Using VEA to assess effectiveness in the development of human capabilities

Panagiotis Ravanos[1] and Giannis Karagiannis[2]

[1]PhD candidate, Department of Economics, University of Macedonia, Thessaloniki, Greece; email: ravanos@uom.edu.gr
[2]Professor, Department of Economics, University of Macedonia, Thessaloniki, Greece; email: karagian@uom.edu.gr; tel. ++30 2310 891 759

ABSTRACT: In this paper we use Value Efficiency Analysis (VEA) to assess effectiveness by means of a model decision-making unit (DMU) which is used as an overall benchmark. The latter is unanimously agreed that is 'doing the right things' and defines a range of operating input/output mixes that is preferred by experts. We then decompose the effectiveness estimates into an efficiency component capturing the extent of 'doing things right' and a mix component capturing the relative distance of the assessed units' operating mix from the preferred range of operating mixes. The latter is residually estimated as the ratio of Data Envelopment Analysis (DEA) and VEA efficiency scores. We use the proposed approach to examine the effectiveness of countries in utilizing their economic prosperity to further develop their citizens' social prosperity or human capabilities using UN data for the year 2015. The empirical results provide useful insights that help detecting countries which may need a shift of focus to relatively neglected aspects, in order to further enhance the capabilities of their citizens.

KEYWORDS: DEA, VEA, efficiency, effectiveness, human capabilities, social efficiency, social prosperity.

AKNOWLEDGMENT: The first author acknowledges financial support by the Hellenic Foundation for Research and Innovation (HFRI) and the Greek General Secretariat for Research and Technology, under the HFRI PhD Fellowship grant (GA. no. *698*).

February 2020

Using VEA to assess effectiveness in the development of human capabilities

1. Introduction

Efficiency and effectiveness are two distinct but related notions of performance evaluation. Efficiency measures the extent to which a decision-making unit (DMU) 'does things the right way', namely whether it produces the maximum possible outputs from given inputs or uses the minimum possible inputs to produce a given amount of outputs. Effectiveness, on the other hand, measures the ability of DMUs to state and achieve desired goals (Cooper *et al.*, 2007, p. 66) i.e. it examines the question of doing the right things. The goals or ''right things'' reflect behavioral or organizational objectives of DMUs or their supervising agency, which can be either monetary or non-monetary. The former refers to economic objectives, such as cost minimization or revenue maximization, the extent of which can be assessed as long as price data are available while the latter refers to managerial preferences about the production process itself as well as targets to be achieved by the constituent DMUs (see e.g. Asmild *et al.,* 2007).

There are four different approaches in the literature to assess effectiveness. The first of them uses a two-stage process (see Førsund (2017) and the references therein) where at the first stage the efficiency of DMUs is assessed by focusing on the process of converting inputs to outputs. Effectiveness, assessed at the second stage, reflects the ability of DMUs to convert outputs to outcomes[1]. Conventional DEA models are used in both stages and the behavioral objectives are expressed through the selection of outcomes. Recently, Førsund (2017) and Hanson (2018) provide innovative refinements of this approach, especially suitable for application related to public sector. In the second, additional constraints reflecting behavioral objectives (see Asmild *et al*., 2007) are introduced into conventional DEA model. If these are related to economic objectives, such as cost minimization or revenue maximization, then effectiveness coincides with the notion of overall efficiency. If the behavioral objectives reflect managerial goals then we need restrictions on the input and/or output multipliers to incorporate them into the conventional DEA model.[2] The resulting model "evaluate(s) both the technical inefficiency that arises from not fully exploiting production possibilities and the inefficiency due either to lack of fulfillment of managerial goals or to the departure from the specified value system of the inputs and outputs'' (Cooper *et al.,* 2011, p. 101). In the third (see Prieto and Zofio, 2001), effectiveness is estimated by means of pure output DEA models. Here the goals of DMUs are considered as given and we concentrate in estimating the extent to which they are achieved regardless of the amount of resources that might be needed to provide them. This follows the idea of the Koopmans' 'helmsman' that attempts to steer all the outputs towards their maximum levels without considering the inputs used (see Lovell *et al.,* 1995). In the fourth approach, effectiveness is related to the distance of DMUs from target points on the existing DEA efficiency frontier (see Golany *et al.,* 1993). Such targets may minimize the distance of DMUs from the DEA frontier, or maximize the outputs of a DMU under a fixed resource allocation.

In this paper we propose an alternative way to incorporate behavioral objectives into conventional DEA in order to assess effectiveness. This is based on Value Efficiency Analysis (VEA), where the behavior objectives reflect the preferences of a Decision Maker or supervising agency, which provides the necessary information regarding the right things to do by simply selecting a model DMU, instead of having to select weight restrictions by means of absolute or relative bounds. According to Korhonen *et al.* (2001), this is an easier method to reflect preferences for both the Decision Maker, who is more keen on picking a "model" DMU rather than engaging to the task of selecting weight restriction bounds, which is a more technical issue. The "model" DMU reflects the most-preferred solution (MPS) from the Decision Maker's point of view and is then used as a global benchmark that determines a range of preferred input and output mixes which comply with her view of 'doing the right things' and provide the base for estimating effectiveness. Efficiency is estimated by means of the conventional DEA model and the two are related by a mix component. The latter serves as a measure of the closeness of the actual input/output mix of DMUs to the most-preferred input/output mix and can aid analysts and Decision Makers identify DMUs with effective operating mixes (which can serve as models) and DMUs which need a mix restructuring in order to comply with managerial preferences, social norms or supervising agency directives.

We use the proposed approach to provide estimates of countries' ability to efficiently and effectively utilize their economic prosperity to enrich the lives of their citizens using 2015 UNDP data. We rely on Amartya Sen's capability approach that views humans as the ultimate ends of the process of economic prosperity and development itself as an expansion of their capabilities, in contrast with the Human Capital approach which views humans as the primary means of economic development. Our empirical models operationalize the differential treatment of income on the capability approach as a means to a number of important ends, rather than an end in itself (Anand and Sen, 2000; Klugman *et al.,* 2011). More specifically, we follow the DEA social efficiency literature pioneered by Despotis (2005a,b) (see also Mariano and Rebelatto, 2014) and use income as an input proxying economic prosperity, with life expectancy, mean and expected years of schooling as the outputs proxying social prosperity. Effectiveness is expressed through an increased as well as a balanced provision of health and education and the empirical results help classify countries into groups displaying high and balanced social prosperity provision (Leaders), countries with a balanced mix but relatively lower achievements, which could use their economic prosperity more efficiently (mix efficient), countries with high but unbalanced provision of health and education (Efficient), which could benefit moving towards a more balanced social prosperity bundle and finally Laggard countries with both low and unbalanced achievements. Such results can prove useful to policymakers in the national level in order to reshape national policies as well as to international organizations to better allocate development or international aid funds.

The rest of the paper is organized as follows: In the next section, we introduce VEA and explain how it can be used to estimate effectiveness. The empirical application is presented in the third section, while concluding remarks follow in the last section.

2. Effectiveness assessment with VEA

VEA, developed by Halme *et al*. (1999), is a performance evaluation method that takes into account Decision Maker preferences about managerial goals by means of a linear value function (i.e., an indifference curve) that become tangent to the DEA efficient frontier at the point of the most-preferred solution (MPS).[3] This point reflects the Decision Maker's choice of a virtual or real DEA-efficient DMU as a model DMU. Then, the VEA frontier is constructed by extending towards the axes the hyperplanes of the DEA efficient facets intercepting at the MPS. In Figure 1(a), by selecting for example DMU B as the MPS, the two efficient facets AB and BC are extended towards the axes, creating the VEA frontier (the blue kinked line). The range of preferred mixes is given between rays OA and OC. All DMUs producing within the preferred range receive a VEA score that is equal to their respective DEA score whereas DMUs producing outside of the preferred range are penalized by receiving VEA scores less than their corresponding DEA scores.

Choosing the model DMU is a crucial step in VEA, as the selected MPS affects the preferred range of input and output mixes and consequently, the resulting value efficiency scores. [4] Although Decision Makers are more inclined to simply choosing a DMU from the set of DEA-efficient ones (see Korhonen *et al.,* 2001 and also Halme *et al.,* 2014 for an empirical application) such as DMU B in Figure 1(a), they also have the freedom to select an MPS that is DEA-inefficient (Korhonen *et al.,* 2002) or propose instead an artificially constructed MPS that may or may not be efficient. In the latter case, the selected DMUs are first projected on to the DEA efficient frontier and then their peers are instead used as the MPS. In Figure 1(a), consider for example DMUs H and G, which are DEA-inefficient and K which is an artificial DMU. H and K are (for ease of presentation) both projected on point B of the DEA efficient frontier. Then, their use as MPS implies instead the use of DMU B, their peer, as the MPS. In a similar fashion, the use of G, which is projected on the efficient facet BC, as the MPS, implies the joint use of DMUs B and C as the MPS. Note that projecting an artificial DMU such as K on the DEA frontier requires solving a super-efficiency DEA model (see Andersen and Petersen, 1993).

