communications sustainability

A Nature Portfolio journal

Article

https://doi.org/10.1038/s44458-025-00006-6

Stronger institutional performance correlates with ecological effectiveness

Check for updates

Tobias Böhmelt

The literature distinguishes between the *institutional* and *ecological effectiveness* of international environmental organizations. The former refers to the compliance with stated objectives, the latter is about measurable improvements in environmental quality. The relationship of both concepts – and, thus, overall effectiveness – remains insufficiently understood, however. Here, we address that gap by examining the impact of institutional effectiveness on ecological performance. Leveraging data on international environmental organizations between 2008 and 2018, the findings reveal a clear link: stronger *institutional* performance correlates with *ecological* effectiveness as measured by reduced particulate pollution. By shedding light on the conditions under which international organizations drive meaningful environmental change, this research offers new insights for strengthening global environmental governance.

This study investigates the relationship between the *institutional* and *ecological* effectiveness of international environmental organizations, drawing on the latest data and employing rigorous econometric techniques. Global environmental challenges such as climate change rank among the most pressing policy issues of our time[1,2]. International institutions are widely regarded as the most promising instruments in global governance for addressing these challenges[3–6]. A substantial body of research explores these organizations' effectiveness and its determinants[7–14].

Miles et al.[15] define effectiveness as "the successful performance of a certain set of function(s) or the solution of problems that motivated the establishment of the institution in the first place"[7,16]. Their definition is arguably the most general and widely used in the study of international environmental organizations' effectiveness. Beyond this conceptual foundation, Miles et al.[15] provide a series of in-depth qualitative case studies examining the key drivers of performance: problem structure (e.g., whether an environmental issue is characterized as a public good or common pool resource) and organizational design (e.g., whether mechanisms are in place that allow to monitor countries' behavior and enforce agreement terms) are central determinants of effectiveness—an insight that has since been explored in several analyses[17–22]. While these and other studies have advanced our understanding of international environmental organizations' effectiveness, the complexities of *overall* institutional performance remain only partially understood[7,22,23].

Specifically, the literature[22–27] disaggregates *overall* impact and differentiates theoretically between *institutional* and *ecological effectiveness*, both of which together constitute an organization's *overall effectiveness*. Jackson and Bührs[23] define institutional effectiveness as implementation of and compliance with an organization's rules and regulations, whereas ecological effectiveness pertains to positive biophysical impact in the form of measurable environmental improvements at the outcome level, e.g., positive conservation impacts, enhanced air quality, reduced atmospheric emissions, or expanded protected areas[28]. These two dimensions of overall effectiveness are thus well defined, and while it is reasonable to expect that institutional effectiveness fosters greater ecological effectiveness[7], there is no guarantee given the self-selection processes at work in this context[22,29,30]: even if an environmental organization meets its stated objectives and implements its goals (institutional effectiveness), these could have been deliberately softened to begin with[31] and, hence, we do not know whether an institution "actually has a positive impact on the environment: is the natural environment better because *of* international regulatory efforts than it would have been otherwise"[32]? Jacobson and Brown Weiss[33] echo this: states could achieve full compliance with an institution, but it may nevertheless not be able to bring about improvements in environmental quality as, for example, states could simply agree on targets that are easy to implement and comply with, but have little to no impact on improving the environment. In the following, I shed more light on this relationship as I provide systematic empirical evidence on how institutional effectiveness shapes ecological effectiveness.

The main contribution of this article is thus to provide a deeper understanding of the links between institutional and ecological effectiveness. En route, I address organizational design and account for countries' self-selection into institutions[22,29,30]. If international environmental organizations demonstrate strong institutional performance yet fail to improve environmental quality on the ground, their "nature and merit" will inevitably be questioned, along with "whether and how they can be changed to enhance their ecological effectiveness"[7,23].

Jackson and Bührs[7] identify two key challenges that have hindered thorough assessment of the institutional-ecological effectiveness nexus.

University of Essex, Colchester, UK. e-mail: tbohmelt@essex.ac.uk

First, there is a lack of data and ambiguity persists regarding how to measure the two components of overall effectiveness[33]. Existing accounts are rather of a theoretical nature, and there is no systematic empirical analysis of how institutional and ecological effectiveness are related to each other. Second, establishing causality between an institution's actions and actual environmental improvements remains a hurdle.

This study seeks to address both challenges by providing an empirical analysis that produces estimates that are likely causal, thus aiming at facilitating our understanding of environmental organizations' overall performance and the conditions under which they can drive meaningful environmental change. The novelty of my work lies in particular in identifying the magnitude of institutional effectiveness on ecological effectiveness, using recently compiled data and by using a rigorous empirical strategy. By offering insights for scholars and policymakers alike, it seeks to enhance global environmental governance, and I offer findings that potentially add to the literature on organizational theory and management: research on the performance of government organizations[34], bureaucratic effectiveness at local-level institutions[35], or, generally, the drivers behind public management efficacy[36], though situated in different contexts, may benefit from the insights I present below.

