

Evaluating governmental policies for the sustainable development goals using hierarchical clustering

Jan Anton van Zanten & Maria Putintseva

To cite this article: Jan Anton van Zanten & Maria Putintseva (06 Jan 2025): Evaluating governmental policies for the sustainable development goals using hierarchical clustering, International Journal of Sustainable Development & World Ecology, DOI: [10.1080/13504509.2024.2448669](https://doi.org/10.1080/13504509.2024.2448669)

To link to this article: <https://doi.org/10.1080/13504509.2024.2448669>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 06 Jan 2025.



[Submit your article to this journal](#)



Article views: 1265



[View related articles](#)



[View Crossmark data](#)

Evaluating governmental policies for the sustainable development goals using hierarchical clustering

Jan Anton van Zanten ^{a,b} and Maria Putintseva^c

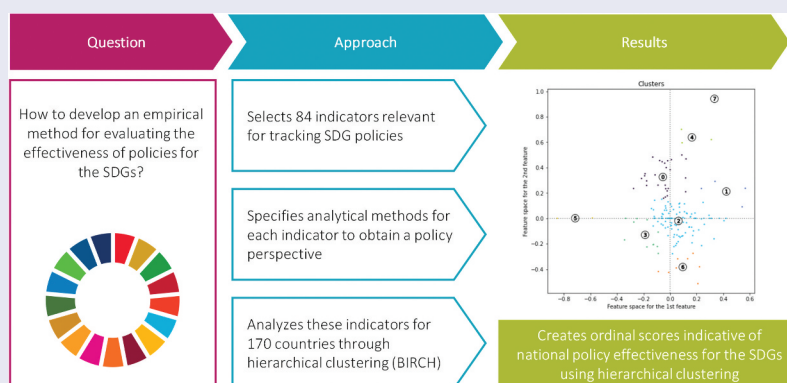
^aRotterdam School of Management, Erasmus University, Rotterdam, The Netherlands; ^bSustainable Investing Center of Expertise, Robeco, Rotterdam, The Netherlands; ^cSI Data Engineering, Robeco Switzerland, Zurich, Switzerland

ABSTRACT

Governments play a crucial role in achieving the Sustainable Development Goals (SDGs), yet many countries fall short in implementing transformative policies to attain them. A significant knowledge gap exists in understanding whether national policies support or hinder sustainable development outcomes. To address this gap, this paper aims to develop an empirical method for evaluating the effectiveness of governmental policies in achieving the SDGs. First, it identifies 84 indicators that directly or indirectly measure the policy effectiveness for each of the 17 SDGs, along with specified evaluation methods for each indicator. Second, it analyzes countries' performance on these indicators using BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) to assign scores that reflect the quality of governmental policies for each SDG. Finally, each country is awarded a total score, indicative of its overall policy-effectiveness for the SDG agenda. The analysis suggests that around one in five countries have policies conducive to SDG progress, while half implement policies that hinder progress, and a third maintain neutral policy. Socio-economic SDGs often receive stronger policy support, whereas goals related to environmental preservation, climate change mitigation, and institutional improvement face significant policy challenges. Policymakers are encouraged to adopt a nexus approach to the SDGs that pursues synergies and avoids trade-offs between the goals; systematically collect globally comparable data to monitor policy-effectiveness for the SDGs; while sustainable investors can use the insights from this study to allocate financing to countries with supportive policies for sustainable development and thereby contribute to closing the SDG financing gap.

HIGHLIGHTS

- We develop an empirical method for evaluating policy-effectiveness for the SDGs
- The method selects and analyzes readily available SDG indicators
- The results reveal potential for improving sustainable development policies globally
- Policymakers can adopt a nexus approach to the SDGs for increased coherence
- Investors can use the method to finance countries with policies that advance SDGs



ARTICLE HISTORY

Received 1 August 2024
Accepted 27 December 2024

KEYWORDS

Sustainable development goals; 2030 agenda; indicators; policy; national development; nexus

1. Introduction

In 2015, the United Nations General Assembly adopted the 2030 Agenda for Sustainable Development. This agenda aims to 'transform our world for the better' (UNGA 2015, p. 40) by urging public, private, and civil

society actors to contribute to sustainable development. The Sustainable Development Goals (SDGs) are at the heart of this agenda. These 17 goals specify universal economic, social, and environmental sustainability objectives that are to be achieved by 2030.

CONTACT Jan Anton van Zanten  vanzanten@rsm.nl; j.van.zanten@robeco.nl  Rotterdam School of Management, Erasmus University, Burgemeester Oudlaan 50, Rotterdam, PA 3062, The Netherlands

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Although the SDGs have been criticized (Swain 2018; HICKEL 2019), they constitute the dominant framework for advancing sustainable development (Sachs 2015; Biermann et al. 2017).

Currently, the world is far off track to achieve the SDGs. While there is variety in the extent to which individual countries are progressing, with those in the Global South typically facing the widest gap, all countries are falling short of meeting most of the goals by 2030 (UN 2023). Moreover, the challenges are mounting. Diverse crises have wiped out years of progress on the SDGs – such as the COVID-19 pandemic, climate-related disasters, and wars and conflicts in diverse regions – causing various SDG targets to show backward progress (Independent Group of Scientists Appointed by the Secretary-General 2023). The sustainable development progress of lower income regions is particularly vulnerable to such crises, revealing global inequalities in pathways for attaining the SDGs (Li et al. 2024).

Despite this bleak outlook, scientists conclude that transformation for the SDGs is possible (Sachs et al. 2019; Independent Group of Scientists Appointed by the Secretary-General 2023). At the same time, they warn that organic changes will be insufficient for reaching the goals. Leaders must remove roadblocks that deteriorate outcomes and implement substantial actions to accelerate progress, requiring scientists to assess the impact of SDG governance policies (Malekpour et al. 2023). Yet, the tracking the indicators of the SDGs is called an *'unprecedented statistical challenge'* (MacFeely 2020), while measuring policies for SDG transformations is a massive data and knowledge gap (Sachs et al. 2019).

Since the SDGs were adopted in 2015 a variety of initiatives emerged to measure the global goals. The High-level Political Forum on Sustainable Development (HLPF) is responsible for overseeing follow-up and review at the global level (UNGA 2015), among others by conducting thematic reviews and through discussing Voluntary National Reviews (VNRs) prepared by UN member states (Beisheim and Fritzsche 2022). Moreover, each year the UN publishes its SDG report, serving as the official report monitoring progress on the 2030 Agenda for Sustainable Development (e.g. UN 2024). Also on an annual basis, the Sustainable Development Solutions Network and the Bertelsmann Stiftung publish their SDG Index, which is an annual ranking of countries on their attainment of the global goals (Sachs et al. 2023). Next to such policy assessments is an array of scientific research on SDG measurement. Scholars dissected the theoretical challenges in measuring the SDGs (Hák et al. 2016; Fukuda-Parr and McNeill 2019; Dang and Serajuddin 2020), explored empirically how countries are performing (Costanza

et al. 2016; Schmidt-Traub et al. 2017; Allen et al. 2018; Guo et al. 2022), and studied how these goals are interlinked, whereby progress on one goal can support, but also deteriorate, advancement on other goals (van Soest et al. 2019; Bennich et al. 2020). Finally, it is noteworthy that there is an extensive body of literature that focuses on tracking progress in individual SDGs, with examples being health (e.g. Asma et al. 2020), energy systems and climate (e.g. Kartal 2022, 2023; Ozcan et al. 2024), or agriculture and biodiversity (e.g. Movilla-Pateiro et al. 2021). These efforts are paramount to understanding how countries are faring on implementing the SDGs and to identify best practices in progressing towards their attainment.

In addition, a stream of literature on policymaking for the SDGs has emerged. Scientists are advising the UN to adopt an integrated approach to the global goals that recognizes their (Independent Group of Scientists Appointed by the Secretary-General 2023). This view is similar to the nexus approach, which argues that because the SDGs are interconnected, policymakers ought to try to take advantage of synergies between the goals, while reducing the risk that policies in support of one goal undermine progress on another (Boas et al. 2016; Bleischwitz et al. 2018; Liu et al. 2018; van Zanten and van Tulder 2021). Moreover, in implementing the SDGs, a multi-level governance approach is needed that promotes cooperation between state and non-state actors at all political levels (global, regional, national, and sub-national) (Hickmann 2021; Weiland et al. 2021). Beyond studying how policymakers can advance all goals simultaneously, researchers investigated policies for individual goals (e.g. Siddiqi et al. 2020) as well as in unique geographic settings (e.g. Xie et al. 2021), or concerning specific policy instruments like environmental taxes (e.g. Kartal 2024).

But in spite of all these efforts, a review of over 3,000 scientific studies on the SDGs revealed that the SDGs have only had a limited impact on governmental policies (Biermann et al. 2022). This finding raises concerns about the potential of the SDG agenda to advance sustainable development (Hickmann et al. 2024). Moreover, it reveals a need for monitoring whether governments are implementing policies that support the SDGs. Along these lines, researchers recently argued that *'more empirical analysis of SDG policies and their implementation is needed to assess the extent, nature and effectiveness of national implementation and integration efforts, and for comparative analysis across countries'* (Hickmann et al. 2024, p. 8).

This paper responds to such calls for research on the effectiveness of implementing the SDGs, by addressing the following research question:

RQ1: How could an empirical method for evaluating the effectiveness of national policies in achieving the SDGs be developed?

To address this question, we adopt a data-driven approach to evaluate to what extent governmental policies support the SDGs. For 170 countries, we scrutinize 84 SDG indicators that are directly or indirectly informative of national policy-effectiveness for the SDGs. Indicators that directly assess policy-effectiveness focus on investments, taxation, or regulation undertaken by countries that can promote or hinder progress on specific SDG targets. Next to such direct measures of policymaking, we analyze how countries perform on SDG indicators over time as an indirect proxy for policy-effectiveness. To illustrate, if a country is progressing/deteriorating on a particular target, like '3.1 ... reduce the global maternal mortality ratio ...', we interpret that as an indication of the country's policies being supportive/reductive of achieving that SDG. The performance of countries on all indicators belonging to an SDG is analyzed using BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies), allowing us to assign scores indicating the quality of governmental policies for each of the SDGs.

Having evaluated governmental policies for each of the SDGs, we also score a country's policy-effectiveness for the entire SDG agenda. The findings reveal that 21% of countries have policies that support SDG progress, while 29% are neutral, and 50% risk undermining the goals. Policies targeting human well-being show the most positive impact, with 39% to 56% of countries supporting progress on poverty eradication (SDG 1), health (SDG 3), education (SDG 4), and gender equality (SDG 5). Conversely, environmental and governance goals face significant challenges; 71% of countries exhibit policies reducing progress on Life on Land (SDG 15), 53% on Peace, Justice, and Strong Institutions (SDG 16), and 52% on Responsible Consumption and Production (SDG 12). Moreover, policy coherence remains a significant issue, as only one-third of SDGs show positive correlations in policy effectiveness, while many exhibit negative correlations. This suggests that governments often implement policies that advance one goal while hindering another.

Overall, the study underscores the need for more coherent and balanced policy frameworks to ensure holistic progress. While acknowledging alternative methods for assessing policy effectiveness, these results aim to advance the empirical monitoring of SDG implementation and stimulate further discussion on achieving sustainable development by 2030.