A variable-returns-to-scale formulation of the VEA model, in its multiplier form is given as (Halme and Korhonen, 2000):

$$
\begin{aligned}
&\min_{v_i^o, u_j^o, u^o} \varphi_{VEA}^o = \sum_{i=1}^{R} v_i^o x_i^o - u^o && \\
&s.t. \quad -\sum_{j=1}^{s} u_j^o\, y_j^k + \sum_{i=1}^{m} v_i^o\, x_i^k - u^o \geq 0 && k \neq MPS \\
&\quad -\sum_{j=1}^{s} u_j^o\, y_j^k + \sum_{i=1}^{m} v_i^o\, x_i^k - u^o = 0 && k = MPS \qquad (1) \\
&\quad \sum_{j=1}^{s} u_j^o\, y_j^o = 1 && \\
&\quad v_i^o \geq 0 && i = 1 \dots m \\
&\quad u_j^o \geq 0 && j = 1 \dots s \\
&\quad u^o free &&
\end{aligned}
$$

where x and y refer to input and output quantities, φ_{VEA} to the (inverse) of the VEA efficiency score, v and u are parameters to be estimated, k is used to index DMUs $(k = 1, \dots, o, \dots, K)$, i is used to index inputs $(i = 1 \dots m)$, and j is used to index outputs $(j = 1 \dots s)$. The above formulation is only slightly different from the conventional DEA model: the restriction corresponding to the MPS is turned from inequality. This affects the optimal values of the input/output multipliers and essentially determines a range of preferred input/output mixes such as those between the rays OA and OB in Figure 1(a)

In this paper we use VEA to estimate effectiveness and compare it to efficiency which is estimated by means of DEA, in the sence that former reflects the behavioral objectives of a Decision Maker or supervising agency, which provides the necessary information regarding the “right things” by means of a model DMU, that determines the MPS and the range of preferred input and output mixes. Then, the distance of a DMU from the VEA frontier is used to measure effectiveness while its distance from the DEA frontier is used to measure efficiency. Consider for example DMU F in Figure 1(a) where $\frac{OF}{OF''}$ measures the extent to which DMU F “does the right things” while $\frac{OF}{OF'}$ measures the extent to which the DMU F does “things the right way”. From that we see that effectiveness and efficiency are related to each other as follows:

$$\underbrace{\frac{1}{\varphi_{VEA}^{F}}}_{effectiveness} = \underbrace{\frac{1}{\varphi_{DEA}^{F}}}_{efficiency} \times \underbrace{\frac{OF'}{OF''}}_{\substack{mix\\component}} \qquad (2)$$

The second term in (2), i.e. the mix component, reflects the extent to which the DMU produces out of the given range of preferred input/output mixes and it is given by the ratio of the effectiveness to the efficiency score, taking values within the [0,1] range.[5,6] When a DMU operates within the preferred mix range, effectiveness and efficiency scores coincide (for example DMU G Figure 1(a)) and the mix component equals unity. This in general indicates that the particular DMU operates in a manner that is in line with the behavioral objectives set out by the manager or the supervising agency but has different implications when it occurs for efficient and inefficient DMUs. Inefficient DMUs with a mix component equal to one are on the "right operating path" and their ineffectiveness is caused only by inefficient utilization of inputs to produce outputs (i.e. inefficiency) while efficient DMUs producing with the preferred mix range are classified as effective and receive a score of one in all three scores. Such DMUs can serve as examples to follow for the rest of the group or models for the creation of new DMUs in the future. On the other hand, when production takes place outside of the preferred mix range (see DMU F in Figure 1(a)) effectiveness is lower than efficiency and the mix component is lower than unity. This indicates that a DMU has diverged from the "right things" norm or mandate and there is a need to change its operating mix, while if the DMU is also inefficient, additional actions are needed to eliminate technical inefficiencies.

The "right thighs" norm or mandate can include directions set out by the management authorities of a corporation to its branches (e.g. in the case of a bank branch), regulations set out by the government agencies supervising a country's sector (e.g. in the case of financial sector regulations set out by Capital Market Commissions) or the international organizations supporting and monitoring a nation's actions (e.g. in the case of a country being part of the European Monetary System, NATO or the UN). This broad definition highlights the generality of our approach and the fact that it can be applied in various real-world cases.

We can now compare the VEA formulation of effectiveness to those of the second approach referred to in the Introduction, namely that of imposing behavioral (e.g. economic or managerial) objectives. The use of economic objectives is depicted in Figure 1(b) and that of managerial objectives by means of weight restrictions in Figure 1(c). In Figure 1(b) the straight blue line refers to a known output price ratio (i.e., iso-revenue line), which defines a single optimal output mix along the ray OB. Effectiveness, which in this case coincides with the notion of overall (revenue) efficiency, of DMU F is given by the ratio $\frac{OF}{OF''}$ while (technical) efficiency is given by the ratio $\frac{OF}{OF'}$. In this case, the mix component, which is given by the ratio $\frac{OF'}{OF''}$, coincides with allocative efficiency. In Figure 1(c) we depict different cases of weight restrictions that are used to reflect "the right things to do". The straight green and red lines tangent to the DEA frontier in points B and C correspond respectively to a common (across DMUs) and an equal weight scheme. Both define a single optimal output mix, although the common-weight scheme reflected into lines AB or BC would define a range of preferred input/output mix ratios. The broken yellow line corresponds to a form of relative weight bounds, which is similar to VEA frontier in Figure 1(a), and both define a preferred range of mix ratios.[7] Taking DMU F as an example, in Figure 1(b) efficiency is defined by the ratio $\frac{OF}{OF^{I}}$. Effectiveness is defined as the ratio $\frac{OF}{OF^{II}}$ for the common weights scheme, the ratio $\frac{OF}{OF^{III}}$ for the relative weights bounds scheme and the ratio $\frac{OF}{OF^{IV}}$ for the equal weights scheme. This indicates that effectiveness estimates for a DMU may differ when different weighting schemes are used to reflect "the right things to do".[8]

3. Estimating effectiveness in the development of human capabilities

3.1. Methods and Materials

In this section, using 2015 UNDP data, we employ VEA to estimate the extent to which countries utilize their economic prosperity efficiently and effectively to enhance the development of human capabilities for their citizens, i.e. to increase their nations social prosperity. Social prosperity is considered within the capability approach which focuses on the ability of people to live the lives they have reason to value (Sen, 1999, p. 293) and views development as a process that is "removing restrictions" (Fukunda-Parr, 2003) and "enlarging people's choices" (UNDP, 1990). People themselves are the primary ends of the process of development, in addition to them being the principal means of economic production and subsequent economic growth.[9] This differs from the human capital literature that tends to concentrate merely on the role of human beings in augmenting production possibilities, i.e. seeks what people "put into" development. According to Sen (1999, p. 293-295) the latter is a narrower view that tells us nothing about why economic growth is sought in the first place and can fit in the more inclusive perspective of human capabilities, which seeks "what people get from development" (Anand and Sen, 2000). The two approaches are of course related to each other in a causal way; see Ranis *et al.* (2000) and Suri *et al.* (2011).

Economic prosperity, which is usually reflected through a country's per capita income, is viewed by the capability approach as being merely a means to the ends of human development rather than an end in itself (UNDP, 1990; Anand and Sen, 2000; Fukunda-Parr, 2003; Alkire, 2005; Klugman *et al.,* 2011). Nevertheless, Sen (1993) noted that means of development such as income "can indirectly influence the evaluation [of human well-being] through their effects on variables included in [the evaluative space of human well-being]" (p. 33).[10] This brings forth the question of whether countries are able to efficiently utilize their economic prosperity to enhance the social prosperity of their people. The need to provide an answer to such a question is necessary because, despite the high correlation of income levels with longevity and education outcomes, "this tight relation does not obtain" (Sen, 2003, p. 3). There exist many examples of countries with similar levels of income that achieve very different outcomes in terms of basic capabilities such as being healthy and receiving adequate education (see e.g. Sen, 1983, pp. 753-754 and Sen, 2003, pp. 3-4) and for that reason, Sen (1983, p. 754) noted that "not merely is it the case that economic growth is a means rather that an end, it is also the case that for some important ends it is not a very efficient means either".[11]

This line of reasoning was operationalized within the DEA framework by what is now referred to as DEA social efficiency model (Despotis, 2005a,b; Marianno and Rebelatto, 2014) where income is treated as an input and life expectancy and educational attainment as outputs.[12,13] In the context of the DEA social efficiency model, efficiency does not have a strictly production-oriented meaning, i.e. it does not explicitly refer to producing a given set of outputs with the minimum possible inputs. Instead, a "socially efficient" country is one which manages to provide to its citizens high social prosperity levels given its current economic prosperity levels in a relative sence, i.e. given the achievements in social prosperity of other countries with similar economic prosperity levels. This definition adheres to our earlier definition of efficiency as "doing the right things" and does not include any considerations about the relative composition of health and education indicator levels.