The first issue, i.e., data quality and measurement ambiguity, is mitigated by recently compiled data covering a range of organizations and ecological performance indicators. Lall[24–26] has collected an extensive data set on institutional performance, which includes six institutions relevant to the environmental domain. These can be directly linked to the ecological performance of their target or program states, for which I draw on van Donkelaar et al.[37] who provide data on $PM_{2.5}$, i.e., fine particulate matter (≤2.5 micrometers) known for its severe health risks. To address the second issue as identified by Jackson and Bührs[23], I employ a rigorous empirical strategy incorporating several modeling approaches in the main text and in the Supplementary Information (SI): two-way fixed effects OLS regressions, a within-between estimator[38], and a general error correction model[39,40]. While each method has certain limitations and cannot fully resolve concerns over causal inference, their combined application instills a higher degree of confidence in the causal nature of the findings. In addition, I present sample-selection models (Supplementary Notes 7, Supplementary Tables 7–8) in the SI, a placebo test (Supplementary Notes 8, Supplementary Table 9), as well as a fixed-effects counterfactual estimator (Supplementary Notes 9, Supplementary Table 10) developed by Liu et al.[41], especially the latter allows for a more direct causal interpretation.

The results presented below and the SI demonstrate that institutional effectiveness is positively associated with and, thus, promotes ecological effectiveness, and this is particularly important as their relationship may well be negative or only weakly pronounced due to the underlying selection problem, e.g., countries that join environmental institutions, especially those with higher effectiveness scores, might be more environmentally committed to begin with[22,29,30].

Results

To examine the relationship between institutional and ecological effectiveness, I compiled quantitative data on six international environmental organizations and their program or target states in 2008–2018 using information from Lall[24–26] and van Donkelaar et al.[37]. I combine these data with standard controls in the literature on environmental quality and I also disaggregate institutional performance by design features. In the main text, I present the results of three different estimation techniques to assess the impact of organizations' institutional performance on ecological effectiveness: two-way fixed effects models, the within-between estimator[38], and a general error correction model[39,40]. In the Methods section, I discuss in detail the data, variables, and estimation strategies used. In the SI, I assess the robustness of my main finding when employing several alternative research design specifications, which includes the explicit consideration of states' self-selection into international organizations (Supplementary Notes 7, Supplementary Tables 7–8) and additional (causal) estimation techniques (Supplementary Notes 8–9, Supplementary Tables 9–10).

Analysis I: two-way fixed effects models

The results of the two-way fixed effects regression models are reported in Table 1. Due to the inclusion of fixed effects for countries and years, these models capture within-country effects, i.e., how ecological effectiveness changes within a state if the environmental organizations that state is targeted by or is a program member of become more institutionally effective. I consider two different units of analysis: Models 1–4 are based on the pairing of each environmental organization with their program/target states; hence,

Table 1 | Two-way fixed effects models

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Lagged Dependent Variable		0.309***	0.313***			0.191***
		(0.050)	(0.051)			(0.067)
Institutional Effectiveness	−0.189***	−0.197***	−0.192***	−0.233***	−0.687***	−0.898***
	(0.053)	(0.042)	(0.043)	(0.046)	(0.202)	(0.201)
Democracy			2.836**			2.303**
			(1.326)			(1.155)
GDP per capita			−44.483*			−10.335
			(25.705)			(31.084)
GDP per capita^2			5.761*			1.479
			(3.235)			(3.837)
GDP per capita^3			−0.242*			−0.066
			(0.133)			(0.156)
Population			1.933			1.541
			(2.076)			(2.149)
Globalization			−0.042*			−0.058**
			(0.024)			(0.023)
Observations	2137	1960	1685	1685	1254	947

Table entries are coefficients; standard errors clustered on country in parentheses; constant, year fixed effects, and country fixed effects included in all models, but omitted from presentation.
* p <0.10; ** p <0.05; *** p <0.01.

I use a dyadic data structure for these estimations. In Models 5–6, I employ monadic data with countries as the unit of analysis: institutional performance scores are averaged across organizations per program/target state. As a result, Models 5–6 have fewer observations than Models 1–4. All models, regardless of the unit of analysis, however, clearly show that institutional effectiveness is significantly and substantively associated with lower particle pollution and, thus, increased ecological effectiveness. On average, the estimated coefficient of *Institutional Effectiveness* in Models 1–4 is –0.203, which translates into a reduction of $PM_{2.5}$ by 0.203 units for every unit increase in an organization's institutional performance. As $PM_{2.5}$ ranges between 3.5 and 83.1 in my sample, the estimated impact is rather small in substance, though. The overall effect somewhat increases in size to –0.793 (on average) in Models 5 and 6, but this is driven by the change in the unit of analysis (monadic data structure). The control variables are discussed in the Methods section.

These effects are further substantiated by the predicted values of *Ecological Effectiveness* (Fig. 1): moving from the minimum to the maximum of *Institutional Effectiveness*, the predicted values of *Ecological Effectiveness* decrease from about 26.3 to almost 25.5. These effects are statistically significant. I also simulated[42,43] the impact of *Institutional Effectiveness* 1000 times. The mean value of the simulated marginal effect of *Institutional Effectiveness* (–0.197 for the dyadic data and –0.781 for the monadic data) is close to what is reported in Table 1. Only 2 out of 1000 simulations (0.2 percent) are positively signed when using the dyadic data structure, while all 1000 simulations exhibit a negative effect according to Models 5–6. Hence, 99.8–100 percent of my simulations produce a negatively signed coefficient for *Institutional Effectiveness*, which follows from the regression results that all display large *t*-statistics.

In summary, the results presented in Table 1 and Fig. 1 consistently point to the conclusion that institutional effectiveness increases ecological effectiveness. Having said that, although the two-way fixed effects models underlying these results provide conservative estimates, they cannot assess the between-country effects, i.e., the differences in ecological effectiveness between countries, and have their limitations in capturing short and long-term effects with precision.

