Stronger institutional performance correlates with ecological effectiveness – Supplementary Information

In this Supplementary Information (SI), I present the following additional information, analyses, and robustness checks that complement the main article's argument and findings:

- I introduce the **six organizations in 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).
- 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**.
- Two different estimators are used to address **sample selection** in Supplementary Notes 7, Supplementary Tables 7-8.
- As a **placebo test** (Supplementary Notes 8, Supplementary Table 9), I examine the **ecological performance in organizations' contributing states**.
- In Supplementary Notes 9, Supplementary Table 10, I present the results of a **counterfactual estimator for causal inference**.

Supplementary Notes 1: Organizations in the Data Set

The data set I analyze in the main text and below in the SI comprises information on six international environmental agreements. First, the Adaptation Fund,[1] which "finances projects and programs that help vulnerable communities in developing countries adapt to climate change. Initiatives are based on country needs, views and priorities," and the institution committed about $1.25 billion since 2010. Second, there is

[1] See online at: `https://www.adaptation-fund.org/`.

the Climate Investment Funds,[2] which is one of the largest active climate finance mechanisms in the world. According to the organization, it supports 362 projects in 10 programs in 83 countries. Third, the Global Environment Facility,[3] serves as a financial mechanism for several environmental conventions, including the Convention on Biological Diversity, the United Nations Framework Convention on Climate Change, or the UN Convention to Combat Desertification, among others. Fourth, the Least Developed Countries Fund,[4] was established in 2001 by the United Nations Framework Convention on Climate Change to help least developed states to adapt to climate change. Fifth, the Multilateral Fund for the Implementation of the Montreal Protocol[5] addresses the depletion of the ozone layer. It has secured more than $4.3 billion in grant funding and assisted more than 144 developing states with that money. Finally, the United Nations Environment Programme is is the leading global environmental authority, with a mandate to provide guidance to UN member states on environmental issues. It is thus responsible for coordinating responses to environmental issues for all UN members.[6]

According to this overview, the organizations I analyze comprise donors or contributing members that are predominantly higher income countries, while target/program states come from the Global South and are lower income. The substantive focus of these institutions is primarily on climate change and the pollutants that drive it. While fine particulate matter ($PM_{2.5}$) is not a greenhouse gas itself, it influences the climate through various mechanisms, while its sources are often linked to activities that contribute to greenhouse gas emissions. For example, vehicle exhaust, power plants, and industrial facilities are sources of $PM_{2.5}$ that also release greenhouse gases. Moreover, burning wood and other biomass for energy releases $PM_{2.5}$ and greenhouse gases. As a result, and also as concluded for example by a World Bank report,[7] addressing the sources of $PM_{2.5}$, like coal combustion and traffic, "would address the most toxic air pollution. Given that these sources are also key contributors to climate warming, tackling air pollution from these sources also mitigates climate change."

Supplementary Table 1. Organizations' Donor/Contributing Countries and Target/Program States

	Donor/Contributing Countries	**Target/Program States**
Adaptation Fund	`https://tinyurl.com/bddfrdtj`	`https://tinyurl.com/53btvk39`
Climate Investment Funds	`https://tinyurl.com/bdeyfknn`	`https://tinyurl.com/nhcc6awv`
Global Environment Facility	`https://tinyurl.com/yk49d5y7`	`https://tinyurl.com/3xjrn3s9`
Least Developed Countries Fund	`https://tinyurl.com/2ar225cj`	`https://tinyurl.com/2jybjkrn`
Multilateral Fund	`https://tinyurl.com/4r7dsk3p`	`https://tinyurl.com/4r7dsk3p`
United Nations Environment Programme	`https://tinyurl.com/46ptvnbe`	`https://tinyurl.com/29y6c9vc`

As indicated above, each organization comprises two types of members. On one hand, there are donor or