The next section introduces the materials and methods used in this study. Section 3 presents the results. These are discussed in Section 4, which also talks about the study's limitations and avenues for future research. Conclusions are provided in Section 5.

2. Materials and methods

The ambition of this paper is to develop an illustrative empirical methodology for gauging policy-effectiveness for the SDGs at a national level. The steps taken to develop this illustrative method are visualized in Figure 1 and explained below.

2.1. Data

The first step in this analysis involved identifying data that could illuminate policy-effectiveness for each of the 17 SDGs. Each of the official SDG indicators was evaluated for its policy relevance. Indicators were classified as directly informative if they measure governmental actions (e.g. taxation, regulation, and investment) and indirectly informative if they track national sustainability outcomes that could proxy policy impact via trends or performance averages. Indicators deemed irrelevant to policymaking, such as those overly influenced by external factors (e.g. indicator 1.5.1 measures the 'number of deaths, missing persons and persons affected by disaster per 100,000 people', which is to a great extent determined by factors outside of the control of policy-makers), were excluded.

After defining a refined set of indicators, data availability was assessed across countries, addressing the SDG agenda's well-documented data gaps. Of the 232 official indicators, 72 are not regularly compiled, and 62 lack collection methodologies (MacFeely 2019). Two strategies mitigated these gaps. First, selected indicators were chosen for widespread availability, with an average of 137 out of 170 countries providing data. Some indicators, like Net Official Development Assistance (SDG 17), applied only to specific contexts, such as high-income countries, thus having lower availability by design. Second, efforts ensured broad data coverage across multiple indicators within each SDG, enabling robust evaluations even when some data were missing. Countries were neither rewarded nor penalized for unavailable indicators, resulting in neutral assessments.

Ultimately, 84 indicators were selected, linked to their respective SDG targets, and categorized as directly or indirectly measuring policy-effectiveness, as shown in Table 1. Timeseries data for these indicators were compiled for 170 countries with sufficient data coverage.

2.2. Analyzing SDG indicators

To evaluate policy-effectiveness for the SDGs, we developed different methods for assessing each indicator.

First, for indicators directly reflecting policy effectiveness, the most recent value was analyzed when

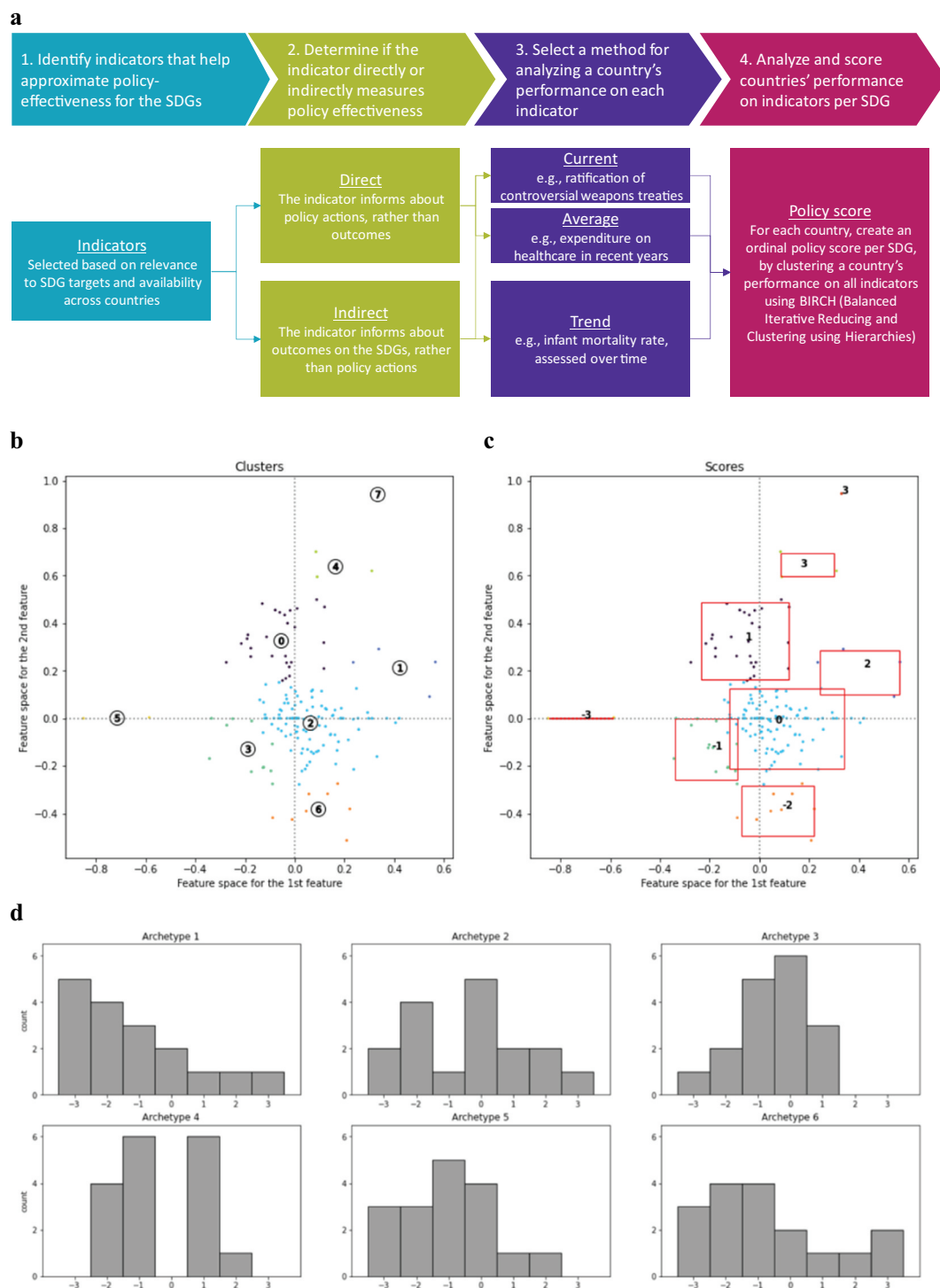


Figure 1. Methodological approach. a, schematic procedure for selecting relevant indicators and assessing them to approximate policy-effectiveness for the SDGs. b, example of clustering countries on their policymaking for SDG 9 – industry, innovation and infrastructure, whereby each cluster is numbers and dots represent plotted countries on features 1 and 2. c, example of assigning scores to each cluster. d, example archetypes for a total SDG policy-effectiveness score of – 3.

historical data added no further insights. For example, the number of international human rights treaties adopted (SDG 16.10) reflects current policymaking and fundamental freedoms, rendering earlier data unnecessary.

Second, for indicators where historical performance was informative, average values over time were evaluated. For instance, average health and education expenditures from 2015 to the present were used to

assess SDGs 3 and 4, acknowledging annual budget fluctuations. This approach provides a stable picture of governmental priorities during the SDG period.

Third, trends over time were analyzed for indicators indirectly tied to policy effectiveness, particularly those measuring national sustainability outcomes. For example, trends in domestic material consumption per capita (SDG 12) were evaluated using regression analysis. A positive trend indicated ineffective policies for

Table 1. Indicators used in this study.

SDG	Indicator	SDG target	Policy assessment	Method	Source
SDG 1 – No Poverty	Poverty Headcount Ratio at National Poverty Lines: % of Population	1.1	Indirect	Trend	World Bank
	Poverty Headcount Ratio at \$2.15 a Day: 2017 PPP: % of Population	1.1	Indirect	Trend	World Bank
	Poverty Headcount Ratio at \$3.65 a Day: 2017 PPP: % of Population	1.1; 1.2	Indirect	Trend	World Bank
	Bank Account Ownership at a Financial Institution or with a Mobile-Money-Service Provider: % of Population Aged 15+	1.4	Indirect	Trend	World Bank
SDG 2 – Zero Hunger	Prevalence of Moderate or Severe Food Insecurity in the Population: % of population	2.1; 2.2	Indirect	Trend	World Bank
	Prevalence of Severe Food Insecurity in the Population: % of population	2.1; 2.2	Indirect	Trend average	World Bank
	Prevalence of Overweight: % of Adults	2.1	Indirect	Trend	World Bank
	Prevalence of Overweight: Weight for Height: % of Children Under 5, Modelled Estimate	2.1	Indirect	Trend	World Bank
SDG 3 – Good Health and Well-being	Births Attended by Skilled Health Staff: % of Total	3.2	Indirect	Trend	World Bank
	Current Health Expenditure: % of GDP	3c	Direct	Average	World Bank
	Hospital Beds: per 1000 People	3.7	Indirect	Trend	World Bank
	Maternal Mortality Ratio: Modelled Estimate: per 100,000 Live Births	3.1	Indirect	Trend	World Bank
	Mortality Rate: Under-5: per 1000 Live Births	3.2	Indirect	Trend	World Bank
	Nurses and Midwives: per 1000 People	3.7	Indirect	Trend	World Bank
	Physicians: per 1000 People	3.7	Indirect	Trend	World Bank
	Cause of Death: by Communicable Diseases and Maternal, Prenatal and Nutrition Conditions: % of Total	3.3	Indirect	Trend	World Bank
	Mortality Caused by Road Traffic Injury: per 100,000 Population	3.6	Indirect	Trend	World Bank
	SDG 4 - Quality Education	Compulsory Education: Duration	4.1; 4.2	Direct	Average (relative)
Government Expenditure on Education: Total: % of Government Expenditure		4c	Direct	Average	World Bank
Trained Teachers in Primary Education: % of Total Teachers		4c	Indirect	Trend Average	World Bank
Trained Teachers in Secondary Education: % of Total Teachers		4c	Indirect	Trend Average	World Bank
Educational Attainment: At Least Bachelor's or Equivalent: Population 25+ Years: % Cumulative		4.3	Indirect	Trend Average	World Bank
Educational Attainment: At Least Completed Lower Secondary: Population 25+ Years: % Cumulative		4.1	Indirect	Trend Average	World Bank
Educational Attainment: At Least Completed Primary: Population 25+ Years: % Cumulative		4.1	Indirect	Trend Average	World Bank
SDG 5 – Gender Equality	Gender Parity Index GPI: Literacy Rate: Youth Aged 15–24	5a; 4.5	Indirect	Trend	World Bank
	Gender Parity Index GPI: Primary and Secondary School Enrolment: Gross	5a; 4.5	Indirect	Trend	World Bank
	Gender Parity Index GPI: Tertiary School Enrolment: Gross	5a; 4.5	Indirect	Trend	World Bank
	Proportion of Seats Held by Women in National Parliaments	5.5	Indirect	Trend Average (relative)	World Bank
	Women Who were First Married by Age 18: % of Women Aged 20–24	5.4	Indirect	Average (relative)	World Bank
	Unmet Need for Contraception: % of Married Women Aged 15–49	5.6	Indirect	Average (relative)	World Bank
	SDG 6 – Clean Water and Sanitation	Level of Water Stress: Freshwater Withdrawal As Proportion of Available Freshwater Resources	6.4	Indirect	Trend
People Using At Least Basic Drinking Water Services: % of Population		6.1	Indirect	Trend	World Bank
People Using Safely Managed Drinking Water Services: % of Population		6.1	Indirect	Trend	World Bank
People with Basic Handwashing Facilities Including Soap and Water: % of Population		6.2	Indirect	Average (relative)	World Bank
SDG 7 – Affordable and Clean Energy	Renewable Energy Consumption: % of Total Final Energy Consumption	7.2	Indirect	Trend (relative) Average (relative)	World Bank
	Access to Electricity: % of Population	7.1	Indirect	Trend (relative)	World Bank
SDG 8 - Decent Jobs and Economic Growth	Wage And Salaried Workers: Modelled ILO Estimate: % of Total Employment	8.3; 8.5	Indirect	Trend	World Bank
	Children's rights in the workplace – enforcement score	8.7	Direct	Average	UNICEF
	Employment To Population Ratio: Modelled ILO Estimate: Aged 15+	8.5	Indirect	Trend	World Bank
	Share of Youth Not in Education, Employment or Training: Total: % of Youth Population	8.6	Indirect	Trend	World Bank
	Labour Force Participation Rate: Modelled ILO Estimate: Ratio of Female to Male	8.7	Direct	Average	World Bank

(Continued)

Table 1. (Continued).