Nevertheless, additional information regarding "the right things to do", i.e., norms about the preferred performance of nations should be considered. Examples include institutional constraints laid down by international bodies, positive or negative externalities pointing towards desirable performance and fairness or social conscience (Golany and Thore, 1997). Such an example may be the intention to simultaneously improve the provision of health and education services. Mishra and Nathan (2018) refer to such a balanced realization of performance as the uniformity axiom and state that it is a desirable property for any index of material well-being and capabilities. Also, from a policy perspective, such a balanced prioritization norm between health and education provision, if followed, would aid the country to exploit possible spillover effects existing between the two.[14] We adopt this equal prioritization norm to define effectiveness and to choose our "model" country for the VEA model. Thus, a country should not only manage to provide to its citizens high social prosperity levels given its current economic prosperity, but also to equally prioritize between the provision of health and education services. We selected Norway which is the country that ranked 1st in the 2015 version of the UN HDI and it is a good example of balanced prioritization among health and education provision. Norway is a DEA-inefficient country and thus the units comprising its reference set, namely the efficient countries Australia, Switzerland and Hong-Kong, are instead used as MPS in its place.[15]

Our empirical models use the natural logarithm of GNI per capita in 2011 PPP $ as the single input. We consider three outputs, the first of which is life expectancy at birth, a proxy for health provision. To proxy educational attainment, we follow Lozano and Gutiérrez (2008) and Sayed *et al.* (2015) and use the indicators of mean and expected years of schooling as two separate outputs instead of taking their arithmetic average, to better reflect their different focus on the future expectations of education versus the current realizations of it. Such a choice is also grounded on recent statistical results by Canning *et al.* (2013) who found that combining the two variables into a composite causes a substantial loss of information. The data were normalized using the distance-to-the-leader scheme, as suggested by Herrero *et al.* (2012). This normalization scheme retains the unit invariance property for our models, while also leads to normalized values that necessarily lie within the [0,1] range.[16] Descriptive statistics of the model variables are given in Table 1.

The decomposition of effectiveness estimates into efficiency and the mix component, as in (2), allows the classification of countries into five groups based on their relative ability to provide an increased *as well as* a balanced provision of health and education to their citizens. The "Leaders" group contains those DMUs which score above 99% in both efficiency and the mix component and therefore are considered as effective. The "mix efficient" group contains those DMUs that have a mix component score higher than 99% but an inefficiency score less than 99%. The reverse occurs for "efficient" DMUs which have efficiency scores higher than 99%, but lower mix component scores. The group of "Laggards" consists of the relatively worst performing countries, which achieve efficiency and mix component scores below a certain threshold, which was set at 80%. Thus, we consider inefficiencies below 80% as significant enough to raise alarms to supervising agencies. The remaining DMUs are relatively inefficient with respect to both measures to some extent, but not as severely as the Laggards, i.e. their mix and efficiency scores are both below 99% but at least one is above 80%. These were altogether grouped as "inefficient".

3.2. Empirical Results

Estimates of effectiveness, efficiency and the mix component by group, income class and geographical region are given in Table 2. The arithmetic average and aggregate values of efficiency scores and the mix component for the full sample of 188 countries are 0.927 and 0.906 respectively, indicating that ineffectiveness is caused more by countries' imbalanced prioritization on health and education provision (captured by the mix component) than by having relatively low achievements relative to their economic prosperity levels (inefficiency).[17] This is clearly reflected in the shape of their kernel distributions (see Figure 2) where the mix component distribution has a higher density that the technical efficiency one for lower values of estimates (below 0.85). Also, from the 20 countries that DEA identifies as offering the highest possible social prosperity relative to their economic prosperity (i.e., the efficient countries) only four (namely Australia, Hong Kong, Japan and Switzerland) are identified by VEA as 100% effective.[18] These countries receive a score of one for all three measures.

Eight countries in total (Australia, Hong Kong, Japan and Switzerland which are efficient and Norway, Denmark, Singapore and Sweden which are inefficient) have a mix component equal to one, i.e. their DEA and VEA scores are equal to each other. The inefficient countries with a mix component equal to one manage to offer to their citizens a high balance in the provision of health and education services, but not the "highest possible" amount relative to their economic prosperity, as there are other countries having slightly higher social achievements with the same levels of economic prosperity. The lowest efficiency score among those four countries is 0.972, by Denmark. Nevertheless, the negative skewness for the three measures (see Figure 2) along with the minimum scores indicates the existence of highly inefficient countries, which are in dire need of restructuring actions. Such actions could include increases in health and education expenditures and a better management in order to decrease resource waste. For countries with low mix component scores, budget redistribution could also be an action leading to increased levels of future social prosperity.

In the second panel of Table 2 we present the results for the clustering of countries according to their efficiency and mix component scores. This clustering is also portrayed depicted in Figure 3 and Table 3 presents in more detail the countries in each cluster. Note that in Table 3 we have split the inefficient group into six sub-groups based on an additional threshold set at 95% for technical efficiency and the mix component. The Leader group of countries outperforms on average all other groups in all three measures. The eight Leader countries (see Table 3) include industrialized and well-performing countries in terms of social prosperity such as Canada, Australia and Japan. These are the ones doing 'the right things' and can be considered by the supervising agency (e.g. the UN) as undeniable best performers whose behavior should be copied by other countries in the future. The 17 "mix efficient" countries include 12 European ones, among which we find most of the Nordic countries along with many EU members such as France, Ireland and Belgium. The group is filled with 2 Asian (Korea, Singapore) and 3 Arabic countries (Qatar, Saudi Arabia and United Arab Emirates). These countries, which provide to their citizens the highest possible balance in social prosperity outcomes (i.e. the operate within the preferred range of mixes set out by the "model" country), are also providing very high levels of social prosperity relative to their economic prosperity (their average efficiency score is 0.966) and consequently they display high effectiveness (average 0.963).

On the contrary, the 22 countries of the "efficient" group, which display the "highest possible" social prosperity achievements relative to their economic prosperity, are not concentrated on a specific region but are scattered across the world, including countries as diverse as USA, Chile and Uzbekistan. Furthermore, this group of countries is well performing only with respect to efficiency while having mediocre average performances in terms of balance in the provision of social prosperity (the groups' mix component varies from the low 59.4% in Central African Republic to the well-performing 98.8% in Italy) and consequently, in terms of effectiveness. This group should focus disproportionately more in improving balance in their social prosperity outcomes through gradual changes in their mix. The inefficient countries slightly outperform the efficient countries in terms of the mix component (average estimate 0.903 compared to 0.841). Lastly, the thee Laggard countries, namely Chad, Lesotho and Sierra-Leone (see Table 3) belong to the Sub-Saharan African region. For those countries there seems to be vast room for future improvement, as they are displaying both very low as well as very unbalanced social prosperity achievements relative to their levels of economic prosperity. The average estimates of efficiency (0.765) and mix component (0.773) are the lowest across groups and lead to an average effectiveness score of 59.1%, also the lowest among all groups.

The performance of countries by income class is depicted in the third panel of Table 2 and in Figure 4 we plot effectiveness, efficiency and the mix component against GNI per capita[19]. Effectiveness and the mix component seem to follow an S-curve with respect to income, which is more intense in the mix component case. As income increases average effectiveness and the mix component also increase, with the highest shift in average values being between low and lower-middle income classes. This suggests that a small initial ''push'' in a country's income can spark significant improvements in the provision of health and education services, i.e. that returns to income are increasing in lower income levels. After a certain level of GNI per capita (around 12.000 $ which is close to the threshold between upper-middle and high-income class) the curvature changes and the returns to income in social prosperity become decreasing, indicating that further economic prosperity increases improve only slightly a country's achievements in terms of social prosperity. There is no high-income country that departs more than 8.7% from the preferred range of health and education mixes, as the minimum high-income country mix component estimate is as high as 0.913 (see Table 2). Thus, the virtually zero gain in human development and well-being when income surpasses a certain threshold that is found by Kahneman and Deaton (2010) can be partly explained by that fact that most high-income countries are already displaying both very high as well as highly balanced achievements in terms of social prosperity relative to their (also high) levels of economic prosperity. They are prioritizing relatively even between health and education provision and exploiting heavily the spillovers between simultaneous improvements in both of them.