Analysis II: within-between estimator

I thus continue with Table 2, which summarizes the findings of a within-between random effects model[38]. I control for country and year fixed effects as well as the lagged dependent variable, but leave out the substantive controls for reasons of parsimony. Including the items for income, population, democracy, and globalization does not affect the core finding, as demonstrated in Table 1 already. For the within-between estimator, I consider variables both in their time-varying form and their over-time average. I thus differentiate within-country over-time effects and between-country effects. Examining Table 2, I obtain evidence for a within-effect, which mirrors the results of Table 1. The effect size is similarly strongly pronounced as in Models 5–6, as Model 7 is also based on a monadic data structure. The between-effect of *Institutional Effectiveness* is not statistically significant, though. Given the results in Table 2, while considering Models 1–6 as well, I obtain additional evidence that institutional effectiveness increases ecological effectiveness, but this effect predominantly occurs within (not between) countries.

Analysis III: general error correction model

Table 3 summarizes the results of a general error correction model to temporally disaggregate the within-country effect. To this end, the dependent variable of Model 8 captures inter-annual *shifts* in *Ecological Effectiveness* (not *levels* of $PM_{2.5}$), while the variable *Institutional Effectiveness* enters the estimation in two forms: as an inter-annual shift in performance and as a temporally lagged level indicator. This allows to uncover both the immediate and the long-term impact of environmental organizations' institutional performance on *Ecological Effectiveness*. According to Model 8 (the substantive controls are left out again although their inclusion does not qualitatively affect the results), there is evidence for a short-term (contemporaneous) and a long-term effect. The short-term marginal effect is estimated at −0.606 (statistically significant at the 5 percent level), while the long-term marginal effect is calculated at −0.919 (statistically significant at the 1 percent level).

Figure 2 leverages the dynamics of the error correction model to plot the (long-term and short-term) effects of institutional effectiveness on ecological effectiveness over time. Model 8 provides the baseline for the calculations to display a forecast of ecological effectiveness when institutional performance is set to a low or high value. Low institutional effectiveness is defined as the 5th percentile of that variable (–0.457 for inter-annual changes and –1.982 for the level of performance); high institutional effectiveness is defined as the 95th percentile of that variable (2.026 for inter-annual changes and 0.612 for the level of performance). For simulating the effects, the initial value for *Ecological Effectiveness* is set to 25.79 (lagged sample mean).

As predicted, an improvement in institutional effectiveness (red bars) lowers $PM_{2.5}$ pollution and, thus, enhances ecological effectiveness steadily over time. The long-term effect is arguably more strongly pronounced, according to Table 2, which is demonstrated by the more substantive reduction effect on pollution over the years (a 10-year window is used for the simulations). Conversely, lower institutional performance (blue bars) is also linked to a decrease in ecological effectiveness as, over time, particulate ($PM_{2.5}$) pollution increases. As the confidence intervals mostly overlap in the "weak-performance scenario" though, there is no statistically significantly different impact across short-term and long-term influences there.

Analysis IV: disaggregation of institutional performance

While the previous analyses establish a positive impact of institutional effectiveness on ecological performance, particularly from a policy perspective could it be of interest which organizational design features drive this result. To this end, I draw on Lall[24] who establishes what the most important factors behind institutional performance are and use these factors instead of *Institutional Effectiveness* in Table 4. As in Table 1, the first two models of Table 4 are based on a dyadic data structure, while Models 11 and 12 rely on monadic-level data.

Lall[24] theoretically and empirically argues for a disaggregation of institutional effectiveness by de facto/de jure autonomy, income of an organization, its partnerships with other stakeholders, and organizational age. I introduce the variables on de facto autonomy and de jure autonomy separately in the models of Table 4 due to their conceptual and empirical overlap. Table 4 demonstrates, also in line with the findings in Lall[24], that especially *De Facto Policy Autonomy* drives my core result. The more autonomous an organization, the less it is tied to state interests, and the more it can shape ecological outcomes more effectively. The item *De Jure Policy Autonomy* is also negatively signed and statistically significant, but its effect is smaller in size than the impact of *De Facto Policy Autonomy*. The financial capabilities of an organization seem to matter, too, as the corresponding variable consistently yields negative and significant effects in Table 4. The other variables, *Partnerships* and *Age*, display inconsistent effects (sign and/or significance) across Models 9-12.

In the SI, I present numerous additional analyses, which, among others, discuss potential confounding factors, provide sensitivity tests, employ alternative estimation procedures, and rule out overfitting of the models. First, I introduce the six organizations of the data set and list the sources for each institution's donor/contributing countries as well as target/program states (Supplementary Notes 1, Supplementary Table 1). Second, I consider three alternative dependent variables (Supplementary Notes 2, Supplementary Table 2). In Supplementary Notes 3 (Supplementary Table 3), I employ a panel-corrected standard-error model. In Supplementary Notes 4 (Supplementary Table 4), I disaggregate institutional performance by quartiles. I re-estimate the main model when excluding certain organizations from the sample in Supplementary Notes 5, Supplementary Table 5. In Supplementary Notes 6, Supplementary Table 6, I control for domestic-level climate policy measures. I also explicitly control for countries' self-selection

