BOD	C2NLS		
<i>Correlation coefficient</i>						
Aggregation over 12 indicators						
BOD	1.0000	0.9518	0.9988	0.9496	0.7639	0.5635
C2NSL	0.9518	1.0000	0.9501	0.9978	0.7353	0.5220
Aggregation over 11 indicators						
BOD	0.9988	0.9501	1.0000	0.9498	0.7670	0.5711
C2NSL	0.9496	0.9978	0.9498	1.0000	0.7423	0.5265
HDI	0.7639	0.7353	0.7670	0.7423	1.0000	0.6963
GDP	0.5635	0.5220	0.5711	0.5265	0.6963	1.0000
<i>Spearman rank correlation coefficient</i>						
Aggregation over 12 indicators						
BOD	1.0000	0.8868	0.9823	0.8842	0.7910	0.7672
C2NSL	0.8868	1.0000	0.8630	0.9925	0.8169	0.7109
Aggregation over 11 indicators						
BOD	0.9823	0.8630	1.0000	0.8731	0.7844	0.7844
C2NSL	0.8842	0.9925	0.8731	1.0000	0.8247	0.7210
HDI	0.7910	0.8169	0.7844	0.8247	1.0000	0.8572
GDP	0.7672	0.7109	0.7844	0.7210	0.8572	1.0000



significantly in the sample.<sup>15</sup> All five countries except Sweden are relatively low-income countries and have large adjusted net savings compared to their lower socio-economic indicators. Thus, the inclusion of the sustainability indicator causes their ranking to rise. Korea significantly raises its ranking in both composite indicators, which reflects that the value of the sustainability indicator is much higher than the values of the other well-being indicators. In addition, Australia, Finland, Germany, Greece, and Japan are the five countries whose values or rankings of their composite indicator decline the most significantly in the sample. All are relatively high-income countries except Greece and have smaller adjusted net savings compared to their higher socio-economic indicators. While only Australia and Finland show lower composite indicators after inclusion of the sustainability indicator, the other three countries lose their ranking under  $CI_{C2NLS}$ .

Table 5 confirms two points: the usefulness of  $CI_{C2NLS}$  in the present application and the relatively modest impact of the inclusion of a sustainability indicator. High correlations between  $CI_{BOD}$  and  $CI_{C2NLS}$  are found in either case of aggregation over 12 or 11 indicators. This suggests that  $CI_{C2NLS}$  differentiates the performance among countries ranked the highest under  $CI_{BOD}$  while hardly changing the ranking of other countries. High correlations between composite indicators aggregating 12 indicators and those aggregating 11 indicators are also observed. This shows that an additional sustainability indicator does not significantly change the score and ranking of composite indicators aggregating 11 indicators.

In addition, it is shown that all composite indicators and HDI, which share a similar pattern of distribution, are highly correlated with each other. This contrasts with relatively lower correlation between composite indicators and GDP per capita. The correlation becomes even lower when we integrate the sustainability indicator into other well-being indicators by aggregating 12 indicators. The quest for an alternative welfare measure stems from an acknowledgement of the limitations of GDP per capita as a welfare measure. Judging from the picture of well-being across countries drawn by  $CI_{C2NLS}$ , GDP per capita is even more problematic as a measure of sustainable well-being than as a measure of current well-being.<sup>16</sup>

## 5 Robustness Check and Discussion

The values of composite indicators adopted by the present study can be considered as productive efficiency. Composite indicators aggregating 12 or 11 well-being indicators of 36 countries corresponds to efficiency measure based on DEA in the case of 12 or 11 outputs and 1 input with 36 observations. Recently, the asymptotic property of efficiency score in DEA has been widely examined by explicitly incorporating data-generating process.<sup>17</sup> Studies, such as Korostelev et al. (1995), show that a much larger number of observations are necessary as the number of inputs and outputs increase, in order to avoid large statistical bias and imprecise estimation with a larger confidence interval. Our 36 observations are relatively small based on their standards.

<sup>15</sup> Table 10 reports the extent of the increase of both the values and rankings of their composite indicators after sustainability concerns are included.

<sup>16</sup> It is also shown that HDI is much better than GDP per capita as a well-being measure. However, there is a certain difference between composite indicators and HDI, which is mainly attributed to the fact that HDI misses a variety of important socio-economic conditions.

<sup>17</sup> See Daraio and Simar (2007).