For efficiency however there does not seem to be a particular pattern. There is no significant difference between low- and the two middle-income classes in terms of efficiency while the high-income countries are on average only 6% more efficient that the low-income countries. Thus, a low level of economic prosperity does not appear to prevent a country from exploiting it to the highest possible extent to provide social prosperity outcomes (i.e. ''doing things right'') as efficiency is realized for a wide range of income levels, from the extremely low-income Central African Republic (GNI per capita 587.474 $ in 2015) to high-income Switzerland (PPP GNI per capita 56,363.958 $ in 2015). On the other hand, equal prioritization and efficient resource use seem to be associated with higher income levels. High-income countries seem to have the know-how about ''doing the right things'' in terms of enhancing social prosperity and further developing the capabilities of their citizens.

The lower panel of Table 2 we present the results by geographical region. From there we can see: *first*, North-Central America and the Caribbean is a rather diverse region whose good average efficiency and mix component performance is mainly supported by the two North American countries, USA and Canada. Most Central American and Caribbean countries are classified as inefficient. *Second*, South American countries on average appears to provide slightly less social prosperity achievements relative to their economic prosperity, compared to their Northern neighbors (average efficiency in South America is 0.842 compared to 0.857 in North America) but on the same time, offering considerably more balanced provision of health and education (the average mix component in South America is 0.879 while that of North America is 0.791). *Third*, South American countries located north or south of the tropical Amazon rainforest seem, to offer a more balanced mix of health and education outcomes compared to countries in the center of South America (e.g. Brazil, Argentina, and Paraguay, see also Table 3). *Fourth*, Europe is the best performing region on average in all three measures, while EU-member countries slightly outperform the non-EU countries in terms of balance in the provision of social prosperity but lag slightly behind in terms of efficiency. *Fifth*, three of the Nordic countries (Norway, Denmark, Sweden) have a mix component equal to one, with the rest of the Nordic group (Iceland and Finland) following suit with scores greater than 0.99. On the opposite, the worst performing European countries in terms of the mix component are three Balkan countries (North Macedonia, Albania, Bosnia and Herzegovina) and two Baltic countries (Estonia and Lithuania). *Sixth*, the Sub-Saharan African (SSA) group of countries is the worst performing region on average in all three measures. The majority of the SSA countries score below 90% in terms of the mix component while half of them (26) score below the threshold of 80%. This poor performance suggests that the SSA region is in urgent need of restructuring actions, such as shifting the allocation of natural resources revenues from recruitment and administrative to health and education expenditures (Raheem *et al.*, 2018) or using the same revenues in order to boost human capital (Oyinlola *et al.*, 2020).

3.3. Robustness checks

We next present a robustness check by (a) considering two alternative MPS choices for the year 2015, namely Australia (the country that ranked second in the 2015 HDI) and an artificial country comprised by the average of the five efficient countries with the highest ranks in the 2015 HDI, namely Australia, Switzerland, Iceland, Hong-Kong and United Kingdom and (b) extending the period under consideration to 2014-2018, using our initial "model" country.

The use of Australia as the "model" country for the year 2015 results into relatively higher effectiveness and mix component scores relative to Norway being the "model" country, mostly because it results to a wider preferred mix range. This is to be expected as Australia was a DEA-efficient country while Norway was not. This is also clear from Figure 1(a) if we consider DMU J as Norway and DMU B as Australia. Using DMU B as the MPS expands both facets AB and BC (blue line) and results into more countries attaining a score of one for the mix component and consequently, to higher effectiveness scores compared to the use of DMU J as the MPS, which expands only facet AB (the red line). The average effectiveness score with Australia as the "model" country was 0.917 (compared to 0.841 with Norway as the "model" country) while the number of "mix efficient" countries was 126 (compared to 17 in the case of Norway). On the other hand, using as "model" country the artificial "average best performing" country was operationalized by using its efficient peers as the MPS, i.e. Australia, Switzerland, Hong-Kong and Japan. This set of peers is the same as that for Norway with the addition of Japan. Adding Japan in the set of MPS brings virtually no changes to effectiveness and mix component estimates with respect to the case of Norway being the "model" country; all average scores as well as the classification of countries remains the same. Referring again to Figure 1(a), let Norway and the artificial average country correspond respectively to DMUs J and I, which both are projected in facet AB and therefore, their use as MPS extends the same facet.

Estimates of effectiveness, technical efficiency and mix components scores for years 2014, 2017 and 2018 using Norway as the "model" country are given in Table 4 and their kernel density distributions are portrayed in Figure 5. Overall, the results across years remain relatively stable. The only notable change occurs in 2017, where the distribution of the mix component became more skewed towards unity compared to other years (see Table 4). This is due to changes in the set of peers for Norway. More specifically, Norway's peers were the same in 2014 and 2015 (namely, Australia, Switzerland and Hong-Kong), while in 2017 changed to Australia, Hong-Kong and Japan. The exclusion of Switzerland, a country with a relatively narrower preferred range of mixes, caused the preferred range of mixes to widen in 2017 relative to other years. Referring to Figure 1(a), this is as the preferred range of mixes temporarily moved from OA-OB to OA-OC. In 2018, Norway's peers are Australia, Switzerland and Japan, highlighting a return to a narrower preferred mix range which is similar to those of years 2014 and 2015, as it can be seen from Table 4.

4. Concluding remarks

In this paper we use VEA to assess effectiveness. VEA captures the extent of ''doing the right things'' through the selection of a "model" DMU which defines a preferred range of input and output mixes. DMUs operating within that range are perceived as effective by managerial authorities, while DMUs operating outside the preferred range should be directed towards mix changes and restructuring. Effectiveness was then decomposed to two measures reflecting the extent of ''doing things the right way'' (efficiency) and producing out of the given range of input and output mixes' (mix component).

The proposed approach could be utilized in many real-world instances where an evaluation of units is sought and managerial preferences need to be taken into account along with efficiency issues. At a micro level, our approach could aid firm owners, CEOs or HR departments to make decisions upon hiring, promoting or allocating personnel based on both their operating efficiency and effectiveness or allocate pay-for-performance funds within a firm. Similarly at a macro level, international organizations such as the International Development Association, Development Banks, the World Bank or the European Union could utilize such a tool with the aim of allocating several kinds of funds towards countries or regions which are either the best (if the funds act as rewards) or the worst performers (if the funds act as aid).[20] A final note regards the sensitivity of the method to the selection of the MPS. As the selected MPS unit defines the preferred range of mixes, it certainly affects the effectiveness scores and their decomposition into efficiency and the mix component.

In the case of the development of human capabilities considered in the empirical application, the use of a "model" country to assess the effectiveness of converting economic to social prosperity allows us to identify the countries providing inappropriate mixes of health and education services and those providing an inappropriate amount of the suggested mixes of health and education services. The former may use policies to correct their deficiencies such as redistributing government expenditures more evenly across health and education in their future balance sheets or directly targeting "priority areas", i.e. the service provision sector which is the relatively most neglected among health and education, through the creation of infrastructure (schools, hospitals) or the implementation of new regulations (e.g. population immunization policies through mandatory vaccination).[21] The latter can benefit from policies that redistribute government expenditures from other uses (e.g. administrative expenditures) to health and education to further improve their achievements, i.e. increase their HD-allocation ratio (see Ranis *et al.*, 2000), as well as from policies that enhance the efficient use of those expenditures, such as better monitoring mechanisms for government officials that handle the relevant contracts.

Table 1: Descriptive statistics of Model's variables.

variable	min	Max	standard deviation	median	average
raw variables					
Life expectancy at birth	48.943 (Swaziland)	84.163 (Hong Kong)	8.297	73.415	71.353
Expected years of schooling	4.872 (South Sudan)	20.433 (Australia)	2.897	13.140	12.983
Mean years of schooling	1.442 (Burkina Faso)	13.370 (Switzerland)	3.097	8.656	8.372
GNI per capita	587.474 (Central African Republic)	129915.601 (Qatar)	19069.312	10415.970	17313.866

Note: The country in parenthesis indicates where the respective minimum/maximum is found.