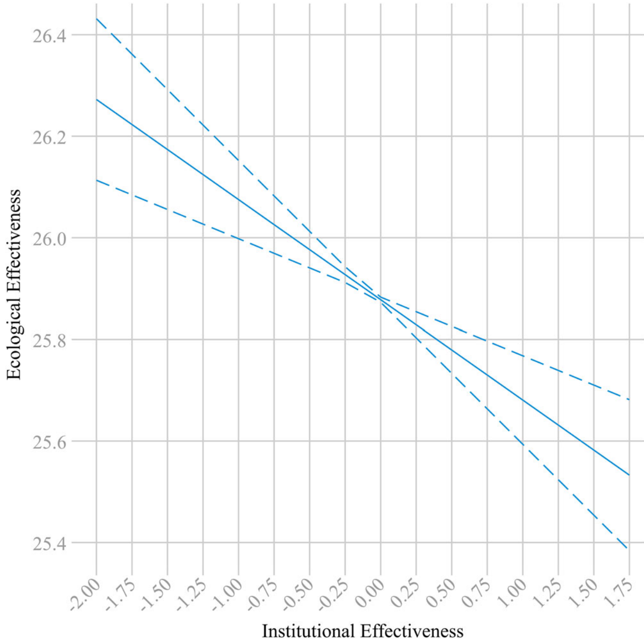

Fig. 1 | Predicted values of *Ecological Effectiveness*—two-way fixed effects. *Notes:* Graph shows the predicted values of *Ecological Effectiveness*. The figure is based on Model 2, which comprises a lagged dependent variable, fixed effects for years, and fixed effects for countries. The dashed lines pertain to the 95 percent confidence interval.

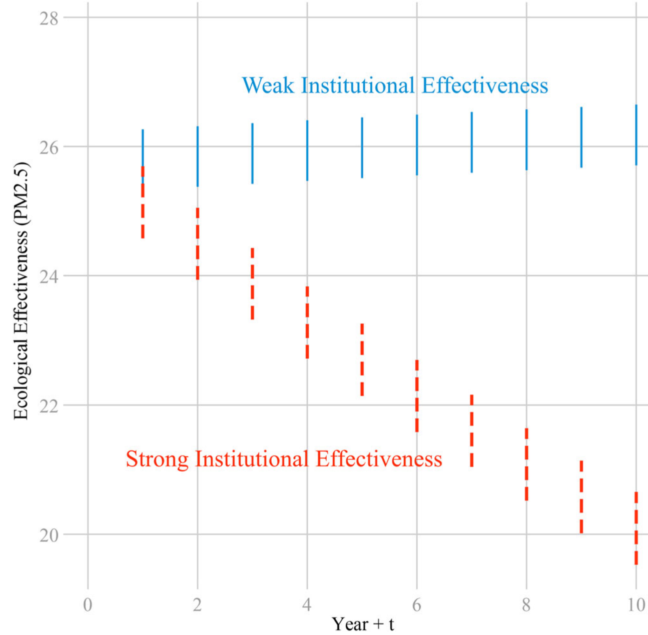

Fig. 2 | Dynamic simulation of *Ecological Effectiveness*. *Notes:* Graph displays expected values of *Ecological Effectiveness* when *Institutional Effectiveness* is set to high and low values. Estimates are based on Model 8 ($N = 1000$ simulations). Vertical bars report 95% confidence intervals.

Table 2 | Within-between random-effects estimator

	Model 7
Lagged Dependent Variable (within)	0.201***
	(0.038)
Lagged Dependent Variable (between)	1.050***
	(0.172)
Institutional Effectiveness (within)	–0.902***
	(0.247)
Institutional Effectiveness (between)	–0.650
	(22.714)
Observations	947

Table entries are coefficients; standard errors in parentheses; constant, year fixed effects, and country fixed effects included, but omitted from presentation.
*** p <0.01.

Table 3 | General error correction model

	Model 8
Lagged Dependent Variable	–1.095***
	(0.074)
Δ Institutional Effectiveness	–0.606**
	(0.237)
Lagged Institutional Effectiveness	–0.919***
	(0.356)
Observations	824

Table entries are coefficients; standard errors clustered on country in parentheses; constant, year fixed effects, and country fixed effects included, but omitted from presentation.
** p <0.05; *** p <0.01.

Table 4 | Two-way fixed effects models

	Model 9	Model 10	Model 11	Model 12
Lagged Dependent Variable	0.310***	0.310***	0.186***	0.188***
	(0.050)	(0.050)	(0.064)	(0.064)
De Facto Policy Autonomy	–0.210**		−1.797*	
	(0.080)		(1.017)	
De Jure Policy Autonomy		–0.084***		–1.364**
		(0.032)		(0.647)
Income	–0.154***	–0.112***	–0.949**	–0.634*
	(0.044)	(0.036)	(0.373)	(0.334)
Partnerships	–0.107*	0.133*	–1.623	0.280
	(0.063)	(0.078)	(1.372)	(0.747)
Age	–0.111**	−0.012	–0.225	1.267*
	(0.054)	(0.035)	(0.574)	(0.717)
Observations	1960	1960	1122	1122

Table entries are coefficients; standard errors clustered on country in parentheses; constant, year fixed effects, and country fixed effects included in all models, but omitted from presentation.
* p <0.10; ** p <0.05; *** p <0.01.

into organizations[22,29,30] using a two-part model and a Heckman selection model (Supplementary Notes 7, Supplementary Tables 7–8). Furthermore, I implement a placebo test (Supplementary Notes 8, Supplementary Table 9) and a newly developed counterfactual model (Supplementary Notes 9, Supplementary Table 10). The results in the SI continue to support the main findings presented above, and they lend further support to the argument developed and the findings presented above.