[2]See online at: `https://www.cif.org/`.
[3]See online at: `https://www.thegef.org/`.
[4]See online at: `https://www.thegef.org/what-we-do/topics/least-developed-countries-fund-ldcf`.
[5]See online at: `https://www.multilateralfund.org/`.
[6]See online at: `https://www.unep.org/`.
[7]See online at: `https://tinyurl.com/3wcbfu9u`.

contributing countries, which provide the funds for the institutions' projects and activities. These members should not directly see an improvement in environmental quality from the organizational activity. On the other hand, there are target or program countries, which are at the receiving end and the places where projects are implemented on the ground. These states should see improvements in environmental quality, especially air pollution and climate change challenges, if there is a causal effect stemming from institutional effectiveness on ecological effectiveness. My analysis focuses on the latter type of institutional members, but I draw upon the sample of donor or contributing countries in a placebo test below. Table 1 in this SI lists my sources for classifying donor/contributing countries as well as target/program states.

Supplementary Notes 2: Alternative Dependent Variables

To capture ecological effectiveness, I rely on 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, as an environmental pollutant, is particularly harmful for human health. The data [1] capture estimations of annual surface-level $PM_{2.5}$ through a combination of satellite data and chemical transport models. 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. In addition, the data provide one of the largest spatio-temporal coverage (countries and years) for a measure of environmental quality at the outcome level, which is also sufficiently general. To assess the robustness of my main finding, however, I consider three alternative outcome measures. First, I rely on the population-weighed concentration [1], which equals the mean of concentrations per grid cell weighed by its population and divided by the total population of a country-year. This variable thus captures the mean exposure of people to a specific $PM_{2.5}$ concentration. The corresponding results are summarized in Model A1 (two-way fixed effects estimator as introduced in the main text).

Second, there is the consumption-based carbon footprint [2,3,4]. The coding descriptions [4,5] state that the carbon footprint "represents the area of forest land required to sequester anthropogenic carbon dioxide emissions." The footprint is calculated "using several parameters including domestic fossil fuel combustion and electricity use, embodied carbon in traded items and electricity, a country's share of global international transport emissions, and non-fossil-fuel sources. The total amount of carbon dioxide allocated to each country is converted into global hectares based on the footprint intensity of carbon. This conversion factor is derived from (a) the yield of the productive land that is required to absorb the carbon dioxide emissions, (b) the amount of carbon absorbed by oceans, (c) an equivalence factor for carbon as a land type, and (d) an adjustment factor for temporal changes in yield from the forest" see also [6]. The consumption focus of the indicator "measures [in global hectares per capita] the human demand on nature by assessing how much biologically productive land and sea area is necessary to maintain a given consumption pattern" [6]. The data are taken from the Global Footprint Network [4,5], which relies on institutions such as the UN, the

FAO, the IPCC, or the IEA as primary sources for the footprint's individual components.[8] This addresses concerns regarding whether data may be systematically missing for countries with limited infrastructure. In addition, the consumption-based carbon footprint addresses directly the causes of climate change, which the six organizations in the data mainly focus on. Model A2 (two-way fixed effects model) presents the corresponding results.

Finally, I explore variation in greenhouse gas emission levels (in kilo tons of CO_2 equivalent, log-transformed) within the territory of each state. The data are provided by the World Development Indicators, which define these emissions as those composed of CO_2 totals excluding short-cycle biomass burning (such as agricultural waste burning and savanna burning), but including other biomass burning (such as forest fires, post-burn decay, peat fires, and decay of drained peatlands), all anthropogenic methane sources, nitrous oxides sources, and fluorinated gases. The greenhouse gas emission variable is conceptually arguably closer to what the six organizations in my data seek to address than the the concentration of $PM_{2.5}$, although the country-year coverage is less strongly pronounced. The quality of the $PM_{2.5}$ data is also likely higher[1]. Model A3 is based on two-way fixed effects as well and summarizes the findings based on the greenhouse gas emissions outcome variable.

