Unraveling the links between public spending and Sustainable Development Goals: Insights from data envelopment analysis

Jorge Cristóbal[*, a], Michael Ehrenstein[b], Antonio Domínguez-Ramos[a], Ángel Galán-Martín[c,d], Carlos Pozo[e], María Margallo[a], Rubén Aldaco[a], Laureano Jiménez[f], Ángel Irabien[a], Gonzalo Guillén-Gosálbez[g]

[a]Chemical and Biomolecular Engineering Department, University of Cantabria, Avd. De los Castros s/n, 39005 Santander, Spain.

[b]Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, SW7 2AZ London, UK.

[c]Department of Chemical, Environmental and Materials Engineering, Universidad de Jaén, Campus Las Lagunillas s/n, 23071 Jaén, Spain.

[d]Center for Advanced Studies in Earth Sciences, Energy and Environment, Department of Chemical, Environmental and Materials Engineering, Universidad de Jaén, Campus Las Lagunillas s/n, 23071 Jaén, Spain.

[e]Departament d'Enginyeria Química, Agrària i Tecnología Agroalimentària, Universitat de Girona, Edifici Politécnica I, C. Maria Aurèlia Capmany 61, 17003 Girona, Spain.

[f]Departament d'Enginyeria Quimica, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Spain.

[g]Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zürich,Vladimir-Prelog-Weg 1, 8093 Zürich, Switzerland.

*Corresponding author: Email: cristobalj@unican.es; Phone: +34 942200931

Abstract

The global agenda is undoubtedly determined by the planetary success of achieving the sustainable development goals (SDGs). Both public and private institutions show great efforts towards the full integration of the SDGs in their own agendas. Ultimately, national governments are responsible for the effective budget allocation for sustainable development. The lack of open, discussed and widely accepted general guidelines related to how to link national public spending (based on the classification of the functions of government) and the achievement of the SDG is reported in the literature. Thus, the aim of this paper is to propose an initial mapping between them, as well as to assess, through data envelopment analysis (DEA), the national public spending efficiency where government expenditure is consumed (inputs) to produce a certain progress in indicators specific to all 17 SDGs (outputs). On the one hand, results were analyzed for each SDG by income groups, unraveling inefficient spending strategies, thus identifying potential weaknesses that should be overcome before some countries can achieve the same level of progress on SDGs as the best performing countries. On the other hand, it was demonstrated that the income groups which deliver higher average public spending efficiency are low income and high income countries. Countries of these two groups are more often deemed efficient, being displayed alongside the efficient frontiers of the DEA. This situation highlights that low middle-income and upper middle-income countries exhibit the major room for improvement in public spending.

Keywords: sustainable development goals; spending efficiency; public spending; government expenditure; data envelopment analysis; linear programming.

1. Introduction

In 2015, the United Nations (UN) adopted the 2030 Agenda for Sustainable Development, a major milestone towards transforming our world. The Agenda addresses major challenges faced by humanity, including ending poverty and other deprivations, improving health and education, reducing inequality, and spurring economic growth – all while tackling climate change and working to preserve aquatic and terrestrial ecosystems. The Agenda is an ambitious and urgent call for action – agreed on by all 193 UN member states – which was translated into 17 sustainable development goals (SDGs). These goals were, in turn, translated into their corresponding internationally agreed 169 SDG targets, whose progress is tracked by 232 unique indicators (UN, 2018).

With just ten years ahead to achieve the SDGs promised, the question that arises is whether the actions implemented so far are on the motion and scale required to meet the goals by 2030. Shedding light on this question requires tracking progress towards all 17 SDGs both individually and collectively. Hence, different assessment tools and approaches have been proposed to monitor the SDGs progress (Allen et al., 2018; Miola and Schiltz, 2019). Among them, the annual SDG Index and Dashboard (SDGI&D) introduced by Sachs et al. (2016) emerges as the most recognized, standardized, quantitative, transparent, and scalable composite measure (Schmidt-Traub et al., 2017; Nature Sustainability editorial, 2018) monitoring the national progress on the SDGs. Results from the latest SDGI&D (Sachs et al., 2020) show that no single country is on track to achieve all SDGs by 2030, which calls for urgent and more ambitious actions everywhere.

Furthermore, the speed of global progress is not keeping pace with the ambitions of the Agenda, necessitating immediate and accelerated actions by countries and stakeholders at all levels, especially after the COVID-19 pandemic, which is having devastating impacts on specific SDGs and targets (UN, 2020).

With this background, it is clear that further efforts must be made within the so-called Decade of Action (i.e., 2020-2030). Closing the gap between the current progress and the SDGs will only be possible through both public and private investments across all SDGs. According to the United Nations Development Programme (UNDP) (2018), meeting the 2030 Agenda will require unprecedented investments in all sectors, in the range of several trillions[1] of dollars (US$). The estimates from different sources (i.e., public and private at all scales) vary depending on the areas of emphasis and the countries considered, being of particular interest the analysis for developing countries (IMF, 2019). Thus, an estimation from the United Nations Conference on Trade and Development (UNCTAD) (2014) placed the world's total annual SDGs investment needs at roughly $5 to $7 trillion per year only concerning infrastructure (water, agriculture, telecommunications, energy, transport, buildings, industrial, and forestry sectors). Those estimates in developing countries alone range from $3.3 to $4.5 trillion per year (with an annual investment gap that ranges from $1.9 to $3.1 trillion per year). Besides, other provisions such as climate change mitigation, biodiversity conservation, communicable disease control, and investment in research and science, etc., are estimated at several trillion more per year (UNDP, 2018).

[1] For this paper, the short scale is used and thus, trillions and billions are intended as 10^{12} and 10^{9}, respectively.

As mentioned before, all society segments are called to take an active role in implementing the SDGs and filling the financial gap. However, ultimately, national governments are primarily responsible for realizing the transformation. As stated by Kharas and McArthur (2019), public spending estimates[2] will be crucial since it is the form of spending most directly falling under policymakers' purview. Hence, public spending (typically between 15-30 percent of gross domestic product (The Economist Intelligence Unit, 2020)) will be critical for most SDGs in development areas such as health, education, water and sanitation (IMF, 2019), considering private investments as a complement for achieving them. Private investment will be decisive in specific SDG sectors such as power, transport or telecommunications. Even so, public spending policies and intervention within those sectors are essential to foster and support a favorable investment climate.

According to the International Monetary Fund (IMF) (2019), a large portion of that spending is lost due to inefficiencies such as misallocation, low quality of public services, waste of resources, the crowding out of private spending, and corruption (Rayp and Van de Sijpe, 2007), leading to estimations of additional SDG-spending by 2030 of about ten percentage points of Gross Domestic Product (GDP) in low income countries (LICs) and two percentage points of GDP in emerging market economies. For that reason, ensuring spending efficiency is gaining attention within the SDG financing debate. For example, the Czech Republic and Greece have similar public spending (around $160 billion in 2017) and similar population (around 10 million

[2] Note that spending estimates have always to be interpeted with caution since it is difficult to account for sinergies and trade-offs between sectors, as well as the difficulty in considering the balance of public and private responsabilities that differs by area.

inhabitants), but according to the SDGI&D the Czech Republic occupies the 7th position in the rank with a score of 80.7 (out of 100) while Greece occupies the 50th position with a score of 71.4. Hence, a deeper understanding of the links between investments and the attainment of SDGs on a cross-national basis may help identify the main drivers of inefficiencies and incentivize countries to take further actions to ultimately achieve the SDGs.

Measuring spending efficiency is not straightforward and, especially for public spending, it remains a conceptual challenge (EC, 2008). The main objective of this paper is to analyze public spending efficiency in achieving the SDGs. This efficiency is given by the relationship between the progress made in a wide range of often disparate indicators and the associated expenditures. Acknowledging that the analysis proposed is challenging due to various factors such as data gaps and underlying uncertainties, several methodological steps are proposed herein to unravel the relationship between resource spending and sustainable development. Firstly, Initial and Exploratory Data Analysis (I&EDA) is used to identify the most appropriate indicators to be included in the study, as well as to preprocess them for further analysis. Secondly, a mapping criterion is proposed to allocate public spending to the individual SDGs. Finally, data envelopment analysis (DEA) is used to measure the relative national performance of spending towards 17 SDGs across 156 countries.

