

The Cost of Well-Being

Vincenzo Patrizii¹ · Anna Pettini¹ · Giuliano Resce²

Accepted: 17 June 2016 / Published online: 24 June 2016
© Springer Science+Business Media Dordrecht 2016

Abstract It is by now common knowledge that in switching from GDP to alternative, multidimensional, measures of collective well-being one can provide a better account of a country's socio-economic conditions. Such a gain, however, comes at the price of losing output-to-input type of link between well-being and the resources necessary to make it available. Since well-being measures are not meant to be only an exercise in documentation, but also to inform policies and priorities, we propose a method to build a measure of well-being in the form of a single index, as for GDP, which takes into account: (1) the social and environmental costs, not considered in the GDP, and (2) the use of conventional resources (capital and labour), not considered in the currently available multidimensional measures of well-being. We use a *Data Envelopment Analysis* type of model, integrated with *Principal Component Analysis*, to evaluate OECD countries' relative efficiency in providing well-being. Our results show that the costs of producing well-being have a large and significant impact on the resulting index of well-being. Therefore, high efficiency in providing well-being and high income cannot be considered a proxy to each other. In addition, it is shown that countries react differently to the different costs of well-being: poor countries are, on average, more efficient in terms of conventional inputs (labour and capital), while rich countries have higher efficiency indices relative to social and environmental costs. The close to zero correlation between GDP and well-being indices for rich countries provides new support to the “Easterlin paradox”.

✉ Giuliano Resce
Giuliano.Resce@uniroma3.it

Vincenzo Patrizii
Vincenzo.Patrizii@unifi.it

Anna Pettini
Pettini@unifi.it

¹ Department of Economics and Management, University of Florence, Via delle Pandette 21, 50127 Florence, Italy

² Department of Economics, Roma Tre University, Via Silvio D'amico 77, 00145 Rome, Italy

Keywords Well-being · Data envelopment analysis · Better life index · GDP · Social efficiency

JEL Classification C61 · E23 · I31

1 Introduction

It is by now common knowledge that in switching from GDP to alternative, multidimensional, measures of collective well-being one can provide a better account of a country's socio-economic conditions. Such a gain, however, comes at the price of losing the link between well-being and the resources necessary to make it available. In the case of GDP, such a link is provided by the accountancy identities (production = final uses = income) contained in the System of National Accounts, and they also provide the conventional, although widely criticized, answer to the question of what is the cost of a country's well-being.

If alternative, multidimensional, measures of collective well-being are to gain a role similar to that so far played by GDP as a social and economic indicator, such a link between the proposed measures and the deployed resources (natural and human) has to be established, grounded on sound economic and statistic theory and made available for policy evaluation.

In the case of multidimensional measures the cost of well-being also includes less conventional types of cost given by the negative externalities (both social and environmental) broadly linked to economic and social activities. These costs are the main battle field for the criticisms to GDP as a reliable measure of well-being. There is no more questioning as to whether they should be included in any reasonable multidimensional measure of well-being. The question is rather that of devising a consistent framework for inclusion.

In short, what are the costs of well-being, when it is measured in a multidimensional setting, is a question that still awaits an answer. The first aim of this paper is that of providing a theory-consistent answer.

The second aim of the paper comes from observing that such an output-to-input link, which relates well-being to its costs (broadly defined as aggregate production function, with the most relevant information condensed into indices referred to as productivity or productive efficiency), varies across and within countries as it comes to be affected by economic, social and institutional arrangements. In devising a set of indicators to assess countries' performance in terms of their use of resources, one also provides essential information useful in judging social and economic policies, in a way similar to that made possible by conventional National Accounts for the GDP measure of well-being. This makes for the second purpose of this paper.

To summarize, we propose a methodology to build a measure of well-being in the form of a single index, as for GDP, which takes account of: (1) the social costs (human and environmental) not considered in the GDP accounting system, and of (2) the use of conventional resources (mainly capital and labour) not considered in the so far available of multidimensional measure of well-being.

The paper is organized as follows. Section 2 summarizes the literature's pros & cons about the criticism to GDP as a measure of well-being and in favour of a move towards multidimensional measures of well-being. Section 3 reviews the literature on the use of

Data Envelopment Analysis (DEA) to evaluate countries and/or other public institutions in terms of their socio-economic performances. Section 4 describes the DEA type of models chosen for the analysis. Section 5 presents the data set been used. Section 6 details the results in terms of a set of efficiency indices. The last section contains concluding remarks and some hints for further research.

2 Towards a Multidimensional Measure

For many decades, western economies have continued to grow: more production and more consumption have pushed income up for an increasing number of people. Growth has been a synonymous of development for a long time. In the 1930's, in response to the information gap revealed by the Great Depression, Simon Kuznets developed a set of national income accounts, and among them GDP became the most widely popular measure of welfare. Kuznets himself launched the measure saying that "*the welfare of a nation can scarcely be inferred from a measure of national income*" (Kuznets 1934). In spite of such a warning, GDP is still used not only as the principal measure of economic performance, but also interpreted as an index of economic well-being. Over the last few decades, both these interpretations have been largely discussed among activists, and policy experts; the debate remains incredibly vivid. Some authors still consider GDP as the most suitable tool for international comparisons, and a measure easier to handle than any other currently available alternative. Others are definitely sceptical and argue that it leads to misguided aspirations, and even to destructive values. Yet others claim that nowadays, besides methodological issues, overcoming GDP is a matter of mainly cultural and political processes, because the new measures of well-being, ultimately, lack legitimacy. Segre et al. (2011) argue that the main problem in parting from GDP lies in that there is no agreement on what well-being and progress are. For an overview of the literature see Fleurbaey (2009), Bleys (2012), Costanza et al. (2009), De Beukelaer (2014), and three interesting and most recent books: Coyle (2014), Karabell (2014), Fioramonti (2013).

Along with the debate on pros and cons of GDP, research quickly developed on alternative measures; the interest burgeoned at the beginning of the new millennium, at both the theoretical and the policy evaluation front. Interestingly, Bandura (2005, 2008) reports that 80 % of new welfare indices available in 2005 has been developed after 2000. Bleys (2012), on the other hand, claims that research has been concentrated more on new measures and on their refinements, than on the comparison of existing ones. For this reason, he proposes a classification of the existing indices. A lot of work has also been done by international institutions. For instance, UN in 2015 launched the upgrade of the Sustainable Development Goals, which is a set of international objectives to improve global well-being. The aim is to define a 'new objective for humanity', namely a summary of the existing indices and of the research linked to them (Costanza et al. 2014).

Very broadly speaking, indicators alternative to GDP are based either on the aggregation of variables or on sets of indicators. In the first group, in turn, we find two subgroups of indices that either extend and correct the computation of GDP, or go beyond the realm of GDP by considering social, economic and environmental variables. The Measure of Sustainable Welfare and the Genuine Progress Indicator belong to the first subgroup. They include variables for non-marketed goods and services (weights being provided by prices, actual or shadow). In the second subgroup, there are indices which combine variables by using weights other than prices; examples are the Human Development Index (firstly

developed by the UNDP in 1990), the Index of Economic Well-being (Osberg 1985 and Osberg and Sharpe 1998), and the Ecological Footprint (Wackernagel and Rees 1998).

In 2009, the Commission on the Measurement of Economic Performance and Social Progress—the Stiglitz-Sen-Fitoussi Commission—released a document on which OECD built the *Framework for Measuring Well-Being and Progress*, and a set of indicators named better life index (hereafter BLI), launched in 2011. According to it, the measure of well-being requires looking not only at the functioning of the economic system, but also at the differences in people’s living conditions. The BLI framework is built around three distinct domains: material conditions, quality of life and sustainability. Each of them includes a number of relevant dimensions (variables), not hierarchically ordered (Kerényi 2011; Lind 2014). That is why OECD has conceived this set of indicators to be an interactive web-based tool, which is what makes it both useful and theoretically interesting. It is useful for international comparisons based either on a subset of dimensions, or on an aggregation made by subjective weights. It is theoretically interesting because it avoids the legitimacy issue (as in Segre et al. 2011): by allowing the end user to select the relevant variables it does not force one to accept a given interpretation of well-being.

Last but not the least for research purposes, BLI has the advantage of including bad-outputs measures, not always present in other indices (e.g. Human Development Index by the UNDP), and in addition it has a convenient variable to case ratio: 24 variables for 36 countries. A well-conceived index such as the Index of Economic Well-being (Osberg 1985) has a narrower coverage.

The OECD initiative has been followed by many national data collections of the same kind as, for example, “Measuring national well-being” from the Office of National Statistics—ONS, Great Britain, 2010; BES Fair and sustainable welfare, ISTAT and CNEL, Italy; the new data collection by the Institut National de la Statistique et des Études Économiques, INSEE, France; ONS (2011); INSEE (2010).

3 Previous Work

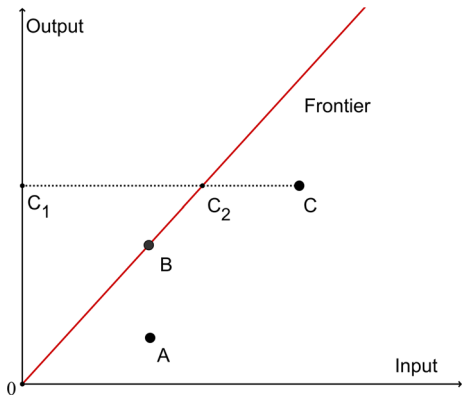
The use of DEA¹ for socio-economic performance evaluation originates in the early 1990s. DEA relies on linear programming methods to estimate a “best practice” production function (i.e.: an efficiency frontier) which is then used as a benchmark to measure the relative performance of Decision Making Units (DMU).² Figure 1 shows the basic of DEA functioning. Given a simple, though extreme case of three DMUs (A, B & C) producing one output with a single input, DMU B being the most productive one (highest output to input ratio) comes to represent the benchmark (the frontier) for assessing the other DMUs’ relative efficiency. For instance, DMU C’s efficiency index is given by the ratio of “efficient input” to “actual input”, that is: $\overline{C_1 C_2} / \overline{C_1 C}$, in Fig. 1. Therefore, the efficiency index ranges from 0 (totally inefficient) to 1 (fully efficient).