Supplementary Table 2. Alternative Dependent Variables

	Model A1	**Model A2**	**Model A3**
Lagged Dependent Variable	0.342***	0.609***	0.589***
	(0.044)	(0.075)	(0.055)
Institutional Effectiveness	−0.198***	−0.006**	−0.004***
	(0.050)	(0.003)	(0.002)
Democracy	3.018**	0.012	0.005
	(1.178)	(0.088)	(0.038)
GDP per capita	−30.442	−5.674**	−7.160***
	(28.539)	(2.671)	(1.793)
GDP per capita^2	4.151	0.709**	0.947***
	(3.574)	(0.355)	(0.235)
GDP per capita^3	−0.183	−0.027*	−0.039***
	(0.147)	(0.015)	(0.010)
Population	2.001	0.429**	0.572***
	(2.106)	(0.208)	(0.119)
Globalization	−0.039	−0.004*	0.001
	(0.025)	(0.002)	(0.001)
Observations	1,685	1,684	1,677

* $p < 0.10$,** $p < 0.05$,*** $p < 0.01$

Standard errors clustered on country in parentheses.

The variable *Institutional Effectiveness* still exerts a negative and statistically significant effect even when considering any of the three alternative dependent variables as proxies for ecological effectiveness. This greatly adds to the confidence in the findings of my research.

[8] Available online at: `https://www.footprintnetwork.org/`.

Supplementary Table 3. Panel-Corrected Standard-Error Model

	Model A4
Lagged Dependent Variable	0.236***
	(0.065)
Institutional Effectiveness	−0.886***
	(0.224)
Democracy	2.377**
	(1.108)
GDP per capita	−13.775
	(31.199)
GDP per capita^2	1.914
	(4.014)
GDP per capita^3	−0.084
	(0.169)
Population	1.485
	(1.921)
Globalization	−0.057**
	(0.025)
Observations	947

* $p < 0.10$,** $p < 0.05$,*** $p < 0.01$

Standard errors in parentheses.

Supplementary Notes 3: Panel-Corrected Standard-Error Model

In Table 3 of this SI, I employ a panel-corrected standard-error model[7] that comprises fixed effects for countries and years. Panel-level heteroskedastic errors address country-specific idiosyncrasies and I specify that, within panels, there is first-order autocorrelation with the coefficient of this process being common to all panels. This approach addresses measurement errors, assuming that the error is "systematically related to the country, but does not change much over time" [8]. The fixed effects allow me to control for any unobserved unit-level effects as well as temporal shocks in a given year. The main finding is robust when using the panel-corrected standard-error model[7]. In fact, all the results are virtually identical to what is presented in the main text. Note that the same assumptions for identifying causal effects as in the case of the main text's two-way fixed effects models must be met.

Supplementary Notes 4: Quartile Disaggregation of Institutional Effectiveness

In Model A5, *Institutional Effectiveness* is divided into four equally sized parts – quartiles. I include the quartiles into the model estimation (two-way fixed effects regression), leaving out the first quartile comprising the lowest scores of the original variable *Institutional Effectiveness*. As a result, I focus on three remaining binary variables, *Low Institutional Effectiveness*, *Medium Institutional Effectiveness*, and *High Institutional Effectiveness*, which are compared to the left-out baseline quartile (*No Institutional Effectiveness*). Table 4 demonstrates that the main finding is robust as – in comparison to *No Institutional Effectiveness* – the other

Supplementary Table 4. Disaggregation of Institutional Effectiveness

	Model A5
Lagged Dependent Variable	0.310***
	(0.051)
Low Institutional Effectiveness	−0.536**
	(0.210)
Medium Institutional Effectiveness	−0.626***
	(0.184)
High Institutional Effectiveness	−0.485***
	(0.145)
Democracy	2.811**
	(1.323)
GDP per capita	−44.947*
	(25.786)
GDP per capita^2	5.811*
	(3.248)
GDP per capita^3	−0.244*
	(0.134)
Population	1.912
	(2.067)
Globalization	−0.041*
	(0.024)
Observations	1,685

* $p < 0.10$,** $p < 0.05$,*** $p < 0.01$

Standard errors clustered on country in parentheses.

three dichotomous variables are linked to negative and statistically significant effects. *Medium Institutional Effectiveness* exerts the strongest influence when looking at the size of the estimated coefficients, although the estimates of *Low Institutional Effectiveness*, *Medium Institutional Effectiveness*, and *High Institutional Effectiveness* are not statistically different from each other.