SDG	Indicator	SDG target	Policy assessment	Method	Source
SDG 9 – Industry, Innovation and Infrastructure	Research and Development Expenditure: % of GDP	9.5	Direct	Average (relative)	World Bank
	Global Innovation Index	9.5	Indirect	Trend (relative)	WIPO
	Mobile Cellular Subscriptions per 100 People	9c	Indirect	Trend (relative)	World Bank
	Internet Users: Individuals: % of Population	9c	Indirect	Trend (relative)	World Bank
SDG 10 – Reduced Inequalities	Greenhouse Gas Emissions: Tonnes of CO ² Equivalent per GDP per Year	9.4; 13.2	Indirect	Trend (relative)	EDGAR
	Fragile States Index: Group Grievance	10.3	Indirect	Trend	The Fund for Peace
	Coverage: Social Safety Net Programs: % of Population	10.4; 1.3	Direct	Trend Average	World Bank
	Coverage: Social Protection and Labour Programs: % of Population	10.4; 1.3	Direct	Trend	World Bank
	Coverage: Social Insurance Programs: % of Population	10.4; 1.3	Direct	Average (relative)	World Bank
	Palma ratio (manually calculated based on income and population statistics provided by World Bank)	10.1	Indirect	Average (relative) Trend	World Bank
	Railways: Passengers Carried	11.2	Indirect	Trend	World Bank
SDG 11 – Sustainable Cities and Communities	Population Living in Slums: % of Urban Population	11.1	Indirect	Trend	World Bank
	Score of adoption and implementation of national DRR strategies in line with the Sendai Framework	11b; 11.5	Direct	Recent	UN Statistics
	PM2.5 Air Pollution: Mean Annual Exposure: Micrograms per Cubic Meter	11.6; 12.4; 3.9	Indirect	Trend	World Bank
	Material Resources: Domestic Material Consumption per Capita: Total	12.1; 12.2	Indirect	Trend Average (relative)	OECD
SDG 12 – Responsible Consumption and Production	Total Environment: Tax Revenue: % of GDP	12c	Direct	Average	OECD
	Fossil Fuel Support: per Capita: Total	12c	Direct	Average	OECD
	Proportion of municipal waste recycled (%)	12.5	Indirect	Trend	UN Statistics
SDG 13 – Climate Action	Greenhouse Gas Emissions: Tonnes of CO ² Equivalent per Capita per Year	13.2	Indirect	Trend Average (relative)	EDGAR
	GHG Emissions p/C Reduction 2015–30 (target)	13.2	Direct	Recent (relative)	Climate Resource
	ECIU Net Zero Tracker	13.2	Direct	Recent (relative)	ECIU Net Zero Tracker
	Fossil Fuel Support: per Capita: Total: Coal	13.2; 12.5	Direct	Average	OECD
SDG 14 – Life Below Water	Marine Protected Areas as % of Total Surface Area	14.2; 14.5; 14c	Direct	Trend Average	World Bank
	Ocean Health Index: Score	14.1	Indirect	Trend	National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara
	Fish Stock Status	14.4	Indirect	Trend	SeaAroundUs
	Fish Trawling	14.4	Indirect	Trend	SeaAroundUs
SDG 15 – Life on Land	Forest Area: % of Land Area	15.1; 15.2; 15.5	Indirect	Trend	World Bank
	Forest Area (Ha)	15.1; 15.2; 15.5	Indirect	Trend	World Bank
	Terrestrial Protected Areas as % of Total Land Area	15.1; 15.2; 15.5	Direct	Trend Average	World Bank
	Proportion of land that is degraded over total land area (%)	15.3	Indirect	Trend	UN Statistics
	CPIA: Transparency: Accountability: and Corruption in The Public Sector Rating: 1=Low To 6=High	16.5; 16.6	Indirect	Average (relative)	World Bank
SDG 16 – Peace, Justice and Strong Institutions	Intentional Homicides: per 100,000 People	16.1	Indirect	Trend	World Bank
	Corruption Perceptions Index	16.5	Indirect	Recent (relative)	Transparency International
	Fragile States Index: State Legitimacy	16.6; 16a	Indirect	Trend	The Fund for Peace
	Freedom in the World Score: Total Score	16.10	Direct	Recent	Freedom House
	Ratification of 18 International Human Rights	16.10	Direct	Recent (relative)	UN Human Rights – Office of the High Commissioner
	Ratification of controversial weapons agreements	16.1	Direct	Recent (relative)	UNODA – UN Office from Disarmament
	Internal and external conflict	16.1	Direct	Recent	PRS Group
Democracy status	16.5; 16.6	Direct	Recent	BTI	

(Continued)

Table 1. (Continued).

SDG	Indicator	SDG target	Policy assessment	Method	Source
SDG 17 – Partnerships for the Goals	Net Official Development Assistance Provided: To The Least Developed Countries: % of Gross National Income	17.2	Direct	Recent	World Bank
	Net Official Development Assistance Provided: % of Gross National Income	17.2	Direct	Recent	World Bank
	Sink Offshore Financial Centre	17.1	Direct	Recent	IFC
	Conduit Offshore Financial Centre	17.1	Direct	Recent	IFC

The table shows all indicators used in the study, aggregated per SDG and linked to their underlying targets. We explain if we view the indicator as a direct or indirect proxy for policy-effectiveness and show the method with which the indicator has been analyzed (which can be multiple).

reducing material use. When fewer than three data points were available, the change between the two values was used as a proxy for the trend. Indicators with insignificant regression coefficients or insufficient data points were neutralized.

These methods are not exclusive, and multiple approaches were sometimes applied to a single indicator. To ensure fairness, countries were assessed in absolute terms or relative to others within their income group when income affected policy effectiveness. The assessment method (s) for each indicator are shown in Table 1 and were selected based on discussions between the authors and a group of advisors. Table 2 explains the selected indicators for SDG 15 to illustrate this approach.

To address variability, we normalized scattered indicator values using log transformation and scaled them with RobustScaler from Python's sklearn library, based on the 2.5% to 97.5% quantile range. Indicators measuring similar constructs, such as healthcare capacity (e.g. nurses, midwives, and physicians per 1,000 people) and outcomes (e.g. infant and maternal mortality), were strongly correlated. To minimize noise and ensure the indicators uniquely and comprehensively represent sustainable development policymaking, we applied 'feature agglomeration' using sklearn's FeatureAgglomeration. This grouped highly similar indicators into two distinct features for each SDG, enabling a more streamlined and effective analysis of policy effectiveness while maintaining the multidimensional nature of the assessment.

2.3. Clustering and scoring countries on the 17 SDGs

The study used BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) to group countries based on their performance across the two features for each SDG. BIRCH, introduced by Zhang et al. (1996), was selected for its efficiency in processing large datasets and its ability to handle non-spherical, unevenly sized clusters. Compared to alternatives like K-means, which assumes uniform cluster shapes and struggles with high-dimensional data, or hierarchical clustering, which yielded less intuitive results during preliminary testing, BIRCH provided better interpretability and adaptability.

Each country's performance on the two features of an SDG was used to form clusters, with an 'anchor' cluster established at the center of the coordinate space. The anchor typically represents countries with neutral performance on both features. Scores ranging from -3 to +3 were assigned to each cluster based on a location metric, which evaluated clusters relative to the anchor. Positive scores indicated balanced progress across both features, while penalties were applied for imbalances, such as improvement in one feature accompanied by deterioration in the other. This scoring ensured alignment with the integrated goals of the SDGs.

The scoring process was fine-tuned using a 'stickiness' parameter, which determined how much change in the location metric warranted a score

Table 2. Example of assessment types for SDG 15.

Indicator	Assessment type(s)	Rationale
Forest Area: % of Land Area	Trend	An increase in forest area relative to total land area suggests that policies are effective in promoting forest ecosystems, while a decline indicates the opposite
Forest Area (Ha)	Trend	An increase in forest area in hectares suggests that policies are effective in promoting forest ecosystems, while a decline indicates the opposite
Terrestrial Protected Areas as % of Total Land Area	Average	A high proportion of protected areas suggest that policymakers are committed to preserving land-based ecosystems, while a low proportion indicates the opposite.
	Trend	Increasing proportions of protected areas suggest that policymakers have growing commitment to preserving land-based ecosystems, while reductions indicate the opposite.
Proportion of land that is degraded over total land area (%)	Trend	A reduction in the proportion of degraded land area suggests that policies are effective in improving the quality of land-based ecosystems, while increases indicate the opposite

This table presents indicators that were selected for SDG 15, the selected assessment type(s), and a rationale for how this can help evaluate policy effectiveness for this goal.

adjustment. Lower stickiness values were applied when clusters were closely spaced, while higher values were used for widely scattered clusters. This ensured an even distribution of scores, avoiding excessive neutrality or extreme scoring. The clustering and scoring process was validated qualitatively to align with expert judgment and SDG principles.

Figure 1(b,c) illustrates this process for SDG 9 (Industry, Innovation, and Infrastructure). Feature one captures trends in infrastructure and environmental sustainability indicators, such as electricity access and emissions per GDP, while feature two evaluates innovation through research and development expenditures and trends in the Global Innovation Index. Using a stickiness parameter of 0.1, countries were clustered into seven groups. Figure 1(b) shows the cluster distribution, and Figure 1(c) details the assigned scores. The anchor cluster (2) received a neutral score (0), representing limited positive or negative performance on both features. Clusters showing increasing progress on both features were scored +1 to +3, while those with declining performance were scored -1 to -3.

By combining quantitative analysis with expert insights, this method provided a robust, nuanced

framework for evaluating country-level SDG policies. Figure 2 explains the process in more detail, showing, for SDG 14 (Life Below Water), how the clustering process was applied.

2.4. Calculating a policy-effectiveness score across all SDGs

After scoring countries on their policy-effectiveness for each of the 17 SDGs, we calculated a total score for each country to indicate how likely they are to advance the entire SDG agenda.