**Table 6** Correlation among composite indicators aggregating sub-aggregates

	Over 12 indicators		Over 11 indicators	
	BOD	C2NLS	BOD	C2NLS
<i>Correlation coefficient</i>				
Aggregation over 3 indicators				
BOD	0.7810	0.7587	0.7874	0.7658
C2NSL	0.7608	0.7669	0.7651	0.7710
Aggregation over 2 indicators				
BOD	0.7643	0.7388	0.7757	0.7524
C2NSL	0.7452	0.7356	0.7568	0.7507
<i>Spearman rank correlation coefficient</i>				
Aggregation over 3 indicators				
BOD	0.8358	0.8497	0.8419	0.8499
C2NSL	0.8230	0.8734	0.8254	0.8759
Aggregation over 2 indicators				
BOD	0.8097	0.8115	0.8421	0.8211
C2NSL	0.7914	0.8235	0.8297	0.8392

Therefore, there is a possibility that the present study suffers from the so-called curse of dimensionality, so that the empirical result in the previous section is merely an artefact of statistical noise. In this section, we verify the robustness of our analysis by employing a model with a much smaller number of inputs and outputs. By reducing the number of inputs and outputs, statistical bias associated with composite indicators can be reduced. This allows us to evaluate more accurately the methodological advancement proposed in the present study, such as the integration of the sustainability indicator as well as the adaptation of C<sup>2</sup>NLS.

Instead of directly aggregating 12 indicators into a single number, the construction of a composite indicator of sustainable well-being is implemented in two stages in this section. First, we aggregate 11 indexes into two sub-aggregates based on simple averaging following the classification of the BLI: an indicator of material living conditions and an indicator of QOL.<sup>18</sup> Second, we aggregate these two sub-aggregates and the sustainability indicator into a composite indicator. Focusing on the second stage of aggregation, we compare two aggregation procedures, BOD and C<sup>2</sup>NLS, as well as investigate the impact of the integration of the sustainability indicator.

Table 6 shows that composite indicators derived in the second stage are highly correlated with the original composite indicators that directly aggregate 12 or 11 indicators.<sup>19</sup> This suggests that the composite indicators constructed by both approaches evaluate countries' relative performances similarly. Thus, as verified below, the conclusion drawn from the empirical analysis in the previous section is expected to still be valid in the setting of the present section, which is characterized by less stochastic noise.

<sup>18</sup> We adopt the procedure of simple averaging as an example. Similar results are obtained by applying C<sup>2</sup>NLS to aggregation of the first stage. They are available upon request.

<sup>19</sup> For example, the correlation between  $CI_{C2NLS}$  applied to 12 indicators and  $CI_{C2NLS}$  applied to 2 sub-aggregate and a single sustainability indicator is 0.7669 and its rank correlation is even larger (0.8734).

**Table 7** Descriptive statistics of composite indicators aggregating sub-aggregates

	Over 3 indicators		Over 2 indicators		HDI	GDP <sup>a</sup>
	BOD	C2NLS	BOD	C2NLS		
<i>Mean</i>						
Overall	8.1649	6.8777	8.1019	7.4456	9.1944	35046
High income	9.7882	8.5531	9.7503	8.7397	9.6892	62975
Middle income	9.0494	7.8847	8.9622	8.5097	9.4793	37798
Low income	6.3727	4.8903	6.3374	5.4923	8.6258	22430
<i>Median</i>						
	8.4964	7.3548	8.2913	7.8043	9.3571	34471
<i>SD</i>						
	1.6521	1.8909	1.6763	1.9421	0.5339	14046
<i>Max</i>						
	10.0000	10.0000	10.0000	10.0000	10.0000	88848
<i>Min</i>						
	4.4329	2.5898	4.1703	2.9034	7.8640	14301
<i># of highest score</i>						
	6	1	4	1	1	1

<sup>a</sup> Unit: constant 2011 PPP US dollar

Table 7 summarizes  $CI_{BOD}$  and  $CI_{C2NLS}$  in two cases, compared with HDI and GDP per capita. The two composite indicators have a smaller mean and larger standard deviation than HDI, which is different from the results in the previous section.<sup>20</sup> However, the impact of the adoption of  $CI_{C2NLS}$  is found to be similar to the previous section. While the scores of composite indicators become larger as per capita income grows, the difference in the score of  $CI_{BOD}$  and  $CI_{C2NLS}$  between high-income and middle-income countries is much smaller than the difference in GDP per capita. The stronger discriminating power of  $CI_{C2NLS}$  is also verified in the case of the relatively small number of inputs and outputs. While four countries are ranked equally under  $CI_{BOD}$  applied to two indicators, six countries are ranked equally under  $CI_{BOD}$  applied to three indicators. Their performances are completely differentiated under  $CI_{C2NLS}$ . This lead to the lower mean of  $CI_{C2NLS}$  than  $CI_{BOD}$ .