Supplementary Notes 5: Exclusion of Specific Organizations

Not all organizations in my sample are created equal – or are equally important for the outcome variable on ecological effectiveness. That is, their design matters and I discuss this in the last analysis of the main text. What is more, while all institutions primarily deal with climate change and related challenges, it has been suggested that the Adaptation Fund, the Climate Investment Funds, and the Multilateral Fund for the Implementation of the Montreal Protocol are not primarily designed to reduce $PM_{2.5}$ pollution. While it is plausible that some of their projects, e.g., those targeting de-carbonization, generate positive externalities for air quality and reduce $PM_{2.5}$ levels as well, they most likely focus on rather unrelated objectives, e.g., adaptation infrastructure. To this end, I re-estimate the main model while excluding these three organizations from the sample to assess whether the findings hold when focusing on other institutions with a more direct connection to air quality outcomes and $PM_{2.5}$. Models A6-A9 summarize the results: Model A6 excludes observations pertaining to the Adaptation Fund, Model A7 omits the Climate Investment Funds, Model A8 drops the Multilateral Fund from the sample, and Model A9 discards observations that

Supplementary Table 5. Exclusion of Specific Organizations

	Model A6	**Model A7**	**Model A8**	**Model A9**
Lagged Dependent Variable	0.305***	0.281***	0.293***	0.235***
	(0.053)	(0.059)	(0.053)	(0.065)
Institutional Effectiveness	−0.224***	−0.200***	−0.225***	−0.298***
	(0.050)	(0.058)	(0.049)	(0.084)
Democracy	2.975**	2.317*	2.981**	2.749**
	(1.298)	(1.366)	(1.327)	(1.333)
GDP per capita	−40.136	−39.856	−41.564	−29.775
	(25.333)	(28.920)	(25.690)	(29.350)
GDP per capita^2	5.156	5.274	5.365*	3.890
	(3.178)	(3.600)	(3.226)	(3.641)
GDP per capita^3	−0.214	−0.227	−0.224*	−0.164
	(0.131)	(0.147)	(0.133)	(0.149)
Population	2.290	1.243	2.122	1.923
	(2.137)	(2.160)	(2.156)	(2.406)
Globalization	−0.042*	−0.048**	−0.042*	−0.050**
	(0.024)	(0.023)	(0.024)	(0.023)
Observations	1,564	1,391	1,567	1,152

* $p < 0.10$,** $p < 0.05$,*** $p < 0.01$

Standard errors clustered on country in parentheses.

are related to any of the three institutions. Having said that, regardless of which organization is omitted, or all three simultaneously, the main effect is robust as *Institutional Effectiveness* remains negatively signed and statistically significant. The overall effect size of this item increases, which may not be surprising as the excluded institutions deal with $PM_{2.5}$ pollution less directly.