To compute the total score, we first determined how many scores a country had at each of the ordinal values from -3 to +3, which we termed the 'realized score distribution.' We then created 44 hypothetical score distributions, called 'archetypes,' to represent how each of the seven types of ordinal scores might look. For example, Figure 1(d) shows six archetypes representing a -3 total score. These archetypes were developed to reflect the integrated and indivisible nature of the SDGs (UNGA, 2015) and to consider the precautionary principle of sustainability (Kriebel et al. 2001). The archetypes show that countries with negative policies across a substantial number of SDGs

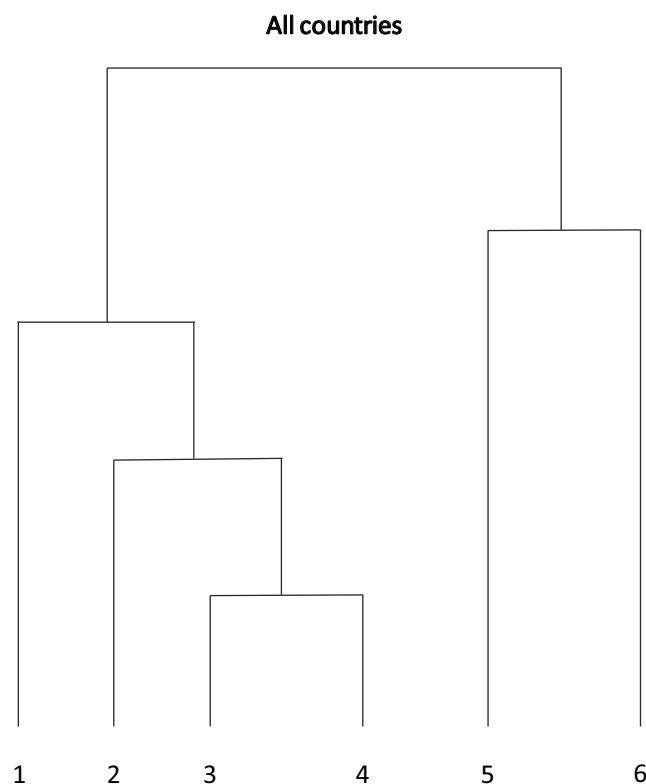


Figure 2. Dendrogram illustrating the application of BIRCH to SDG 14. The figure illustrates how countries are separated into six clusters through a step-by-step process. Countries that score poorly on all indicators related to this SDG are allocated to cluster 1 and are assigned a -2 score. Countries with more neutral performance on all indicators are allocated to cluster 2 and awarded a 0 score. Cluster 3 contains countries with decent performance on all indicators and are assigned a +1 score, while cluster 4 indicates countries with slightly poor performance on indicators, which receive a -1 score. The right-hand branch of the dendrogram, with clusters 5 and 6, contain countries that score poorly on the second feature of this SDG, which contains indicators that gauge fish capture. Countries that perform highly negative on this feature end up in cluster 6 and are awarded with a -3 score, while those with medium negative performance score -2 through cluster 5.

cannot achieve a neutral or positive total score. Conversely, countries with positive policies for many SDGs but no substantial negative policies receive a positive total score. Countries with neither substantially positive nor negative policies receive a neutral score.

These archetypes formed the training set for different regression and SVM models. We also used total scores from prominent countries as a testing set. Various models were trained with different sets of explanatory variables, such as realized probabilities, sample moments, and L-moments (Hosking 1990). To enhance robustness, each model was trained on the set with one observation removed, and the difference between the predicted and expert-assigned total scores was computed for each observation. These errors were aggregated to assess model performance. The best model that minimized the prediction errors was the ν -SVC model (Scholkopf et al. 2000) with L-moments. Although SVC does not directly model ordinal data, it works by finding the optimal hyperplane that separates data points into different classes. For ordinal data, the categories are treated as separate classes, and SVC finds boundaries to distinguish them. Kernel methods in SVC map data into higher dimensions to help find a hyperplane that separates the classes more effectively. This is particularly useful for ordinal data as it captures complex relationships between ranks, improving predictive accuracy. By using appropriate kernel functions such as linear, polynomial, or RBF, SVC adapts well to ordinal data, even without explicit ordinal regression techniques.

Other models considered included linear regression, ridge classifier, logistic regression, and SVC models with linear, polynomial, and RBF kernels. L-moments and realized score frequencies were used as inputs for these models. The ν -SVC model performed well with L-moments and score frequencies, as L-moments capture distribution properties more robustly than regular moments. L-moments also have advantages, such as being more robust and requiring only a finite mean for higher-order L-moments. Comparing the results of the ν -SVC model with expert-assigned scores, the model did not miss any score by more than one point on the -3 to $+3$ scale. Logistic regression also delivered similar but slightly less accurate performance.

After selecting and estimating the best model, we computed total scores for countries by matching their realized score distributions to the archetypes. This approach allows us to capture the diversity of possible scores across the 17 SDGs.

2.5. Correlations between SDG policy scores

We calculated Pearson correlation coefficients between SDG policy scores, Gross National Income, and a country's SDG Index score (Sachs et al. 2023)

using a bootstrap approach, which is useful for estimating statistics without assuming normality (Efron and Tibshirani 1994). The bootstrap method involves drawing samples with replacement from a bivariate sample, calculating the correlation for each sample, and repeating the process 10,000 times to build a bootstrap distribution. We then used the 'basic bootstrap' method to calculate confidence intervals at a 5% level. Significant correlations between SDGs were visualized as network graphs using Python's networkX package.

3. Results

The method explained above was applied to a sample of 170 countries using data from 2023, in order to estimate illustrative SDG policy scores. This section presents the results of this exercise, starting with a view on policy effectiveness for the SDG agenda as a whole, commencing with an analysis of policies for individual SDGs and concluding by shedding light on policy coherence across the 2030 Agenda.

3.1. Advancing the SDG agenda

Of the 170 countries analyzed, 35 (21%) have policies that are supportive of the SDGs. Twenty-one and fourteen of these countries' policies are, respectively, marked as 'low positive' ($+1$ score) and 'medium positive' ($+2$ score). No country received the highest score of $+3$ ('high positive') for their policy effectiveness across all 17 SDGs. Another 50 countries (29%) have policies that we label as 'neutral': they do not substantially advance, nor deteriorate, further progress on the SDG agenda. The remaining 85 countries (50%) have policies that detract from the SDGs. This comprises countries whose policies are marked as low (41 countries; -1 score), medium (28 countries; -2 score), and highly (16 countries; -3 score) negative for advancing the global goals.

Figure 3(a) shows the global distribution of policy scores for the SDGs. All continents host countries with positive and negative policy scores. Europe has the lowest portion of countries with negative scores, making it the continent with the most supportive policy agenda for attainment of the SDGs. We find a slight positive correlation between a country's SDG policy score and its Gross National Income (GNI) (0.21, 95% confidence interval [0.07, 0.36]), visualized in Figure 3(b). We also document a moderate positive correlation (0.39, 95% confidence interval [0.28, 0.51]) between country's policies for the SDGs and its current level of SDG attainment, measured by the SDG Index (Sachs et al. 2023) (Figure 2c). We find countries at all income and SDG-attainment levels with very positive scores (e.g. France and Ghana both receive

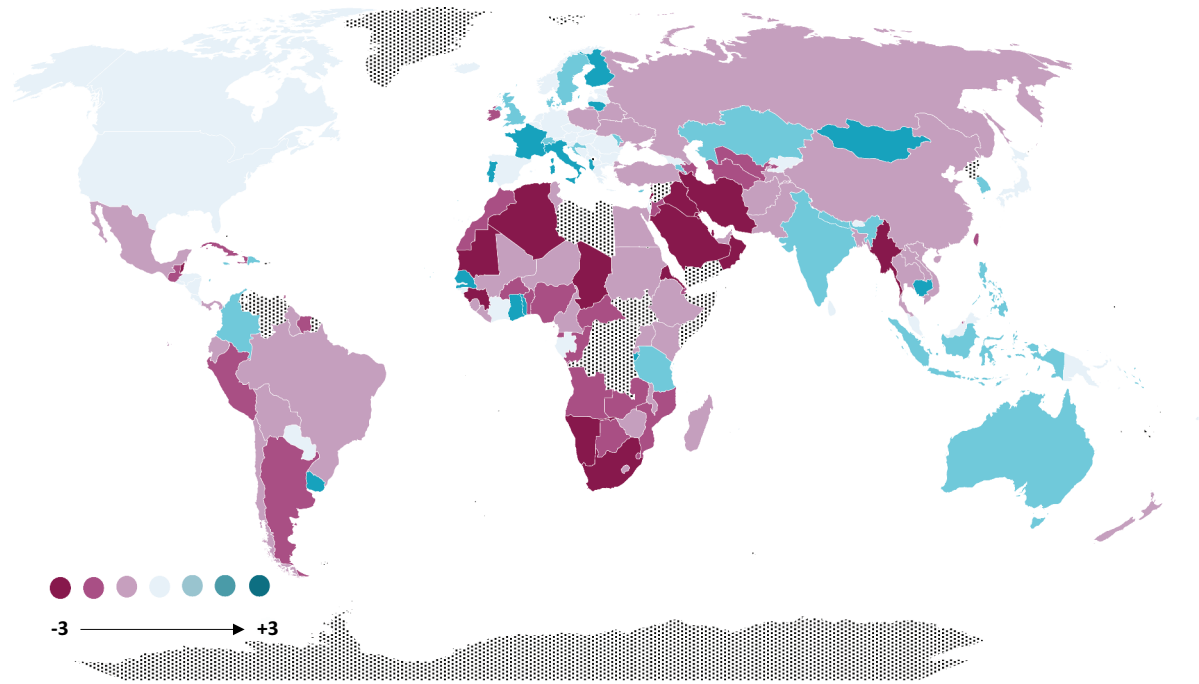
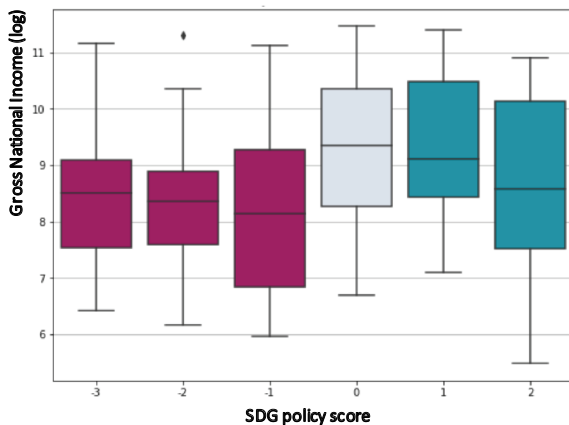
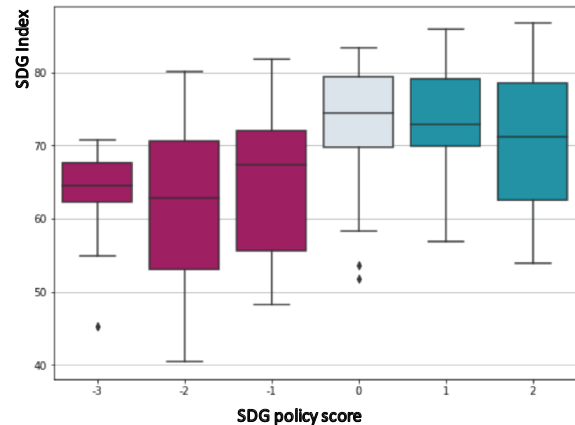
a**b****c**

Figure 3. Overview of governmental SDG policy scores. a, distribution of scores that assess governmental policies for the entire SDG agenda around the world. b, comparison between a country's SDG policy score and its gross national income (log). c, comparison between a country's SDG policy score and its position on the SDG Index. Blue and pink colors represent positive and negative policies, respectively. Darker shades in panel a indicate increasingly positive or negative policies.