Table 2: Estimates of effectiveness, efficiency and the mix component.

		effectiveness	efficiency	mix component
World (188 countries)	maximum	1.000	1.000	1.000
	minimum	0.587	0.704	0.594
	average	0.841	0.927	0.906
	aggregate	0.849	0.930	0.913
by cluster				
leaders	maximum	1.000	1.000	1.000
	minimum	0.989	0.996	0.991
	average	0.995	0.999	0.996
	aggregate	0.995	0.999	0.996
mix efficient	maximum	0.989	0.989	1.000
	minimum	0.885	0.893	0.991
	average	0.963	0.966	0.996
	aggregate	0.963	0.967	0.996
efficient	maximum	0.988	1.000	0.988
	minimum	0.594	0.991	0.594
	average	0.840	0.998	0.841
	aggregate	0.850	0.998	0.852
inefficient	maximum	0.968	0.988	0.990
	minimum	0.588	0.704	0.706
	average	0.822	0.910	0.903
	aggregate	0.829	0.911	0.909
laggards	maximum	0.595	0.789	0.790
	minimum	0.587	0.743	0.749
	average	0.591	0.765	0.773
	aggregate	0.591	0.766	0.772
by income class				
high income	maximum	1.000	1.000	1.000
	minimum	0.845	0.857	0.913

	average	0.948	0.963	0.984
	aggregate	0.949	0.964	0.984
upper-middle income	maximum	0.951	1.000	0.980
	minimum	0.588	0.704	0.829
	average	0.855	0.908	0.941
	aggregate	0.855	0.909	0.940
lower-middle income	maximum	0.897	1.000	0.961
	minimum	0.595	0.707	0.779
	average	0.794	0.911	0.871
	aggregate	0.796	0.913	0.872
low income	maximum	0.801	1.000	0.826
	minimum	0.587	0.743	0.594
	average	0.691	0.918	0.755
	aggregate	0.692	0.918	0.754

(Table 2 continued)

		effectiveness	efficiency	mix component
	by geographical region			
North America and the Caribbean	maximum	0.990	1.000	0.994
	minimum	0.728	0.857	0.791
	average	0.884	0.940	0.940
	aggregate	0.886	0.940	0.942
Europe (all)	maximum	1.000	1.000	1.000
	minimum	0.890	0.916	0.913
	average	0.949	0.968	0.980
	aggregate	0.949	0.968	0.980
Europe (EU)	maximum	0.988	1.000	1.000
	minimum	0.890	0.916	0.913
	average	0.948	0.966	0.981
	aggregate	0.948	0.966	0.981
Europe (non-EU)	maximum	1.000	1.000	1.000
	minimum	0.892	0.934	0.945
	average	0.951	0.972	0.978
	aggregate	0.952	0.973	0.979
North Africa	maximum	0.879	0.948	0.949
	minimum	0.833	0.885	0.909
	average	0.855	0.917	0.933
	aggregate	0.856	0.916	0.934
South, East Asia and Oceania	maximum	1.000	1.000	1.000
	minimum	0.720	0.856	0.784
	average	0.861	0.944	0.912
	aggregate	0.868	0.947	0.916
South America	maximum	0.969	1.000	0.969
	minimum	0.785	0.842	0.879
	average	0.874	0.924	0.945

	aggregate	0.874	0.925	0.945
Sub-Saharan Africa	maximum	0.883	1.000	0.977
	minimum	0.587	0.704	0.594
	average	0.700	0.872	0.807
	aggregate	0.702	0.862	0.814
North, West and Central Asia	maximum	0.991	1.000	0.991
	minimum	0.694	0.812	0.789
	average	0.861	0.935	0.921
	aggregate	0.864	0.936	0.923

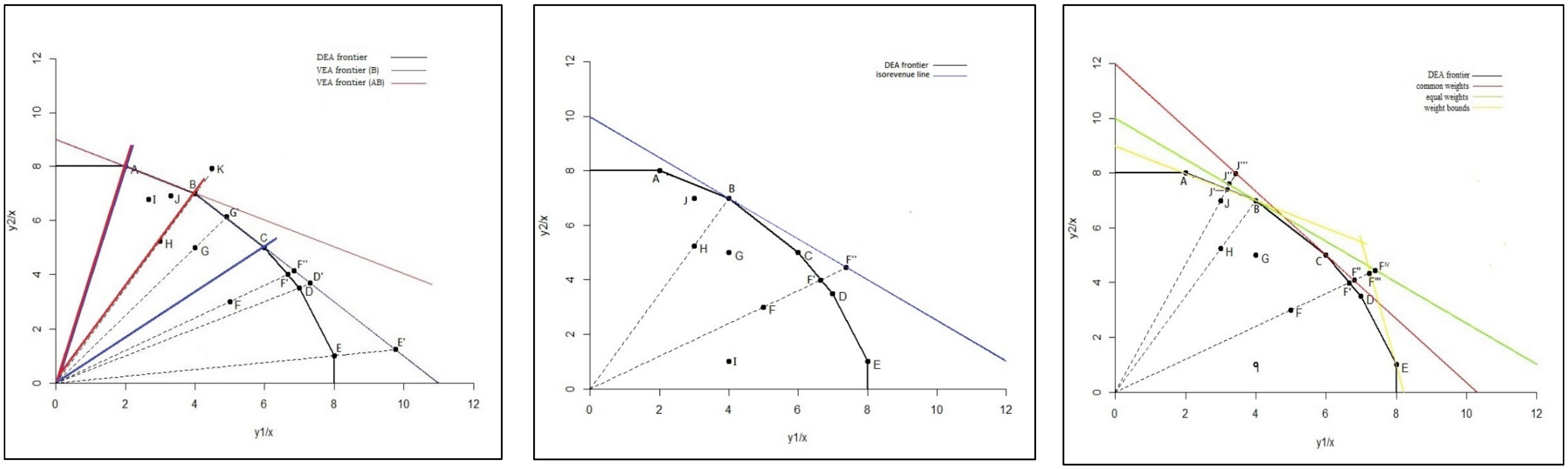
Table 3: Classifying countries by means of efficiency and the mix component.

		mix component			
		[1-0.99]	(0.99-0.95]	(0.95-0.8]	(0.8-0]
efficiency	[1-0.99]	Australia, Canada, Hong Kong (China, SAR), Iceland, Israel, Japan, New Zealand, Switzerland	Chile, Cuba, Germany, Italy, Spain, United Kingdom, United States	Georgia, Kyrgyzstan, Moldova (Republic of), Nepal, Tajikistan, Tonga, Uzbekistan, Vanuatu	Burundi, Central African Republic, Congo (Democratic Republic of), Liberia, Malawi, Solomon Islands, Togo
	(0.99-0.95]	Austria, Belgium, Cyprus, Denmark, Finland, France, Ireland, Korea (Republic of), Liechtenstein, Luxembourg, Netherlands, Norway, Singapore, Sweden	Andorra, Argentina, Costa Rica, Czech Republic, Estonia, Greece, Malta, Portugal, Slovenia	Albania, Bangladesh, Belarus, Bosnia and Herzegovina, Cabo Verde, Dominica, Fiji, Grenada, Honduras, Lebanon, Lithuania, Maldives, Micronesia (Federated States of), Nicaragua, Palau, Palestine (State of), Samoa, Syrian Arab Republic, Ukraine, Viet Nam	Comoros, Ethiopia, Madagascar, Niger, Rwanda
	(0.95-0.8]	Qatar, Saudi Arabia, United Arab Emirates	Antigua and Barbuda, Azerbaijan, Bahamas, Bahrain, Barbados, Brunei Darussalam, Bulgaria, Croatia, Hungary, Iran (Islamic Republic of), Jordan, Kuwait, Latvia, Malaysia, Mauritius, Mexico, Montenegro, North Macedonia, Oman, Panama, Peru, Poland, Romania, Saint Kitts and Nevis, Serbia, Seychelles, Slovakia, Sri Lanka, Suriname, Thailand, Trinidad and Tobago, Turkey, Turkmenistan, Uruguay, Venezuela (Bolivarian Republic of)	Algeria, Armenia, Belize, Bhutan, Bolivia, Brazil, Cambodia, China, Colombia, Congo, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Ghana, Guatemala, Guyana, India, Indonesia, Iraq, Jamaica, Kazakhstan, Kenya, Kiribati, Lao People's Democratic Republic, Libya, Mauritania, Mongolia, Morocco, Myanmar, Namibia, Pakistan, Papua New Guinea, Paraguay, Philippines, Russian Federation, Saint Lucia, Saint Vincent and the Grenadines, Sao Tome and Principe ,South Africa, Sudan, Tanzania (United Republic of), Timor-Leste, Tunisia, Yemen, Zambia	Afghanistan, Benin, Burkina Faso, Eritrea, Gambia, Guinea, Guinea-Bissau, Haiti, Mali, Mozambique, Senegal, South Sudan, Uganda, Zimbabwe
	(0.8-0]	-	Botswana, Equatorial Guinea, Gabon	Angola, Cameroon, Côte d'Ivoire, Nigeria, Swaziland	Chad, Lesotho, Sierra Leone

Table 4: Distribution of effectiveness, technical efficiency and mix component scores, 2014, 2015, 2017 and 2018 with the same MPS choice (Norway).