Discussion

How can we thoroughly assess the effectiveness of international environmental organizations? How do institutional and ecological effectiveness intersect? Despite their critical relevance to the academic discourse on global environmental governance and for policymakers, comprehensive answers to these questions have remained elusive. This article seeks to bridge that gap, contributing to the literature by addressing a key challenge: systematically identifying an association between higher institutional effectiveness and greater ecological effectiveness.

Jackson and Bührs[23] have identified major obstacles to resolving this question: the scarcity of high-quality data and measurement ambiguity, as well as issues of causality. By leveraging the most complete, high-quality data on ecological[37] and institutional[24–26] effectiveness, and employing distinct econometric techniques in the main text and the SI, we have now gained new insights into organizations' effectiveness from both a theoretical and an empirical point of view. I have also shown that de facto autonomy and financial strength, being core drivers of institutional effectiveness, matter most for ecological effectiveness. From a policy perspective, this insight is particularly consequential: plans to reduce financial contributions to global environmental organizations, such as those proposed by the Trump administration, would, according to this study's results, diminish ecological effectiveness and, thereby, degrade environmental quality.

Another academic contribution of this research is its validation of Lall's[24–26] data. If his indicator failed to precisely measure institutional performance, the statistical results presented here would not have emerged. This underscores the value Lall's[24–26] data as a foundation for future research on international organizations, their institutional design, and their effectiveness. Updating these data as new source information becomes available will be an endeavor of considerable importance.

Future research could further explore the broader implications of institutional effectiveness – for other environmental institutions or in domains beyond environmental governance. Additionally, the six international organizations examined in this study primarily operate in lower-income countries. Investigating whether the relationship between institutional and ecological effectiveness holds in higher-income contexts would be a compelling avenue for future inquiry.

Methods

I have compiled a time-series cross-sectional data set using data from Lall's[24–26] "Performance of International Institutions Project (PIIP)." The final version of the data set and all replication materials are available at the Harvard Dataverse[44]. Lall's[24] defines institutional performance as "the extent to which they [organizations] achieve their stated objectives and do so in a manner that is cost-effective and responsive to a wide range of (public and private) stakeholders." This definition, especially the first part, largely mirrors the definition in Jackson and Bührs[23] who focus on implementation and compliance. These data make use of stakeholders' reports and ratings, e.g., the Multilateral Aid Review of the UK Department for International Development (DFID), the Australian Agency for International Development as well as the Danish, Dutch, and Swedish Ministries of Foreign Affairs, to code international organizations' institutional performance.

In more detail, based on the general definition of institutional effectiveness outlined in the main text, Lall[24–26] develops a measure that incorporates the views of a variety of groups, supplements these with "objective" data, and is multi-dimensional in nature. First, the different views are incorporated as the DFID review considers the Multilateral Organization Performance Assessment Network survey, the Survey on Monitoring the Paris Declaration, the Heavily Indebted Poor Countries Capacity Building Project evaluations, as well as stakeholder consultations, workshops, interviews, and written submissions. Reports from the other stakeholders supplement this. Second, "objective" information is given by quantitative data from multiple sources, including the Quality of Official Development Assistance Assessment, the Publish What You Fund Aid Transparency Index, the Heavily Indebted Poor Countries Capacity Building Project, and the Common Performance Assessment System. Third, multi-dimensionality is taken into account as Lall[24–26] focuses on institutions' achievements of stated objectives, their responsiveness to stakeholders, and their cost-effectiveness. Against this background, Lall[24–26] suggests coding six performance indicators: (1) delivery of results, (2) contribution to meeting the international community's objectives, (3) cost and value consciousness, (4) financial resources management, (5) accountability and transparency, and (6) strategic/performance management. The first two variables pertain to the first dimension (goal attainment), the third and fourth items stand for the second dimension (cost-effectiveness), and the last two variables capture the third dimension (responsiveness to diverse stakeholders).

While the codings in Lall[24] are based on earlier reports like the 2011 DFID report, the DFID and other stakeholders issued updates of their reports, which Lall[25,26] used to update his original data. Eventually, I use Lall's[24–26] most recent data and merge the six performance indicators (Australian, Danish, MOPAN, Dutch, Swedish, and British) into a general institutional effectiveness variable using principal component analysis, as it is done in Lall[24]. The final item is represented by scores on the first principal component in the analysis, and this is the institutional effectiveness variable used in the main text. It is continuously scaled and originally ranges in [-5.95; 3.65], with higher values standing for more institutional effectiveness.

While Lall[24–26] codes more than 40 international organizations, only six of them pertain to the environmental sector: the Adaptation Fund, the Climate Investment Funds, the Global Environment Facility, the Least Developed Countries Fund, the Multilateral Fund for the Implementation of the Montreal Protocol, and the United Nations Environment Program. For most of these organizations, I have data in every year between 2008 and 2018. The unit of analysis in my final data is either the organization-country year or the country-year. Specifically, first, for the former, each of the six organizations is paired with each program or target state, i.e., those countries that are targeted by the institutions for project implementation. For example, the Climate Investment Funds distinguish between contributor and program countries. There are 15 contributor countries, which are all higher-income (and mostly Western democratic) countries; additionally, there are more than 80 (mostly lower-income) program countries—states that do not contribute financially to the Climate Investment Funds, but the organization targets them for the implementation of policies. I thus pair environmental organizations only with the *program states* in the case of the Climate Investment Funds and employ the same approach for the other five organizations. Note that the vast majority of member/target cases in my sample is indeed less developed and has a lower-than-average income. Second, for the country-year analysis, I collapse the organization-country year cases and calculate the average institutional effectiveness score across all organizations a country belongs to the program countries (i.e., it is a target state). More information on the organizations and the member states can be found in the SI.