Next, we consider how the inclusion of the sustainability indicator changes the composite well-being indicators. Since the present section deals with 2 well-being sub-aggregates instead of 11 indicators, the impact of the inclusion of the sustainability indicator becomes relatively large compared with the previous section. While the mean and standard deviation of  $CI_{C2NLS}$  hardly change in the previous section, both decrease by around 5 % in the present section, as shown in Table 7.

However, Table 8 shows significantly high correlations between composite indicators aggregating three indicators and those aggregating two indicators. Thus, the inclusion of the sustainability indicator still has little impact on the evaluations on countries' relative performances in terms of overall well-being, as we found in the previous section. In addition, the inclusion of the sustainability indicator assures the usefulness of  $CI_{C2NLS}$ . High correlations between  $CI_{BOD}$  and  $CI_{C2NLS}$  are found in either case of aggregation over three or two indicators. This suggests that  $CI_{C2NLS}$  differentiates the performance among countries ranked the highest under  $CI_{BOD}$  without certainly changing the ranking of other countries.

<sup>20</sup> The two composite indicators and HDI have a similar mean and standard deviation in the previous section.

**Table 8** Correlation among composite indicators of different procedures

	Over 3 indicators		Over 2 indicators		HDI	GDP
	BOD	C2NLS	BOD	C2NLS		
<i>Correlation coefficient</i>						
Aggregation over 3 indicators						
BOD	1.0000	0.9829	0.9895	0.9700	0.9068	0.7195
C2NSL	0.9829	1.0000	0.9660	0.9716	0.8930	0.7001
Aggregation over 2 indicators						
BOD	0.9895	0.9660	1.0000	0.9837	0.8920	0.7180
C2NSL	0.9700	0.9716	0.9837	1.0000	0.8717	0.6523
<i>Spearman rank correlation coefficient</i>						
Aggregation over 3 indicators						
BOD	1.0000	0.9637	0.9856	0.9447	0.9047	0.8761
C2NSL	0.9637	1.0000	0.9399	0.9714	0.8650	0.8440
Aggregation over 2 indicators						
BOD	0.9856	0.9399	1.0000	0.9540	0.8816	0.8717
C2NSL	0.9447	0.9714	0.9540	1.0000	0.8240	0.8110

Thus, the present section verifies two main conclusions drawn in the previous section under the setting of a smaller number of underlying indicators: the usefulness of  $CI_{C2NLS}$  and the relatively modest impact of the inclusion of a sustainability indicator. However, there is a difference between the empirical results of the two sections, as shown in Table 8. While both composite indicators show higher correlation with HDI than GDP per capita, the inclusion of a sustainability indicator has a different impact on composite indicators. It lowers the correlations with HDI and GDP per capita in the previous section but raises those in the present section.<sup>21</sup> Which finding is a more accurate picture of reality? To answer this question, we consider it necessary to obtain larger observations or to investigate the method in the first-stage aggregation further.

## 6 Conclusion

Well-being is a multidimensional concept. The OECD recently specified 11 topics that are essential to people's well-being and released 11 corresponding well-being indicators. However, the OECD leaves the aggregation of the data to the user and a sustainability indicator is not included among the 11 indicators. Thus, the present study introduces an additional sustainability indicator from the World Bank's adjusted net savings and aggregates the 11 well-being indicators and the sustainability indicator using composite indicators. We adopt two composite indicators based on the BOD and C<sup>2</sup>NLS approaches. Unlike HDI, both approaches aggregate individual indicators by investigating country-specific weights that favour each country.