	2014	2015	2017	2018	2014	2015	2017	2018	2014	2015	2017	2018
	effectiveness				efficiency				mix component			
(zero)	0	0	0	0	0	0	0	0	0	0	0	0
(0,0.1)	0	0	0	0	0	0	0	0	0	0	0	0
[0.1,0.2)	0	0	0	0	0	0	0	0	0	0	0	0
[0.2,0.3)	0	0	0	0	0	0	0	0	0	0	0	0
[0.3,0.4)	0	0	0	0	0	0	0	0	0	0	0	0
[0.4-0.5)	0	0	0	0	0	0	0	0	0	0	0	0
[0.5,0.6)	2	5	0	3	0	0	0	0	0	1	0	0
[0.6,0.7)	16	19	7	23	1	0	0	0	1	2	1	3
[0.7,0.8)	37	37	33	37	11	11	6	6	20	26	5	26
[0.8,0.9)	65	71	64	69	41	40	43	46	42	40	36	43
[0.9,1)	64	52	82	53	117	117	125	124	119	112	134	108
1	4	4	3	4	18	20	15	13	6	7	13	9
maximum	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
minimum	0.596	0.587	0.644	0.571	0.685	0.704	0.704	0.714	0.610	0.594	0.652	0.658
average	0.850	0.841	0.871	0.838	0.924	0.927	0.925	0.928	0.918	0.906	0.941	0.900
aggregate	0.859	0.849	0.880	0.847	0.928	0.930	0.929	0.932	0.925	0.913	0.947	0.909
standard deviation	0.104	0.107	0.086	0.108	0.067	0.066	0.060	0.058	0.078	0.085	0.064	0.086
median	0.875	0.867	0.891	0.867	0.938	0.939	0.934	0.944	0.951	0.943	0.969	0.920

Figure 1: Effectiveness assessment based on different approaches.



Panel (a): Reflecting managerial preferences through VEA

Panel (b): Known prices, effectiveness coincides with overall efficiency

Panel (c): Approximating prices with weight restrictions

Figure 2: Kernel density estimates for the efficiency, effectiveness and mix component, 2015.

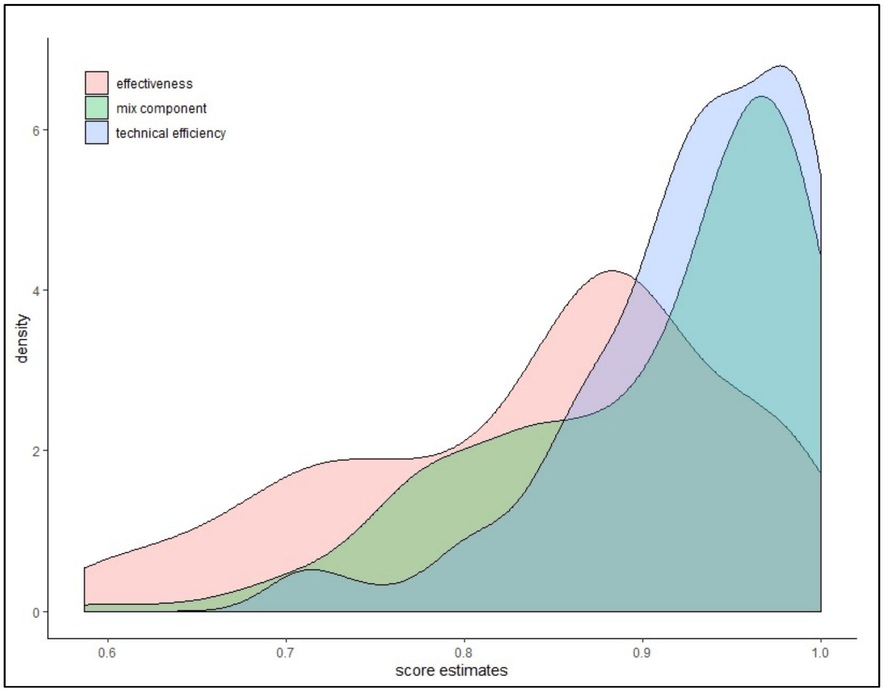


Figure 3: Country clusters by means of efficiency and the mix component.

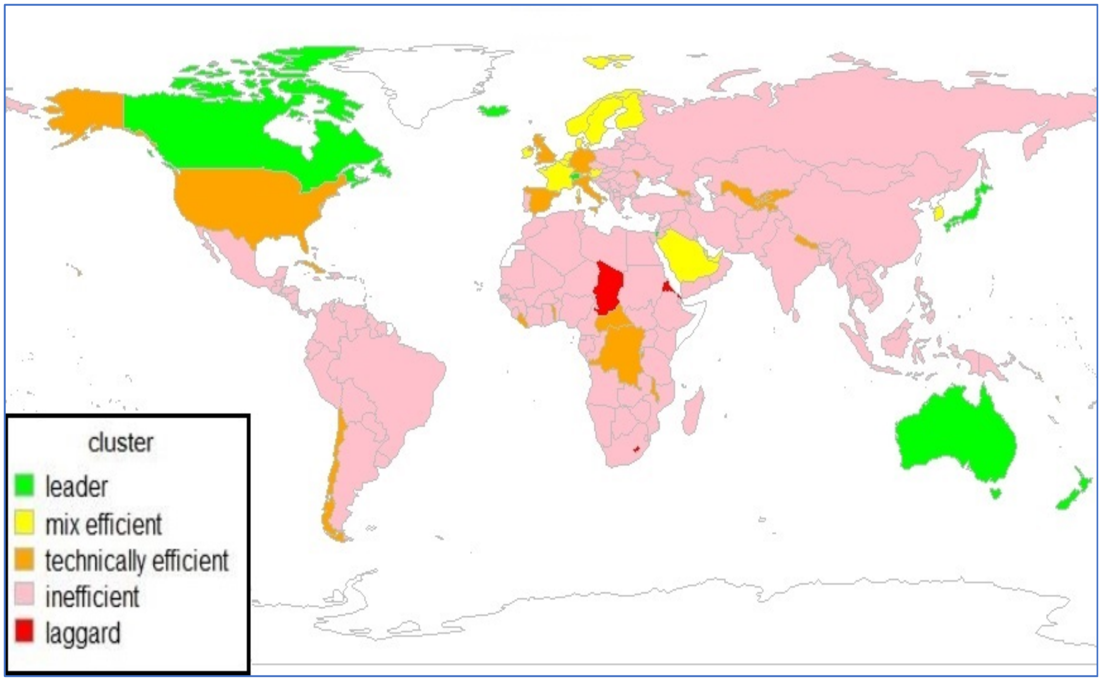


Figure 4: Relation between income and performance.