DEA’s strong feature is that of allowing for efficiency measurement in a multi-dimensional setting, where there is more than one input and/or more than one output. That is why DEA is often used as a procedure to get at a single index out of set of variables, as it is the case of multi-dimensional measures of well-being (see Sect. 2).

¹ Dea is a non-parametrical technique, originates from the work of Farrel (1957) and further developed by Charnes et al. (1978), see Cooper et al. (2007) for a comprehensive treatment of the subject.

² The term Dmu in is used Dea literature to generally indicate the decision center responsible for converting inputs into outputs.

Fig. 1 Productivity benchmarking



The earliest work we are aware of in this field is that by Hashimoto and Ishikawa (1993) which estimates citizens' well-being for 47 Japanese Prefectures. The literature on the subject has thereafter grown rapidly around three major issues: (1) Whether or not conventional inputs (labour and capital) are considered³; (2) The choice of multidimensional measures of well-being; iii) The discriminating power of the model (see Table 1).

As for conventional inputs, DEA models retain an informative advantage even in cases when such inputs measures are not considered because only an aggregate index of well-being is sought. This is because DEA does not require any predetermined functional form to be specified for the implied production function, as it assigns the most favourable shadow-price to each output measures in order to get the efficiency index.⁴ Within this framework the efficiency index coming from DEA models rank countries according to their level of (multidimensional) well-being irrespective of the resources necessary to make that well-being available.

On the other side, if conventional inputs are included, the resulting efficiency index takes the meaning of a productivity ratio, because it gives a higher ranking to countries with a higher index of well-being relatively to inputs.

In the choice of conventional inputs reliance is made to the traditional economic literature on production function (e.g. Dasgupta 2001; Arrow et al. 2004). Golany and Thore (1997) are the first who include conventional inputs. As a proxy for capital they select three measures over a 15 years period: investments and public spending (the latter being split into educational and non-educational). In order to make the assessment not dependent on countries' relative dimension, they choose to scale inputs and outputs by GDP. Another contribution that by Zaim et al. (2001), although primarily concerned with methodological issues, is to be singled out as the first attempt to include both of conventional inputs. Recently, Mizobuchi (2014) chooses as inputs physical, natural and human capital (from World Bank 2011), each in per-capita values, and as outputs 11 well-being indices from the BLI database (named "topics" in OECD 2013a, b).

³ Conventional inputs measures have been considered in Golany and Thore (1997), Zaim et al. (2001), Despitos (2004), Mizobuchi (2014), while have not been considered in Hashimoto and Ishikawa (1993), Hashimoto and Kodama (1997), Zhu (2001), Mahlberg and Obersteiner (2001), Ramanathan (2006), Murias et al. (2006).

⁴ Lovell et al. (1995) use the same quantity of input for all Dmus, or Hashimoto and Ishikawa (1993) treat only bad outputs as inputs.

Table 1 Dea literature on socio-economic performance evaluation

Reference	Dmu	Inputs	Outputs
Hashimoto and Ishikawa (1993)	47 prefectures of Japan	Suicide, Crime, Traffic accidents, Unemployment	Life expectancy, National income, Forest area, Water services
Lovell et al. (1995)	19 Oecd Countries	Constant Unit	Gdp, Uninflation, Employment, Trade balance, (Carbon) ⁻¹ , (Nitrogen) ⁻¹
Hashimoto and Kodama (1997)	Time series of Japan 1956–1990	Suicide, Crime, Traffic accidents, Unemployment	Life expectancy, National income, Forest area, Water services
Golany and Thore (1997)	72 Countries	15 years of: domestic investments, Government consumption, expenditure on education	Growth rate, 1-Infant mortality, Enr. secondary education, Welfare payments
Zhu (2001)	20 Cities (15 Usa, 5 international)	High-end and Lower-end housing rental, Cost of French bread and of Martini, Class A office rental, Violent crimes	Household income, Population with degree, Doctors, Museums, Libraries, Golf courses
Mahlberg and Obersteiner (2001)	174 Countries	Constant Unit	Life expectancy, Adult literacy rate, Combined enr. ratio, Adj. income
Zaim et al. (2001)	55 Countries	Capital stock, Labor	Infant survival rate, Life expectancy, Enr. primary school, Enr. secondary school, Gdp
Murias et al. (2006)	50 Spanish provinces	Gini index, Population minimum income, Unemployment	Disposable income, Size of dwelling, Net capital, Human capital, Long-term contracts
Ramanathan (2006)	18 Mena Countries	Age dependency ratio, Adult female illiteracy rate, Infant mortality rate	Employment, Life expectancy, Primary education, Gnp
Mizobuchi (2014)	Oecd Countries	Produced capital, Natural capital, Intangible capital	Bli

Coming to the issue about the selection of output (i.e. well-being) measures one observes that earlier work (Hashimoto and Ishikawa 1993) felt obliged to justify the selected measures, while subsequent work (Mahlberg and Obersteiner 2001; Despotis 2005; Mizobuchi 2014) can afford to rely on the results of long lasting research projects lead by international institutions such as Human Development Index by UNDP, the *better life index* by OECD.

The third issue concerns the discriminatory power of DEA models referred to as the “dimensionality curse”: as the number of inputs and outputs increases, the discriminating power of any DEA model is reduced and so is its usefulness. To face this problem, most research papers (e.g. Hashimoto and Kodama 1997) use the Assurance Region Analysis (Thompson et al. 1986) which introduces a set of constraints on the shadow prices of inputs and outputs. That amounts to imposing an order of importance (trade-off) on inputs and outputs. In this way the ranking over countries is more selective, although that comes at the

cost of losing the core attribute of DEA, namely its least requirement in terms of “a priori” constraints on variables.⁵

The results are very assumption sensitive. Limiting attention to work similar to ours (that is to models which simultaneously incorporate conventional inputs—labour and capital-, social and environmental types of cost, in addition to measures for well-being) one can observe that the inclusion/exclusion of conventional inputs can alter the results significantly. Mizobuchi (2014) finds positive correlation between GDP and the aggregate index of well-being through a DEA model when conventional inputs are not included; while correlation becomes negative when they are included.

Golany and Thore (1997) include as inputs reproducible capital and public expenditure. They find that both “young” countries (e.g. Morocco, Tunisia, Zaire, Peru and Uruguay) and “mature” ones (U.K., the Scandinavian countries, Australia and New Zealand) are plagued by inefficiency. The first due to increasing returns to scale, the latter because of decreasing returns relatively to public expenditure.

While the aim of our paper is very much in the line of Golany and Thore (1997) and Mizobuchi (2014) we part from them in many respects. As for Golany and Thore (1997), we are in the lucky position of being able to take advantage of the BLI dataset available from 2011. In addition, we avoid including public expenditure items as inputs because they are more likely to represent proxies for services provided to citizens, hence belonging to the outputs measures already contained in BLI.⁶ As far as Mizobuchi (2014) is concerned, while we rely on same dataset for the output variables (OECD 2013a, b), the major source of differentiation rests on the methodology we use to aggregate outputs and the more DEA-consistent procedure we use to deal with bad outputs (see Sect. 4. for details).⁷

It is worth mentioning that in order to avoid the dangers posed to DEA models by data transformation & aggregation, we decided to use as outputs all of the 24 measures contained in the BLI database. To countervail the effects that such an abundance of outputs variables could have in hampering the discriminating power of our DEA model, we integrate it with Principal Component Analysis in order to reduce the dimensions while retaining most of the variance contained in the database of input–output measures.

In short, our model provides a consistent methodology to be used in building a well-being index which takes into account the role of resources—conventional inputs (labour and capital) and externalities type of costs (social and environmental)—necessary to provide a given level of well-being, while keeping at a minimum, thanks to DEA being a non-parametric method, the requirement in terms of “a priori” constraints on variables.

⁵ Interestingly, Zhu (2001) uses the three best cities selected by Fortune magazine to describe the efficient frontier. It is clear, however, that the results are strongly affected by the choice made by the analyst.

⁶ On the other hand we include labour inputs, not considered in Golany and Thore (1997) who include as *bad-output*, “Infant mortality”, by taking the complement to one of the variable. Against this procedure a wide range of Dea literature (for instance Färe and Grosskopf 2004), has made it clear that such a transformation can lead to absurd conclusions. Moreover, Lovell and Pastor (1999) show that even the value relative to which the complement is taken affects the results.

⁷ Mizobuchi (2014) in using the 11 Bli indices (topics) from OECD (2013a, b) ends up dealing with bad output as in Golany and Thore (1997). A further source of differentiation concerns labour inputs. While Mizobuchi (2014) uses “Intangible wealth” World Bank (2011) we choose to use “per capita worked hours” (Oecd 2015a). Our choice relies on two observations: i) “intangible capital” estimates contain the effects on human capital of variables such as, for instance, education, health expenditure, role of institutions, consumption and wealth which are well-being measures contained in the Bli database; ii) in the World Bank’s own words intangible capital estimates contain some risk of the “black box” type (World Bank 2011, p. 94).