Supplementary Notes 6: Domestic-Level Policy Measures

Models A10 and A11 control for national climate policies, which may also shape a country's air quality, but are not directly related to what an international organization does. To begin with, national policies could confound the observed relationship between institutional effectiveness and pollution levels, especially if countries with high institutional performance also undertake ambitious domestic-level reforms. To this end, I incorporate additional controls to account for national environmental policies using the Climate Policy Database (CPDB).[9] The CPDB comprises information on more than 6,000 policies in 198 countries. I consider all six sectors (agriculture, buildings, electricity, general, industry, and transport) and all nine policy instrument types (barrier removal, climate strategy, economic instruments, information and education, policy support, regulatory instruments, research and development, target, and voluntary approaches) for the variable I add to Model A10. According to the CPDB, "policy instruments provide a link between desired policy outcomes and implementation. The term refers to instruments used to put the policy in practice, e.g., subsidies or carbon pricing." In Model A11, I focus on policy instruments that only deal with air pollution according to the CPDB. On one hand, the results show that the effect associated with *Institutional Effectiveness* is again negatively signed and statistically significant. On the other hand, the policy measures

[9]See online at: `https://climatepolicydatabase.org/`.

Supplementary Table 6. Domestic Climate Policies

	Model A10	**Model A11**
Lagged Dependent Variable	0.310^{***}	0.311^{***}
	(0.052)	(0.052)
Institutional Effectiveness	-0.186^{***}	-0.188^{***}
	(0.044)	(0.044)
Climate Policies	0.049	
	(0.032)	
Air Pollution Policies		0.326
		(0.242)
Democracy	2.853^{**}	2.818^{**}
	(1.368)	(1.380)
GDP per capita	-49.303^{*}	-50.117^{*}
	(26.853)	(26.881)
GDP per capita^2	6.320^{*}	6.425^{*}
	(3.352)	(3.355)
GDP per capita^3	-0.263^{*}	-0.267^{*}
	(0.137)	(0.137)
Population	2.145	2.103
	(2.112)	(2.119)
Globalization	-0.042^{*}	-0.041^{*}
	(0.024)	(0.024)
Observations	1,648	1,648

$^{*}\ p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Standard errors clustered on country in parentheses.

exert a positive impact, but are not significant at conventional levels. This mirrors the existing literature in that the relationship between policies (outputs) and environmental quality (outcomes) may be rather ambiguous and several, partly rather long causal channels are at play in this context see [9,10,11,12].

Supplementary Notes 7: Sample-Selection Models

The discussion in the main text points to a two-stage decision-making process, which must be considered jointly in the empirical analysis [13]. The first stage (selection equation) is the selection into environmental organizations, while the second stage (outcome or regression equation) pertains to their ecological effectiveness. This selection process features prominently in the literature [14,15,16]. To deal with this issue comprehensively, I employ two different estimation procedures.

First, I present the results of a two-part model (2PM) [17]. This model comprises a probit model and an OLS regression, omits the inverse Mills ratio from the latter, and estimates the two models jointly using seemingly unrelated regression. The main purpose of this model is to explicitly consider sample selection: ecological effectiveness of an institution is only observed when countries strategically select themselves into such organizations as members or targets. In light of this, the probit part of the 2PM addresses the selection stage or the determinants of censoring. In my case, this refers to states' strategically selecting themselves into institutions. The OLS component of the 2PM is about the estimation of the outcome of interest based on the non-censored observations. Countries' "non-selection" into organizations as members or target

states is not actually observed, however, and I thus create a selection sample and a selection variable based on *Institutional Effectiveness*[18]. Specifically, I contend that "active" membership in an organization or participation in an institution's programs is only given when *Institutional Effectiveness* is greater than one standard deviation than its mean value. I use this threshold to create a binary selection variable distinguishing between countries participating in or being targeted by an environmental organization and those that are not. This item is the dependent variable in the probit part of them 2PM, and I add the control variables as predictors in this estimation next to the cubic polynomial approximation[19] to address temporal dependencies. The outcome stage of the 2PM in Model A12 is based on OLS and I merely include the lagged dependent variable, *Institutional Effectiveness* as such, and year as well as country fixed effects as predictors. According to Model A12, *Population* and the temporal controls are solid predictors for states' selection into environmental organizations. Incorporating the information from this selection stage into the outcome equation lowers possible bias due to countries' strategic considerations for participation. Still, the main result is robust as *Institutional Effectiveness* remains negatively signed and statistically significant at the 1 percent level. Hence, institutional effectiveness drives ecological effectiveness also when employing the 2PM[17].

