

Forecasting NDC non-compliance: Economic structure, national capacity and the limits of climate finance

Michelle Márcia Viana Martins

`michelle.viana@ufv.br`

Universidade Federal de Viçosa

Rodrigo Fagundes Cezar

Fundação Getulio Vargas

Débora Aparecida Carrer

Universidade de São Paulo

Alan Leal

Fundação Getulio Vargas

Research Article

Keywords: climate finance, disbursement predictability, machine learning, land use

Posted Date: October 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11011032/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Forecasting NDC non-compliance: Economic structure, national capacity and the limits of climate finance

Abstract: Two thirds of the countries outside Annex I of the United Nations climate convention are not on track to meet the 2030 decarbonization targets declared in their Nationally Determined Contributions (NDCs), and this article identifies them before the deadline. We build the NDC Compliance Gap Indicator for 2030 (NCGI), the distance between projected emissions and each country's target, recalculated yearly from 2013 to 2024 for 104 countries, and test whether the design of international public climate finance, not only its volume, helps anticipate who falls short. The forecast is validated outside the training period with seven machine learning algorithms; each variable's contribution is decomposed through Shapley additive explanations (SHAP); and the effect of finance is estimated by double machine learning. Three results stand out: non-compliance risk is predictable six years ahead; including or excluding land use reclassifies 37% of countries between compliance and non-compliance, exposing the fragility of self-declared targets; and climate finance, decomposed into predictability, concessionality and fragmentation, neither anticipates non-compliance nor shows an identifiable effect on the gap. The two estimates that cross the conventional threshold in a single coverage reflect donor selection, shown through a placebo in which finance still to be received anticipates the present gap more strongly than finance already received. What relates best to compliance is productive structure, land endowment, the energy profile and the quality of public institutions, rather than the international support received, moving the focus of climate policy from the volume of finance to the structural conditions of implementation.

Keywords: climate finance; disbursement predictability; machine learning; land use.

JEL codes: C53; Q56; F35.

1. Introduction

The Paris Agreement established Nationally Determined Contributions (NDCs) as a central element of climate governance, requiring each country to declare transparent and progressively more ambitious targets for reducing greenhouse gas (GHG) emissions (UNFCCC, 2015). Holding warming below 2°C, with efforts toward 1.5°C, depends on what each country actually delivers, so NDCs have become verifiable objects tracked through inventories and periodic reports until 2030. What decides the outcome is not declared ambition but the national capacity to sustain it, which runs through coordination across agencies, regulatory design and resource allocation, and extends beyond government to how producers, the financial sector and subnational entities respond to the regulation put in place (Peterson & van Asselt, 2025).

International public climate finance is a central instrument for making implementation viable in developing countries, yet its volume falls far short of need and its composition is uneven. In 2022 the volume provided and mobilized by developed countries reached US\$ 115.9 billion, exceeding the US\$ 100 billion goal for the first time (OECD, 2024), against estimated global needs of US\$ 8.1 to 9 trillion per year through 2030 (Buchner et al., 2023). Loans accounted for 69% of the public total and grants for 28%, but that average hides a composition effect: over 2016 to 2022 loans reached 87% of what upper middle-income countries received, while grants made up 64% of what low-income countries received (OECD, 2024). Because middle-income countries concentrate 70% of the volume, credit dominates the aggregate, and the cost of support falls unevenly, since middle-income economies accumulate debt service that competes with other priorities while low-income countries, financed mainly by grants, do not.

Countries with similar targets deliver different results, owing to financial and institutional constraints. On the financial side, difficulty concentrates in conditional NDCs, in which part of the promised reduction is to be achieved only if international support arrives, alongside an unconditional target undertaken with domestic resources (Pauw et al., 2020). On the domestic side, execution depends on the quality of public institutions to design instruments, coordinate ministries and monitor results, and faces resistance from emission-intensive sectors (Peterson & van Asselt, 2025; Northrop et al., 2016). Together these feed the implementation gap, in which a country announces targets but does not sustain compatible policies (Roelfsema et al., 2020).