The composite indicator based on BOD is now a standard tool for evaluating multi-faceted concepts, such as well-being. However, since more than half of countries are

<sup>21</sup> This difference is more evident in  $CI_{C2NLS}$ .

ranked the highest under the application of BOD in the present study, BOD fails to distinguish their performances. The composite indicator based on  $C^2NLS$  we first introduced here gives a similar cross-country ranking to that based on BOD. Moreover, it even allows us to differentiate completely the performance of countries that are ranked equally under BOD. Thus,  $C^2NLS$  enables a complete cross-country comparison of overall well-being, improving on the BOD approach.

We quantify the impact of the inclusion of the sustainability indicator into other well-being indicators by using composite indicators based on  $C^2NLS$ . The inclusion of the sustainability indicator has a rather modest effect and does not significantly change the score and ranking of composite indicators for many countries. However, the composite indicators of some countries are affected significantly by integrating the sustainability indicator. Each of these countries has a large gap between its sustainability indicator and other well-being indicators. While the composite indicators of countries whose sustainability indicator is much larger than their other well-being indicators increase their ranking, such as Korea, the composite indicators of countries whose sustainability indicator is much smaller than their other well-being indicators lose their ranking, such as Australia.

Our results verify that  $C^2NLS$  is an indispensable tool for integrating sustainability concerns into a composite well-being indicator. However, the greater discriminatory power of the composite indicator based on  $C^2NLS$  compared with BOD does not mean that the former more accurately captures the state of the sustainable well-being of each country than the latter.<sup>22</sup> Our proposal to introduce  $C^2NLS$  is justified merely from a practical standpoint of being able to completely distinguish the level of sustainable well-being among countries. Future research should investigate the theoretical framework for evaluating composite indicators.<sup>23</sup>

**Acknowledgments** The author is grateful for three anonymous referees and Shigemi Kagawa and Knox Lovell for their helpful comments and suggestions. This article was completed when the author visited Center for Efficiency and Productivity Analysis in the School of Economics at the University of Queensland. The good research environment offered by the institute and the department is greatly appreciated. This research was financially supported by Grant-in-Aid for Scientific Research (KAKENHI 25870922). All remaining errors are the author's responsibility.

## Appendix

See Tables 9 and 10.

<sup>22</sup> Both BOD and  $C^2NLS$  are shown to be consistent estimators of the production frontiers. See Banker (1993) for BOD and Kuosmanen and Johnson (2010) for  $C^2NLS$ .

<sup>23</sup> Index number theory proposes desirable axioms that plausible price indices need to satisfy (Balk 2008). Axiomatic justification might be applicable to studies on composite indicators.