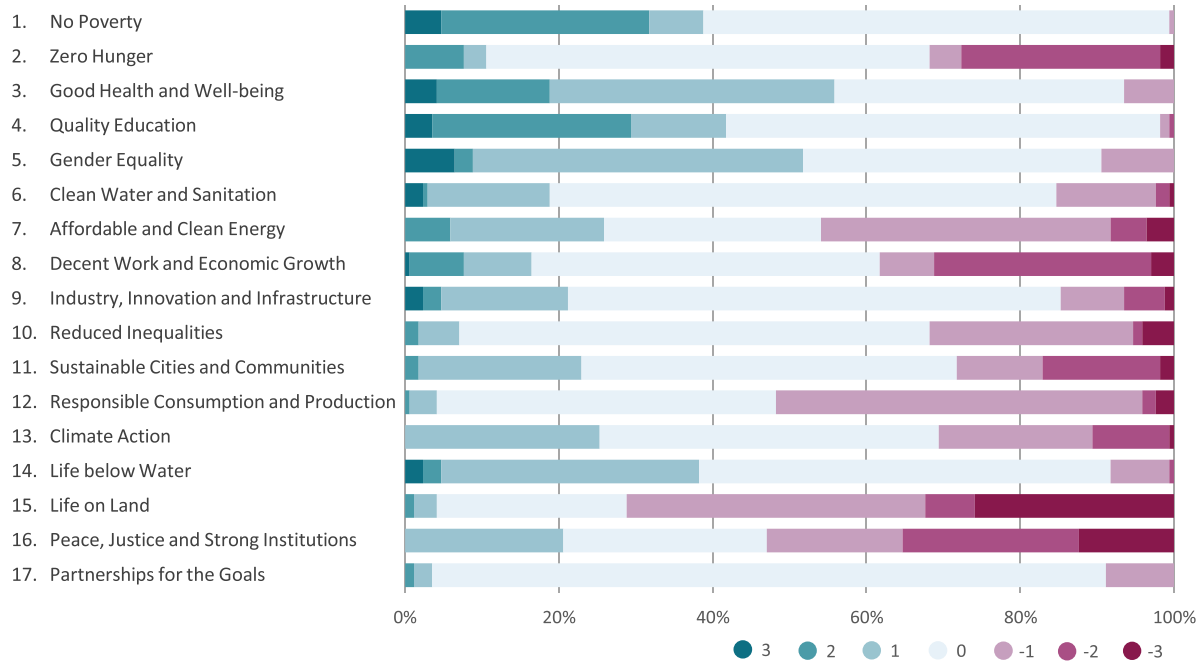
a 'medium positive' score) and very negative ones (e.g. Myanmar and Qatar both have 'high negative' scores).

3.2. Policies for individual SDGs

The analysis presented above on countries' policies for the entire SDG agenda is based on an assessment of their policies for individual SDGs. Figure 4(a) shows the proportion of countries with positive, neutral and negative policies for each SDG. Figure 4(b) exemplifies how 13 countries score across the 17 goals and are given a total score for their policy-effectiveness across the whole SDG agenda.

Countries are most likely to have policies that support the attainment of SDGs aimed at improving human well-being. Ninety-five countries (56%) have policies that support further progress on Good Health and Well-being (SDG 3). Eighty-eight countries (52%) are advancing Gender Equality (SDG 5). Quality Education (SDG 4) is being promoted by 71 countries (42%), and 66 countries (39%) have policies that help eradicate poverty (SDG 1). An example of a country that is having supportive policies for all these SDGs is Ghana. Additionally, relatively few countries, between 1% and 9% of our sample, have policies that actively detract from attaining these socially oriented goals.

This contrasts policy-effectiveness for SDGs aimed at safeguarding the natural environment, where many

a**b**

Country	SDG policy score	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Ghana	+2	+2	0	+2	+2	+1	+1	+1	0	0	0	+1	-1	0	0	0	+1	0
France	+2	0	0	0	0	+1	0	0	-1	+1	0	0	0	+1	+3	0	+1	0
Jamaica	+1	0	0	0	+2	+1	0	0	0	0	-1	-2	0	+1	+1	0	+1	0
Indonesia	+1	+3	0	+3	+2	0	-1	-1	0	0	0	0	0	0	+1	-3	-1	0
Sweden	+1	0	0	0	0	+1	0	+1	0	+1	-1	-2	-1	+1	+1	-1	+1	+2
South Korea	+1	0	0	+1	0	0	0	-1	+2	+3	0	+2	-1	-1	+1	-1	-1	0
Canada	0	0	0	0	0	0	0	+1	0	0	0	-1	-1	-2	+1	-1	+1	0
Japan	0	0	0	0	0	-1	0	0	0	+1	0	0	0	+1	+2	-1	+1	0
Ethiopia	-1	0	-2	+3	+2	+3	-1	-1	0	0	0	+1	-1	0	0	-1	-3	0
Ireland	-2	0	0	0	0	0	-1	0	0	0	0	-2	-3	-1	+1	0	+1	-1
Argentina	-2	0	-2	+1	0	+1	0	0	-2	-1	0	-1	-1	-1	+2	-1	0	0
Qatar	-3	0	0	0	+2	0	-1	-1	-2	-2	0	0	-1	-3	+1	-1	-3	0
South Africa	-3	+2	-2	+1	0	+1	-1	-1	-2	0	0	-2	-1	-1	+1	-3	-1	0

Figure 4. Overview of SDG policy scores per goal. a, distribution of scores that assess governmental policies for individual SDGs, whereby the chart shows the proportion of countries ($n = 170$) that have highly positive (+3) to highly negative (−3) policies for an individual SDG. b, examples of SDG policy scores for selected countries, both in terms of the total score and the sub-scores per each SDG. Blue and pink colors represent positive and negative policies, respectively. Darker shades indicate increasingly positive or negative policies.

countries are found to have policies that are likely to reduce progress. For instance, 121 (71%), 88 (52%) and 52 (31%) countries, respectively, have policies that deteriorate Life on Land (SDG 15), Responsible Consumption and Production (SDG 12), and Climate Action (SDG 13). The Republic of Korea is an example of a country that, despite having positive policies for various socio-economic SDGs, scores negatively on these SDGs. Additionally, 90 countries have policies that reduce progress on Peace, Justice, and Strong Institutions (SDG 16), for reasons such as war (e.g.

Russia), limited freedoms and democracy (e.g. Myanmar, China), or corruption (e.g. Panama, Hungary, Guyana).

To further illustrate our assessment, for each SDG **Figure 5** shows how a country performs on a relevant indicator for that goal. The indicators were selected to display variety in terms of direct and indirect proxies of policy-effectiveness. Countries were selected that are quintessential examples of (positive and negative) policy scores for each of the displayed indicators.













Kyrgyzstan	1. No Poverty	Bank Account Ownership at a Financial Institution or with a MobileMoney-Service Provider: % of Population Aged 15+	Indirect	Trend	+144%		+3
Afghanistan	2. Zero hunger	Prevalence of Severe Food Insecurity in the Population: % of population	Indirect	Average	18		-2
Ghana	3. Good health and well-being	Maternal Mortality Ratio: Modelled Estimate: per 100,000 Live Births	Indirect	Trend	-4%		+2
Indonesia	4. Quality education	Government Expenditure on Education: Total: % of Government Expenditure	Direct	Average	18		+2
Nigeria	5. Gender equality	Proportion of Seats Held by Women in National Parliaments	Indirect	Trend	-35%		-1
Kuwait	6. Clean water and sanitation	Level of Water Stress: Freshwater Withdrawal As Proportion of Available Freshwater Resources	Indirect	Trend	+13%		-3
Vietnam	7. Affordable and clean energy	Renewable Energy Consumption: % of Total Final Energy Consumption	Indirect	Average	24		-3
Cuba	8. Decent work and economic growth	Wage And Salaried Workers: Modeled ILO Estimate: % of Total Employment	Indirect	Trend	-4%		-3
South Korea	9. Industry, innovation and infrastructure	Research and Development Expenditure: % of GDP	Direct	Average	4		+3
Colombia	10. Reduced inequalities	Palma Ratio	Indirect	Average	2		-3
Jamaica	11. Sustainable cities and communities	Score of adoption and implementation of national DRR strategies in line with the Sendai Framework	Direct	Current	0.1		-2
Ireland	12. Responsible consumption and production	Fossil Fuel Support: per Capita: Total	Direct	Average	597		-3
Denmark	13. Climate action	Greenhouse Gas Emissions: Tonnes of CO2 Equivalent per Capita per Year	Indirect	Trend	-7%		+1
France	14. Life below water	Marine Protected Areas as % of Total Surface Area	Direct	Trend	+90%		+3
Burundi	15. Life on land	Forest Area: % of Land Area	Indirect	Trend	+7%		+2
China	16. Peace, justice and strong institutions	Ratification of 18 International Human Rights treaties	Indirect	Current	8		-3
Norway	17. Partnerships for the goals	Net Official Development Assistance Provided: To The Least Developed Countries: % of Gross National Income	Direct	Current	0.2		+2

Figure 5. Examples of policy-effectiveness for SDG indicators. The figure shows how selected countries perform on examples of indicators used in the assessment to gauge policies across the 17 SDGs. The table lists if the indicator is a direct or indirect proxy of policy-effectiveness and what assessment method was applied. The performance charts specify the country's performance since 2015. Each country's policy score for the relevant SDG is given. Note that this score is not only driven by the country's performance on the selected indicator. Rather, the score is a result of a country's performance on all indicators belonging to an SDG.

3.3. Policy coherence for the SDGs

To shed light on the degree of coherence among policies for the SDGs, Figure 6 shows correlations between scores for individual goals. Positive correlations between pairs of SDGs indicate that countries score similarly (positive or negative) for their policy-effectiveness towards both goals. Negative correlations indicate diverging (positive and negative) governmental policies for pairs of SDGs.

In total, there are 136 pairs of SDGs that could have correlated policy scores. Across all 170 countries (Figure 6a), 37 pairs of SDGs are significantly correlated at a 95% confidence interval. 24 and 13 pairs are

positively and negatively associated, respectively. We find high centrality among some SDGs, in the sense that they are correlated with many others. This includes SDGs around hunger (SDG 2), healthcare (SDG 3), energy (SDG 7), inequality (SDG 10), climate change (SDG 13), and life on land (SDG 15). These goals correlate with six or seven other SDGs. Goals with relatively low centrality, correlating with one or two others, are SDGs 6 (water), 9 (industrialization), and 14 (life below water).