In the last two decades, many cross-country evaluations of public spending efficiency have been performed, mostly at the sectoral level (e.g., health, education, infrastructure). This was done via a range of approaches and methodologies (e.g., indexes and performance indicators, parametric and non-parametric methods). Afonso et al. (2005) developed composite indicators to measure the public sector efficiency in member states of the Organization for Economic Co-

Operation and Development (OECD), considering administration, education and health, and income distribution, among others. For Europe and OECD countries, many examples (mainly from the IMF) appear for the health and education sectors, mainly using DEA (Jafarov and Gunnarsson, 2008; Verhoeven et al., 2007). Lavado and Domingo (2015) performed an efficiency analysis of public spending on health, education, and social protection in a broad group of Asian countries with varying development levels using DEA. Lavado and Domingo (2015) also included a literature review of previous studies in the health and education sectors specifying the methodologies used: mainly DEA, Free Disposal Hull (FDH), which is a non-parametric method (Afonso and St. Aubyn, 2005), and Stochastic Frontier Analysis (SFA), which is a parametric approach (Sampaio de Sousa and Stosic, 2005). More recently, Halaskova et al. (2018) evaluated public expenditure efficiency in five areas of public services (i.e., general public services, health, education, social protection, and recreation, culture and religion) in relation to two economic indicators for the EU.

To the best of the authors' knowledge, here for the first time the link between public spending and the level of achievement of the SDGs is studied by means of a three-step methodological framework including I&EDA, a mapping between public spending and SDGs, and DEA. This analysis provides insight into the ability of countries belonging to different income groups to meet the SDGs, pinpointing areas for improvement to reduce inefficiencies and ultimately meet the goals at minimum public spending. Furthermore, the main contributions of this paper include (i) putting forward criteria to allocate public spending (input) on the individual SDGs (output), and (ii) identifying inefficient public expenditure by income group.

2. Materials and methods

The general methodological approach is outlined in Fig. 1. In the first step, data on the level of attainment of the SDGs, as well as public spending, is collected and pre-processed for 156 countries. Next, government expenditures to SDG progress are mapped based on similar themes (e.g. by linking expenditure on health to SDG 3 for good health and well-being), assuming that such expenditures contribute to SDG progress. Subsequently, DEA is applied to evaluate the efficiency with which public expenditure is translated into progress towards the achievement of the SDGs. For each SDG, an independent input-oriented DEA model is solved, with government expenditure serving as DEA input, while a number of SDG indicators (between 3 and 17) serve as outputs based on the aforementioned mapping. For the sake of better understanding and readability, the logical order of the methodological steps summarized in Fig. 1 has been altered within this section: First, the employed DEA model is introduced in Section 2.1, followed by a description of the data and sources in Section 2.2. The mapping between expenditures and SDGs is discussed in Section 2.3, before presenting the results of the analysis in Section 3.

[FIGURE 1]

2.1 Data envelopment analysis

DEA is employed to assess the efficiency of countries' budgets in progressing towards the 17 SDGs. Here, efficiency can generally be defined as achieving the most progress for a given level of expenditure or spending the least for a certain level of progress. DEA is a non-parametric linear programming (LP) method used to assess the relative efficiency of a set of alternatives known as decision-making units (DMUs), each of which consumes some inputs while producing a number of outputs. In the present context, public expenditure in different spending categories serves as the input, while outputs are represented by the progress attained in SDG indicators.

Initially proposed in 1978 by Charnes et al. (1978), DEA has become a popular benchmarking method applied in a myriad of areas such as energy (Ewertowska et al., 2016), sustainability (Zhou et al., 2018), waste management (Cristóbal et al., 2016), food supply (Lucas et al., 2020), and human development (Mariano et al., 2015). It is important to highlight that, throughout these last years, an increasing number of studies have combined the use of Life Cycle Assessment (LCA) and DEA methodologies to assess the eco-efficiency of different environmental systems (Vásquez-Ibarra et al., 2020). Furthermore, DEA has been applied in sustainable development to assess the performance of a country (or another entity) in several indicators simultaneously. Santana et al. (2014) applied DEA to assess sustainable development in the BRIC (Brazil, Russia, India and China) countries via economic, environmental, and social efficiency, using gross fixed capital formation, employed population, and R&D expenditure as inputs, and national GDP, CO_2 emissions, and life expectancy as outputs. Bruni et al. (2011) employed DEA to benchmark sustainable development in Italian regions, considering energy consumption, CO_2 emissions, poverty rates, and regional GDP in their analysis. In pursuing investment schemes promoting sustainable development in Chinese provinces, Chen et al. (2017) applied DEA while considering employment, energy consumption, regional GDP, chemical oxygen demand, and CO_2 and SO_2 emissions. Yan et al. (2018) addressed sustainable urban development in Chinese cities through DEA, considering water consumption, constructed land, and fossil energy consumption as inputs, and several social, economic, and environmental indicators as outputs. Pozo et al. (2019) performed a similar analysis for London boroughs covering the three pillars of sustainability through different indicators. The efficiency of seven Brazilian industrial sectors in the context of sustainable development was assessed by Camioto et al. (2014), considering energy consumption, CO_2 emissions, sectoral GDP, employment, and personnel expenses as indicators. Finally, He et al. (2016) evaluated sustainable development over time in a single Chinese province, using annual data over ten years as DMUs, while considering resource consumption, emissions, and economic indicators, whereas Ehrenstein et al. (2020) performed a cross-national analysis of environmental impacts and well-being for 151 countries.

Enlarging the mostly regional scopes of the above works, here DEA is employed to assess a wide range of countries' economic efficiency in the context of sustainable development. In this study, each country is modelled as a DMU consuming government expenditure (inputs) to produce a certain progress in indicators specific to all 17 SDGs (outputs). A country is classified as economically efficient in an SDG if, for the same progress in the corresponding indicators, there is no other country spending less in the relevant budgets. Efficient countries are assigned an efficiency score of 1 and form the so-called "efficient frontier" in the input-output space. In contrast, inefficient countries would get a positive score strictly below 1 and closer to 0 the larger the degree of inefficiency. Thus, in this study, economic efficiency is quantified from a country's distance to the efficient frontier in the "input-direction", since governments have greater control over the input (public expenditures) than over the output (the SDGs achievement) (Baciu and Botezat, 2014), providing a direct measure of how much public spending has not been translated into additional progress in the SDGs (compared to efficient countries). In other words, this input-oriented approach assesses how much money could be saved by following the best practices exhibited by efficient countries. The goal is to ensure no backtrack in achieving the SDGs while alleviating the burden on the countries' economies. In this context, inefficiencies should not provide an opportunity for Governments to reduce their public investment but prompt them to reallocate it to other budgets enhancing their progress towards the SDGs. This is illustrated in Fig. 2 for the simplified example of a single input (expenditure on health) and output (Universal Health Coverage Index). Here, countries A, B, and C are deemed efficient, while D and E are inefficient. As an example, country D could reduce its public spending on health by half, freeing an additional share of the country's budget that could be invested in other categories, thus improving additional progress towards other SDGs.

[FIGURE 2]

Alternatively, an output-oriented analysis would assess the extent to which the output can be increased without modifying the inputs. Hence, it would provide insights about what level of SDGs could be achieved with the given public expenditure. It is important to highlight that both input and output-oriented models will identify the same set of efficient DMUs, however they would most likely lead to different efficiency scores for the inefficient units (Afonso et al., 2005). Acknowledging the significance of the output-oriented approach, the results derived from this additional analysis are briefly discussed in section 3.3.

2.1.1 Models for efficiency assessment

Since its inception in 1978, a variety of DEA models have been introduced in the literature, including the original Charnes-Cooper-Rhodes (CCR) model, the additive model, and the slacks-based measure (SBM). These different approaches may be divided into radial and non-radial methods. The former assesses efficiency based on the proportional reduction of inputs (for the input-oriented case), while the latter does not (Cook and Seiford, 2009). Non-radial models generally have the advantages of higher discriminatory power, as well as not relying on the assumption that inputs (or outputs) can always be reduced (increased) proportionally (Chang et al., 2013). Concerning the returns-to-scale, a variable return-to-scale (VRS) model is selected due to the large set of countries that includes countries operating at very different scales (i.e., showing very different spending profiles, for example, the spending capacity of Germany is far greater than in Ghana). This assumption ensures that the DMUs are benchmarked against others of similar scale. Hence, this study relies on the input-oriented non-radial SBM under VRS, first introduced by Tone (2001).