4 The Model

The basic idea is that in assessing a country's level of well-being one has to provide a model capable of taking into account: (1) the use of conventional inputs (such as labour and capital); (2) the production of social and environmental (negative) externalities (we shall refer to them as: bad-outputs); (3) the production of welfare enhancing goods and services (referred to as: good-outputs). In addition, the model has to be capable to aggregate the above input–output measures in a single index, in a way similar to what is done by GDP. Figure 2 provides a representation of what can be thought of as an aggregate (country) well-being production function.⁸

We propose to rely on a DEA type of model for the following reasons: (a) being grounded on the economic theory of production, it is capable to deal with input and output measures according to an efficient use of resources; (b) because it does not require the imposition of a functional form to the production function to be estimated; (c) can deal with multi-outputs and multi inputs technology while keeping relatively (to statistical regression type of estimation) low requirements on the quantity of available data; d) because the efficient frontier is made of actual countries, rather than hypothetical estimated points, policy recommendations gain in terms of operational relevance.⁹

Using *better life index* (BLI) data in a DEA type of model poses two kind of problems. On the one hand the large number of variables contained in BLI greatly limits the selective power of the model. There are 24 outputs and 2 inputs variables, with 35 DMUs (countries) yielding a variables to cases ratio well above the conventional standard (Dyson et al. 2001) and exposing the analysis to the “*dimensionality curse*” of DEA. On the other hand some of the variables have the nature of “*bad-outputs*” (8 out of 24 output variables) and they pose a difficult problem to any DEA model because there is no fully satisfactory method, so far available, to deal with such type of variables.

In order to overcome the dimensionality problem we follow the idea, first put forward by Ueda and Hoshiai (1997), Adler and Golany (2001, 2007), Adler and Yazhemy (2010) of integrating Principal Component Analysis (PCA) into DEA models. PCA makes it possible to reduce the number of variables while retaining the most of data variability (70–80 % is the commonly accepted threshold). Because earlier procedures employed for integrating principal components into Dea face the problem of negative data brought about by the possibility of negative loadings, which usually result from the single value decomposition of the correlation matrix of the data, we follow the procedure proposed by Yap et al. (2013) and select only positive loadings. The resulting principal components are necessarily positive and still represent the recommended (70–80 %) of the original data variability. By following this procedure in loading selection, the original 16 *good- output* variables get reduced to 5 latent variables containing 73 % of the original data variability. Likewise, the original 8 *bad-output* variables are reduced to 4 latent variables with 78 % of explained variance. The selected loadings and principal components are the same for all the models we use in the analysis. Hence, explained variance is constant across models. On the input side, as there are only 2 variables (capital and labour) we decided not to use principal components.¹⁰

⁸ Input and output measures in Fig. 2 anticipate the contents of the Bli database described further in Sect. 5.

⁹ For a detailed discussion of these points see Cooper et al. (2007).

¹⁰ We did actually run Pca on the two input variables. However the eigenvectors showed loadings of similar size with opposing signs, suggesting that labour and capital original variables tend to be orthogonal (higher capital input correspond to lower labour input). In this case, limiting the choice to one latent variable would heavily reduce explained variation.

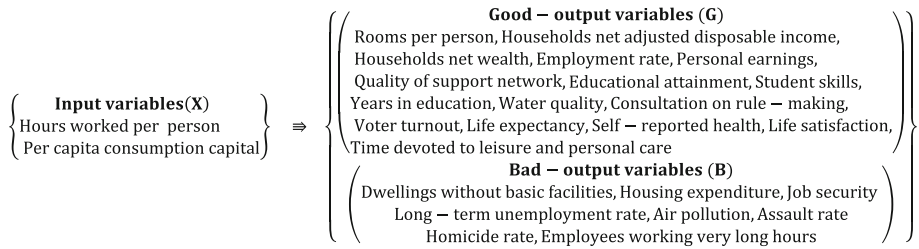


Fig. 2 Well-being production function

To deal with the problem posed by the presence of *bad-outputs* and taking into account that, on this matter, there is not in Dea’s literature a procedure which dominates the others, we propose to rely on two alternative models. One proposed by Tone and Tsutsui (2006) based on the idea that bad-outputs are better seen as “*non-conventional*” inputs and are part of a single production system, which includes *bad* and *good-outputs* besides conventional inputs (capital & labour). Therefore the model attaches a higher efficiency index to countries where production of *good-outputs* is carried out with relatively less use of conventional inputs (capital and labour) or/and *non-conventional* inputs (*bad-outputs*). The model is of the Slack Based Measure (Sbm) type, capable, therefore, to isolate both the traditional, radial, component of inefficiency, and its mix component¹¹ (Tone 2001).

The Principal Component modified linear model of the Tone and Tsutsui (2006) Sbm type is for each country.¹²

$$\begin{aligned}
 \min_{t^k, z^k, s_G^k, s_B^k, \lambda^k} \quad & Ew^k = t^k - p_X^k z^k \\
 & t^k + p_G^k s_G^k + p_B^k s_B^k = 1 \\
 & t^k L_G s^k = L_G G \lambda^k - s_G^k \\
 & t^k L_B b^k = L_B B \lambda^k + s_B^k \\
 & t^k L_X x^k = L_X X \lambda^k + z^k \\
 & t^k, z^k, s_G^k, s_B^k, \lambda^k \geq 0
 \end{aligned} \tag{1}$$

where X is the m by n matrix of m conventional input variables (labour and capital) and n countries, G is the h by n matrix of h *good-output* variables, B is the j by n matrix of j *bad-output* variables; t is a scalar, z^k , s_G^k and s_B^k , are, respectively, $(m \times 1)$, $(h \times 1)$ and $(j \times 1)$ column vectors of conventional inputs, *good-outputs* and *bad-outputs* efficiency slacks; p_X^k , p_G^k and p_B^k are row vectors of conventional inputs, *good-outputs* and *bad-outputs* weights relative to the Dmu (country) being evaluated. In addition, we have introduced L_G , L_B and L_X matrices, made of (row) eigenvectors obtained from single value decomposition of X , G and B correlation matrices, respectively, with loadings selected according to Yap et al. (2013).¹³

¹¹ Mix inefficiency measures the inefficiency due to the “wrong” combination of inputs and outputs.

¹² For the sake of generality in model presentation we keep including the eigenvector matrix for input variables (L_X) despite the fact that, as mentioned above, in the model we actually run input variables are left unchanged.

¹³ A detailed list of input and output variables contained in X , B and G matrices is in Fig. 2.

The main advantage of this model is that it provides a single index which simultaneously embodies the role of all the three sets of variables: conventional inputs, *good-outputs* and *bad-outputs*. We shall refer to this index of efficiency as: “*Efficient Well-being*” (Ew).

In addition to Ew index from model (1), it is of interest, for further comparisons of results between this model and the subsequent ones, to get a single measure of well-being without considering the role played by conventional inputs (labour and capital). In short, we want to reduce to a single dimension (scalar) all the different measures of well-being contained in the sets of *good-outputs* and *bad-outputs* variables (see Fig. 2). To this end we proceed by setting all conventional inputs equal to a uniform value (say 1) across countries. Running model (1) with such a uniform value of conventional input we get an index we shall refer to as “*Index of Well-being*” (Iw).

We now observe that a drawback of model (1) is that it fails to distinguish between sources of inefficiency. There is no way to attribute inefficiency to the production of *good-outputs* and *bad-outputs*. It can happen that a country is overall efficient because it performs well in producing good outputs while doing poorly for bad-outputs, or the other way round. In cases where *bad-outputs* are an important part of the problem, this feature of the model limits its information contents.

To overcome this problem we rely on Luptacik (2000) to split the analysis into three steps. On the first, the efficiency index of producing *good-outputs* with conventional inputs (capital and labour) is evaluated. For obvious reasons, we shall refer to it as *Technical Efficiency* (Te).

Then, as a second step, one computes the efficiency index of producing *good-outputs* out of *bad-outputs*, which come to play the role of *non-conventional* inputs. We name such an index: *Social Efficiency* (Se).

The last step proceeds to summarise these two indices into a single one by a standard Sbm (Tone 2001) model, which takes Te and Se indices as outputs while a uniform input (set equal to one) is imputed to all the Dmus. We shall name the index thus obtained as *Social and Technical Efficiency* (Ste).

Following the idea of Luptacik (2000), but using the SbmDea model (Tone 2001) amended as in model (1) by the integration of Principal Components, in order to compute the Te index we have for each country, $k(1, \dots, n)$:

$$\begin{aligned}
 \min_{t^k, z^k, s_G^k, \lambda^k} \quad & \text{Te}^k = t^k - p_X^k z \\
 & t^k + p_G^k s_G^k = 1 \\
 & t^k L_G g^k = L_G G \lambda^k - s_G^k \\
 & t^k L_X x^k = L_X X \lambda^k + z^k \\
 & t^k, z^k, s_G^k, \lambda^k \geq 0
 \end{aligned} \tag{2}$$

While the model for the Se index, always for $k(1, \dots, n)$, is:

$$\begin{aligned}
 \min_{t^k, s_B^k, s_G^k, \lambda^k} \quad & \text{Se}^k = t^k - p_B^k s_B^k \\
 & t^k + p_G^k s_G^k = 1 \\
 & t^k L_G g^k = L_G G \lambda^k - s_G^k \\
 & t^k L_B b^k = L_B B \lambda^k + s_B^k \\
 & t^k, s_B^k, s_G^k, \lambda^k \geq 0
 \end{aligned} \tag{3}$$

And the model for the Ste index, which does not require Principal Components, for each country, $k(1, \dots, n)$, is:

$$\begin{aligned}
 \min_{t^k, s_{Se}^k, s_{Te}^k, z^k, \lambda^k} \quad & Ste = t^k - p_{Tz}^k z^k \\
 & t^k + p_{Se}^k s_{Se}^k + p_{Te}^k s_{Te}^k = 1 \\
 & t^k Se^k = Se \lambda^k - s_{Se}^k \\
 & t^k Te^k = Te \lambda^k - s_{Te}^k \\
 & t^k 1 = 1 \lambda^k + z^k \\
 & t^k, s_{Se}^k, s_{Te}^k, z^k, \lambda^k \geq 0
 \end{aligned} \tag{4}$$

where symbols are as in model (1), and variables super and sub scripts are amended as required by each model.