The dependent variable in the main text's estimation pertains to ecological effectiveness, which I capture via the concentration of $PM_{2.5}$ per country-year, measured in micrograms per cubic meter. $PM_{2.5}$ consists of particles of a size smaller than 2.5 micrometers and has been found to be particularly harmful for human health. This makes this data particularly relevant as a general indicator for overall environmental quality. The data are taken from van Donkelaar et al.[37] who estimate annual surface-level $PM_{2.5}$ through a combination of satellite data and chemical transport models, for a grid size of $0.01 \times 0.01°$.

Van Donkelaar et al.[37] deliver their estimates in two versions. The first one is a population-weighed concentration, which equals the mean of concentrations per grid-cell weighed by its population and divided by the total population of a country-year, whereby it captures the mean exposure of people to a specific $PM_{2.5}$ concentration. The second version of the variable is a geographic mean, which is the average of grid-cell concentrations per country year divided by the number of grid-cells, whereby it captures the average exposure of a geographic unit to a certain $PM_{2.5}$ concentration. My main analysis focuses on the geographic-mean measure, but the results are virtually identical when using the population-weighed item (see SI).

As discussed in detail in van Donkelaar et al.[37], the advantage of this measure is that it avoids potential biases induced through differential in-situ monitoring activity and reporting of air pollution levels across the set of included countries. Having said that, the data from van Donkelaar et al.[37] have potentially several shortcomings. First, data accuracy depends on calibration with ground-based monitors. In areas with few or no in-situ monitors (e.g., large parts of Africa), satellite-model estimates are less reliable. This means that although the bias is reduced, measurement error increases, especially in data-poor regions. Second, satellite retrievals depend on clear-sky observations. Cloud cover, aerosol mixing, and vertical column-to-surface conversion introduce uncertainty, particularly in tropical or mountainous regions. Third, the data are a combination of satellite data and chemical transport models, which require assumptions about emission inventories and involve modeling physical and chemical processes in the atmosphere. Fourth, the data provide concentration levels, not the composition or origin of the pollution. For policy targeting, the data are not sufficient without additional modeling. Van Donkelaar et al.[37] discuss these issues in more detail and conclude that there is "evidence that global-monitored locations tend to be in cleaner regions […], with large measurement gaps in the Global South. Uncertainty estimates exhibit regional consistency with observed differences between ground-based and satellite derived $PM_{2.5}$. The evaluation of uncertainty for agglomerated values indicates that hybrid $PM_{2.5}$ estimates provide precise regional scale representation, with residual uncertainty inversely proportional to the sample size." To address concerns along those lines, I also present models using alternative dependent variables in the SI, most importantly on the consumption-based environmental footprint and greenhouse gas emissions.

I consider a number of standard controls based on the literature of the Environmental Kuznets Curve[45]. First, there are income and population. Both variables are log-transformed and derived from the World Bank Development Indicators. More populous countries have an overall higher demand for energy and burning fossil fuels is necessary for meeting all citizens' demands. According to the World Bank, population is defined as a country's midyear total population, which counts all residents regardless of legal status or citizenship (except for refugees not permanently settled). States usually tend to "become rich first" before "cleaning up later"[46]—economic wealth is generally more important than environmental performance. GDP per capita (in constant 2017 international dollar) is defined as the gross domestic product (GDP)—the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products—divided by midyear population. I control for a non-linear impact by adding income raised to the power of 2 and raised to the power of 3, which follows the argument of the literature of the Environmental Kuznets Curve[45] on a curvilinear impact effect of income on environmental quality. By adding income to the power of 3, I allow for more flexibility in the functional form and the variable's impact than by simply incorporating income squared.

Second, there is there is the Varieties of Democracy (V-Dem) "electoral democracy" index, which comprises five components based on Dahl's[47] polyarchy: an elected executive, free and fair elections, universal suffrage, freedom of association, and freedom of expression. Boese[48] argues that the variable's performance is superior due to the coherence of its definitions, measurement strategies, and aggregation procedures.

My last control variable is *Globalization*, which is the KOF globalization index[49]. A higher value of this variable signifies a greater embeddedness in the global political, economic, and social network. Countries could be subject to transnational influences pushing them to more effective policy outcomes considering their trading partners' efforts[50,51].

The empirical models in the main text are based on three different modelling approaches, including two-way fixed effects OLS regressions, a within-between estimator, and a general error correction model. First, the fixed effects in the first set of regression models (OLS) are located at the level of countries and years, respectively, and thus control for unobserved (time-invariant unit-level) influences and common temporal shocks.

Second, while the approach of a two-way fixed effects setup has many advantages, Imai and Kim[52] propose several alternatives. I thus consider the results of a within-between random effects model[38], which comprises variables both in their time-varying form and their over-time average. This allows to differentiate within-country over-time effects from the between-country effects, i.e., the latter of which addresses differences in climate policies and outcomes between states, which are usually "controlled away" in two-way fixed effects models. For the within-between estimator, each explanatory variable enters the models in two forms: demeaned (time-varying) for the within-country effect and as a country-average, which captures the between-country effect.

Third, I present the results from a general error correction model. This estimation procedure is extremely flexible as I can assess both the immediate and long-term impact of a shock for an explanatory variable on climate policies and outcomes. The dependent variable in this setup is based on shifts (not levels), which addresses potential issues of nonrandom error structures[39]. Moreover, as the core explanatory variables enter the estimation as shifts and temporally lagged variants, I directly identify the contemporaneous impact of a shock to an explanatory variable as well as its cumulative impact[40]. The error correction model also includes year and country fixed effects.