A general assumption in DEA is that generating more outputs while consuming fewer inputs is a mark of efficiency. However, situations can arise where a DMU produces undesirable outputs, such as CO_2 emissions and groundwater depletion. In this case, a reduction in undesirable outputs should increase the efficiency level. An approach to include both desirable and undesirable outputs in the same SBM DEA model has been put forward by Tone (2004) for the case of a non-oriented DEA, which is extended here to the input-oriented case as follows:

Model M1

$$\text{eff} = \min_{\lambda, s^-, s^g, s^b} 1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{io}}$$

$$\text{s.t.} \quad x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- \quad \forall i$$

$$y_{ro}^g = \sum_{j=1}^{n} y_{rj}^g \lambda_j - s_r^g \quad \forall r = 1 \dots s_1$$

$$y_{ro}^{b} = \sum_{j=1}^{n} y_{rj}^{b}\lambda_j + s_r^{b} \ \ \forall r = 1 \dots s_2$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, s_i^{-}, s_r^{g}, s_r^{b} \geq 0$$

Model M1 is solved for each DMU (i.e., country) to obtain the respective efficiency (eff). Here, set i represents inputs, set r represents outputs that are here split into two groups (i.e., desirable outputs, denoted by superscript g ("good"), and undesirable, denoted by superscript b ("bad")), and set j represents DMUs. Variables λ_j are the weights assigned to DMU j, s_i^- the slacks for input i, s_r^g the slacks for desirable output r, and s_r^b the slacks for undesirable output r. Finally, parameter m represents the cardinality of input set i, x_{ij} represents the inputs i for DMU j, and y_{rj} represents the outputs r for DMU j. Furthermore, the optimal values of the input slacks s_i^{-*} directly represent the magnitude of excess spending.

However, the SBM DEA model M1 cannot handle negative values for inputs and outputs. This shortcoming is relevant in the context of several SDG indicators considered in this work. To overcome this limitation, the base point translation method developed by Tone et al. (2020) for SBM DEA models is employed here. Following this approach, indicators with negative values are scaled by an amount such that all indicator values are strictly greater than zero. As proven by Tone, this approach is consistent with the ordinary SBM model while retaining the properties of unit invariance and monotonicity.

2.2 Dataset development and analysis

2.2.1 Sustainable development goals data

Data on SDG progress for 156 UN member states were obtained from the 2019 SDGI&D (Sachs et al., 2019) in which 85 global indicators were reported. This dataset excludes 37 countries for which at least 20% of the indicators' data was not available (see Table SM1 in the supporting material). The remaining data gaps were filled using average values within income groups based on the classification reported in the 2019 SDGI&D (i.e., HIC, UMIC, LMIC, LIC[3]). Most indicators have been retained in their original format and classified within the individual SDGs according to the 2019 SDGI&D. In contrast, other indicators required reallocation to a different SDG in order to better align with the official SDG indicators from the UN, or transformation in order to conform with the isotonicity requirement in the DEA methodology[4]. Transformation methodologies are applied to ensure that outputs are measured in such a way that "more is better" (Afonso et al., 2005). To this end, there are alternative approaches, e.g., taking complements (i.e., subtracting the quantity from the base measurement), using reciprocals (i.e., the reciprocal of the selected number), or calculating related ratios. As mentioned (see section 2.1.1), certain indicators for which "less is better" were considered key for the assessment and modeled as undesirable outputs.

Moreover, 17 new indicators were added to the existing SDGI&D list to enhance the coverage level of certain SDGs underrepresented in number (e.g., SDG7, SDG10, SDG17). According to the I&EDA performed, from the total list of 102 indicators, eight indicators were excluded from the analysis. The indicator "Government health and education spending" was discarded as an output as it is already modeled as an input on the expenditure side of the analysis. This was

[3] HIC – High income countries; UMIC - Upper middle-income countries; LMIC – Lower middle-income countries; LIC – Low income countries. Classification based on Gross National Income (GNI) per capita (current US$) calculated using the Atlas method with thresholds as of July 2018 (LIC < 996; LMIC between 996 and 3895; UMIC between 3896 and 12055; HIC > 12055).

[4] "an increase in any input should result in some output increase and not a decresase in any output (Bowlin, 1998)"

done under the rationale that this indicator represents a means and not a goal *per se*. The "net open position in foreign exchange to capital", an indicator of sensitivity to market risk (Sugiyarto, 2015), is also omitted because reported values may be unreliable for specific countries (Grolleman et al., 2019). Indicators for energy intensity, CO_2 emissions from fuel consumption and electricity output, particulate matter, and imported CO_2 emissions were removed because they do not conform with the isotonicity requirement. Finally, a refinement criterion was included to exclude indicators with missing data for more than 35% of countries. Thus, two indicators, "food insecurity" and "fish stocks" were excluded for that reason. The whole list of indicators initially considered in this study, including the reference data source, transformation method (if applied), and rejection decision and criteria (if applied), as well as their role in the input-output structure for the DEA, can be found in Table SM2 in the supporting material.

2.2.2 Public spending

National budgets reflect the governments' expenditures and revenues, as well as governments' policy priorities. Budget classification systems are key for policy decision-making (e.g., allocating resources efficiently among sectors) and accountability. The IMF's Government Finance Statistics Manual (IMF, 2001) provides a standard framework for developing a budget classification structure in which the expenditures are typically recorded following an administrative, economic and functional classification. Among the classifications of expenditure according to purpose included in the System of National Accounts (SNA)[5] (UN, 1993), and

[5] There are four classifications: COFOG – Classification of the Functions of Government; COICOP – Classification of Individual Consumption According to Purpose; COPNI – Classification of the Purposes of Non-Profit Institutions Serving Households; and COPP – Classification of the Outlays of Producers According to Purpose.

revised in 1999 (UN, 1999), the classification of the functions of Government (COFOG) is regarded as the appropriate basis to examine the structure of the total government expenditure (TGE). This framework provides a detailed classification based on ten functions[6], or socio-economic objectives, that general government units aim to achieve through various expenditures.

Data on COFOG (usually reported as % of GDP) was compiled as follows:

First, the TGE was obtained for 187 countries from the IMF database (IMF, 2020) for the period 2015-2018. In order to consider the lagged effect of public spending, this study considers the average data for that period. The consideration of time evolution of TGE and SDG achievement is clearly of great interest in the context of sustainable development, and multiple DEA methods have been introduced for the measurement of intertemporal efficiency change, including window analysis (Wang et al., 2013), the Malmquist index (Färe et al., 1994), and dynamic DEA (Tone, 2010). However, as currently there is a lack of sufficient data for both expenditures and SDG progress for multiple periods, the consideration of the time evolution of TGE is beyond the scope of this work.

Next, based on detailed COFOG data as reported by the IMF (for 65 countries) within the period 2015-2018 (IMF, 2020), average values were calculated for the said period. Data for the rest of the countries and COFOG categories were collected from other sources (using the last value available within the period 2010-2018):

[6] 1. General public services; 2. Defence; 3. Public order and safety; 4. Economic affairs; 5. Environmental protection; 6. Housing and community amenities; 7. Health; 8. Recreation, culture and religion; 9. Education; 10. Social protection

- For COFOG 2 (Expenditure on defense), data were obtained from the World Bank database (World Bank, 2020).
- COFOG 4 (Expenditure on economic affairs) was further divided, since specific data were reported for COFOG 4.2 (Expenditure on agriculture, fishing, forestry, and hunting) by the International Food Policy Research Institute (IFPRI, 2020).
- For COFOG 7 (Expenditure on health), data were obtained from the World Health Organization (WHO, 2020).
- For COFOG 9 (Expenditure on education), data were obtained from the World Bank database (World Bank, 2020).
- For COFOG 10 (Expenditure on social protection), data were obtained from the International Labor Organization (ILO, 2020).