As a last step, there is one more index worth computing for comparison purposes with the above indices, and we shall refer to it as “Efficient Gdp” (Egdp), that is how efficiently Gdp is produced by each country relatively to the use of conventional inputs (labour and capital). To this end we use a standard Sbm model (Tone 2001) without Principal Components and with just one output variable (per capita Gdp) and the two conventional inputs: labour and capital.

5 Data Description

The data on output variables all come from the Bli database, provided by Oecd for all of its 34 member countries (and for Brazil and Russia), and made of 11 “topics”, that is groups of output variables.¹⁴ Each topic is made of one or more of the 24 output variables, of which 16 are *good outputs* (Bli gets higher as they increase: e.g. *life expectancy*) and 8 are *bad-outputs* (Bli gets smaller as they increase: e.g. *Dwellings without basic facilities*). Table 2 details the 24 output variables and how they are grouped into the 11 topics.

In order to group 24 output variables into 11 topics, Oecd, firstly, normalizes the value each variable takes, so that they all are in the 0–10 range (*min max method* Oecd 2008, p. 30):

$$index = \left(\frac{observed\ value - minimum\ value}{maximum\ value - minimum\ value} \right) \times 10 \tag{5}$$

Secondly, *bad-outputs* variables undergo a unit translation (1 – index) in order to make the complement to one comparable with the good-outputs variables.

Thirdly, the indices so obtained are aggregated into 11 “topics” by simple average:

$$Topicindex_j = \left(\frac{\sum_{i=1}^N index_i}{N} \right); (j = 1, \dots, 11) \tag{6}$$

We observe that the above transformation of raw variables would be a cause of serious problems for any Dea type of model. Firstly, most models are not *translation-invariant*¹⁵ and the *bad-output* transformation in (5) is a translation. Secondly, (6) implicitly assigns equal weights to different indices, while one of the most valuable properties of Dea is that of assigning to input and output variables the weights most favourable to the Dmu being evaluated. Therefore, and differently from Mizobuchi (2014), we avoid using as output

¹⁴ This is how Oecd (2011) refers to the aggregate output variables.

¹⁵ Pastor (1996) shows that in the cases where there are translated variables, it is only safe the use of the old version of Bcc input oriented or the additive Vrs models.

Table 2 Topics and Related variables in better life index

Topics	Related variables
Housing	Dwellings without basic facilities, Housing expenditure, Rooms per person
Income	Households net adjusted disposable income, Households net financial wealth
Jobs	Employment rate, Job security, Long-term unemployment rate, Personal earnings
Community	Quality of support network
Education	Educational attainment, Student skills, Years in education
Environment	Air pollution, Water quality
Civic engagement	Consultation on rule-making, Voter turnout
Health	Life expectancy, Self-reported health
Life Satisfaction	Life satisfaction
Safety	Assault rate, Homicide rate
Work-Life Balance	Employees working very long hours, Time devoted to leisure and personal care

variables the 11 *topics* indices of Bli and use instead the 24 original indices split into two sets: 16 *good-output* variables and 8 *bad-output* variables (see Fig. 2). We can afford doing so, without jeopardising the discriminating power of models (1–4) thanks to Pca which makes possible to reduce dimension while keeping most of the variance.

As for the choice of input variables we observe that available data at international level lead to a natural distinction between capital and labour. From a very abstract point of view there are many similarities between capital and labour inputs. As capital goods, purchased or rented, can be seen like a carrier of capital services which are the actual input to production, in very much the same way hired employees can be seen as carriers of human capital services which are the actual labour inputs. In either cases actual input service is not a stock measure but rather a flow, which then comes to be exhausted in the production process. Differently from labour, the difficulty with capital comes from it being, in most cases, owned by the entrepreneur. This implies that in the provision of capital services one does not observe any market transaction. There lies the difficulty in capital inputs measurement for production and for productivity analysis (Schreyer 2004). Capital services originate from the stock of past investment (net or gross capital stock), but the two do not coincide, nor are they related in a simple manner, short of very special cases. Most textbooks present the production function defined in terms of capital stock, besides labour inputs. They contain the implicit assumption that capital service flows are proportional to capital stocks. In general such an assumption does not hold and the use of capital stock as a proxy for input services leads to a bias because: (1) assets have different life span, (2) they do not have a uniform vintage distribution, and (3) capital services vary with the age of the asset (Oecd 2009, Chps. 2 and 9; Chp. 5 in ONS 2007).¹⁶

These are in short the reasons which, in recent times, have lead international organizations as Oecd to provide of a Volume index of capital services (Vics). However, coverage is far from complete.¹⁷ Only few countries provide their own estimates. For this

¹⁶ As succinctly stated by Oecd: “It is also noteworthy that neither the net nor the gross stock are the conceptually correct measure to capture capital inputs into production—these are best reflected through measures of the flow of capital services” Oecd (2013a, b).

¹⁷ Out of the 35 countries in the Bli database only 20 can be found in the Oecd Vics database.

reason, in research work the compulsory choice is to select a proxy.¹⁸ The most frequent choice is for “Consumption of fixed capital”, a measure widely available as it belongs to the System of National Accounts.

As far as labour inputs is concerned, we have at a first stage considered to use some sort of quality adjusted labour input (Qali) in line with what Ons (2007) does for the UK. However, besides the coverage problem, due to the fact that such data have only recently been provided in some of the major industrialized countries, we noticed that the main quality indicator used in Qali is Educational Attainments which in turn belongs to the set of output measures included in Bli. Therefore, to avoid double counting on the effects of education, we decided to revert to the more traditional *Hours worked* (Oecd 2015a).

To summarise the choice in terms of input and output variables, we have 35 countries (all the Oecd members and Russia¹⁹). On the output side there are 24 output variables (all from Bli database, OECD 2013a, b, 2015 edition). As for inputs, we have two variables: Consumption of fixed capital (World Bank 2015) and *Hours worked* (Oecd 2015a). The reference year for all input variables is 2012, because most of the outputs variables refer to 2012/2013 (Table 3). In order to get scale-independent results, we use per-capita values for all input measures. Table 3 shows some summary statistics for inputs and outputs.

6 Results

Let us first look at how efficiently Gdp is produced (Egdp) by a Sbm model with one output (per-capita Gdp) and two inputs (per-capita worked hours and per-capita capital consumption, see Sect. 4).²⁰ Figure 3 clearly shows that efficiency in Gdp production, Egdp, does not imply a high level of Gdp: Russia, Ireland and Luxemburg are efficient (they are on the frontier), but they have very different per-capita Gdp (they belong to the first e forth quartile in terms of per capita Gdp). Countries most distant from the frontier, that is least efficient, belong to very different quantile in terms of per-capita Gdp: Japan and Switzerland are inefficient as they are Greece and Mexico, much poorer than the former.

Differences among rich and poor countries remain: the former use too much capital (North-West of Fig. 3), the latter too much work (South-Est of Fig. 3). Altogether Fig. 3 shows that the conventional ranking of countries according to their per capita Gdp does change substantially if the ranking is made according to how efficiently (productive) they are in producing Gdp, that is if the ranking takes into account the resources they use.

6.1 Efficiency in Producing Well-Being

We are now ready for the assessment of countries’ relative efficiency in producing well-being. We consider a production process that uses two inputs (per-capita hours worked and per-capita consumption of capital) to produce a set of *good* and *bad* outputs, see Fig. 2. By model (1) we get a measure of how efficiently well-being is produced, the Efficient Well-

¹⁸ For instance, Ons (2004, 2006, 2008a, b) and Atkinson (2005) choose to rely on “Consumption of fixed capital” in their assessment of UK public sector productivity. We thank a referee for calling our attention on the need to clarify the point about capital input selection.

¹⁹ Bli is available for 36 countries. As Brasil lacks data on labour, we exclude it from our analysis.

²⁰ Gdp data is in the form of purchasing power parities (http://stats.oecd.org/index.aspx?DatasetCode=PDB_LV).

Table 3 Summary Statistics for Input and Output Measures *Source:* OECD (2013a, b), 2015 edition; Oecd (2015a), World Bank (2015)

Good Outputs	Year	Average	SD	Max	Min
Rooms per person	2013	1.66	0.44	2.50	0.90
Household net adj. disposable income (.000)	2012	24.48	7.05	41.36	13.09
Household net financial wealth (.000)	2012	41.23	32.40	145.77	3.25
Employment rate	2013	66.37	7.70	82.00	49.00
Personal earnings (.000)	2013	36.59	12.83	56.34	16.19
Quality of support network	2014	89.63	5.17	96.00	72.00
Educational attainment	2012	76.23	15.96	94.00	34.00
Student skills	2012	496.69	26.34	542.00	417.00
Years in education	2012	17.53	1.30	19.80	14.40
Water quality	2014	83.00	10.60	97.00	56.00
Consultation on rule-making	2008	7.16	2.66	11.50	2.00
Voter turnout	2010–2015	69.91	12.26	93.00	49.00
Life expectancy	2012	79.91	2.93	83.20	70.20
Self-reported health	2013	67.89	14.59	90.00	30.00
Life satisfaction	2014	6.57	0.80	7.50	4.80
Time devoted to leisure and personal care	1999–2011	14.88	0.55	16.06	13.42
Bad Outputs	Year	Average	SD	Max	Min
Dwellings without basic facilities	2013	2.41	3.74	15.10	0.00
Housing expenditure	2012	20.83	2.88	26.00	11.00
Job security	2013	5.70	2.80	17.80	2.40
Long-term unemployment rate	2013	3.58	3.94	18.39	0.01
Air pollution	2010	19.94	8.31	46.00	9.00
Assault rate	2010	3.94	2.17	12.80	1.30
Homicide rate	2012	2.18	4.35	23.40	0.30
Employees working very long hours	2013	9.12	8.46	40.86	0.16
Inputs	Year	Average	SD	Max	Min
Average hours worked per person (.000)	2012	1.73	0.21	2.23	1.38
Per capita consumption of fixed capital (.000)	2012	6.46	3.88	17.24	0.61

being index (Ew), to be compared with the simple Index of Well-being (Iw), which is computed with constant input for all countries (see Sect. 4).