We find moderate (0.40) positive correlation between SDGs 7 (Affordable and Clean Energy) and SDG 16 (Peace, Justice and Strong Institutions). This is mainly driven by the 54 countries that score

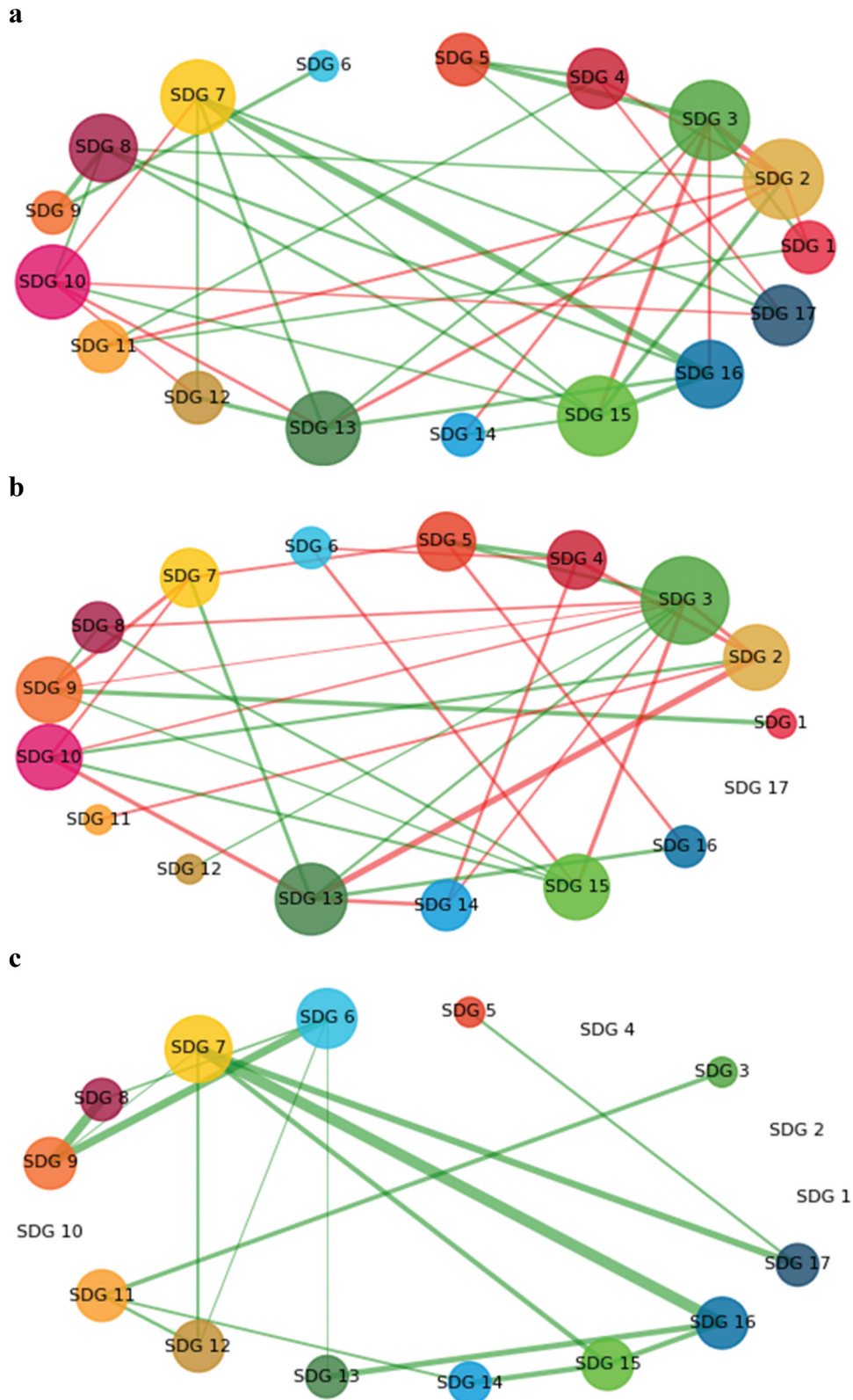


Figure 6. Network graphs showing positive and negative correlation between pairs of SDGs for countries at different income levels. a, all countries ($n = 170$). b, low-, lower middle- and upper middle-income countries ($n = 115$). c, high-income countries ($n = 55$). Nodes, and their colors, refer to SDGs, whereby their size indicates each SDG's centrality (how many SDGs they correlate with). Edges indicate statistically significant correlation between SDGs, whereby their width signals the degree of correlation between two SDGs. Edges colors show positive (green) and negative (red) correlation.

negatively on both goals. These primarily are countries that use relatively little, and in various cases declining, renewable energy in their total final energy consumption, while at the same time having

weak (democratic) institutions. Examples include Russia, Kazakhstan, Qatar, Algeria, Mauritania, and Saudi Arabia. At the same time, 19 countries have positive policies for this pair of goals. This includes

countries that have substantial (and often increasing) renewables in their energy mix, combined with solid (and often improving) institutions that guarantee personal freedoms (e.g. Denmark, Costa Rica, Slovenia, New Zealand, Iceland).

Moreover, policies for Good Health and Well-being (SDG 3) positively correlate (0.31) with Gender Equality (SDG 5). Fifty-eight countries have positive policies for both goals. Countries that do well on different dimensions of gender equality (women representation in parliament; low rates of women being married before aged 18; low unmet need for contraceptives) often also do well on health policies such as increasing the number of births attended by skilled staff, and declining infant and maternal mortality rates.

However, policy scores for Good Health and Well-being (SDG 3) negatively correlate with four other SDGs. The strongest negative correlation (-0.34) is with policies for Zero Hunger (SDG 2). Forty-three countries have positive policies for SDG 3 and negative ones for SDG 2. This is primarily driven by an increase in food insecurity in recent years, suggesting that policies are insufficiently addressing hunger, while having good policies in terms of reducing mortality of infants, mothers, and reducing deaths of communicable diseases. The robustness of this result may be vulnerable over time: rising hunger threatens mortality. Further to adversely correlating with hunger, SDG 3 also negatively relates (-0.27) to Life On Land (SDG 15). Seventy-eight countries have policies that support health but detract from land-based ecosystems. These countries see strong reductions in mortality rates, but increases in deforestation and land degradation.

We note that policy coherence for the SDGs varies for countries at different income levels. Figure 6(b,c) show correlations between SDG pairs for emerging and developed countries. Among emerging countries, 30 pairs of SDGs are correlated. In high-income countries this number drops to 17. Emerging market countries display a substantial number of negative correlations, implying that policies for two SDGs diverge. The correlations between SDGs in developed countries are all positive, suggesting that they move in the same direction (positive or negative).

4. Discussion

The adoption of the SDGs has been hailed as a landmark initiative in global sustainable development governance. The SDGs provide an ambitious, inclusive, and globally relevant sustainability agenda that covers a broad range of challenges and requires the involvement of both the public and private sectors (Biermann et al. 2017). However, the SDGs are non-binding aspirations with weak institutional oversight, allowing high levels of discretion for national governments in implementing them (Biermann et al. 2017).

Therefore, the success of the SDGs largely depends on how governments design policies to meet the goals they have committed to achieving.

This study addresses the growing calls for developing approaches to evaluate the policy effectiveness of SDGs (Hickmann et al. 2024). A major flaw in the 2030 Agenda's design is the lack of standardized reporting on governmental actions taken to advance the SDGs. The High-Level Political Forum (HLPF), tasked with overseeing and reviewing SDG implementation – including by coordinating and discussing Voluntary National Reviews (VNRs) submitted by countries – has been criticized for providing only superficial oversight to date (Beisheim and Fritzsche 2022). This gap in systematic evaluation of national policies for the SDGs hinders the success of the 2030 Agenda by complicating efforts to identify effective approaches and share lessons learned.

While countries have been providing a wealth of data on their progress toward meeting the SDG targets, these data do not directly reveal whether countries are implementing policies that advance the SDGs. This paper contends that these data can be analyzed to gain insights into policy effectiveness for sustainable development. It presents an empirical method that can be used to assess national policy effectiveness for the SDGs. The method is intended to be illustrative and serve as an inspiration for further monitoring of the SDGs.

The method introduced in this paper consists of four key steps relevant for analyzing governmental sustainability policies empirically. The first step involves identifying indicators that can shed light on policy effectiveness for the SDGs. While many indicators are linked to the SDGs, it is important to focus on those that have direct policy relevance. For example, the percentage of government spending on education is highly relevant to SDG 4 (Quality Education), while indicators such as the number of deaths from natural disasters are less directly influenced by policy and are more dependent on external factors. The second step emphasizes that empirical indicators can reflect policy effectiveness either directly or indirectly. For example, government expenditure on education is a direct metric, while educational attainment of students is an indirect indicator, which can still provide valuable insights into policy effectiveness.

Third, based on the distinction between direct and indirect indicators, the paper suggests that certain analysis techniques are better suited to assessing policy effectiveness. For instance, indicators such as government spending on education, the number of deaths from diseases, and the adoption of treaties on cluster weapons can inform policy effectiveness for SDGs 4 (Quality Education), 3 (Good Health and Well-being), and 16 (Peace, Justice, and Strong Institutions), respectively. The key insight is that not all indicators are

equally useful for monitoring SDG policies. Rather, specific analytical methods must be applied to extract meaningful insights regarding the effectiveness of the policies.

Finally, once indicators have been analyzed using the appropriate methods, the assessments must be aggregated and compared across countries to gauge how effective each country's policies are in advancing the SDGs. For this purpose, the paper applies BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) to cluster countries based on their performance on each SDG, enabling the development of initial ordinal country SDG policy scores.

When this method was applied to the data from 170 countries, the findings indicated limited evidence that countries' policies are likely to support the achievement of the SDGs. This result aligns with other research suggesting that the SDGs are insufficiently advancing sustainable development (Biermann et al. 2022; Hickmann et al. 2024). While some governments are advancing socio-economic objectives, many are undermining environmental sustainability. This highlights the trade-off between socio-economic development and environmental sustainability, which exists not only in terms of outcomes (e.g. Zeng et al. 2020) but also in policy-making.

Additionally, the study found that there is significant room for improvement in the coherence of policy-making across the SDGs. There is a low correlation between governments' policy scores across the 17 SDGs. Out of the 136 pairs of SDGs analyzed, only 37 pairs exhibited correlated policies. A third of these (13) were negatively related, indicating that governments have diverging policies on certain pairs of goals. Twenty-four pairs were positively related, but in some cases, this applied to governments that had negative policies for both goals. These correlations varied across countries of different income levels. High-income countries exhibited more coherent policies than low and middle-income countries.

This analysis presents a viable approach to evaluating government policies for the SDGs, responding to calls for more systematic evaluations (Sachs et al. 2019; Hickmann et al. 2024). The method complements existing research on estimating overall SDG attainment (Schmidt-Traub et al. 2017; Allen et al. 2018), developing modeling tools to support national SDG planning (Allen et al. 2016; Collste et al. 2017), and analyzing policies for individual goals (e.g. Siddiqi et al. 2020) or countries (e.g. Xie et al. 2021). Moreover, this study's analysis goes beyond examining the current sustainability performance of countries and instead assesses their likelihood of making faster progress, a key consideration when evaluating the effectiveness of global goal-setting initiatives like the SDGs (Fukuda-Parr et al. 2013). The findings also support studies showing that

the SDGs have had a limited transformative impact on policy-making (Biermann et al. 2022).

The analysis underscores that unless governments dramatically transform their policies, achieving most of the SDGs by 2030 is unlikely. The results can inspire policymakers and other stakeholders to take further steps toward implementing the SDG agenda more effectively.

4.1. Robustness

To further test for the method's robustness, we recalculate SDG policy scores for the 170 countries using 2024 data. This means that for nearly all indicators an extra year was added to the analysis, while various indicators also have received new values for past years (thus having been rewritten), and multiple countries received data for indicators that previously lacked them.

With a correlation coefficient of 0.66, we find a moderately strong positive relationship between the countries' 2023 and 2024 SDG policy scores. Sixty-nine of the countries received the exact same score in 2024 as in 2023. Of the countries that saw a score change, the vast majority (61 countries; 36% of total) witnessed only a one-notch upgrade or downgrade. Thirty-five countries (21%) saw their score in/decrease by two notches. And the remaining five countries (3%) had a score that changed by three notches. Using 2024 data, we find that 30% of the countries have a positive, 26% a neutral, and 44% a negative policy score. These results are slightly more positively distributed than the 2023 assessment, potentially indicating that countries are improving their sustainable development policies. Because of the correlation between countries' 2023 and 2024 scores, and due the fact that most countries only had limited changes to their scores, we believe there is considerable robustness in the model. A different conclusion would have to be reached if the updated scores show lower correlation and larger jumps.