Two-way fixed effects can uncover causal effects if there are no time-varying confounders and if the treatment effect is constant. However, fixed effects alone cannot resolve self-selection bias if the bias arises from heterogeneous treatment effects that are correlated with selection into treatment. This is a well-known limitation: while fixed effects absorb time-invariant unobserved heterogeneity, they do not eliminate selection on unobservables that vary with treatment status. That said, causal effects can still be estimated under the assumption that, conditional on fixed effects and observed covariates, the remaining selection into treatment is uncorrelated with potential outcomes. This corresponds to a "selection-on-observables" assumption within the fixed effects framework[53]. Under these conditions, the models in Table 1 identify an average treatment effect. But if treatment effect heterogeneity is systematically related to treatment assignment beyond what fixed effects can account for, estimates will be biased.

The within-between estimator[38] has the same requirements and underlying assumptions, though it sheds light on the within-unit and between-units effects. Finally, the general error correction model is helpful for identifying short and long-term effects, but can disclose causal relationships only when the variables are cointegrated, the model is well-specified, and exogeneity holds. The specific assumptions relevant for each estimation strategy in order to identify causal effects may be too strong and not applicable. Each method has certain limitations and cannot fully resolve concerns over causal inference on its own, but their combined application allows for a higher degree of confidence in the causal nature of the findings. In addition, I present sample-selection models (Supplementary Notes 7, Supplementary Tables 7–8) in the SI as well as a fixed-effects counterfactual estimator (Supplementary Notes 9, Supplementary Table 10) developed by Liu et al.[41], especially the latter allows for a direct causal interpretation.

Data availability

The replication materials for this article are available online at: https://doi.org/10.7910/DVN/OIVWTJ.

Received: 8 May 2025; Accepted: 30 September 2025;
Published online: 09 January 2026

References

1. Koubi, V. Climate change and conflict. *Annu. Rev. Political Sci.* **22**, 343–360 (2019).
2. Cornell, S. & Gupta, A. *Is climate change the most important challenge of our times? Contemporary Climate Change Debates*. London: Routledge (2019).
3. Biermann, F. & Pattberg, P. Global environmental governance: Taking stock, moving forward. *Annu. Rev. Environ. Resour.* **33**, 277–294 (2008).
4. Biermann, F. & Pattberg, P. (eds.). *Global environmental governance reconsidered*. (Cambridge, MA: MIT Press, 2012).