Figure 4a compares different indices according to quartiles of per-capita Gdp distribution, with per-capita Gdp index normalized relative to the highest value (Luxemburg). Three main observations are in order.

- a. Within the first two quartiles of the (per-capita) Gdp distribution, Iw is generally positively correlated to (per-capita) Gdp. However 3rd and 4th quartile have the same Iw (Austria and Netherlands have a lower Iw than countries with similar or lower per capita Gdp). In other words, marginal contribution of Gdp to Iw is somehow decreasing, and it seems that it gets to zero as Gdp reaches median per-capita Gdp (see

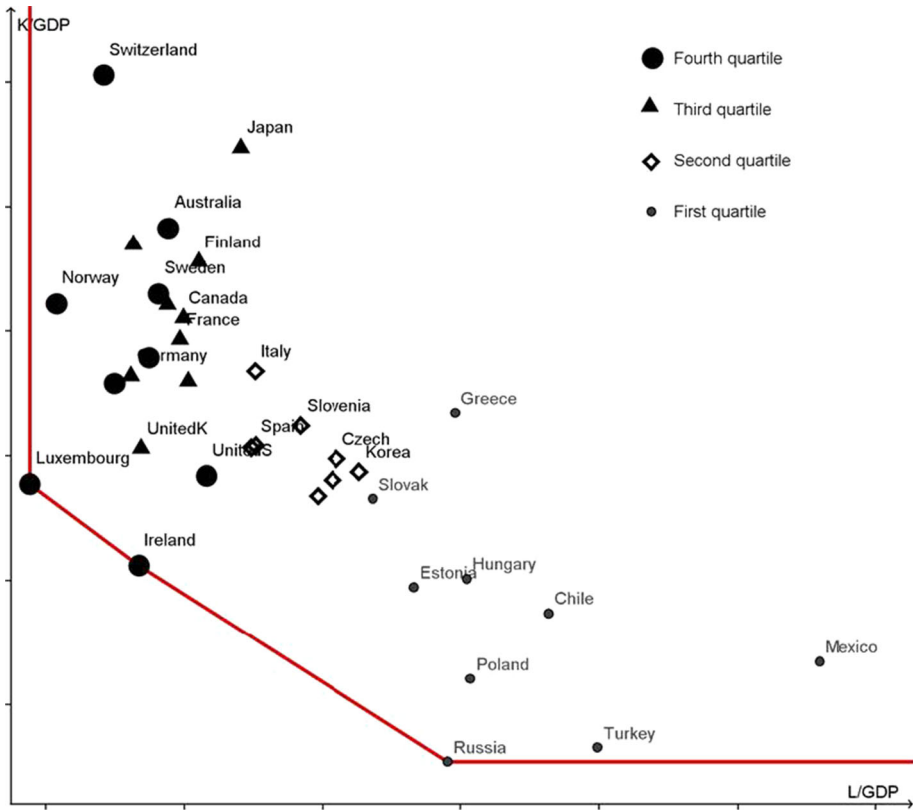


Fig. 3 Efficient Frontier in Terms of Gdp (2013) Inputs (Labour and Capital). *Original data source:* Oecd (2015a, b), World Bank (2015)

Tables 6 and 7 in the Appendix). This is a result in line with others' according to which well-being is only partially related to Gdp. For example, from the research work in late 80's (Fuà 1993), we know that life expectancy at birth, increases at a decreasing rate with respect to per-capita Gdp. The same holds in the happiness literature, which shows that economic prosperity, as measured by per-capita Gdp, is not systematically associated with greater happiness (the "happiness paradox", Easterlin 1974).

- b. Countries' ranking in terms of Iw differ from that in terms of Ew, i.e. countries perform differently when the use of inputs to produce well-being is taken into account. Figure 4A shows that the relative position of "poor" countries (1st quartile) greatly improve in moving from Iw ranking to the Ew ranking. In the move, poor countries show, on average, an efficiency gain (with the exception of Greece), while rich ones experience a loss (Netherlands being the only exception). Thus, Fig. 4a provides evidence of the important role played by the use of resources (conventional inputs) in the well-being rankings of countries.
- c. In absolute terms, however, it is still the case that rich countries (4th quartile) perform better than the others in any of the four dimension of Fig. 4a. This point deserves further investigation and will be picked up again in relation to social and technical indices (see Sect. 6.2).

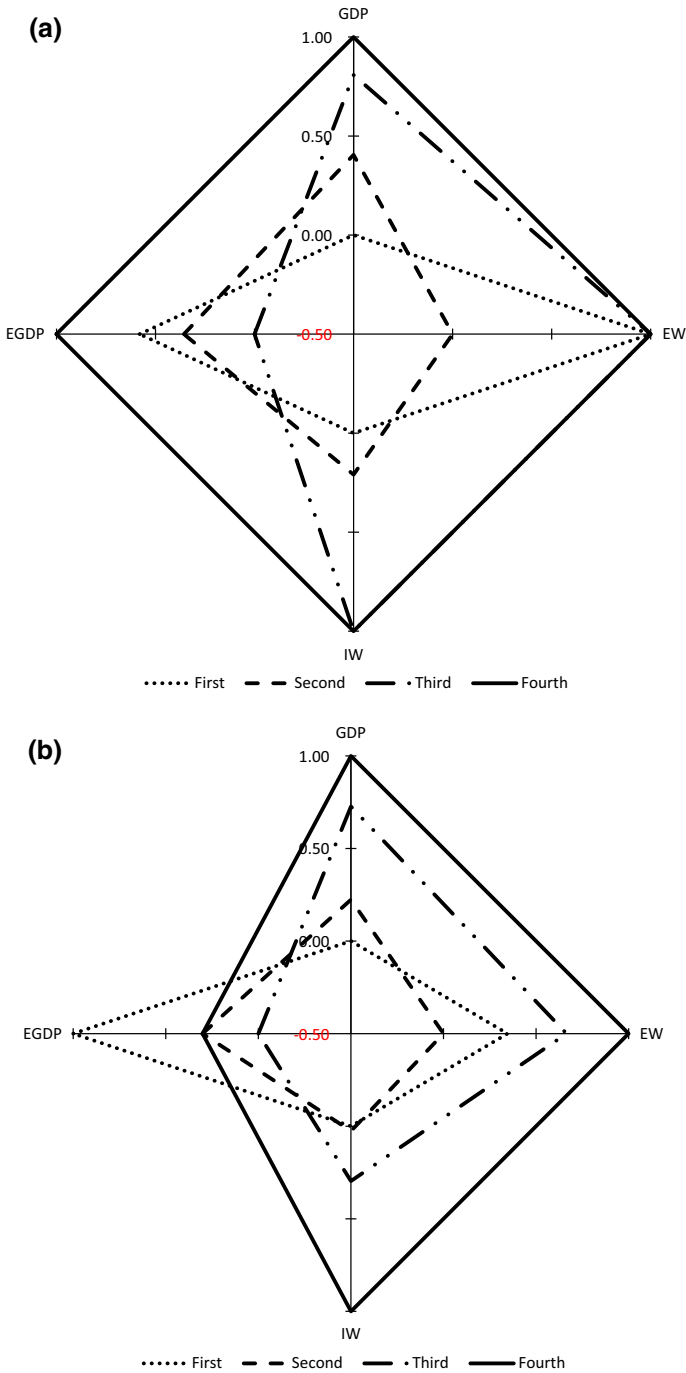


Fig. 4 **a** Quartile median of Gdp-per capita (Gdp), Efficient Gdp (Egdp), Index of well-being (Iw), and Efficient well-being (Ew) indices; **b** Cumulative quartile median. *Note* data has been normalized according to the observed-minimum value divided by the range

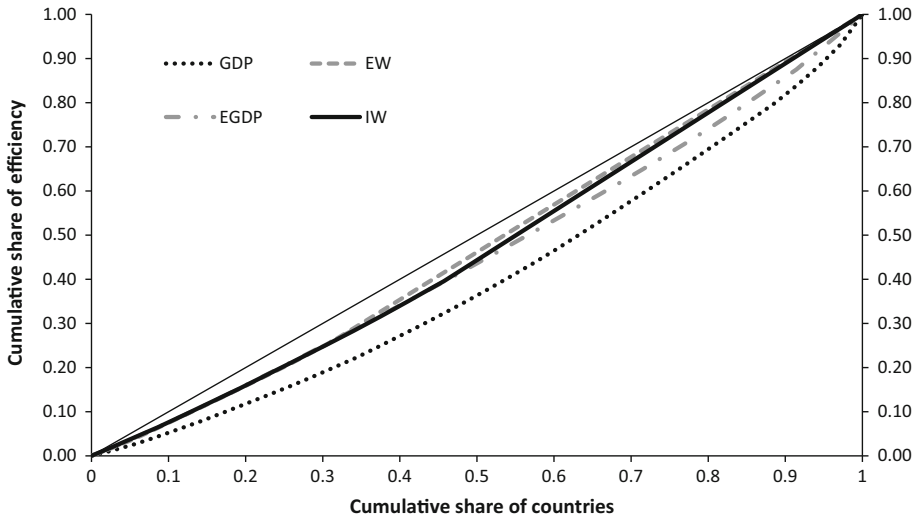


Fig. 5 Lorenz concentration curve for efficiency indices

The different behaviour of countries belonging to different quartile in terms of per-capita Gdp, is better seen in Fig. 4b, which shows the cumulative distribution of quartile median for the various indices of well-being already depicted in Fig. 4a. The overall median (4th quartile) can be seen as a result of a path going from high efficient Gdp index (Egdp), by relatively poor country belonging to the first quartile, to higher efficient well-being (Ew) by relatively richer countries belonging to higher quartiles. Along this path Egdp index is reduced while Ew increases hinting to the presence of a trade-off away from Gdp and towards well-being as countries become richer in terms of per-capita Gdp.