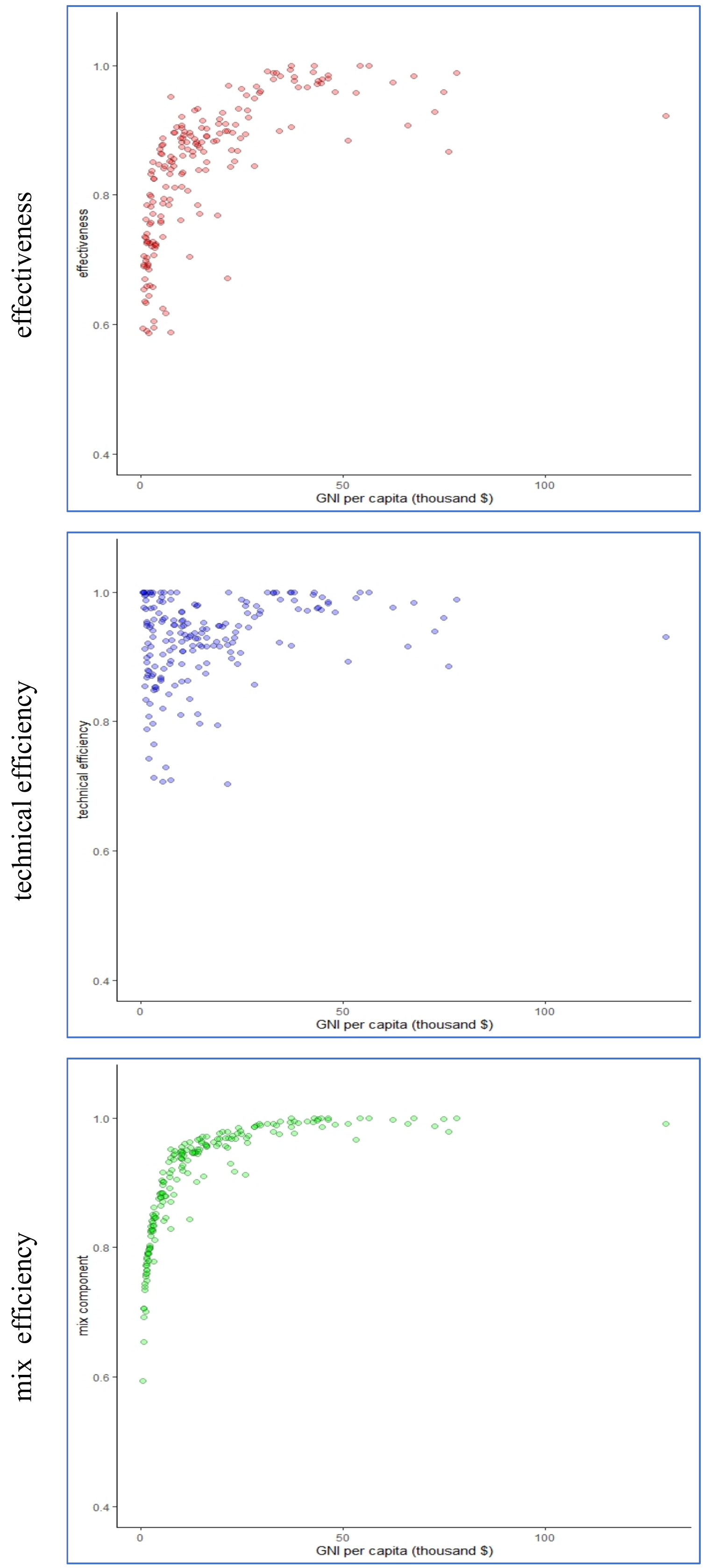
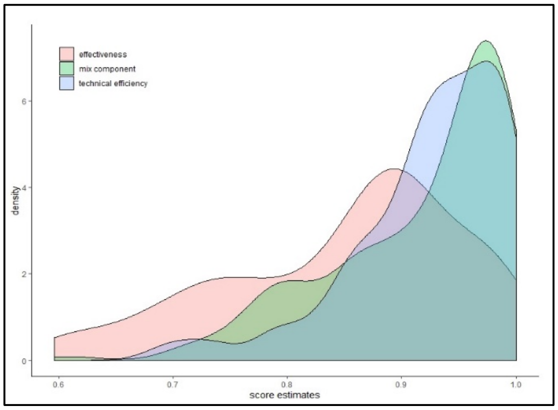
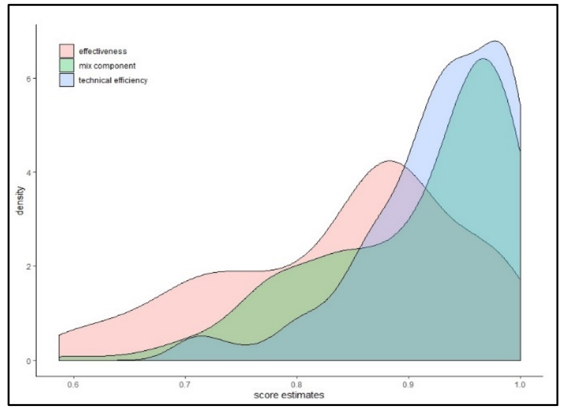


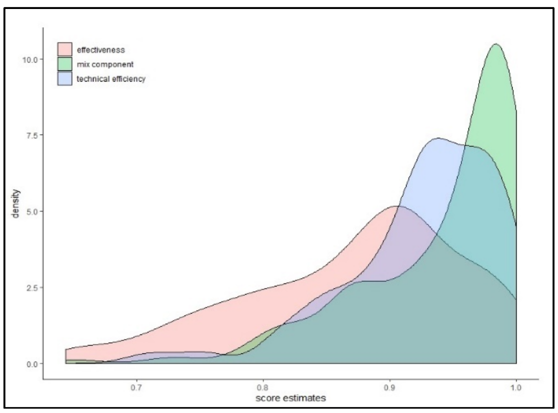
Figure 5 Kernel density estimates for the efficiency, effectiveness and mix component, 2014, 2015, 2017 and 2018 with the same MPS choice (Norway).



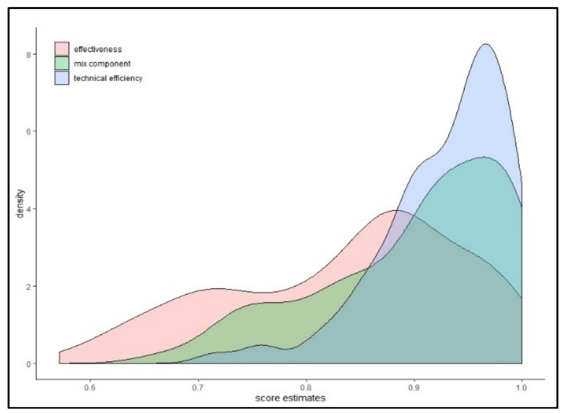
(a) 2014



(b) 2015



(c) 2017



(d) 2018

References

Alkire S (2005). Why the Capability Approach? *Journal of Human Development and Capabilities* 6:115-135.

Anand S, Sen A (2000). Human development and economic sustainability. *World Development* 28:2029-2049.

Andersen P, Petersen NC (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science* 39:1261-1264.

Asmild M, Paradi JC, Reese DN, Tam F (2007). Measuring overall efficiency and effectiveness using DEA. *European Journal of Operational Research* 178:305–321.

Canning D, French D, Moore M (2013). Non-parametric estimation of data dimensionality prior to data compression: the case of the human development index. *Journal of Applied Statistics*, 40:1853-1863.

Cooper WW, Ruiz JL, Sirvent I (2011). Choices and uses of DEA weights. In: WW Cooper, LW Seiford, J Zhu (Eds.), Handbook on data envelopment analysis (pp.93–126) Springer.

Cooper WW, Seiford LM, Tone K (2007). Data envelopment analysis: a comprehensive text with models, applications, references and DEA-Solver software, 2nd ed. Kluwer Academic, Boston.

Despotis DK (2005a). A reassessment of the Human Development Index via Data Envelopment Analysis. *Journal of Operational Research Society* 56:969-980.

Despotis DK (2005b). Measuring human development via data envelopment analysis: the case of Asia and the Pacific. *Omega* 33:385-390.

Färe R, Karagiannis G (2017). The denominator rule for share-weighting aggregation. *European Journal of Operational Research* 260:1175-1180.

Färe R, Zelenyuk V (2003). On aggregate Farrell efficiencies. *European Journal of Operational Research* 146:615-620.

Filippetti A, Peyrache A (2011). The Patterns of Technological Capabilities of Countries: A Dual Approach using Composite Indicators and Data Envelopment Analysis. *World Development* 39:1108–1121.

Førsund FR (2017). Measuring effectiveness of production in the public sector, *Omega* 73:93-103.

Fukuda-Parr S (2003). The Human Development paradigm: Operationalizing Sen's ideas on capabilities. *Feminist Economics* 9:301-317.

Golany B, Phillips FY, Rousseau JJ (1993). Models for improved effectiveness based on DEA efficiency results, *IIE Transactions* 25:2-10.

Golany B, Thore S (1997). Restricted best practice selection in DEA: An overview with a case study evaluating the socio-economic performance of nations. *Annals of Operations Research* 73:117–140.

Halme M, Joro T, Korhonen P, Salo S, Wallenius T (1999). A value efficiency approach to incorporating preference information in data envelopment analysis. *Management Science* 45:103-115.

Halme M, Korhonen P (2000). Restricting weights in value efficiency analysis. *European Journal of Operational Research* 126:175-188.

Halme M, Korhonen P, Eskelinen J (2014). Non-convex value efficiency analysis and its application to bank branch sales evaluation. *Omega* 48:10-18.

Hanson T (2018). Estimating output mix effectiveness: An applied scenario approach for the Armed Forces. *Omega* 83:39-49.

Herrero C, Martínez R, Villar A (2012). A Newer Human Development Index. Journal of *Human Development and Capabilities* 13:247-268.

Joro T, Korhonen P (2015). Extension of data envelopment analysis with preference information: Value efficiency. International Series in Operations Research and Management Science 218, Springer.

Kahneman, D, Deaton A (2010). High income improves evaluation of life but not emotional well-being. Psychological and cognitive Sciences. *Proceedings of National Academy of Sciences* 107:16489–16493.

Klugman J, Rodríguez F, Choi H-J (2011). The HDI 2010: new controversies, old critiques. *Journal of Economic Inequality* 9:249–288.

Korhonen P, Soismaa M, Siljamaki A (2002). On the use of value efficiency analysis and some further developments. *Journal of Productivity Analysis* 17:49-65.

Korhonen P, Tainio R, Wallenius J (2001). Value efficiency analysis of academic research. *European Journal of Operational Research* 130:121-132.