Supplementary Table 7. Two-Part Model (2PM)

	Model A12
Lagged Dependent Variable	0.706^{***}
	(0.073)
Institutional Effectiveness	-0.633^{***}
	(0.180)
Democracy	-0.100
	(0.184)
GDP per capita	1.397
	(4.235)
GDP per capita^2	-0.171
	(0.525)
GDP per capita^3	0.007
	(0.021)
Population	-0.052^{**}
	(0.023)
Globalization	-0.002
	(0.004)
Time	-0.297^{***}
	(0.115)
Time^2	0.126^{**}
	(0.051)
Time^3	-0.008
	(0.006)
Observations	1,881

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$
Standard errors in parentheses.

Moreover, I consider a Heckman selection model. Here, the estimated mean function in the outcome stage is conditioned on the first-stage selection process and thereby provides a consistent estimate for the truncated distribution of the second-stage sample[20]. This estimator is frequently seen as more robust than competing approaches, but there is the exclusion restriction: the selection equation has to include at least

one regressor that is not part of the outcome equation. To estimate the Heckman selection model, I have constructed another data set comprising "true" 0s of non-participation. That is, the unit of analysis is the treaty-country-year in 2008-2018. This mirrors the setup of the data used in the main text, though I now add states' non-participation to this by combining each of the six organizations with *any* state across these organizations. After having compiled all the information on institutions' target or program states, I compare membership across the organizations. Assuming that a country in organization (a) could, at least in principle, also be a target of or program country in institution (b), I code non-participation of that country in the latter if it is not listed as a target or program state there. To this end, the pool of countries potentially able to join and work with an organization comprises all target or program states across all six environmental organizations combined in the data. The final data set comprises 3,867 observations, which has 2,137 "participations" in environmental organizations and 1,730 "non-participations."

The first stage of the model has this binary selection variable as the outcome item, while I consider the variables for democracy, population, and the temporal controls[19] as predictors. These variables are not considered in the outcome stage for theoretical and empirical reasons. These items seemed to perform best at the selection stage and, at least, in terms of democracy, are theoretically more likely to shape selection into organizations than ecological effectiveness. That is, both variables appear largely unrelated to or are inconsistently associated with the dependent variable of the outcome stage in various model specifications I explored. For the other variables that I eventually consider for the outcome stage in Model A13, the effects are more consistently statistically significant. Moreover, "theoretically," the existing literature offers mixed evidence on the overall effect of democracy on environmental quality outcomes (i.e., ecological effectiveness), but findings are more consistent see[9,21,22,12,23] with regard to the role of regime type in shaping environmental policy outputs at the domestic level or at the international level (and this relates to organizational selection processes in the form of participation in or ratification of international environmental agreements). In light of this, the outcome equation has the main text's item for ecological effectiveness as the dependent variable and I include a lagged dependent variable, institutional effectiveness, the income variables, *Globalization*, as well as year and country fixed effects as predictors.

In the selection equation, democracies are more likely to participate in international environmental organizations as target and program countries. This mirrors earlier research on states' general participation in international environmental organizations[22]. Population size is also significantly linked to countries being target or project states in such institutions. In the outcome equation, all variables are statistically significant at conventional levels – most importantly, *Institutional Effectiveness* exerts again a negative effect on environmental pollution. Hence, the Heckman selection model further increases the confidence in the finding that institutional effectiveness is strongly linked to the ecological effectiveness of international environmental organizations.

Supplementary Table 8. Heckman Selection Model

	Model A13
Lagged Dependent Variable	0.317^{***}
	(0.053)
Institutional Performance	-0.164^{***}
	(0.053)
GDP per capita	-46.780^{*}
	(25.883)
GDP per capita^2	6.273^{*}
	(3.312)
GDP per capita^3	-0.272^{**}
	(0.139)
Globalization	-0.041^{*}
	(0.024)
Democracy	0.463^{**}
	(0.204)
Population	0.143^{***}
	(0.027)
Time	0.010
	(0.136)
Time^2	0.014
	(0.101)
Time^3	-0.020
	(0.019)
Observations	3,063

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Standard errors clustered on country in parentheses.