Due to the mentioned lack of detailed data on the COFOGs, only specific data on 4.2, 7, 9 and 10 are available, and thus the remaining COFOGs are grouped in a common category named "REST". Further data was retrieved from the database of Government Spending Watch (GSW, 2020), a joint initiative by Development Finance International and Oxfam. Data gaps were covered with the averages of the respective income groups previously mentioned. COFOG data was then transformed from % of GDP to constant $2017 purchasing power parity (PPP) per capita using the GDP (in constant 2017 international $ - PPP) reported by the World Bank. This adjustment considers the cost-of-living differences across countries and the population (reported by the UN) in the last period of the interval (i.e., 2018).

2.3 Allocation criterion

Since the Millennium Development Goals (MDGs) definition, and continuing with the SDGs, there has been a debate about the importance of coordinating them with the countries' budget processes. Budget planning and execution processes are usually driven by political decisions, not necessarily linked to sustainability matters. Thus, the previously mentioned classifications were not designed to allocate budgets to either the MDGs or the SDGs. Although UNDP is considering developing a code structure that incorporates all the SDGs, there is no universal SDG budget classification yet. This would allow the automatic presentation of budget allocations linked to the SDGs. Some efforts are being made by, for example, the Mexican government that is linking its budgetary programs (i.e., spending categories based in groups of goods or supporting services with a common objective) to the SDGs (UNDP and Ministry of Finances and Public Credit, 2017).

In this work, a new allocation criterion is presented linking the public expenditure assigned to the governments' functions with the SGDs based on their similar nature and assuming that all public expenditure influences the SDGs.

Herein, the allocation consists of two steps:

1. First of all, an *ex post* mapping of both COFOG and SDG at a high classification level (i.e., ten division levels for COFOG and 17 goal levels for SDGs) is performed. As shown in Fig. 3, links between the two categories are classified as either direct or indirect. The former (wide lines in the chord diagram) apply when the COFOG level matches the SDG goal perfectly (e.g., COFOG 7 (Health) and SDG 3 (Good health and well-being)). In contrast, indirect links (narrow lines in the chord diagram) model a clear relation but not a perfect match (e.g., COFOG 9 (Education), and SDG 10 (Reduced inequalities)). Otherwise, no link is considered,

even though a further analysis based on a more detailed classification level might reveal additional links.

2. The allocation rule within this exercise assumes that any COFOG linked with an SDG, even if indirectly, will contribute to progress in the said SDG. Thus, the respective COFOGs are used as input for that SDG.

[FIGURE 3]

3. Results and discussion

The results of applying Model M1 to assess the spending efficiency of 156 countries with respect to progress in the 17 SDGs are next presented and discussed. For each SDG, independent input-oriented DEAs are run, where government expenditures serve as DEA inputs based on the mapping introduced in Fig. 3 and a number of indicators of SDG progress (between 3 and 17) serve as outputs. Results highlight which countries exhibit best practices (i.e., are efficient) and which do not (i.e., are inefficient). The latter nations fail to efficiently allocate their public expenditure in the evaluated areas in relation to their transformation to outputs (i.e., SDG indicators).

3.1 Efficiency scores

Of all the 156 evaluated countries, no single one is deemed efficient for all the SDGs. Across all countries, the highest average efficiency is exhibited by the Central African Republic, followed by Burundi, the Democratic Republic of the Congo, and Sweden. Conversely, South Africa, Botswana, Eswatini and Angola show the lowest average efficiency. Fig. SM1 in the supporting material shows a heatmap of individual input-oriented efficiencies for the 156 countries considered in this analysis across all 17 SDGs.

[FIGURE 4]

Herein, an analysis clustering the evaluated countries into their respective income group is performed to shed light on the efficiency patterns. Fig. 4 shows the number of efficient countries segmented by income classification for each SDG (regional trends are shown in Fig. SM2 in the supporting material). Among the 156 countries considered, 28 are classified as LIC, 39 as LMIC, 42 as UMIC, and 47 as HIC (see Table SM3 in the supporting material for the whole list). First of all, results show that the number of efficient countries varies enormously between SDGs, ranging from 130 for SDG 3 to only seven for SDG 7. This is likely related to the number of SDG progress indicators used, which is very high for certain SDGs (e.g., 17 indicators for SDG 3), enabling a larger number of countries to be efficient along some dimensions. Nonetheless, the distribution of efficient countries is not balanced across income categories. HICs and LICs form the efficient frontier for most of the SDGs, accounting for between 53% (for SDGs 3 and 15) and 87% (for SDG 9) of the efficient countries, except for SDG 1 where UMICs have the highest representation. UMICs are less represented within the efficient frontiers, appearing mostly in SDG 14 with 32% of the countries. LMICs have a significant presence in the efficient frontier of SDGs 7 and 4. It is important to highlight that there is at least one efficient country in each income group except for SDG 7, where there is no efficient UMIC.

Important information can also be gathered from the prevalence of countries in the efficient frontiers for the different SDGs. Fig. 5 shows, by income group, the number of times that countries appear as efficient in all SDGs, being clearly LICs and HICs the most prevalent (results for regional aggregation are shown in Fig. SM3 in the supporting material). All evaluated countries have been deemed efficient at least for one SDG, except for one LMIC (i.e., Eswatini) and four UMICs (i.e., Botswana, Colombia, Venezuela, and South Africa). On the other hand, two LICs exhibit efficiency in 14 and 13 SDGs (i.e., Central African Republic and Burundi, respectively), followed by Estonia (i.e., HIC) and the Democratic Republic of Congo (i.e., LIC) deemed efficient for 10 SDGs.

[FIGURE 5]

3.2 Potential public expenditure reduction

DEA quantifies how much the respective inputs (i.e., expenditures) could be reduced if inefficient countries followed the best practices shown by efficient countries. These potential reductions are herein evaluated across SDGs and income groups (see Fig. 6), and are discussed in the ensuing subsections for different groups of SDGs.

[FIGURE 6]

3.2.1 Potential public expenditure reduction for SDGs 1-4

For SDG 1 (end poverty), all income groups show similarly distributed efficiency scores. Results for LICs show that, on average, this set of countries overspend by 56% compared to best practices, being the most efficient group. For LMICs, UMICs and HICs, the average overspending is 58%, 60%, and 64%, respectively. It is important to note that expenditures of COFOG 10 on social protection are the only contributor to this SDG and, therefore, to this efficiency measure. They account for expenditures on services and transfers provided to individual persons and households (i.e., the provision in the form of cash benefits and benefits in kind for sickness and disability, old age, survivors, family and children, unemployment, and housing), and expenditures on services provided on a collective basis (such as grants, loans and subsidies for research, formulation of policies, enforcement of legislation, cash benefits and benefits in kind for victims of disasters). Thus, the average inefficient public expenditure would account for up to around \$0.46 billion per year for LICs, \$8.3 billion per year for LMICs, \$52.6 billion per year for UMICs, and \$121 billion per year for HICs.

As mentioned before, HICs appear as the most inefficient group of countries, being the highest projected reductions for Luxembourg of around \$21,000 per capita, followed by Denmark, Finland and Norway with projected reductions of around \$12,000 per capita. HICs performing so poorly might be due to the difficulty in reaching the poorest people or the higher cost in acting on the poorest people (technical and scale efficiency). Generally, countries may be performing poorly in ensuring that their social spending expenses benefit their poorest citizens more than the wealthy (i.e., money is spent on people already above the poverty thresholds). Meanwhile, the most needed people have no means to access that social expenditure. Besides, the general tendency of increasing requirements on living standards and the quality of the provided services might consume further expenses by wealthy citizens.

It is important to note that people affected by natural-related disasters are also considered within this inefficiency measure although these depend mostly on external factors. Some countries are more exposed to extreme events than others, being LICs and LMICs the most affected, but also some HICs such as the United States of America and Japan.

SDG 2 (zero hunger) presents an average overspend of 41% for LICs, 40% for LMICs, 52% for UMICs, and 35% for HICs. In this case, the input contributing to the efficiency measure is COFOG

4.2 on agriculture, fishing, forestry, and hunting. This spending accounts for services provided on a collective basis in these areas (e.g., compensation, grants, loans or subsidies to farmers in connection with agricultural activities, supervision and regulation of forest operations and issuance of tree-felling licenses, protection, propagation and rationalized exploitation of fish and wildlife stocks).