5. Andonova, L. & Mitchell, R. The rescaling of global environmental politics. *Annu. Rev. Environ. Resour.* **35**, 255–282 (2010).
6. Hale, T. Transnational actors and transnational governance in global environmental politics. *Annu. Rev. Polit. Sci.* **23**, 203–220 (2020).
7. Gutner, T. & Thompson, A. The politics of IO performance: A framework. *Rev. Int. Organ.* **5**, 227–248 (2010).
8. Young, O. Effectiveness of international environmental regimes: Existing knowledge, cutting-edge themes, and research strategies. *Proc. Natl. Acad. Sci.* **108**, 19853–19860 (2011).
9. Grundig, F. & Ward, H. Structural group leadership and regime effectiveness. *Political Stud.* **63**, 221–239 (2015).
10. Böhmelt, T. & Pilster, U. Zur Problematik kollektiven Handelns. Eine quantitative Studie internationaler Umweltregime. *Z. f. ür. Int. Beziehungen* **18**, 63–90 (2011).
11. Johns, C., Thorn, A. & VanNijnatten, D. Environmental regime effectiveness and the North American Great Lakes Water Quality Agreement. *Int. Environ. Agreem.* **18**, 315–333 (2018).
12. Mitchell, R. et al. What we know (and could know) about international environmental agreements. *Glob. Environ. Politics* **20**, 103–121 (2020).
13. Goyes, D. National legislative adoption of international wildlife law after treaty ratification. *Crime. Law Soc. Change* **81**, 143–162 (2023).
14. Helm, C. & Sprinz, D. Measuring the effectiveness of international environmental regimes. *J. Confl. Resolut.* **44**, 630–652 (2000).
15. Miles, E. et al. *Environmental regime effectiveness: Confronting theory with evidence*. (Cambridge. MA: MIT Press, 2001).
16. Böhmelt, T. & Pilster, U. International environmental regimes: Legalisation, flexibility and effectiveness. *Aust. J. Political Sci.* **45**, 245–260 (2010).
17. Hovi, J., Sprinz, D. & Underdal, A. The Oslo-Potsdam solution to measuring regime effectiveness: Critique, response, and the road ahead. *Glob. Environ. Politics* **3**, 74–96 (2003).
18. Steinar A. International regime effectiveness. *The Handbook of Global Climate and Environment Policy*: 304-319 (2013).
19. Breitmeier, H., Arild, U. & Young, O. The effectiveness of international environmental regimes: Comparing and contrasting findings from quantitative research. *Int. Stud. Rev.* **13**, 579–605 (2011).
20. Kim, Y., Tanaka, K. & Matsuoka, S. Institutional mechanisms and the consequences of international environmental agreements. *Glob. Environ. Polit.* **17**, 77–98 (2017).
21. Kim, Y., Tanaka, K. & Matsuoka, S. Environmental and economic effectiveness of the Kyoto Protocol. *PLOS One* **15**, e0236299 (2020).
22. Mitchell, R. The Influence of International Institutions: Institutional Design, Compliance, Effectiveness, and Endogeneity. In *Power, Interdependence and Non-state Actors in World Politics: Research Frontiers*. (ed. Milner, H. V.) 66–83 (Princeton University Press, Princeton, NJ, 2009).
23. Jackson, W. & Bührs, T. International environmental regimes: Understanding institutional and ecological effectiveness. *J. Int. Wildl. Law Policy* **18**, 63–83 (2015).
24. Lall, R. Beyond institutional design: Explaining the performance of international organizations. *Int. Organ.* **71**, 245–280 (2017).
25. Lall, R. The financial consequences of rating international institutions: Competition, collaboration, and the politics of assessment. *Int. Stud. Q.* **65**, 343–359 (2021).
26. Lall, R. *Making International Institutions*. Work: The Politics of Performance. Cambridge University Press: Cambridge, 2023.
27. Mirumachi, N. & Hurlbert, M. Reflecting on twenty years of international agreements concerning water governance: Insights and key learning. *Int. Environ. Agreem.* **22**, 317–332 (2022).
28. Mitchell, R. Evaluating the performance of environmental institutions: What to evaluate and how to evaluate it. *Inst. Environ. Chang. Princ. Find. Appl. Res. Front.* **79**, 83–84 (2008).
29. Lupu, Y. The informative power of treaty commitment: using the spatial model to address selection effects. *Am. J. Political Sci.* **57**, 912–925 (2013).
30. Spilker, G. & Böhmelt, T. The impact of preferential trade agreements on governmental repression revisited. *Rev. Int. Organ.* **8**, 343–361 (2013).
31. Bernauer, T., Kalbhenn, A., Koubi, V. & Spilker, G. Is there a "depth versus participation" dilemma in international cooperation? *Rev. Int. Organ.* **8**, 477–497 (2013).
32. DeSombre, E. *The Global Environment and World Politics*. London: Bloomsbury Publishing (2005).
33. Jacobson, H. K. & Brown Weiss, E. Strengthening compliance with international environmental accords: Preliminary observations from Collaborative Project. *Glob. Gov.* **1**, 119 (1995).
34. Lynn, L. & Robbie, W. R. Governance and organisational effectiveness: towards a theory of government performance. *J. Public Policy* **33**, 201–228 (2013).
35. Krause, R. M., Feiock, R. C. & Hawkins, C. V. The administrative organization of sustainability within local government. *J. Public Adm. Res. Theory* **26**, 113–127 (2016).
36. Tompkins, J.R. *Organization theory and public management*. Long Grove, IL: Waveland Press (2023).
37. Van Donkelaar, A. et al. Monthly global estimates of fine particulate matter and their uncertainty. *Environ. Sci. Technol.* **55**, 15287–15300 (2021).
38. Bell, A. & Jones, K. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Sci. Res. Methods* **3**, 133–153 (2015).
39. Tromborg, M. Bringing the median voter back in. *J. Eur. Soc. Policy* **24**, 107–121 (2014).
40. De Boef, S. & Keele, L. Taking time seriously. *Am. J. Political Sci.* **52**, 184–200 (2008).
41. Liu, L., Wang, Y. & Xu, Y. A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. *Am. J. Political Sci.* **68**, 160–176 (2024).
42. King, G., Tomz, M. & Wittenberg, J. Making the most of statistical analyses: Improving interpretation and presentation. *Am. J. Political Sci.* **44**, 347–361 (2000).
43. Rittmann, O., Neunhoeffer, M. & Gschwend, T. How to improve the substantive interpretation of regression results when the dependent variable is logged. *Political Sci. Res. Methods* **13**, 203–211 (2025).
44. Böhmelt, T. Replication Data for: Stronger institutional performance correlates with ecological effectiveness. *Harv. Dataverse*. https://doi.org/10.7910/DVN/OIVWTJ (2025).
45. Dasgupta, S., Laplante, B., Wang, H. & Wheeler, D. Confronting the Environmental Kuznets Curve. *J. Econ. Perspect.* **16**, 147–168 (2002).
46. Spilker, G. *Globalization, Political Institutions, and the Environment in Lower-income Countries*. (London: Routledge 2013).
47. Dahl, R.A. *Polyarchy: Participation and Opposition*. New Haven, CT: Yale University Press (1971).
48. Boese, V. How (not) to measure democracy. *Int. Area Stud. Rev.* **22**, 95–127 (2019).
49. Dreher, A. Does globalization affect growth? Evidence from a new index of globalization. *Appl. Econ.* **38**, 1091–1110 (2006).
50. Neumayer, E. Does trade openness promote multilateral environmental cooperation? *World Econ.* **25**, 815–832 (2002).
51. Aklin, M. Re-exploring the trade and environment nexus through the diffusion of pollution. *Environ. Resour. Econ.* **64**, 663–682 (2016).
52. Imai, K. & Kim, I. S. On the use of two-way fixed effects regression models for causal inference with panel data. *Political Anal.* **29**, 405–415 (2021).
53. Angrist, J. & Pischke, J.-S. *Mostly harmless econometrics: An empiricist's companion*. (Princeton University Press, Princeton, NJ, 2009).

Author contributions

T.B. designed and executed the analysis, wrote the first draft, and revised the manuscript during the review process.

Competing interests

The author declares no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44458-025-00006-6.

Correspondence and requests for materials should be addressed to Tobias Böhmelt.

Peer review information *Communications Sustainability* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Yann Benetreau. A peer review file is available.