Even where support is present, physical vulnerability competes with governance and recipient merit in defining who receives which modality, since donors prioritize contexts able to turn resources into executable projects (Weiler et al., 2018). Betzold and Weiler (2017) find biases toward small and middle-income countries, and Pickering et al. (2015) show that disputes among agencies within donor countries impair coordination across channels.

Despite progress in understanding barriers and allocation, the capacity to forecast NDC non-compliance remains limited. Rogelj et al. (2016) show that the submitted contributions fall short of the trajectory compatible with 2°C, implying median warming of 2.6 to 3.1°C by 2100. Fekete et al. (2021) document that effective mitigation policies already exist in the large emitters and that replicating them would cut emissions by around 20% by 2030 without closing the deficit. Peterson and van Asselt (2025) examine 21 countries and the European Union and find that three of four risk profiles face threats to full compliance, with institutional quality operating as a structural constraint. The available diagnosis, however, classifies few countries on judgement or scenarios, without out-of-sample validation and without a measure recalculable every year, so that models resting only on income and general governance miss what lies closest to implementation.

This study extends that diagnosis by treating the architecture of climate finance as an explanatory variable for non-compliance risk. The object is international public climate finance reported to the OECD Creditor Reporting System (CRS), and the hypothesis is that observable features of the flows received proxy for execution and coordination capacity. Predictability, the ratio of disbursed to committed, captures whether a country can draw on promised resources; concessionality, the grant share, separates support that generates no debt service from support that burdens future budgets; and fragmentation, the dispersion of finance across donors, approximates the coordination cost imposed on the recipient. Concessionality and predictability move together (rank correlation 0.60) while fragmentation is close to orthogonal (0.15 and 0.06), which is why results are read dimension by dimension.

To measure risk we propose the NDC Compliance Gap Indicator for 2030 (NCGI), built in two steps. The first projects each country's emissions to 2030 by extrapolating the 2013 to 2024 trend, a continuity scenario whose simplicity matches its purpose: country-by-country sectoral projections cover 25 to 40 economies and are not recalculated annually, whereas an indicator comparable across 104 countries calls for a replicable, auditable rule. The second step compares the projection with the conditional and unconditional targets declared in each NDC, defining the indicator as the percentage difference between trajectory and target. Positive values indicate non-compliance risk and negative values slack; the indicator takes a continuous form and a binary form, classifying a country as at risk or not according to a threshold.

The predictors of the NCGI comprise controls for economic structure, governance, energy and vulnerability, plus the three finance dimensions. We adopt a predictive approach based on machine learning, prioritizing out-of-sample performance and temporal validation over parametric models that tend to miss interactions between finance and context (Athey & Imbens, 2019; Breiman, 2001); seven algorithms are compared under an identical split, training through 2021 and testing on later years. Two distinct questions then require distinct methods. Whether finance helps forecast the NCGI is answered by each variable's contribution to performance, measured through Shapley additive explanations (SHAP) on tree models (Lundberg & Lee, 2017), which captures association rather than effect. Whether altering finance design would change compliance is a causal question, answered by estimating the effect of each dimension through double machine learning, which separates it from the remaining observed determinants (Chernozhukov et al., 2018). Decomposition shows whether finance predicts, double machine learning whether it produces an effect, and the two answers need not coincide.

The article therefore offers an early diagnosis of NDC non-compliance risk through 2030, treating finance design as a predictor of the compliance trajectory and estimating how far that design affects it, on a country-year panel joining emissions, quantified targets, climate finance and structural controls. Four choices set the study apart: treating finance as an object of design rather than a single volume figure; measuring compliance with and without land use, revealing how far that sector repositions each country; separating association from effect for the finance variables; and extending the diagnosis to 104 countries with annual updating, against the 21 countries and the European Union examined by Peterson and van Asselt (2025). The remainder

reviews the literature on climate finance and predictive models, presents the data and empirical strategy, reports the results from the country diagnosis to the role of finance, and concludes with the implications.