4.2. Policy implications

The insights presented in this study have practical implications. First, policymakers are advised to adopt a nexus approach to the SDGs. There is increasing consensus that sustainability goals are best advanced through a nexus approach that seeks to advance multiple SDGs at the same time, while reducing the risk that policies advancing one goal undermine progress on another (Boas et al. 2016; Bleischwitz et al. 2018; Liu et al. 2018; van Zanten and van Tulder 2021). However, there is no boilerplate for implementing a nexus approach to the SDGs. Weitz et al. (2018) underscore that the interaction between the SDGs are contextual

and that policymakers therefore ought to develop a 'cross-impact matrix' that offers insights into how progress towards one SDG targets affects a second target within a particular geographic scope. The insights of this study can help to define policy indicators that are likely to advance an SDG.

For example, from a nexus perspective policymakers might be interested in advancing Life Below Water (SDG 14), as improved marine ecosystem sustainability could indirectly help to reduce hunger (SDG 2) through better managed and long-term sustainable fisheries, which, in turn, can promote health and well-being (SDG 3). The indicators used in this study suggest that a combination of policy actions can be initiated to achieve this objective, from increasing marine protected areas, to reducing pollution, and banning harmful fisheries practices like bottom trawling. As such, the indicator selection used here can deliver levers that provide inspiration for making the nexus approach actionable.

Second, policymakers are encouraged to build a standardized dataset that can help track the actions that governments undertake to promote specific SDGs. Such a dataset can comprise information on a government's investments into specific SDG topics, subsidies and taxation to advance the goals, target-setting and tracking of such targets, and adoption of regulation for promoting sustainability outcomes. This will enable better tracking of the implementation of the 2030 Agenda and facilitate research on policy-coherence for sustainable development.

Third, investors stand to benefit from a better understanding of policy-effectiveness for the SDGs. Finance is one key lever for attaining the SDGs (Sachs et al. 2019), yet achievement of the goals faces a massive financing gap (e.g. Gaspar et al. 2019). A main way for countries to raise financing is through issuing debt in the form of government bonds. Investors can support countries in achieving the SDGs by prioritizing investments in government bonds issued by countries that have policies that promote the SDGs. A discussion about the relevance of government bonds for promoting the SDGs, combined with a practical illustration of how investors can support sustainable development is provided by Swinkels et al. (2024).

4.3. Limitations and future research

In terms of limitations, we acknowledge that measuring sustainable development is highly complex (Hák et al. 2016; Dang and Serajuddin 2020) that creating scoring methodologies for numerous environmental and social sustainability dimensions requires making subjective decisions (Stiglitz et al. 2009), and that resulting scores must be interpreted as giving guidance rather than being objective truths (Kanbur

et al. 2018). As such, the methodology proposed in this paper and the resulting outcomes are best understood as one, but not the only, viable way to approximate policy-effectiveness for the SDGs. At the same time, we believe that novel insights can emerge when existing data is analyzed in new ways that transparency about methodological decisions helps interpret the results, and that the scoring developed in this paper can inform debates about how governments can implement policies that help advance sustainable development.

In addition, we note that our analysis covers a significant part, but not all, of the 2030 Agenda. This is a result of a lack of data for specific SDG targets, either in general or across a sufficiently broad sample of countries (Hák et al. 2016), and the absence of a central dataset of indicators tracking governmental policies for sustainable development. Moreover, the study can be affected by measurement errors in the indicators. This could particularly apply to countries with less reliable data collection and processing capacity. Additionally, there is a time lag in the data, whereby new events are only reflected in the numbers after one or 2 years. These limitations were mitigated to the greatest extent possible by selecting established and reputable data sources. Nevertheless, it is important to remember that the results can be altered by the emergence of new data. This is similar to limitations of measuring progress on the SDGs in general (e.g. MacFeely 2019), and in the creation of other rankings for these goals, like the SDG Index (e.g. Diaz-Sarachaga et al. 2018). However, the methodological approach developed here itself is not affected by new data, as new data can be added and updated to track governmental policies for the SDGs over time.

A related limitation concerns the fact that not all dimensions of policy-effectiveness can be captured in readily available data, such as in the data that were used in this study. This is driven by time lags in data provision as well as shortcomings in the nature of indicators. The case of Israel's war on Gaza is a good example. Israel receives a policy score for SDG 16 (Peace, Justice, and Strong Institutions) of -1 . This negative score is driven by the country insufficiently ratifying international agreements related to the use of controversial weapons, being flagged for involvement in internal and external conflict, and having an increase in corruption. Nevertheless, the severity of the country's warfare, which the International Court of Justice has labelled as plausibly genocidal (ICJ 2024), is not reflected in the data. While our analysis includes an indicator on 'intentional homicides: per 100,000 people', these data exclude killing in armed conflicts and so does not include for the tens of thousands of people murdered by Israeli forces. In addition, the broader destruction of housing, infrastructure, water supply, education and healthcare facilities, and so forth, is

not captured in data, nor are the indirect deaths that followed from causes like diseases or shortages of food and water (Khatib et al. 2024). As such, the -1 score, despite being negative, is arguably a severe underestimation of Israel's policies for SDG 16, driven by data limitations.

To deal with these limitations, future research can investigate governmental policies for the SDGs using novel methods. There are numerous avenues that will be worthwhile to pursue. Large language models can help to evaluate policies from a qualitative angle. Such an approach can screen news articles around the world and thus contribute to real-time insights, overcoming time lags associated with the official SDG indicators. Scholars can furthermore investigate the interaction effects among policies for sustainable development in more detail, across geographic regions and countries at varying income levels. Such efforts will help to understand how policies for the SDGs play out in different contexts. And, together with policy-makers, scholars can work on standardized methods for monitoring governmental policies for the SDGs, whereby this study could provide inspiration.

4.4. Conclusion

Transforming the world through the 2030 Agenda requires transformative actions, yet a major knowledge gap exists in assessing whether national policies support sustainable development. This paper developed a method to evaluate policy effectiveness for the 17 SDGs using 84 indicators. It analyzed countries' performance with BIRCH, creating SDG policy scores to gauge policy quality. Applying this method to 170 countries, the study found that few nations have policies likely to advance sustainable development, while about half risk hindering progress. Additionally, global SDG policies lack coherence. These findings inform governance improvements to enhance sustainable development by 2030 and aim to inspire further thinking about the empirical monitoring of sustainability policies.

Acknowledgments

We thank Rikkert Scholten, Laurens Swinkels, Johan Duyvesteyn, Max Schieler and Paul Ruijs for their advice during the development of the methodology presented in this paper. We are grateful to Bruno Rein for his research support. We also thank Hai-Anh Dang and Masja Zandbergen-Albers for valuable feedback on an earlier version of this research. Finally, we appreciate the constructive feedback of four anonymous reviewers as well as the thoughtful guidance of our editor, Guofan Shao.

Disclosure statement

The authors are employed by Robeco, an international asset management firm headquartered in the Netherlands. The

views expressed in this paper are not necessarily shared by Robeco.

ORCID

Jan Anton van Zanten  <http://orcid.org/0000-0002-2968-6339>

Data availability statement

The data used in this paper can be obtained from the organizations listed as data providers in Table 1. All information is available online for free, with the exception of internal and external conflict metric provided by PRS Group.

References

- Allen C, Metternicht G, Wiedmann T. 2016. National pathways to the sustainable development goals (SDGs): a comparative review of scenario modelling tools. *Environ Sci Policy*. 66:199–207. doi: [10.1016/j.envsci.2016.09.008](https://doi.org/10.1016/j.envsci.2016.09.008).
- Allen C, Metternicht G, Wiedmann T. 2018. Initial progress in implementing the sustainable development goals (SDGs): a review of evidence from countries. *Sustain Sci*. 13(5):1453–1467. doi: [10.1007/s11625-018-0572-3](https://doi.org/10.1007/s11625-018-0572-3).
- Asma S, Lozano R, Chatterji S, Swaminathan S, de Fátima Marinho M, Yamamoto N, Murray CJ, Misganaw A, Ryan M, Dandona L. 2020. Monitoring the health-related sustainable development goals: lessons learned and recommendations for improved measurement. *Lancet*. 395(10219):240–246. doi: [10.1016/S0140-6736\(19\)32523-1](https://doi.org/10.1016/S0140-6736(19)32523-1).
- Beisheim M, Fritzsche F. 2022. The UN High-Level political forum on sustainable development: an orchestrator, more or less? *Global Policy*. 13(5):683–693. doi: [10.1111/1758-5899.13112](https://doi.org/10.1111/1758-5899.13112).
- Bennich T, Weitz N, Carlsen H. 2020. Deciphering the scientific literature on SDG interactions: a review and reading guide. *Sci Total Environ*. 728:138405. doi: [10.1016/j.scitotenv.2020.138405](https://doi.org/10.1016/j.scitotenv.2020.138405).
- Biermann F, Hickmann T, Sénit CA, Beisheim M, Bernstein S, Chasek P, Wicke B, Kim RE, Kotzé LJ, Nilsson M. 2022. Scientific evidence on the political impact of the sustainable development goals. *Nat Sustain*. 5(9):795–800. doi: [10.1038/s41893-022-00909-5](https://doi.org/10.1038/s41893-022-00909-5).
- Biermann F, Kanie N, Kim RE. 2017. Global governance by goal-setting: the novel approach of the UN sustainable development goals. *Curr Opin Environ Sustainability*. 26:26–31. doi: [10.1016/j.cosust.2017.01.010](https://doi.org/10.1016/j.cosust.2017.01.010).
- Bleischwitz R, Spataru C, VanDeveer SD, Obersteiner M, van der Voet E, Johnson C, Van Vuuren DP, Boersma T, Hoff H, van Vuuren DP. 2018. Resource nexus perspectives towards the united nations sustainable development goals. *Nat Sustain*. 1(12):737–743. doi: [10.1038/s41893-018-0173-2](https://doi.org/10.1038/s41893-018-0173-2).
- Boas I, Biermann F, Kanie N. 2016. Cross-sectoral strategies in global sustainability governance: towards a nexus approach. *Int Environ Agreements*. 16(3):449–464. doi: [10.1007/s10784-016-9321-1](https://doi.org/10.1007/s10784-016-9321-1).
- Collste D, Pedercini M, Cornell SE. 2017. Policy coherence to achieve the SDGs: using integrated simulation models to assess effective policies. *Sustain Sci*. 12(6):921–931. doi: [10.1007/s11625-017-0457-x](https://doi.org/10.1007/s11625-017-0457-x).
- Costanza R, Daly L, Fioramonti L, Giovannini E, Kubiszewski I, Mortensen LF, Wilkinson R, Ragnarsdottir KV, De Vogli R,