The difference in countries’ performance is confirmed even in terms of inequality in the distribution of well-being indices. Figure 5 reports the Lorenz curves for the four indices of Fig. 4a, b. It is easily seen that the least equal distribution of indices is for the per-capita Gdp and to a lesser extent for the Egdp. While well-being indices (Ew and Iw) follow a more even distribution. Once again this happens because relatively poor countries perform, on average, not too bad in terms of well-being indices due to their lower use of conventional inputs.

Our model also yields additional information in the form of *slacks*: a measure of the inefficiency in the use of each input and output. In its last column Table 4, contains major slacks for each of the inefficient countries. Greece, for example, could achieve the same level of well-being by lowering the use of capital by 41 % and hours worked by 12 %. Table 4 also contains, in the Peers column, those countries which are the terms of reference for non-efficient countries in the row. For example, Greece’s peers are Poland, Ireland and Hungary.²¹ The second column reports an efficiency score of 0.62 for Greece, that is on average the country is 62 % as efficient as its peers (i.e. 38 % inefficient). The third column shows the Returns to Scale: Irs means that scale of production should be increased

²¹ Slacks are indeed computed over a linear combination of input and outputs of the peer countries, as if we got an ideal country to be compared with each inefficient unit.

Table 4 Summary of efficiency indicators for non-efficient countries

Country (Quart)	Score	Scale	Peers	Input slacks	Output slacks ($\geq 0.5\%$)*
Greece (1st)	0.62	IRS	Poland, Ireland, Hungary	Capital, Labour	Homicide, Ed. Attainment, Job Sec. Unemployment
Italy (2nd)	0.67	IRS	Ireland, Poland, N. Zealand	Capital, Labour	Homicide, Facilities, Long Work, Consultation
Chile (1st)	0.77	IRS	Russia, Poland, UK	Capital	Ed. Attainment, Consultation, Long Work, Facilities
Spain (2nd)	0.77	IRS	Ireland, Poland	Capital	Facilities, Homicide, Wealth, Ed. Attainment, Unemployment
Belgium (3rd)	0.77	IRS	Germany, Ireland, N. Zealand	Capital	Homicide, Facilities, Consultation, Ed. Attainment
Israel (2nd)	0.78	IRS	N. Zealand, Russia, Poland	Capital	Consultation, Homicide, Facilities
Switzerland (4th)	0.79	DRS	Norway, Germany, N. Zealand	Capital	Homicide
France (3rd)	0.80	IRS	Ireland, Germany, UK	Capital	Homicide, Consultation, Facilities
Portugal (2nd)	0.80	IRS	Ireland, Poland, Russia	Capital	Homicide, Ed. Attainment
Austria (4th)	0.86	IRS	Germany, N. Zealand, USA	Capital	Facilities
Finland (3rd)	0.88	DRS	Canada, Ireland, Germany	Capital	Wealth

for the country to become efficient, Drs that the country should decrease its scale of production (Switzerland and Finland).²²

A few general observations can be made about data in Table 4. Firstly, one sees that poor countries have poor countries as Peers and this goes to support the plausibility of the results and their operational contents because reference can be made to actual cases (countries) and not to hypothetical reference points. Secondly, input slacks increase in number and dimensions (not shown Table 4) as efficiency (scores) decreases. That is, inefficient countries tend to have slacks on both (conventional) inputs. Moreover, one observes that the number of slacks on output variables increases as efficiency decreases. That is, inefficiency never comes as a factor specific problem but rather as a system problem. Of some interest, although marginal, is the case of slacks for “*Homicide rate*” and “*Facilities*” (i.e. *Dwelling without basic facilities*, in Table 2; Fig. 2). They are almost constantly present for both relatively efficient and largely inefficient countries.

6.2 Social and Technical Efficiency

The previous section leaves one important question open: do rich and poor countries perform equally in terms of efficient well-being (Ew)? Fig. 4a above, seems to suggest an affirmative answer. To answer the question, however, one has to gain a better insight into the components of the Ew index. Relatively poor countries and relatively rich countries have almost the same performance in terms of Ew, cfr. Fig. 4a. This results, however,

²² Decreasing returns for “mature” countries is a result we have in common with Golany and Thore (1997).

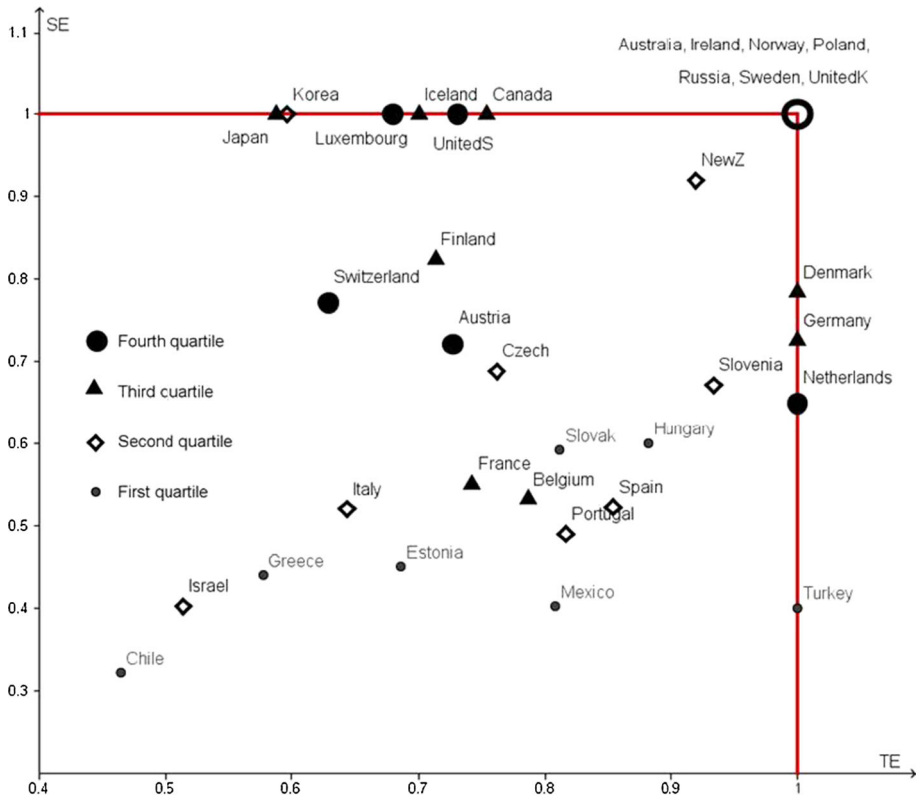


Fig. 6 Efficient frontier in terms of technical and social efficiency

sheds no light on how countries perform in terms of the different components of Ew. We know that in producing well-being countries face two types of costs: conventional inputs, and socio-environmental costs. Ew index from model (1) (hence Fig. 4a) includes them both without the possibility of separating the effects.

The model by Luptacik (2000), as described in Sect. 4—models (2)–(4), make it possible to split efficiency into two components: Social and Technical (respectively, Se and Te). Se refers to the social costs of well-being,²³ namely the amount of bad outputs. Te relates to the amount of conventional inputs (capital and labour) used to get the set of good outputs (see Fig. 2).

Figure 6, with Se on the vertical axis and Te on the horizontal, shows the countries’ position in terms of technical and social efficiency mix. Some of them are only efficient on the technical side (like Denmark and Germany), others are only socially efficient (as Canada and United States), and seven of them are efficient under both criteria. Inside the frontier, there are countries not efficient in any of the two dimensions.

Two facts are worth underlining: (1) relatively poor countries (up to the second quartile) show low Se, while inefficiency in terms of Te does not discriminate among quartiles of the distribution. For example Netherlands, Germany and Denmark have higher Technical than Social efficiency, while Luxembourg, Switzerland, Australia and Canada have better

²³ The *bad-outputs* in Fig. 2. The undesirable effects of economic activity.

social than technical efficiency; (2) among countries efficient in both dimensions there are some from first (Poland and Russia) third & fourth quartile (Australia, Ireland, Sweden and U.K.).²⁴

Figure 7a contains what can be considered the most relevant finding of the paper: when bad outputs (social and environmental costs) and conventional inputs (labour and capital) are both considered in assessing countries' performance in well-being production, it turns out that poor and rich countries perform differently in terms of the components of final well-being.

Namely, poor countries outperform richer ones in terms of T_e (Technical efficiency), while the other way round happens in terms of S_e (Social efficiency). For instance, countries from the first quartile, such as Slovenia and Turkey are champions in terms of T_e , but perform badly in terms of S_e . On the other side, countries from the fourth quartile, such as Luxembourg and the U.S are on the frontier in terms of S_e , but perform poorly in terms of T_e . Figure 8 shows that T_e is more equally distributed than S_e .

Figure 7a provides also a further result. In as far as it shows that Gdp and T_e are badly correlated,²⁵ it provides support to the widely held point that in pursuing well-being countries do not necessarily need to rely on conventional (capital and labour) input. There are other types of cost (social and environmental) that can equally contribute to well-being.²⁶

One more result from Fig. 7a deserves attention because it is surprising. Ranking countries according to per-capita Gdp or according to social and technical index of well-being (Ste) makes little difference as far as quartiles are concerned²⁷ (even clearer in Fig. 7b). In short, despite the criticisms, Gdp happens to be still a good proxy for well-being. It might not be so much surprising if one considers that these findings are in line with the interpretation of per capita Gdp as an indicator of development (World Economic Forum 2013). At earlier stages of economic development, countries are somehow more interested in technical efficiency, while more developed countries are prepared to sacrifice some technical efficiency to gain in social efficiency (Fig. 7a, b). In other words, we observe a different use of resources as the social cost of well-being gains importance. Thus, the choice of spending part of conventional inputs to reduce *bad outputs* may cause a relative decline of *good outputs* and consequently a loss in technical efficiency, but not in Gdp because of the way expenditures are accounted for in the System of National Accounts.