Supplementary Notes 8: Placebo Test

As discussed above and in the main text, a common concern in the literature on the causal effects of international environmental organizations is the issue of self-selection [14,15,16]. For example, countries join environmental institutions, especially those with higher effectiveness scores, if they are already strongly committed environmentally. I address this issue via several estimation strategies, including the 2PM and the Heckman selection model above, but I also considered a straightforward placebo test that an anonymous referee suggested: I examine the impact of institutional effectiveness on contributor countries. If self-selection is not a major concern, the estimated effects on contributors, which are not expected to benefit directly from the institutions' activities, should be statistically insignificant. I have conduced such a placebo test, which is summarized in Table 9: focusing on the same models that are presented in Table 1 of the main text, I now combine each of the six environmental organizations in my sample not with the target or program countries, but with the donor and contributing states. As shown across Models A14-A19, *Institutional Effectiveness* is now statistically insignificant. This further increases the confidence in the causal nature of the main finding.

Supplementary Table 9. Placebo Tests

	Model A14	Model A15	Model A16	Model A17	Model A18	Model A19
Lagged Dependent Variable		0.073	0.007			−0.091
		(0.153)	(0.136)			(0.121)
Institutional Effectiveness	0.079	0.063	0.025	0.074	0.341	0.151
	(0.058)	(0.058)	(0.052)	(0.052)	(0.266)	(0.260)
Democracy			−4.778			−4.382
			(4.747)			(5.083)
GDP per capita			185.302**			205.201***
			(80.249)			(73.518)
GDP per capita^2			−19.570**			−21.815***
			(8.872)			(8.017)
GDP per capita^3			0.682**			0.761**
			(0.321)			(0.288)
Population			16.897*			17.401*
			(9.462)			(9.947)
Globalization			0.012			−0.009
			(0.070)			(0.065)
Observations	557	496	496	496	331	291

* $p < 0.10$,** $p < 0.05$,*** $p < 0.01$
Standard errors clustered on country in parentheses.

Supplementary Notes 9: Fixed Effects Counterfactual Estimator

As a final robustness check, I implement a new counterfactual estimator for causal inference [24]. This is a "simple framework of counterfactual estimation for causal inference with time-series cross-sectional data" [24], in which the average treatment effect on the treated is estimated by directly imputing counterfactual outcomes for treated observations." The corresponding estimation procedures outperform "traditional" two-way fixed effects models and allow for a more precise estimation of causal treatment effects – in my case, the effect stemming from institutional effectiveness. I focus on the fixed effects counterfactual (FEct) estimator. This approach is more efficient than two-way fixed effects models, and unbiased as well as consistent in the estimation of parameters. It is essentially a weighting estimator in which "each treated observation is matched with its predicted counterfactual [...], which is the weighted sum of all untreated observations" [24].

The FEct model I present in Table 10 includes both unit and time fixed effects and I obtain uncertainty estimates using the non-parametric bootstrap procedure (100 runs). For the treatment, I use the last (highest) quartile of the disaggregated analysis above (Table 4) and code this as a value of 1 in a binary variable, and thus allocate less performing environmental institutions into the control group (value of 0). The main quantity of interest calculated by the FEct estimator is the dynamic treatment effect, "or the average treatment effects on the treated (ATT) over different periods, without assuming treatment effect homogeneity of any kind" [24]. According to the results in Model A20, there are no strong pre-trend leads towards the onset of the treatment and no treatment estimates in the pre-treatment period are significantly different from zero. In the post-treatment period, however, I obtain evidence for a significant effect of *Institutional Effectiveness*, suggesting that environmental organizations have a causal impact on ecological effectiveness, i.e., environmental quality at the domestic level.