- Wilkinson R. 2016. Modelling and measuring sustainable wellbeing in connection with the UN sustainable development goals. *Ecol Econ.* 130:350–355. doi: [10.1016/j.ecolecon.2016.07.009](https://doi.org/10.1016/j.ecolecon.2016.07.009).
- Dang HAH, Serajuddin U. 2020. Tracking the sustainable development goals: emerging measurement challenges and further reflections. *World Devel.* 127:104570. doi: [10.1016/j.worlddev.2019.05.024](https://doi.org/10.1016/j.worlddev.2019.05.024).
- Diaz-Sarachaga JM, Jato-Espino D, Castro-Fresno D. 2018. Is the sustainable development goals (SDG) index an adequate framework to measure the progress of the 2030 agenda? *Sustain Devel.* 26(6):663–671. doi: [10.1002/sd.1735](https://doi.org/10.1002/sd.1735).
- Efron B, Tibshirani RJ. 1994. *An introduction to the bootstrap*. New York: CRC press.
- Fukuda-Parr S, McNeill D. 2019. Knowledge and politics in setting and measuring the SDGs: introduction to special issue. *Global Policy.* 10(S1):5–15.
- Fukuda-Parr S, Greenstein J, Stewart D. 2013. How should MDG success and failure be judged: faster progress or achieving the targets? *World Devel.* 41:19–30. doi: [10.1016/j.worlddev.2012.06.014](https://doi.org/10.1016/j.worlddev.2012.06.014).
- Gaspar V, Amaglobeli D, Garcia-Escribano M, Prady D, Soto M. 2019. *Fiscal policy and development: human, social, and physical investment for the SDGs*. Washington, D.C: IMF.
- Guo H, Liang D, Sun Z, Chen F, Wang X, Li J, Shirazi Z, Bian J, Wei Y, Huang L. 2022. Measuring and evaluating SDG indicators with big earth data. *Sci Bull.* 67(17):1792–1801. doi: [10.1016/j.scib.2022.07.015](https://doi.org/10.1016/j.scib.2022.07.015).
- Hák T, Janoušková S, Moldan B. 2016. Sustainable development goals: a need for relevant indicators. *Ecol Indic.* 60:565–573. doi: [10.1016/j.ecolind.2015.08.003](https://doi.org/10.1016/j.ecolind.2015.08.003).
- Hickel J. 2019. The contradiction of the sustainable development goals: growth versus ecology on a finite planet. *Sustain Devel.* 27(5):873–884. doi: [10.1002/sd.1947](https://doi.org/10.1002/sd.1947).
- Hickmann T. 2021. Locating cities and their governments in multi-level sustainability governance. *Polit Governance.* 9(1):211–220.
- Hickmann T, Biermann F, Sénit CA, Sun Y, Bexell M, Bolton M, Weiland S, Censoro J, Charles A, Coy D. 2024. Scoping article: research frontiers on the governance of the sustainable development goals. *Glob Sustain.* 7:1–26. doi: [10.1017/sus.2024.4](https://doi.org/10.1017/sus.2024.4).
- Hosking JRM. 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J R Stat Soc, Ser B.* 52(1):105–124.
- ICJ. 2024. Application of the convention on the prevention and punishment of the crime of genocide in the gaza strip (South Africa v. Israel). [accessed Dec 16]. <https://www.icj-cij.org/sites/default/files/case-related/192/192-20240126-ord-01-00-en.pdf>.
- Independent Group of Scientists appointed by the Secretary-General. 2023. *Global sustainable development report 2023: times of crisis, times of change: science for accelerating transformations to sustainable development*. New York: United Nations.
- Kanbur R, Patel E, Stiglitz J. 2018. Sustainable development goals and measurement of economic and social progress. *For Good Measure: Advancing Res On Well-Being Metrics Beyond GDP.* 33–49.
- Kartal MT. 2022. The role of consumption of energy, fossil sources, nuclear energy, and renewable energy on environmental degradation in top-five carbon producing countries. *Renewable Energy.* 184:871–880. doi: [10.1016/j.renene.2021.12.022](https://doi.org/10.1016/j.renene.2021.12.022).
- Kartal MT. 2023. Production-based disaggregated analysis of energy consumption and CO2 emission nexus: evidence from the USA by novel dynamic ARDL simulation approach. *Environ Sci Pollut Res.* 30(3):6864–6874. doi: [10.1007/s11356-022-22714-5](https://doi.org/10.1007/s11356-022-22714-5).
- Kartal MT. 2024. Impact of environmental tax on ensuring environmental quality: quantile-based evidence from G7 countries. *J Cleaner Production.* 440:140874. doi: [10.1016/j.jclepro.2024.140874](https://doi.org/10.1016/j.jclepro.2024.140874).
- Khatib R, McKee M, Yusuf S. 2024. Counting the dead in Gaza: difficult but essential. *Lancet.* 404(10449):237–238. doi: [10.1016/S0140-6736\(24\)01169-3](https://doi.org/10.1016/S0140-6736(24)01169-3).
- Kriebel D, Tickner J, Epstein P, Lemons J, Levins R, Loechler EL, Stoto M, Rudel R, Schettler T, Stoto M. 2001. The precautionary principle in environmental science. *Environ Health Perspect.* 109(9):871–876.
- Li Y, Zhou T, Jia N, Chen R. 2024. Investigating the decoupling effects between global sustainable development and multi-disaster crises. *Int J Sustain Devel World Ecol.* 1–14. doi: [10.1080/13504509.2024.2436916](https://doi.org/10.1080/13504509.2024.2436916).
- Liu J, Hull V, Godfray HCJ, Tilman D, Gleick P, Hoff H, Li S, Xu Z, Chung MG, Sun J. 2018. Nexus approaches to global sustainable development. *Nat Sustain.* 1(9):466–476. doi: [10.1038/s41893-018-0135-8](https://doi.org/10.1038/s41893-018-0135-8).
- MacFeely S. 2019. The big (data) bang: Opportunities and challenges for compiling SDG indicators. *Global Policy.* 10:121–133.
- MacFeely S. 2020. Measuring the sustainable development goal indicators: an unprecedented statistical challenge. *J Off Stat.* 36(2):361–378. doi: [10.2478/jos-2020-0019](https://doi.org/10.2478/jos-2020-0019).
- Malekpour S, Allen C, Sagar A, Scholz I, Å P, Miranda JJ, Al-Ghanim K, Dube OP, Kanie N, Madise N. 2023. What scientists need to do to accelerate progress on the SDGs. *Nature.* 621(7978):250–254. doi: [10.1038/d41586-023-02808-x](https://doi.org/10.1038/d41586-023-02808-x).
- Movilla-Pateiro L, Mahou-Lago XM, Doval MI, Simal-Gandara J. 2021. Toward a sustainable metric and indicators for the goal of sustainability in agricultural and food production. *Crit Rev Food Sci Nutr.* 61(7):1108–1129. doi: [10.1080/10408398.2020.1754161](https://doi.org/10.1080/10408398.2020.1754161).
- Ozcan B, Depren SK, Kartal MT. 2024. Impact of nuclear energy and hydro electricity consumption in achieving environmental quality: evidence from load capacity factor by quantile based non-linear approaches. *Gondwana Res.* 129:412–424. doi: [10.1016/j.gr.2023.05.014](https://doi.org/10.1016/j.gr.2023.05.014).
- Sachs J. 2015. *The age of sustainable development*. New York: Columbia University Press.
- Sachs J, Lafortune G, Fuller G, Drumm E. 2023. *Sustainable development report 2023. Implementing the SDG stimulus*. Cambridge: Cambridge University Press.
- Sachs J, Schmidt-Traub G, Mazzucato M, Messner D, Nakicenovic N, Rockström J. 2019. Six transformations to achieve the sustainable development goals. *Nat Sustain.* 2(9):805–814. doi: [10.1038/s41893-019-0352-9](https://doi.org/10.1038/s41893-019-0352-9).
- Schmidt-Traub G, Kroll C, Teksoz K, Durand-Delacre D, Sachs J. 2017. National baselines for the sustainable development goals assessed in the SDG index and dashboards. *Nat Geosci.* 10(8):547–555. doi: [10.1038/ngeo2985](https://doi.org/10.1038/ngeo2985).
- Scholkopf B, Smola A, Williamson R, Bartlett PL. 2000. New support vector algorithms. *Neural Computation.* 12(5):1207–1245. doi: [10.1162/089976600300015565](https://doi.org/10.1162/089976600300015565).
- Siddiqi S, Aftab W, Siddiqui FJ, Huicho L, Mogilevskii R, Friberg P, Bhutta ZA, Causevic S, Khamis A, Shah MM. 2020. Global strategies and local implementation of health and health-related SDGs: lessons from consultation in countries across five regions. *BMJ Glob Health.* 5(9):e002859.

- Stiglitz JE, Sen A, Fitoussi JP. 2009. Report by the commission on the measurement of economic performance and social progress.
- Swain R.B. 2018. A critical analysis of the sustainable development goals. In: Leal Filho W., editor. Handbook of sustainability science and research. World sustainability series. Cham: Springer. doi: [10.1007/978-3-319-63007-6_20](https://doi.org/10.1007/978-3-319-63007-6_20).
- Swinkels L, van Zanten JA, Rein B, Scholten R. 2024. Financing the sustainable development goals: exploring the role of government bond investors; [Accessed 2024 Dec 3]. doi: [10.2139/ssrn.4365013](https://doi.org/10.2139/ssrn.4365013).
- UN. 2023. The sustainable development goals report 2023 – special edition. New York: United Nations.
- UN. 2024. The Sustainable Development Goals Report 2024. New York: United Nations.
- UNGA. 2015. Transforming our world: the 2030 agenda for sustainable development Doc (New York: United Nations). A/RES/70/1.
- van Soest HI, van Vuuren DP, Hilaire J, Minx JC, Harmsen MJ, Krey V, Popp A, van Soest HL, Luderer G, Riahi K, Luderer G. 2019. Analysing interactions among sustainable development goals with integrated assessment models. *Global Transitions*. 1:210–225. doi: [10.1016/j.glt.2019.10.004](https://doi.org/10.1016/j.glt.2019.10.004).
- van Zanten JA, van Tulder R. 2021. Towards nexus-based governance: defining interactions between economic activities and sustainable development goals (SDGs). *Int J Sustain Devel World Ecol*. 28(3):210–226. doi: [10.1080/13504509.2020.1768452](https://doi.org/10.1080/13504509.2020.1768452).
- Weiland S, Hickmann T, Lederer M, Marquardt J, Schwindenhammer S. 2021. The 2030 agenda for sustainable development: transformative change through the sustainable development goals? *Polit Governance*. 9 (1):90–95.
- Weitz N, Carlsen H, Nilsson M, Skånberg K. 2018. Towards systemic and contextual priority setting for implementing the 2030 Agenda. *Sustainability Sci*. 13:531–548.
- Xie H, Wen J, Choi Y. 2021. How the SDGs are implemented in China——A comparative study based on the perspective of policy instruments. *J Cleaner Production*. 291:125937. doi: [10.1016/j.jclepro.2021.125937](https://doi.org/10.1016/j.jclepro.2021.125937).
- Zeng Y, Maxwell S, Runting RK, Venter O, Watson JE, Carrasco LR. 2020. Environmental destruction not avoided with the sustainable development goals. *Nat Sustain*. 3 (10):795–798. doi: [10.1038/s41893-020-0555-0](https://doi.org/10.1038/s41893-020-0555-0).
- Zhang T, Ramakrishnan R, Livny M. 1996. BIRCH: an efficient data clustering method for very large databases. *SIGMOD Rec*. 25(2):103–114.