**Table 9** Well-being Indicators and Sustainability Indicator

	Housing	Income	Jobs	Community	Education	Environment	Civic engagement	Health	Life satisfaction	Safety	Work-life balance	Sustainability
Australia	7.6 (4)	4.9 (12)	8.5 (6)	8.9 (10)	7.8 (8)	9.1 (2)	9.5 (1)	9.4 (1)	8.7 (11)	9.6 (5)	5.3 (29)	3.94 (23)
Austria	5.8 (21)	5.0 (9)	8.6 (5)	9.7 (5)	6.7 (21)	7.3 (20)	5.6 (15)	7.7 (19)	9.1 (5)	9.1 (12)	5.9 (26)	6.14 (8)
Belgium	7.3 (5)	5.9 (4)	7.3 (18)	8.1 (19)	7.5 (14)	7.2 (22)	5.8 (13)	7.8 (17)	7.8 (15)	7.5 (30)	8.8 (3)	4.92 (13)
Brazil	4.2 (27)	0.1 (36)	5.7 (27)	7.8 (22)	1.9 (35)	6.0 (27)	4.4 (24)	4.9 (33)	8.1 (13)	2.1 (35)	6.7 (21)	2.54 (31)
Canada	7.8 (1)	5.7 (5)	8.0 (9)	9.2 (7)	7.7 (12)	8.6 (10)	5.8 (13)	9.3 (3)	9.4 (3)	9.7 (4)	6.2 (25)	4.37 (17)
Chile	3.7 (32)	1.1 (32)	5.8 (26)	6.1 (31)	4.4 (32)	3.3 (35)	4.3 (25)	5.8 (25)	6.1 (23)	6.7 (33)	5.1 (32)	1.34 (34)
Czech Republic	4.6 (25)	1.8 (27)	6.3 (23)	6.7 (29)	7.8 (8)	7.5 (19)	3.7 (30)	5.8 (25)	6.5 (21)	9.3 (8)	7.2 (16)	2.97 (27)
Denmark	6.2 (16)	4.0 (15)	8.0 (9)	10.0 (1)	7.8 (8)	9.0 (3)	7.1 (5)	7.4 (21)	9.4 (3)	8.8 (16)	9.8 (1)	5.91 (12)
Estonia	4.4 (26)	0.9 (33)	5.6 (28)	7.4 (25)	8.1 (3)	8.4 (11)	2.3 (34)	4.5 (34)	2.1 (32)	7.3 (32)	7.4 (14)	4.45 (11)
Finland	6.3 (12)	3.5 (19)	7.5 (16)	9.0 (9)	9.2 (1)	9.0 (3)	5.9 (12)	7.5 (20)	8.9 (7)	9.3 (8)	7.4 (14)	4.47 (20)
France	6.4 (11)	5.0 (9)	6.5 (21)	8.2 (17)	5.9 (27)	8.4 (11)	4.3 (25)	7.9 (16)	6.4 (22)	8.3 (24)	7.6 (12)	4.34 (14)
Germany	6.3 (12)	5.3 (7)	8.3 (7)	8.9 (10)	8.0 (4)	8.8 (8)	3.9 (27)	7.2 (22)	7.3 (18)	8.9 (14)	7.9 (7)	5.94 (7)
Greece	3.8 (30)	1.9 (26)	2.2 (36)	0.0 (36)	6.1 (25)	4.6 (32)	3.9 (27)	8.2 (13)	0.0 (36)	8.8 (16)	7.2 (16)	0.00 (36)
Hungary	3.8 (30)	1.3 (29)	4.8 (31)	6.9 (28)	6.7 (21)	7.3 (20)	4.8 (22)	4.3 (35)	0.6 (34)	8.8 (16)	7.8 (11)	3.66 (15)
Iceland	5.9 (19)	3.6 (18)	8.7 (3)	10.0 (1)	7.3 (16)	8.8 (8)	5.3 (16)	8.8 (6)	9.1 (5)	9.2 (11)	5.7 (27)	1.95 (35)
Ireland	7.3 (5)	3.3 (20)	5.9 (25)	9.9 (4)	7.1 (19)	8.3 (13)	6.0 (11)	8.7 (8)	6.7 (20)	9.4 (7)	7.9 (7)	4.57 (18)
Israel	4.2 (27)	3.8 (17)	6.7 (20)	7.7 (23)	5.4 (28)	5.4 (29)	2.3 (34)	9.0 (5)	7.8 (15)	7.4 (31)	4.9 (33)	5.09 (9)
Italy	5.1 (23)	4.4 (14)	5.6 (28)	8.4 (16)	5.1 (30)	6.8 (26)	4.5 (23)	7.8 (17)	4.2 (25)	8.4 (23)	7.5 (13)	2.72 (32)
Japan	4.9 (24)	5.6 (6)	7.9 (11)	7.9 (21)	7.9 (7)	6.9 (24)	3.9 (27)	5.0 (30)	4.1 (27)	9.9 (1)	5.2 (31)	2.81 (30)
Korea	5.9 (19)	2.3 (23)	7.6 (14)	3.1 (34)	8.0 (4)	5.3 (30)	7.5 (3)	5.0 (30)	4.2 (25)	9.5 (6)	4.2 (34)	7.04 (3)
Luxembourg	6.2 (16)	6.5 (3)	8.3 (7)	7.3 (26)	4.5 (31)	8.0 (16)	6.8 (7)	8.0 (15)	7.9 (14)	8.3 (24)	7.9 (7)	7.22 (4)
Mexico	3.7 (32)	0.7 (34)	6.0 (24)	2.0 (35)	1.2 (36)	4.5 (33)	5.3 (16)	5.0 (30)	8.9 (7)	0.4 (36)	2.6 (35)	3.01 (24)
Netherlands	6.9 (7)	5.3 (7)	8.7 (3)	8.6 (15)	7.8 (8)	6.9 (24)	5.0 (20)	8.3 (12)	8.8 (10)	8.3 (24)	8.8 (3)	6.59 (6)
New Zealand	6.6 (9)	2.1 (25)	7.5 (16)	10.0 (1)	7.3 (16)	9.0 (3)	7.2 (4)	9.4 (1)	8.4 (12)	9.3 (8)	6.3 (23)	3.81 (26)