7 Concluding Remarks

The starting point of this paper is that in producing a well-being index from a multidimensional set of measures one has to (1) rely on a consistent system for aggregating dimensions; (2) consider the resources necessary to make the well-being available to citizens (such as conventional inputs—such as labour, capital—and less conventional inputs such as social and environmental costs).

²⁴ To be noticed that Russia excellent performance is affected, in the principal component analysis, by a large loading in "Housing expenditure" which is relatively low, compared to western standards, due to the high frequency of council houses.

²⁵ Table 5 shows that zero rank correlation is within the bounds of the 95 % confidence interval.

²⁶ We thank a referee for calling our attention on this way to look at the results.

²⁷ Gdp and Ste are positively rank correlated at 95 % (see further down Table 5).

Table 5 Rank correlation among efficiency indices (95 % bootstrap upper and lower bounds)

	GDP	IW	EGDP	EW	TE	SE	STE
GDP	1.000						
l.b.	0.200						
IW	0.555	1.000					
u.b.	0.795						
l.b.	-0.370	-0.265					
EGDP	0.065	0.092	1.000				
u.b.	0.426	0.422					
l.b.	-0.052	0.181	-0.136				
EW	0.255	0.449	0.175	1.000			
u.b.	0.534	0.667	0.465				
l.b.	-0.326	-0.152	0.091	-0.008			
TE	0.065	0.219	0.470	0.336	1.000		
u.b.	0.371	0.536	0.768	0.618			
l.b.	0.273	0.696	-0.233	0.235	-0.168		
SE	0.587	0.867	0.175	0.527	0.208	1.000	
u.b.	0.820	0.949	0.527	0.732	0.542		
l.b.	0.170	0.655	-0.106	0.262	0.242	0.724	
STE	0.546	0.833	0.310	0.572	0.587	0.882	1.000
u.b.	0.782	0.918	0.631	0.762	0.815	0.958	

Bootstrap with 1000 replicates, using R package by Hervé (2015)

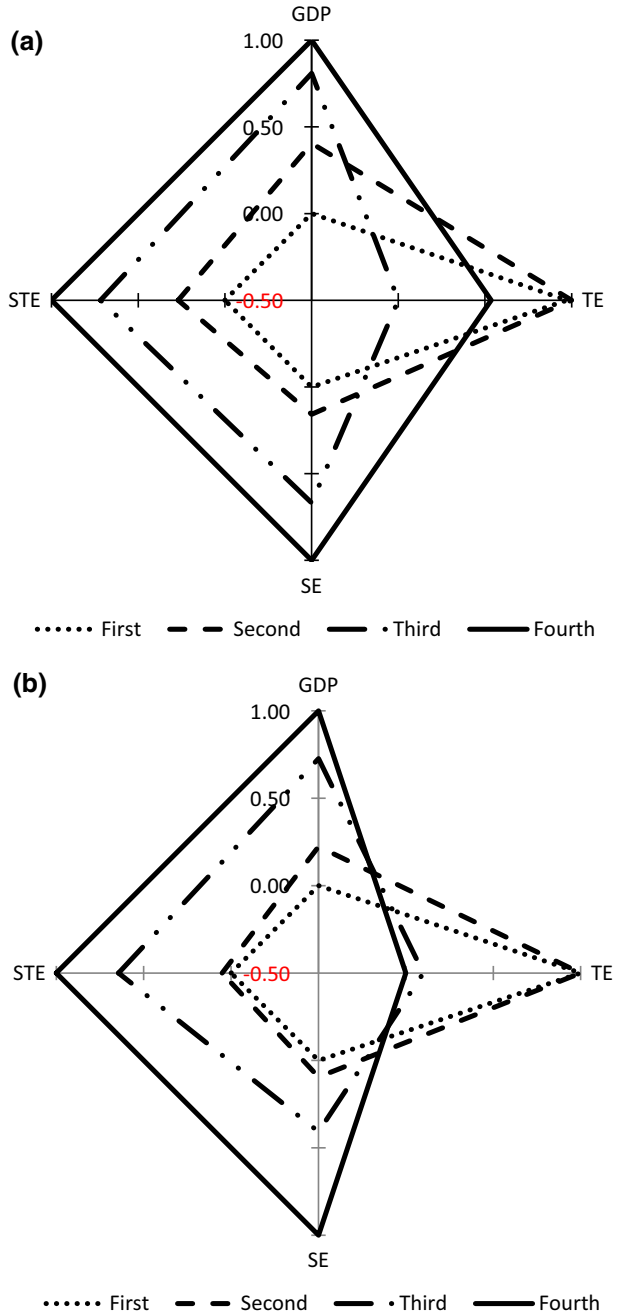
Bold value indicates the correlation and the non-bold value indicates the confidence interval

We propose a methodology which combines Principal Components Analysis and Data Development Analysis. The main result of the paper is that the costs of producing well-being have a large and significant impact on the resulting index of well-being. This is true for both the cases of Gdp relatively to the efficient Gdp, and for the Index of well-being (Iw) relatively to Efficient well-being. It remains true under different model specifications, as it holds also for Technical efficiency (Te) and Social efficiency (Se).

A second result is that countries react differently to the different costs of well-being. Relatively poor countries seem to worry more about conventional input costs (labour and capital), while relatively rich countries are more concerned about the social and environmental costs. Our results show that the index of well-being (Iw) is correlated to the Gdp for first and second quartile. This is in line with the commonly held position according to which beyond a certain threshold gains Gdp are no longer correlated with increases in well-being (Easterlin 1974). Rank correlation indices in Table 5 provide a further support to these results. The high correlation between Gdp and Iw, confirms that part of well-being is still related to Gdp (Sect. 6.1). It can also be noted that there is a high correlation between Gdp and Se, confirming a lower eco-social impact by rich countries in providing well-being.

As a final remark one can observe that the role of inputs, both conventional and in the form of *bad outputs*, has been shown to be of primary importance in qualifying results in

Fig. 7 a Quartile median of Gdp per capita (Gdp) Social and Technical Efficient Well-being (Ste) Social Efficient Well-being (Se), Technical Efficient Well-being (Te); **b** Cumulative quartile median. *Note* Data has been normalized according to the observed-minimum value divided by the range



terms of well-being measurements. One would expect more attention by future applied research to the role input costs (conventional, social and environmental) play in defying and measuring collective well-being.

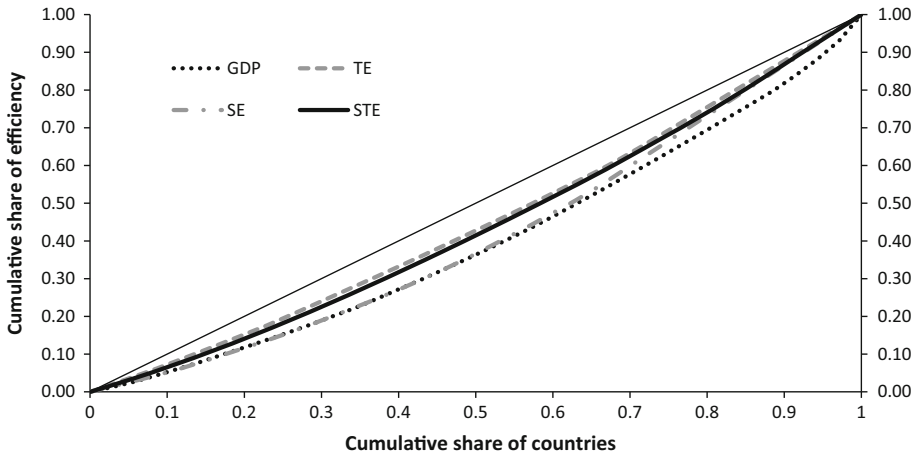


Fig. 8 Lorenz concentration curve for efficiency indices

Acknowledgments The authors would like to acknowledge comments made by the referees which have produced a much improved paper.

Appendix

See Tables 6 and 7.