Supplementary Table 10. Fixed Effects Counterfactual Estimator (FEct)

	Model A20
Lagged Dependent Variable	0.280***
	(0.090)
High Institutional Effectiveness (Treatment)	−0.768*
	(0.425)
Democracy	−0.303
	(1.099)
GDP per capita	−17.191
	(44.437)
GDP per capita^2	2.241
	(5.754)
GDP per capita^3	−0.094
	(0.246)
Population	1.374
	(2.306)
Globalization	−0.025
	(0.022)
Observations	909

* $p < 0.10$,** $p < 0.05$,*** $p < 0.01$

Bootstrapped standard errors in parentheses.

Supplementary References

1. Van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C., et al. *Environmental Science & Technology* **55**(22), 15287–15300 (2021).

2. Wackernagel, M. and Rees, W. *Our Ecological Footprint: Reducing Human Impact on the Earth.* Gabriola Island, BC: New Society Publishers, (1998).

3. Caviglia-Harris, J., Chambers, D., and Kahn, J. *Ecological Economics* **68**(4), 1149–1159 (2009).

4. Ewing, B., Reed, A., Galli, A., Kitzes, J., and Wackernagel, M. *Oakland, CA: Global Footprint Network* (2010).

5. Lin, D., Hanscom, L., Murthy, A., Galli, A., Evans, M., Neill, E., Mancini, M. S., Martindill, J., Medouar, F.-Z., Huang, S., et al. *Resources* **7**(3), 58 (2018).

6. Wiedmann, T., Minx, J., Barrett, J., and Wackernagel, M. *Ecological Economics* **56**(1), 28–48 (2006).

7. Beck, N. and Katz, J. N. *American Political Science Review* **89**(3), 634–647 (1995).

8. Neumayer, E. *Journal of Peace Research* **40**(6), 619–640 (2003).

9. Ward, H. *Environmental Politics* **17**(3), 386–409 (2008).

10. Anderson, B., Böhmelt, T., and Ward, H. *Environmental Research Letters* **12**(11), 114011 (2017).

11. Böhmelt, T., Ward, H., and Bernauer, T. *Journal of Public Policy* , 1–21 (2025).

12. Böhmelt, T., Böker, M., and Ward, H. *Democratization* **23**(7), 1272–1291 (2016).

13. Böhmelt, T. and Spilker, G. In *The SAGE Handbook of Research Methods in Political Science and International Relations*, 701–716. SAGE Publications Ltd (2020).

14. Lupu, Y. *American Journal of Political Science* **57**(4), 912–925 (2013).

15. Ringquist, E. J. and Kostadinova, T. *American Journal of Political Science* **49**(1), 86–102 (2005).

16. Miles, E. L., Andresen, S., Carlin, E. M., Skjærseth, J. B., and Underdal, A. *Environmental regime effectiveness: confronting theory with evidence.* Cambridge, MA: MIT Press, (2001).

17. Vance, C. and Ritter, N. *Journal of Peace Research* **51**(4), 528–540 (2014).

18. Chiba, D., Johnson, J. C., and Leeds, B. A. *The Journal of Politics* **77**(4), 968–982 (2015).

19. Carter, D. B. and Signorino, C. S. *Political Analysis* **18**(3), 271–292 (2010).

20. Heckman, J. J. *Econometrica* **47**(1), 153–161 (1979).

21. Bernauer, T. and Koubi, V. *Ecological Economics* **68**(5), 1355–1365 (2009).

22. Bernauer, T., Kalbhenn, A., Koubi, V., and Spilker, G. *British Journal of Political Science* **40**(3), 509–538 (2010).

23. Povitkina, M. *Environmental Politics* **27**(3), 411–432 (2018).

24. Liu, L., Wang, Y., and Xu, Y. *American Journal of Political Science* **68**(1), 160–176 (2024).