Inefficiencies might be explained similarly as in SDG 1, by countries failing to ensure that expenses reach the citizens with the highest needs, and helping them overcome undernourishment and malnutrition. This might be the case for Bhutan, a LMIC that presents the highest reduction projected by DEA of around $435 per capita. However, the second and third highest reductions projected are for Switzerland and Iceland, two HICs, with values of around $400 per capita. It is important to recall that this SDG commits to ending hunger, as well as improving nutrition and promoting sustainable agriculture. Thus, inefficiencies might also be due to a lack of policies promoting healthy nutrition that avoids obesity, or sustainable agriculture that manages the efficient application of fertilizers. The average inefficient public expenditure is around $0.2 billion per year for LICs, $1.6 billion per year for LMICs, $1.9 billion per year for UMICs, and $0.5 billion per year for HICs.

As shown in the literature review, special attention is usually given to the health and education-related goals, SDG 3 and SDG 4, respectively. Concerning SDG 3 (good health and well-being), average efficiency scores are very high within the four income categories, being 95% for LICs, 89% for LMICs, 90% for UMICs, and 98% for HICs. SDG 3 aims mainly to reduce mortality, as well as its causes, and targets universal health coverage. The input for SDG 3 is COFOG 7 on health, which includes expenditures on services provided to individual persons (i.e., medical products, appliances and equipment, outpatient services, hospital services, and public health services) and services provided on a collective basis (mainly research and development (R&D)). Even if HICs perform well in SDG 3 (only three out of 47 are deemed inefficient), the two highest reductions projected by DEA results are for HICs, France and Belgium, of around $1,800 and $1,320 per capita, respectively. These are followed by two UMICs (Namibia and South Africa, with reductions of $530 and $440 per capita, respectively) and one LMIC (Eswatini with projected reductions of $400 per capita).

The case of France and Belgium is surprising since they usually appear as first quartile rated countries on health system performance classifications (France being even the first positioned) (Murray et al., 2000). The inclusion of indicators such as subjective well-being, traffic deaths, and suicide mortality might be a possible explanation for the inefficiencies found in these countries. A possible explanation of inefficiencies for UMICs and LMICS might be related with skewed spending towards more affluent areas, providing healthcare for the wealthier segments of society. Depending on the country, this can also be true for the urban-rural dichotomy, with the former benefiting from higher healthcare expenditure even if most of the population lives in the latter regions. The average inefficient public expenditure translates in around $57.5 million per year for LICs, $0.1 billion per year for LMICs, $0.9 billion per year for UMICs, and $2.8 billion per year for HICs.

For SDG 4 (quality education), efficiency scores are on average medium-low, i.e., 50% for LICs, 48% for LMICs, 33% for UMICs, and 26% for HICs. In this case, the only input contributing to education is COFOG 9, which includes expenditures on services provided to individual persons (i.e., pre-primary and primary education, secondary education, post-secondary non-tertiary education, tertiary education, and other education not definable by level, as well as subsidiary services to education) and services provided on a collective basis (mainly R&D). Again, HICs appear as the most inefficient group of countries, with the highest potential spending reductions. Luxembourg once again is found at the top of this list, with reductions of $5,160 per capita, followed by Iceland, the United States of America, and Denmark, with reductions of around $3,500 per capita. The highest potential reduction for UMICs, LMICs and LICs would be for Costa Rica ($1,171 per capita), Bhutan ($653 per capita), and Senegal ($140 per capita), respectively. Efficiencies as a whole decrease with rising income categories, potentially owing to the fact that education spending in countries in higher-income groups strongly outpaces that of their lower-income counterparts. A possible explanation of inefficiencies might be that spending is skewed towards tertiary education, impacting a segment of the population already educated (GSW, 2015) and not captured in the indicators used within this study. The average inefficient public expenditure is around $1.1 billion per year for LICs, $7.3 billion per year for LMICs, $23.2 billion per year for UMICs, and $28 billion per year for HICs.

3.2.2 Potential public expenditure reduction for SDGs 5-17

The remaining SDGs (i.e., SDG 5 – SDG 17) are next analyzed, considering that no specific data on the detailed COFOGs were available. Hence a general category called "REST" (that includes COFOG 1, 2, 3, 4 (except 4.2), 5, 6 and 8) is used (except for SDG 5, where the “REST” category is complemented with COFOG 10 on social protection, and for SDG 10 where the “REST” category is complemented with COFOG 9 on education and COFOG 10 on social protection). It is important to highlight that inefficient expenditures on these SDGs are not reported within this section since they present high uncertainty and might be overestimated (due to possible double-counting). The main reason is that those specific expenditures mapped in Fig. 3 are diluted in the “REST” category and relations are no longer univocal.

In SDG 5 (gender equality), the trend across income groups is similar to SDG 4, with an efficiency score of 58% for LICs, 29% for LMICs, 21% for UMICs, and 21% for HICs. Gender inequality remains a significant challenge in countries across all income groups, and the lack of specific data on, e.g., general labor affairs (COFOG 4.12) means that the “REST” category was used as input, and thus, higher expenditures are most likely the cause for the observed pattern. More detailed data through gender-responsive budgeting might help assessing the efficiency of different measures in closing the gender gaps present in health, education, working environments, politics, and other spheres. Generally, areas of improvement are related to women's access to the market as a labor force and their role in national decision-making processes, with schooling and family planning access better addressed according to the results.

For SDG 6 (clean water and sanitation) all income groups present approximately the same efficiency scores, being around 65% for LICs, 68% for LMICs, 64% for UMICs, and 65% for HICs. This might be because this SDG is closely related to basic needs satisfaction, and therefore a priority for Governments across the globe. HICs have a remarkable prevalence of efficiency (almost 50% of them are deemed efficient), compared to LICs group in which only 24% are efficient. The highest reduction potentials are projected for countries on the Arabian Peninsula (i.e., Qatar, Kuwait, the United Arab Emirates, and Oman), where water scarcity naturally leads to higher supply costs for the least endowed regions. This SDG calls for ensuring universal access to safe and affordable drinking water, sanitation and hygiene, and expenditures in this line seem to be more efficient. Generally, areas of improvement are related to resource depletion and wastewater treatment.

SDG 7, SDG 9 and SDG 11 (affordable and clean energy, industry innovation and infrastructure, and sustainable cities and communities, respectively) are mostly focused on technology, infrastructure and innovation. They follow the same pattern across the different income groups. Notable, higher average efficiencies are found for LICs and HICs with a similar distribution of efficient countries, owing to low investment for LICs (since it is not a top priority) and more progress for HICs, followed by LMICs and UMICs. These latter income groups strive to progress through higher spending, but with inefficiencies due to lagging progress in their actions. Generally, for SDG 7, universal access to electricity seems to be better achieved compared to the expected provision of clean fuels and renewable energy. Concerning SDG 9 on fostering innovation and building resilient infrastructure, countries are focusing on providing universal access to the internet, being R&D the main area of improvement. Finally, for SDG 11 promoting inclusive, safe, resilient and sustainable cities, a clear area of improvement is related to transport systems.

SDG 8 (decent work and economic growth) promotes inclusive and sustainable economic growth, employment and decent work. It presents similar patterns as SDGs 7 and 9 but with higher efficiencies and a higher number of efficient countries across all income categories. Generally, indicators related to decent work conditions and child labor avoidance present better achievement than those related to employment generation.

SDG 10 (reducing economic inequalities), which aims to reduce inequality within and among countries, presents a high average efficiency score for LICs (i.e., 85%, with most countries above

50%), followed by HICs (67%), LMICs (57%) and UMICs (36%). Generally, areas of improvement concern the regulation and monitoring of financial markets and institutions, while indicators related to reducing refugees and asylum seekers are better achieved.

SDG 12 (responsible consumption and production), SDG 13 (climate action), SDG 14 (life below water), and SDG 15 (life on land) are considered as "environmental SDGs". Average values follow similar patterns across income categories, with higher efficiency scores for LICs descending towards HICs, which present the lowest values except for SDG 13 and 15. Thus, the efficiency trend in SDGs 12 and 13 might be explained by rising consumption levels (that lead to higher quantities of wastes and emissions, the undesired outputs) and the fact that expenditures increase with the income category. Generally, LICs also present lower fossil fuel exports and energy-related CO_2 emissions. Regarding SDG 14 and 15, countries generally perform better in indicators related to the protection of important areas and ecosystems, particularly the overexploitation of resources and biodiversity. An explanation might be that countries give higher value to their ecosystems and that high environmental protection can be achieved even with little investment. It is cheaper and more efficient to protect nature and ecosystems than restore them, and effective policies should follow that rationale.