Table 6 Decile median of efficiency indices

Decile	GDP	EGDP	IW	EW	TE	SE	STE
First	0.22	0.67	0.68	1.00	0.78	0.38	0.52
Second	0.28	0.81	1.00	1.00	1.00	1.00	1.00
Third	0.30	0.66	0.79	1.00	0.81	0.55	0.65
Fourth	0.36	0.64	0.85	1.00	0.65	0.63	0.74
Fifth	0.38	0.64	0.91	0.88	0.75	0.72	0.70
Sixth	0.42	0.63	1.00	0.88	0.79	0.84	0.78
Seventh	0.46	0.61	1.00	0.94	0.75	0.84	0.78
Eighth	0.48	0.60	1.00	1.00	1.00	1.00	0.88
Ninth	0.50	0.73	0.88	1.00	1.00	0.72	0.80
Tenth	0.68	0.80	1.00	1.00	0.70	1.00	0.82

Table 7 Cumulative decile median of efficiency indices

Decile	GDP	EGDP	IW	EW	TE	SE	STE
First	0.22	0.67	0.68	1.00	0.78	0.38	0.52
Second	0.26	0.68	0.74	1.00	0.81	0.43	0.55
Third	0.28	0.66	0.78	1.00	0.81	0.51	0.63
Fourth	0.28	0.66	0.79	1.00	0.80	0.55	0.65
Fifth	0.30	0.66	0.79	1.00	0.80	0.56	0.66
Sixth	0.32	0.66	0.80	1.00	0.79	0.60	0.68
Seventh	0.34	0.66	0.83	1.00	0.79	0.60	0.68
Eighth	0.36	0.65	0.92	1.00	0.79	0.67	0.74
Ninth	0.38	0.66	0.88	1.00	0.79	0.71	0.75
Tenth	0.40	0.66	1.00	1.00	0.79	0.79	0.76

References

- Adler, N., & Golany, B. (2001). Evaluation of deregulated airlines network using data envelopment analysis combined with principal component analysis with an application to Western Europe. *European Journal of Operational Research*, *132*, 260–673.
- Adler, N., & Golany, B. (2007). Pca-Dea. In J. Zhu & D. W. Cook (Eds.), *Modeling data irregularities and structural complexities in data envelopment analysis*. New York: Springer.
- Adler, N., & Yazhemsy, E. (2010). Improving discrimination in data envelopment analysis: PCA-DEA or variable reduction. *European Journal of Operational Research*, *202*, 273–284.
- Arrow, K., Dasgupta, P., Goulder, L., Daily, G., Ehrlich, P., Heal, G., et al. (2004). Are we consuming too much? *Journal of Economic Perspectives*, *183*, 147–172.
- Atkinson, A. (2005). *Measurement of government output and productivity for the national accounts: Final report*. Palgrave Macmillan: HMSO.
- Bandura, R. (2005). *Measuring country performance and state behavior: A survey of composite indices*. Undp/Ods background paper. New York, NY: United Nations Development Program, Office of Development Studies.
- Bandura, R. (2008). *A survey of composite indices measuring country performance: 2008 update*. Working paper, New York, NY: United Nations Development Program, Office of Development Studies.
- Bleys, B. (2012). Beyond Gdp: Classifying alternative measures for progress. *Social Indicators Research*, *109*, 355–376.
- Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*, 429–444.
- Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software*. Berlin: Springer.
- Costanza, R., Hart, M., Posner, S., & Talberth, J. (2009). *Beyond Gdp: The need for new measures of progress*, The Pardee Papers no. 4, Boston University.
- Costanza, R., Kubiszewski, I., Giovannini, E., Lovins, H., McGlade, J., Pickett, K. E., Ragnarsdóttir, K., Roberts, D., De Vogli, R., & Wilkinson, R. (2014). Development: Time to leave GDP behind. *Nature*, *505*(7483), 283–285.
- Coyle, D. (2014). *GDP: A brief but affectionate history*. Princeton: Princeton University Press.
- Dasgupta, P. (2001). *Human well-being and the natural environment*. Oxford: Oxford University Press.
- De Beukelaer, C. (2014). Gross domestic problem: The politics behind the world's most powerful number. *Journal of Human Development and Capabilities*, *15*(2-3), 290–291.
- Despotis, D. K. (2005). A reassessment of the human development index via data envelopment analysis. *Journal of the Operational Research Society*, *56*(8), 969–980.
- Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, *132*(2), 245–259.
- Easterlin, R. A. (1974). Does economic growth improve the human lot? In A. D. Paul & W. R. Melvin (Eds.), *Nations and households in economic growth: Essays in honour of Moses Abramovitz*. New York: Academic Press Inc.
- Färe, R., & Grosskopf, S. (2004). Modeling undesirable factors in efficiency evaluation: Comment. *European Journal of Operational Research*, *157*, 242–245.

- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A (General)*, 120(3), 253–290.
- Fioramonti, L. (2013). *Gross domestic problem: The politics behind the world's most powerful number*. Zed Books.
- Fleurbaey, M. (2009). Beyond GDP: The quest for a measure of social welfare. *Journal of Economic Literature*, 47(4), 1029–1075.
- Fuà, G. (1993). *Crescita economica: Le insidie delle cifre*. Bologna: Il Mulino.
- Golany, B., & Thore, S. (1997). The economic and social performance of nations: Efficiency and returns to scale. *Socio-Economic Planning Sciences*, 31(3), 191–204.
- Hashimoto, A., & Ishikawa, H. (1993). Using DEA to evaluate the state of society as measured by multiple social indicators. *Socio-Economic Planning Sciences*, 27, 257–268.
- Hashimoto, A., & Kodama, M. (1997). Has livability of Japan gotten better for 1956–1990? A DEA approach. *Social Indicators Research*, 40, 359–373.
- Hervé, M. (2015). RVAideMemoire: Diverse basic statistical and graphical functions, R package version 0.9-52. <http://CRAN.R-project.org/package=RVAideMemoire>.
- INSEE. (2010). *Stiglitz report: The French national statistical agenda*. Paris: INSEE.
- Karabell, Z. (2014). *The leading indicators: A short history of the numbers that rule our world*. Simon and Schuster.
- Kerényi, Á. (2011). The better life index of the organization for economic co-operation and development. *Public Finance Quarterly*, 56(4), 518–538.
- Kutznets, S. (1934). National income 1929–1932. *Letter from the Acting Secretary of Commerce transmitting in Response to Senate Resolution*, 22, 7.
- Lind, N. C. (2014). Better life index. *Encyclopedia of quality of life and well-being research*. Springer, Netherlands, pp. 381–382.
- Lovell, C. A. K., & Pastor, J. T. (1999). Radial DEA models without inputs or without outputs. *European Journal of Operational Research*, 118, 46–51.
- 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.
- Luptacik, M. (2000). Data envelopment analysis as a tool for measurement of eco-efficiency. In Dockner, E. J., Hartl, R. F., Luptacik, M., & Sorger, G. R. *Optimization, dynamics, and economic analysis*. Physica-Verlag.
- 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.
- Murias, P., Martínez, F., & Miguel, C. (2006). An economic wellbeing index for the Spanish provinces: A data envelopment analysis approach. *Social Indicators Research*, 77(3), 395–417.
- OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. Paris: OECD Publishing.
- OECD. (2009). *Measuring Capital: OECD Manual 2009—Second edition*. OECD Publishing, Paris. doi:10.1787/9789264068476-en.
- OECD. (2011). *Compendium of Oecd well-being indicators*. Paris: OECD Publishing.
- OECD. (2013a). *How's life?*. Paris: OECD Publishing.
- OECD. (2013b). Net capital stock. In *National accounts at a glance 2013*, OECD Publishing, Paris. doi:10.1787/na_glance-2013-25-en.
- OECD. (2015a). *Labour force statistics*. Paris: OECD Publishing.
- OECD. (2015b). *Productivity archives*. Paris: OECD Publishing.
- ONS. (2004). *Public sector productivity: Health*. UK: Office of National Statistics.
- ONS. (2006). *Public sector productivity: Health*. UK: Office of National Statistics.
- ONS. (2007). *Productivity handbook*. London: Office for National Statistics, Palgrave Macmillan.
- ONS. (2008a). Capital inputs in public sector productivity: Methods, issues and data. In *UK Centre for the Measurement of Government Activity*. Office of National Statistics, UK.
- ONS. (2008b). *Public sector productivity: Health care*. UK: Office of National Statistics.
- ONS. (2011). *Measuring what matters, National Statistician's reflections on the national debate on measuring national well-being*. Office of National Statistics, UK.
- Osberg, L. (1985). The measurement on economic well-being. In D. Laidler (ed.) *Approaches to economic well-being*, volume 36, MacDonald Commission, University of Toronto Press, Toronto. <http://www.csls.ca/iwb/macdonald.pdf>.

- Osberg, L., & Sharpe, A. (1998). *An index of economic well-being for Canada*, Research Report, Applied Research Branch, Human Resources Development Canada.
- Pastor, J. T. (1996). Translation invariance in data envelopment analysis: A generalization. *Annals of Operations Research*, 66, 93–102.
- Ramanathan, R. (2006). Evaluating the comparative performance of countries of the Middle East and North Africa: A DEA application. *Socio-Economic Planning Sciences*, 40(2), 156–167.
- Schreyer, P. (2004). Capital stocks, capital services and multi-factor productivity measures. *OECD Economic Studies*, 37, 163–184.
- Segre, E., Rondinella, T., & Mascherin, M. (2011). Well-being in Italian Regions. Measures, civil society consultation and evidence. *Social Indicators Research*, 102(1), 47–69.
- Thompson, R. G., Singleton, F. D., Thrall, R. M., & Smith, B. A. (1986). Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces*, 16, 35–49.
- Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130, 498–509.
- Tone, K., & Tsutsui, M. (2006). *An efficiency measure of goods and bads in DEA and its application to US electric utilities*. Presented at Asia Pacific Productivity Conference 2006, Korea.
- Ueda, T., & Hoshiai, Y. (1997). Application of principal component analysis for parsimonious summarization of DEA inputs and/or outputs. *Journal of the Operations Research Society of Japan*, 40(4), 466–478.
- Wackernagel, M., & Rees, W. (1998). *Our ecological footprint: Reducing human impact on the earth* (No. 9). New Society Publishers.
- World Bank. (2011). *The changing wealth of nations: Measuring sustainable development in the new millennium*. Washington, DC: World Bank Publications.
- World Bank. (2015). *World development indicators*. Washington, DC: World Bank Publications.
- World Economic Forum. (2013). *The global competitiveness report 2013–2014*. Geneva: WEF.
- Yap, G. L. C., Ismail, W. R., & Isa, Z. (2013). An alternative approach to reduce dimensionality in data envelopment analysis. *Journal of Modern Applied Statistical Methods*, 12, 1–17.
- Zaim, O., Fare, R., & Grosskopf, S. (2001). An economic approach to achievement and improvement indexes. *Social Indicators Research*, 56, 91–118.
- Zhu, J. (2001). Multidimensional quality-of-life measure with an application to Fortune's best cities. *Socio-Economic Planning Sciences*, 35, 263–284.