Li SK, Zhao L (2015). The competitiveness and development strategies of provinces in China: a data envelopment analysis approach. *Journal of Productivity Analysis* 44:293-307.

Lind N (2019). A Development of the Human Development Index. *Social Indicators Research* (in press) DOI: 10.1007/s11205-019-02133-9

Lovell CAK, Pastor JT, Turner JA (1995). Measuring Macroeconomic Performance in the OECD: A Comparison of European and Non-European Countries, *European Journal of Operational Research* 87:507-518.

Lozano S, Gutiérrez E (2008). Data envelopment analysis of the human development index *International Journal of Society Systems Science* 1:132-150.

Mariano EB, Rebelatto DAN (2014). Transformation of wealth produced into quality of life: analysis of the social efficiency of nation-states with the DEA's triple index approach *Journal of the Operational Research Society* 65:1664-1681.

Mishra S, Nathan HSK (2018). A MANUSH or HUMANS characterization of the Human Development Index. *Journal of Human Development and Capabilities* 19:398-415.

Oyinlola MA, Adedeji AA, Bolarinwa MO (2020). Exploring the nexus among natural resource rents, human capital and industrial development in the SSA region. *Economic Change and Restructuring* 53:87-111.

Prieto AM, Zofio JL (2001). Evaluating effectiveness in public provision of infrastructure and equipment: the case of Spanish municipalities. *Journal of Productivity Analysis* 15:41-58.

Raheem ID, Isah KO, Adedeji AA (2018). Inclusive growth, human capital development and natural resource rent in SSA. *Economic Change and Restructuring* 51:29–48.

Ranis G, Stewart F and Ramirez A (2000). Economic growth and human development. *World Development* 28: 197–219.

Sayed H, Hamed R, Ramadan MA-G, Hosny S (2015). Using Meta-goal Programming for a New Human Development Indicator with Distinguishable Country Ranks. *Social Indicators Research* 123:1–27.

Sen A (1983). Development: Which way now? *The Economic Journal* 93:745-762

Sen A (1993). Capability and well-being. In: Nussbaum, M., Sen, A. (Eds.), The Quality of Life. Clarendon, Oxford.

Sen A (1999). Development as Freedom. Oxford University Press, Oxford.

Sen A (2003). Development as Capability Expansion. In: Fukuda-Parr S, Shiva Kumar AK (eds.) Readings in Human Development, Concepts, Measures and Policies for a Development Paradigm. Oxford University Press, New York.

Suri T, Boozer MA, Ranis G and Stewart F (2011). Paths to success: The relationship between human development and economic growth. *World Development* 39: 506–522.

UNDP (1990). Human Development Report 1990. New York: Oxford University Press.

UNDP (2015). Human Development Report 2015, New York: Palgrave Macmillan.

UNDP (2016). Human Development Report 2016, New York: Palgrave Macmillan.

UNDP (2018). Human Development Report 2018, New York: Palgrave Macmillan.

UNDP (2019). Human Development Report 2019, New York: Palgrave Macmillan.

Yu MM, Lin ETJ (2008). Efficiency and effectiveness in railway performance using a multi-activity network DEA model. *Omega* 36:1005 – 1017.

Endnotes

[1] For example, in assessing effectiveness in transport industry, inputs usually refer to number of vehicles, fuels and labor, outputs refer to the produced transport capacity (e.g. seat-miles) while outcomes refer to the extent that produced capacity is consumed by customers (e.g. passenger-kilometer and ton-kilometer) (Yu and Lin, 2009). Another example provided by Hanson (2018) is the assessment of military forces effectiveness, where inputs refer to resources such as personnel and equipment, outputs to countable services or goods such as the number of military units and the quality of their training, and outcomes to country-wide valued states and public goods such as peace, sovereignty or freedom.

[2] Different types of weight restrictions may be used, such as absolute or relative bounds on the multiplier weights, resulting in a set of equal, common across DMUs, or DMU-specific input and output multipliers.

[3] A detailed presentation of VEA can be found in Joro and Korhonen (2015).

[4] See Korhonen *et al*. (2002) for more details regarding the several alternatives underlying the choice of the model DMU.

[5] The mix component is similar (but not the same) to Filippetti and Peyrache (2011) compositional index and to the Li and Zhao (2015) dimension mix index, with the main difference being that their non-DEA frontiers result from a set of common (across DMUs) weights which in terms of Figure 1 implies a linear frontier; see Figure 1(c).

[6] Effectiveness scores are never higher than efficiency scores, as the VEA frontier envelops the DEA frontier.

[7] VEA can also lead to a common set of weights. If for example both DMUs A and B were selected as MPS units in panel (a) of Figure 1, the VEA frontier would extend only facet AB towards the axes, thus creating a common set of weights that nevertheless defines again a range of preferred mix ratios. The same would occur if the inefficient DMUs I or J were selected to be the MPS, as for both of them the efficient peers identified by DEA are DMUs A and B.

[8] DMUs may be 'favored' by specific weight restrictions more or less than others, as e.g. DMUs J and F: the former is more (less) favored by the green (red) line of common (equal) weights while the opposite holds for the latter. However, the same holds for effectiveness by means of the VEA model, as some DMUs are favored by

the chosen MPS more or less than others: with DMU B as the MPS in Figure 1(a), DMU E is ineffective while if DMU D is chosen as the MPS the DMU E would be effective instead.

[9] The capability approach is the underpinning of the construction of the Human Development Index (HDI), which concentrates in a set of basic and universally valued capabilities-longevity and education as well as gross national income.

[10] “the income of a person can tell us a good deal about her ability to do things that she has reason to value” (Anand and Sen, 2000, p. 100).

[11] Anand and Sen (2000, p. 101) also referred to outlier countries that are “doing much more to enhance life expectancy than their GNP per capita would suggest”. These outlier countries need to be identified and used as benchmarks for other countries.

[12] DEA is a non-parametric methodology for estimating production frontiers and measuring efficiency. Compared to its parametric counterpart, Stochastic Frontier Analysis (SFA), there are advantages and disadvantages. The main advantage of using DEA is that it does not require any information more than input and output quantities, while SFA requires an explicit specification of a functional form for the production function and an explicit distributional assumption for the inefficiency terms. Also, in DEA all deviations from the frontier are readily attributed to inefficiencies, i.e. it does not incorporate stochastic noise in the data as is done by SFA. The latter is a particularly important advantage when additional restrictions are incorporated in the model (as is the case of present paper), as the extension of the DEA frontier by the extra restrictions (see e.g. Figure 1(a)) is not guaranteed to take place in the presence of stochastic noise.

[13] All previous studies using this model assumed variable returns to scale, in order to reflect the diminishing returns as income increases and used an output orientation to gauge efficiency. Output orientation displays a focus towards increasing the current provision of health and education given the resources currently available. It also reflects the views of Ranis *et al.* (2000) and Suri *et al.* (2011) that improving levels of education and health should have priority or at least move together with direct efforts to enhance growth.

[14] Ranis *et al.* (2000, p. 200) offer an example of such a spillover effect, citing studies that provide evidence that "education, especially female, tends to improve infant survival and nutrition".

[15] In terms of Figure 1, Norway corresponds to DMU J.

[16] This normalization scheme also avoids the process of truncating normalized values to unity, which is criticized by Lind (2019, p, 410) since it "suggests that human development has an upper limit".

[17] According to Färe and Karagiannis (2017) denominator rule, the aggregate values are computed using potential output shares. However, as we have more than one outputs for which there are no market prices, we have to approximate their "market" shares. Here we follow the approximation suggested by Färe and Zelenuyk (2003) that assumes that the value of the total amount of any output is the same as the value of the total amount of any other output. This implies that the aggregation weights are equal to the unweighted average of the shares of the individual countries corresponding to each output, i.e. $\frac{1}{K}\sum_{j=1}^{J}\left(y_j^k / \sum_{k=1}^{K} y_j^k\right)$

[18] The efficient countries are (in alphabetical order) Australia, Burundi, Central African Republic, Cuba, Georgia, Hong Kong, Iceland, Italy, Israel, Japan, Kyrgyzstan, Liberia, Moldova, Nepal, Republic of Congo, Solomon Islands, Switzerland, Tajikistan, United Kingdom and Uzbekistan.

[19] The respective information for the 2015 country clustering by income class was retrieved from the World Bank.

[20] A case of reward-funds is set out by Golany and Thore (1997): the evaluation by the World Bank or some UN agency of loan requests made by developing countries.

[21] Ranis *et al.* (2000) refer to the proportion of government expenditures for sectors related to human development that is attributed to such priority areas as HD priority ratio and argue that the latter is affected positively by the extent of government decentralization.