The efficiency levels attained in SDGs 16 (peace, justice and strong institutions) are higher as an average for LICs and HICs (81% for both) compared to LMICs and UMICs (65% and 69%, respectively). This SDG promotes peaceful and inclusive societies, provides access to justice, and builds effective and accountable institutions. Generally, areas of improvement concern the perception of citizens of safety and corruption, the speed of justice, and press freedom.

Finally, SDG 17 (global partnerships) present average efficiencies of 81% for LICs, 55% for LMICs, 52% for UMICs, and 65% for HICs, being herein mostly focused on strengthening domestic resource mobilization.

3.2.3 Total potential public expenditure reductions and contribution to the SDG financing gap

Averaging the public spending efficiency scores across all SDGs, efficiencies of 66% for LICs, 49% for LMICs, 41% for UMICs, and 51% for HICs are obtained (see Fig. 6). According to the data used in this article, the average inefficient public expenditure (calculated using the TGE) would amount to around $3.6 billion per year for LICs, $70 billion per year for LMICs, $155 billion per year for UMICs, and $264 billion per year for HICs. Considering that the mapping between COFOGs and SDGs is not unique, inefficient expenditures of certain COFOGs might be double-counted. Consequently, the calculated total inefficient expenditure might be overestimated and it must be seen as an indication for maximum potential inefficiency. Thus, this study estimates that the aggregated inefficient public money spent per year could be around $102 billion for LICs (considering an aggregated GDP of $1.2 trillion), $2.74 trillion for LMICs (considering an aggregated GDP of $20.6 trillion), $6.5 trillion for UMICs (considering an aggregated GDP of $40.6 trillion), and $12.4 trillion for HICs (considering an aggregated GDP of $58.8 trillion).

For the above-mentioned reason, it is difficult to compare the numbers obtained with estimated financial gap figures proposed in the literature. Furthermore, uncertainties affecting the calculations should also be accounted for. Besides, the system boundaries applied for each study are different. For example, Kharas and McArthur (2019) estimated that the aggregated SDG-related public spending gap by 2025 would be around $1 trillion per year (i.e., $549 billion for LMICs, $223 billion for UMICs, and $150 billion for LICs). Nonetheless, in Kharas and

McArthur (2019), not all public expenditure is considered as contributing to SDG achievement, and the authors considered a rise in GDP for future periods (the period compared differs from the one in this study). Furthermore, the number of countries and their classification (according to income category) are not the same as those used herein.

A fairer comparison of the herein calculated inefficient expenditures can be made at the sectoral level for LICs, since UNCTAD (2014) estimated an annual investment gap by 2030 of $260 billion in food security and agriculture, $140 billion in health infrastructure, and $250 billion in education infrastructure. Thus, the inefficient expenditure estimated with DEA in SDG 2 zero hunger) for LICs would cover $6.3 billion (2% of the gap), $1.6 billion in SDG 3 (good health and well-being, 1% of the gap), and $30.8 billion in SDG 4 (quality education, 12% of the gap).

3.2.4 SDG efficiency vs. accomplishment

Public spending efficiency in achieving the SDGs has to be understood as an enabler to the final aim of the SDG framework that is the total accomplishment of the aspirational and global targets set for 2030. For that reason, in order to fix the context for further discussion, it is important to first analyze the actual level of the SDGs' achievement. This has been measured here using data from the 2019 SDGI&D (Sachs et al., 2019) as a proxy, and the analysis is performed using average values by income groups for each SDG (additional regional trends are shown in Fig. SM4 in the supporting material). Results from Fig. 7 show that, as expected, the higher the income, the better the achievement for SDGs 1 to 9, SDG 11, and SDG 16. On the other hand, for SDGs 12 and 13, this order is reversed, i.e., the lower the income, the higher the achievement, while for SDGs 14 and 15, the achievement level is almost the same. Finally, for

SDG 10, UMICs and LMICs perform worse than LICs, and for SDG 17, HICs perform worse than UMICs. According to these results, within this Decade of Action, LICs would need to increase efforts related to SDGs 1, 7, and 9, while HICs should focus on SDGs 12 and 14. Finally, for UMICs and LMICs, more effort is required in SDGs 9 and 14.

[FIGURE 7]

In view of the actual level of SDGs' achievement, results of public spending efficiency by SDG (see Fig. 8) can help to pinpoint possible weaknesses and opportunities on the public funding strategies within income groups. It is important to acknowledge that each country has different historical background, national circumstances, and priorities, and faces specific challenges to achieve the SDGs, which affects the conclusions drawn.

First of all, income groups should maintain the actual strategy and level of public spending for SDGs categorized with both high achievement and efficiency. As shown in Fig. 8, ten SDGs for HICs present high efficiency and achievement, as well as eight SDGs for LICs, six for UMICs, and five for LMICs. Secondly, those income groups that have SDGs categorized with high achievement and low efficiency could further evaluate their priority and consider the opportunity to reallocate public funding to other SDGs. Thus, HICs present seven SDGs that could be liable to further evaluation, being nine in the case of both UMICs and LMICs, and one for LICs. This reallocated spending should foster other SDGs categorized, if possible, with high efficiency and low achievement, being one SDG for LMICs and seven SDGs for LICs. These are usually not a priority for that income group (e.g., SDGs 9 and 7 for LICs) and most probably, the low degree of achievement is related to low investments.

Finally, those income groups that present SDGs with low achievement and efficiency (i.e., three for UMICs, one for LMICs, and also one for LICs) would require a clear change in the public funding strategy since some are a priority (e.g., SDG 1 in LICs). Structural changes in national governance might be needed to improve the efficiency of transforming public spending focused on achieving certain SDGs. When not properly tackled, lack of efficiency could be seen as a leak of resources and a possible barrier for future foreign aid (i.e., official development assistance - ODA).

[FIGURE 8]

3.3 Output-oriented model results

As mentioned before, the set of efficient/inefficient DMUs shown in section 3.1 will be the same regardless of the choice of an input or output-oriented model. Nevertheless, the analysis of the results changing the model orientation allows evaluating the efficiency scores of the DMUs that differ under VRS. From an output-oriented perspective, efficiency scores show the level of SDGs that could be achieved with the given public spending.

Fig. SM5 in the supporting material shows the efficiency scores for the output-oriented model. The average output efficiency score for LICs equals 0.74 – with the same public spending, the average LIC country is achieving 26% less SDG level than if it were efficient. This average output efficiency score in LICS by SDG varies between 0.23 for SGG 7 (worst efficiency score) and 0.99 for SDGs 3 and 12 (best efficiency scores). It is important to highlight that the derived efficiency score is relative to the best performer DMUs, which might not necessarily meet the final target established in the SDGs for 2030 (no aspirational or total goal scores have been included within the study). Thus, results from the output-oriented analysis have to be interpreted with care, since they do not measure directly the absolute level of SDGs achievement for each income group. In order to translate to the absolute best possible SDG outcome, the 2019 SDGI&D values (Fig. 7) are used as a proxy, as done in section 3.2.4. According to that index, the average LIC achievement of SDGs level is 52%, and according to the results of the output-oriented model herein presented, the average LIC could reach 65% with the same public spending. For SDGs 3 and 7, the average SDG level achieved for LICs according to the 2019 SDGI&D is 43% and 28%, respectively. And the level achievable according to the output-oriented results with the same public spending would be 44% and 50%, respectively.

The same analysis is done for the rest of income groups. The average output efficiency score for HICs, UMICs and LMICs equals 0.85, 0.75, and 0.72, respectively. Thus, public spending by HICs yields an average SDG level 15% lower than the one under efficient conditions, and the worst and best average efficiency scores are 0.5 for SDG 7, and 0.99 for SDGs 1 and 3, respectively. UMICs and LMICs could increase their SDG level by 25% and 28% using the same resources, respectively, and both present a very low average efficiency score for SDG 9 (i.e., 0.16 and 0.14, respectively). Thus, according to the 2019 SDGI&D values, HICs, UMICs, and LMICs present an average SDG achievement level of 76%, 68%, and 62%, respectively, and they could achieve according to the output-oriented model results, with the same public spending, 87%, 85%, and 79%, respectively.