Table 9 continued

	Housing	Income	Jobs	Community	Education	Environment	Civic engagement	Health	Life satisfaction	Safety	Work-life balance	Sustainability
Norway	7.7 (3)	4.0 (15)	9.2 (2)	8.9 (10)	7.3 (16)	8.9 (7)	6.5 (8)	8.1 (14)	9.7 (2)	8.8 (16)	8.7 (5)	10.00 (1)
Poland	3.5 (34)	1.3 (29)	5.2 (30)	7.7 (23)	8.5 (2)	4.8 (31)	5.3 (16)	5.2 (28)	3.4 (30)	9.8 (2)	5.6 (28)	3.49 (21)
Portugal	6.6 (9)	2.5 (22)	4.5 (32)	6.2 (30)	4.4 (32)	7.8 (18)	3.3 (33)	5.9 (24)	1.4 (33)	8.0 (28)	6.8 (20)	1.83 (33)
Russian Federation	3.3 (35)	1.2 (31)	6.8 (19)	5.8 (32)	6.1 (25)	4.3 (34)	2.1 (36)	0.6 (36)	2.7 (31)	6.5 (34)	7.9 (7)	4.16 (10)
Slovak Republic	4.1 (29)	1.4 (28)	4.5 (32)	7.3 (26)	6.2 (24)	8.1 (15)	3.6 (31)	5.4 (27)	4.0 (29)	9.1 (12)	7.2 (16)	3.00 (25)
Slovenia	5.8 (21)	2.2 (24)	6.5 (21)	8.8 (13)	7.6 (13)	7.1 (23)	6.2 (9)	6.6 (23)	4.1 (27)	8.8 (16)	6.6 (22)	3.87 (19)
Spain	6.8 (8)	2.9 (21)	2.6 (35)	8.7 (14)	5.4 (28)	6.0 (27)	5.0 (20)	8.7 (8)	4.7 (24)	8.7 (21)	9.4 (2)	3.15 (28)
Sweden	6.3 (12)	5.0 (9)	7.8 (13)	8.2 (17)	8.0 (4)	9.8 (1)	8.7 (2)	8.8 (6)	8.9 (7)	8.2 (27)	8.1 (6)	7.54 (5)
Switzerland	6.3 (12)	7.3 (2)	9.6 (1)	9.5 (6)	7.4 (15)	8.3 (13)	3.4 (32)	9.3 (3)	10.0 (1)	8.7 (21)	7.1 (19)	8.07 (2)
Turkey	2.2 (36)	0.6 (35)	4.3 (34)	3.8 (33)	2.5 (34)	2.9 (36)	6.2 (9)	5.2 (28)	0.5 (35)	7.8 (29)	0.0 (36)	3.25 (16)
United Kingdom	6.0 (18)	4.8 (13)	7.6 (14)	9.2 (7)	6.4 (23)	9.0 (3)	6.9 (6)	8.4 (11)	7.2 (19)	9.8 (2)	6.3 (23)	3.53 (29)
United States	7.8 (1)	10.0 (1)	7.9 (11)	8.0 (20)	7.1 (19)	7.9 (17)	5.3 (16)	8.5 (10)	7.5 (17)	8.9 (14)	5.3 (29)	4.41 (22)

Parentheses indicate country ranking

**Table 10** Difference in value and ranking of composite indicators

	BOD		C2NLS		Sustainability indicator	Average of 11 indicators
	Value	Ranking	Value	Ranking		
<i>Countries increasing composite indicators the most significantly</i>						
Estonia	0.1254	1	0.0609	0	4.45 (11)	5.3 (28)
Israel	0.0408	1	0.0948	0	5.09 (9)	5.9 (24)
Korea	0.1402	20	0.2245	4	7.04 (3)	5.7 (25)
Russian Federation	0.0532	1	0.1460	0	4.16 (10)	4.3 (33)
Sweden	0.0000	0	0.1100	1	7.54 (5)	8.0 (2)
<i>Countries decreasing composite indicators the most significantly</i>						
Australia	0.0000	0	-0.0478	-1	3.94 (23)	8.1 (1)
Finland	0.0000	0	-0.0048	-3	4.47 (20)	7.6 (8)
Germany	0.0000	0	0.0119	-3	5.94 (7)	7.3 (14)
Greece	0.0000	-1	0.0417	-2	0.00 (36)	4.2 (34)
Japan	0.0000	0	0.0259	-2	2.81 (30)	6.3 (21)

Parentheses indicate country ranking

## References

- Arrow, K., et al. (2004). Are we consuming too much? *Journal of Economic Perspectives*, 18(3), 147–172.
- Balk, B. M. (2008). *Price and Quantity Index Numbers*. New York: Cambridge University Press.
- Banker, R. D. (1993). Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. *Management Science*, 39(10), 1265–1273.
- Cherchye, L. & Kuosmanen, T. (2004). Benchmarking sustainable development: A Synthetic Meta-Index approach, Research Paper No. 2004/28, World Institute for Development Economics Research, United Nations University, Helsinki.
- Cherchye, L., Moesen, W., & Van Puyenbroeck, T. (2004). Legitimately diverse, yet comparable: On synthesizing social inclusion performance in the EU. *JCMS: Journal of Common Market Studies*, 42(5), 919–955.
- Cherchye, L., et al. (2007). An introduction to ‘Benefit of the Doubt’ composite indicators. *Social Indicators Research*, 82(1), 111–145.
- Costantini, V., & Monni, S. (2005). Sustainable human development for European Countries. *Journal of Human Development*, 6(3), 329–351.
- Daraio, C., & Simar, L. (2007). *Advanced Robust and nonparametric methods in efficiency analysis: Methodology and applications*. New York, NY: Springer.
- Dasgupta, P. (2001). *Human well-being and the natural environment*. Oxford: Oxford University Press.
- Desai, M. (1995). Greening of the HDI? In A. McGillivray & S. Zadek (Eds.), *Accounting for change* (pp. 21–36). London: The New Economics Foundation.
- Despotis, D. K. (2005). Measuring human development via data envelopment analysis: The case of Asia and the Pacific. *Omega*, 33(5), 385–390.
- Korostelev, A. P., Simar, L., & Tsybakov, A. B. (1995). Efficient estimation of monotone boundaries. *Annals of Statistics*, 23(2), 476–489.
- Kuosmanen, T. (2008). Representation theorem for convex nonparametric least squares. *Econometrics Journal*, 11(2), 308–325.
- Kuosmanen, T., Cherchye, L., & Sipiläinen, T. (2006). The law of one price in data envelopment analysis: Restricting weight flexibility across firms. *European Journal of Operational Research*, 170(3), 735–757.
- Kuosmanen, T., & Johnson, A. L. (2010). Data envelopment analysis as nonparametric least-squares regression. *Operations Research*, 58(1), 149–160.
- Kuosmanen, T., Johnson, A., & Saastamoinen, A. (2015). Stochastic nonparametric approach to efficiency analysis: A unified framework. In W. D. Cook & J. Zhu (Eds.), *Data envelopment analysis international series in operations research and management science* (pp. 191–244). Boston, MA: Springer.



- Kuosmanen, T., & Kortelainen, M. (2012). Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints. *Journal of Productivity Analysis*, 38(1), 11–28.
- Lovell, C. A. K., Pastor, J. T., & Turner, J. A. (1995). Measuring macroeconomic performance in the OECD: A comparison of European and Non-European Countries. *European Journal of Operational Research*, 87(3), 507–518.
- Mahlberg, B. & Obersteiner, M. (2001). Remeasuring the HDI by data envelopment analysis. *International Institute for Applied Systems Analysis Interim Report*, 01–069.
- Mizobuchi, H. (2014). Measuring world better life frontier: A composite indicator for OECD Better Life Index. *Social Indicators Research*, 118(3), 987–1007.
- Neumayer, E. (2001). The Human Development Index and sustainability—A constructive proposal. *Ecological Economics*, 39(1), 101–114.
- Organisation for Economic Co-operation and Development. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development. (2011). *How's life?*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development. (2013). *How's life? 2013*. Paris: OECD Publishing.
- Ray, M. (2014). Redefining the Human Development Index to account for sustainability. *Atlantic Economic Journal*, 42(3), 305–316.
- Stiglitz, J.E., Sen, A., & Fitoussi, J.-P., (2009). *Report by the commission on the measurement of economic performance and social progress*. [http://www.stiglitz-sen-fitoussi.fr/documents/rapport\\_anglais.pdf](http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf).
- World Bank. (2011). *The changing wealth of nations*. Washington DC: World Bank Publications.