For further information, Fig. SM6 in the supporting material provides a heatmap of individual output-oriented efficiencies for the 156 countries considered in this analysis across all 17 SDGs.

4. Limitations

This study is affected by some limitations. Since the DEA approach employed here only assesses countries' relative spending efficiency, a country which is efficient does not necessarily achieve the associated SDG, or perform sustainably. Additionally, due to the nature of input-oriented DEA, the study assesses how much spending would need to be reduced to follow best practices and become efficient. This means that the countries with lowest overall spending will be deemed efficient, even if they exhibit little actual SDG progress.

The analysis and results are further constrained by the availability of data on SDG progress and governmental spending. Whereas data availability for SDG indicators is good overall (see Table SM1 for the list of countries missing 20% or more indicator data), this is not necessarily the case for expenditures, as many countries do not provide detailed expenditure data for all COFOGs. The issue of data availability further limits the approaches which may be used to analyze SDG spending efficiency: taking a dynamic, period-oriented DEA approach would certainly shed more light on nations' progress over the years. However, data limitations would restrict this analysis to a small number of countries at most, and severely limit the strengths of conclusions drawn. Finally, as has been mentioned previously, the mapping of COFOG expenditures to SDG progress assumes that the respective expenditures contribute fully to SDG progress. While this assumption for all mappings and nations is a simplification, it is nonetheless believed that this approach serves as a reasonable proxy for the actual relationships between spending and SDG progress.

5. Conclusions

This work applied DEA to assess public spending efficiency with regards to achieving the SDGs. The assessment covered 156 countries and results were analyzed for each SDG by income group. Acknowledging that the general aim is too convoluted, results shed light on inherent patterns found among income groups, and unveil inefficient spending strategies that should be further analyzed. This study performs a relative efficiency analysis, so efficient countries should not be wrongly deemed as sustainable. Indeed, the SDGs' final goals are not achieved in any single nation; this situation calls for urgent and more ambitious strategies from all countries, including those deemed efficient. Despite this, inefficiencies are calculated here to identify potential weaknesses that should be overcome before some countries can achieve the same level of progress on SDGs as the best performing countries.

Results revealed that the income groups presenting higher average public spending efficiency are LICs and HICs, particularly in SDGs related with health and reducing economic inequalities. This highlights that the major room for improvement in public spending is on LMICs and UMICs. The high number of inefficient countries, between 92% and 17% depending on the SDG, leads to billions of dollars squandered that could be properly allocated to enhance the progress made towards the SDGs. For sure, this money could be helpful to reduce taxes and/or reduce the financial gap estimated to achieve the SDGs in 2030.

Results also suggest that attracting external investment (both public and private) might be essential for LICs since the money saved by implementation of observed efficient practices in these countries would still be insufficient to cover the estimated SDG financial gap. A more precise analysis of specific SDGs for LICs revealed that reducing public spending inefficiency could free a budget worth between 1% and 12% of the estimated financial gap for areas such as health or education, respectively.

There is still a long way to fully unravel the links between public spending and the SDGs, being the lack of detailed data and methodological shortcomings the main challenges to be addressed. Further research is needed to study the trade-offs and synergies among SDGs, and the standardization of budget allocation and partitioning criteria to evaluate efficiency on SDG progress more accurately. The former topic is now receiving increasing attention (Kroll et al., 2019; Tremblay et al., 2020; Zhao et al., 2021). The latter is being addressed by some countries such as Spain, which along with the 2021 budget, presented a report addressing the alignment with the SDGs.

Overall, this work reinforces the need to open up new research lines, such as explaining why some countries are more efficient than others when it comes to the progress towards SDGs, or predicting which level of SDG progress would be achieved when increasing public spending. These studies should be underpinned by systematic tools and methodologies like the one applied herein.

6. Acknowledgments

One of the authors (J. Cristóbal) acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities through the "Beatriz Galindo" grant (BEAGAL18/00035).

References

Afonso, A., St. Aubyn, M., 2005. Non-parametric approaches to education and health efficiency in OECD countries. Journal of Applied Economics, Vol. VIII, 2: 227-246.

Allen, C., Metternicht, G., Wiedmann, T., 2018. Initial progress in implementing the Sustainable Development Goals (SDGs): A review of evidence from countries. Sustainability Science, 13: 1453–1467.

Baciu, L., Botezat, A., 2014. A comparative analysis of the public spending efficiency of the new EU member states: A DEA approach. Emerging markets finance and trade, 50: 31 – 46.

Bowlin, W.F., 1998. Measuring performance: An introduction to Data Envelopment Analysis (DEA). Journal of Cost Analysis, 15: 3 – 26.

Bruni, M.E., Guerriero, F., Patitucci, V., 2011. Benchmarking sustainable development via data envelopment analysis: an Italian case study. International Journal of Environmental Research, 5(1), 47-58.

Camioto, F.d.C., Mariano, E.B., Rebelatto, D.A.d.N., 2014. Efficiency in Brazil's industrial sectors in terms of energy and sustainable development. Environmental Science & Policy, 37, 50-60.Chang, Y.T., Zhang, N., Danao, D., Zhang, N., 2013. Environmental efficiency analysis of transportation system in China: A non-radial DEA approach. Energy policy, 58, 277-283.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European journal of operational research, 2(6), 429-444.

Chen, L., Wang, Y., Lai, F., Feng, F., 2017. An investment analysis for China's sustainable development based on inverse data envelopment analysis. Journal of Cleaner Production, 142, 1638-1649.

Cook, W.D., Seiford, L.M., 2009. Data envelopment analysis (DEA)–Thirty years on. European journal of operational research, 192(1), 1-17.

Cristóbal, J., Limleamthong, P., Manfredi, S., Guillén-Gosálbez, G., 2016. Methodology for combined use of data envelopment analysis and life cycle assessment applied to food waste management. Journal of Cleaner Production, 135: 158 – 168.

EC – European Commission, 2008. European Economy – The effectiveness and efficiency of public spending. European Communities. Last access December 2020: https://ec.europa.eu/economy_finance/publications/pages/publication11902_en.pdf

Ehrenstein, M., Calvo-Serrano, R., Galán-Martín, Á., Pozo, C., Zurano-Cervelló, P., Guillén-Gosálbez, G., 2020. Operating within planetary boundaries without compromising well-being? A data envelopment analysis approach. Journal of Cleaner Production, p.121833.

Ewertowska, A., Galán-Martín, A., Guillén-Gosálbez, G., Gavaldá, J., Jiménez, L., 2016. Assessment of the environmental efficiency of the electricity mix of the top European economies via data envelopment analysis. Journal of cleaner production, 116, 13-22.

Färe, R., Grosskopf, S., Norris, M., Zhang, Z., 1994. Productivity growth, technical progress, and efficiency change in industrialized countries. The American economic review, 66-83.

Grolleman, D.J., Blache, D., Ben Rahal, A., Godeffroy, J.M., el Radi, A., Vandepeute, A., 2019. Guinea : Financial Sector Stability Review. IMF country report No. 20/42. Available online: https://www.imf.org/en/Publications/CR/Issues/2020/02/11/Guinea-Financial-Sector-Stability-Review-49041

GSW – Government Spending Watch, 2015. Financing the Sustainable Development Goals – lessons from government spending on the MDGs. Research report.

GSW – Government Spending Watch, 2020. Data hub. Last access December 2020: https://www.governmentspendingwatch.org/spending-data

Halaskova, M., Halaskova, R., Prokop, V., 2018. Evaluation of efficiency in selected areas of public services in European Union countries. Sustainability, 10: 4592 – 4609.

He, J., Wan, Y., Feng, L., Ai, J., Wang, Y., 2016. An integrated data envelopment analysis and emergy-based ecological footprint methodology in evaluating sustainable development, a case study of Jiangsu Province, China. Ecological Indicators, 70, 23-34.

IFPRI – International Food Policy Research Institute, 2020. Statistics on public expenditures for economic development. Last access December 2020: https://www.ifpri.org/publication/statistics-public-expenditures-economic-development-speed.

ILO – International Labor Organization, 2020. World Social Protection Data Dashboards. Last access December 2020: https://www.social-protection.org/gimi/gess/WSPDB.action?id=19.

Kroll, C., Warchold, A., Pradhan, P., 2019. Sustainable Development Goals (SDGs): Are we successful in turning trade-offs into synergies? Palgrave Communications, 5(1): 140.

IMF – International Monetary Fund, 2001. Government Finance Statistics Manual (Washington: International Monetary Fund).

IMF – International Monetary Fund, 2019. Fiscal policy and development: human, social, and physical investment for the SDGs. IMF staff discussion note SDN/19/03.

IMF – International Monetary Fund, 2020. Government Finance Statistics (GFS). Last access December 2020: https://data.imf.org/?sk=5804C5E1-0502-4672-BDCD-671BCDC565A9

Jafarov, E., Gunnarsson, V., 2008. Government spending on Health care and Education in Croatia: Efficiency and Reform options. IMF Working paper WP/08/136.

Kampstra, P., 2008. Beanplot: A boxplot alternative for visual comparison of distributions. Journal of statistical software, 28(1), 1-9.

Kharas, H., McArthur, J., 2019. Building the SDG economy: Needs, spending and financing for universal achievement of the Sustainable Development Goals. Global economy and development at Brookings - Working paper 131.

Lavado, R.F., Domingo, G.A., 2015. Public service spending: efficiency and distributional impact – lessons from Asia. ADB Economics working paper series Nº 435.

Lucas, E., Galán-Martín, Á., Pozo, C., Guo, M., Guillén-Gosálbez, G., 2020. Global environmental and nutritional assessment of national food supply patterns: Insights from a data envelopment analysis approach. Science of The Total Environment, 142826.

Mariano, E.B., Sobreiro, V.A., Rebelatto, D.A.d.N., 2015. Human development and data envelopment analysis: A structured literature review. Omega, 54, 33-49.

Miola, A., Schiltz, F., 2019. Measuring sustainable development goals performance: How to monitor policy action in the 2030 Agenda implementation? Ecological Economics, 164: 106373.

Murray, C.J.L., Lauer, J., Tandon, A., Frenk, J., 2000. Overall health system achievement for 191 countries. World Health Organization. Discussion paper series: No. 28.

Nature Sustainability Editorial, 2018. Tracking progress on the SDGs. Nat Sustain 1, 377 (2018).

Pozo, C., Limleamthong, P., Guo, Y., Green, T., Shah, N., Acha, S., Sawas, A., Wu, C., Siegert, M., Guillén-Gosálbez, G., 2019. Temporal sustainability efficiency analysis of urban areas via Data

Envelopment Analysis and the hypervolume indicator: Application to London boroughs. Journal of Cleaner Production, 239, p.117839.

Rayp, G., Van de Sijpe, N., 2007. Measuring and explaining government efficiency in developing countries. The journal of development studies, 43: 360 – 381.

Sachs, J., Schmidt-Traub, G., Kroll, C., Durand-Delacre, D., Teksoz, K., 2016. An SDG Index and Dashboards – Global Report. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).

Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., Fuller, G., 2019. Sustainable Development Report 2019. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).

Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., Fuller, G., Woelm, F., 2020. The Sustainable Development Goals and COVID-19. Sustainable Development Report 2020. Cambridge: Cambridge University Press

Sampaio de Sousa, M., Stosic. B., 2005. Technical Efficiency of the Brazilian Municipalities: Correcting Non-parametric Frontier Measurements for Outliers. Journal of Productivity Analysis, 24 (2): 157–81.

Santana, N.B., Rebelatto, D.A.d.N., Périco, A.E., Mariano, E.B., 2014. Sustainable development in the BRICS countries: an efficiency analysis by data envelopment. International Journal of Sustainable Development & World Ecology, 21(3), 259-272.

Schmidt-Traub, G., Kroll, C., Teksoz, K., Durand-Delacre, D., Sachs, J.D., 2017. National baselines for the Sustainable Development Goals assessed in the SDG Index and Dashboards. Nature Geoscience, 10: 547 – 556.

Sugiyarto, G., 2015. Financial Soundness Indicators for Financial Sector Stability: A Tale of Three Asian Countries. Asian Development Bank (ADB) Reports RPT157598-2.

The Economist Inteligence Unit, 2020. The future of public spending – Why the way we spend is critical to the Sustainable Development Goals.

Tone, K., 2001. A slacks-based measure of efficiency in data envelopment analysis. European journal of operational research, 130(3), 498-509.

Tone, K., 2004. Dealing with undesirable outputs in DEA: A slacks-based measure (SBM) approach. Presentation at NAPW III, Toronto, 44-45.

Tone, K., Tsutsui, M., 2010. Dynamic DEA: A slacks-based measure approach. Omega, 38 (3-4), 145-156.

Tone, K., Chang, T.S., Wu, C.H., 2020. Handling negative data in slacks-based measure data envelopment analysis models. European Journal of Operational Research, 282(3), 926-935.

Tremblay, D., Frotier, F., Boucher, J.F., Riffon, O., Villeneuve, C., 2020. Sustainable development goal interactions: An analysis based on the five pillars of the 2030 agenda. Sustainable Development, 28: 1584 – 1596.

UN – United Nations, 1993. Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations and World

Bank, System of National Accounts 1993. Sales No. E.94.XVII.4. Last access December 2020: https://unstats.un.org/unsd/nationalaccount/sna1993.asp

UN – United Nations, 1999. Department of economic and social affairs statistics division. Classifications of expenditure according to purpose. ST/ESA/STAT/SER.M/84. Sales No. E.00.XVII.6. Last access December 2020: https://unstats.un.org/unsd/publication/SeriesM/SeriesM_84E.pdf

UN – United Nations, 2018. Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agebda fir Systaubabke Development. Last access December 2020: https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%20refinement_Eng.pdf

UN – United Nations, 2020. The Sustainable Development report 2020. Last access December 2020: https://unstats.un.org/sdgs/report/2020/

UNCTAD – United Nations Conference on Trade and Development, 2014. World Investment Report – Investing in the SDGs: An action Plan. Last access December 2020: https://unctad.org/en/PublicationsLibrary/wir2014_en.pdf

UNDP – United Nations Development Programme, 2018. Financing the 2030 Agenda – An introductory guidebook for UNDP Country offices.

UNDP and Ministry of Finance and Public Credit, 2017. Investing for sustainable Develoment: How does Mexico invest in the Sustainable Development Goals? Mexico: Ministry of Finance and Public Credit and UNDP. Last access December 2020:

https://www.transparenciapresupuestaria.gob.mx/work/models/PTP/Presupuesto/Documentos_anteriores/mexico_sdg.pdf

Vázquez-Ibarra, L., Rebolledo-Leiva, R., Angulo-Meza, L., González-Araya, M.C., Iriarte, A., 2020. The joint use of life cycle assessment and data envelopment analysis methodologies for eco-efficiency assessment: A critical review, taxonomy and future research. Science of the Total Environment, 738: 139538.

Verhoeven, M., Gunnarsson, V., Carcillo, S., 2007. Eeducation and Health in G7 Countries: Achievieng better outcomes with less spending. IMF Working paper WP/07/263.

Wang, K., Yu, S., Zhang, W., 2013. China's regional energy and environmental efficiency: A DEA window analysis based dynamic evaluation. Mathematical and Computer Modelling, 58(5-6), 1117-1127.

WHO – World Health Organization, 2020. Global Health Expenditure Database. Last access December 2020: https://apps.who.int/nha/database.

World Bank, 2020. World Bank Open data. Last access December 2020: https://data.worldbank.org/

Yan, Y., Wang, C., Quan, Y., Wu, G., & Zhao, J., 2018. Urban sustainable development efficiency towards the balance between nature and human well-being: Connotation, measurement, and assessment. Journal of Cleaner Production, 178, 67-75.

Zhao, Z., Cai, M., Wang, F., Winkler, J.A., Connor, T., Chung, M.G., Zhang, J., Yang, H., Xu, Z., Tang, Y., Ouyang, Z., Zhang, H., Liu, J., 2021. Synergies and tradeoffs among Sustainable

Development Goals across boundaries in a metacoupled world. Science of The Total Environment, 751: 141749.

Zhou, H., Yang, Y., Chen, Y., Zhu, J., 2018. Data envelopment analysis application in sustainability: The origins, development and future directions. European Journal of Operational Research, 264(1), 1-16.