2. Literature review

2.1 Climate finance: determinants and patterns

Research on international climate finance developed around what determines the allocation of flows and how large the deficit is between funds and needs. The distribution is not defined by climate need alone, and the volume mobilized falls well short of the requirement.

Halimanjaya (2015), on 180 developing countries between 1998 and 2010, finds that donors favor higher emission intensity per unit of GDP, larger carbon sinks, lower per capita income and good governance, so that selection responds to abatement potential rather than emitter size, while lower middle-income countries received 64% of mitigation finance against 6.4% for least developed countries. Betzold and Weiler (2017), on bilateral adaptation aid across 144 countries, find that more exposed countries receive more, though per capita aid peaks near US\$ 1,000 of GDP and favors small, better-governed states. Weiler et al. (2018) show that allocation follows both need and recipient merit, with adaptive capacity carrying no weight and donors favoring well-governed countries, and Nor (2025), on multilateral fund disbursements between 2003 and 2022, finds that approved volume, not pledged volume, explains disbursement, with conversion more efficient in high-income recipients.

Concentration marks both recipients and channels. China, Indonesia and India received more than half of mitigation resources between 1998 and 2010 (Halimanjaya, 2015), and energy (44%) and transport (29%) dominate mitigation flows while agriculture and industry receive under 4%, despite holding larger mitigation potential through 2030 (IPCC, 2022; Buchner et al., 2023). Institutional fragmentation inside donor governments amplifies transaction costs: aid agencies retain control over execution while environment and finance ministries exert asymmetric influence over the split between mitigation and adaptation (Pickering et al., 2015), and development finance institutions provided most public finance in 2021 to 2022 while dedicated multilateral climate funds contributed only 0.5%, so most flows bypass the mechanisms created under the Convention (Buchner et al., 2023).

The limitation of this literature is that it treats finance as a dependent variable, investigating why finance flows to certain countries and how large the deficit is, without exploring how observable patterns of these flows, such as disbursement predictability, concessionality and fragmentation, affect the implementation capacity of recipients. Assessment of such effects remains restricted to aggregate magnitudes or promised ambition: Rashid et al. (2023) find a positive but statistically insignificant effect of climate finance on mitigation ambition, and Wei (2025), on 30 developing countries between 2010 and 2021, estimates that a 1% rise in total climate finance is associated with a 0.03% reduction in emissions. Fragmentation, the bias toward well-governed countries and the predominance of loans are described as features of the system, not mobilized as predictors of outcomes.

2.2 Predictive models and machine learning in climate policy

Machine learning spread in economics and public policy from 2010, driven by data availability and by the recognition that, for some problems, predictive accuracy is a distinct and at times more relevant objective than identifying causal effects (Athey & Imbens, 2019). Traditional econometrics estimates structural parameters under identification assumptions, whereas machine learning prioritizes out-of-sample performance without pre-specifying the form of the relation (Breiman, 2001).

Gradient boosting combines decision trees sequentially, each correcting the errors of the previous ones, and captures nonlinearities and interactions without a pre-specified functional form; its main implementations, XGBoost (Chen & Guestrin, 2016), LightGBM (Ke et al., 2017) and CatBoost (Prokhorenkova et al., 2018), compete with random forests (Breiman, 2001) and neural networks. For tabular data, tree families match or outperform neural alternatives with no single implementation prevailing, which recommends comparing models under the same split (Grinsztajn et al., 2022), while overfitting is contained

by regularization and by validation that respects chronological order (Bergmeir et al., 2018). The opacity of flexible algorithms is addressed by SHAP, which uses Shapley values to attribute to each variable its marginal contribution to each forecast (Lundberg & Lee, 2017), with a tree-specific implementation that makes computation feasible and supports reading by country and year (Lundberg et al., 2020).

SHAP explains where accuracy comes from, not what would change if a variable were altered. Double machine learning serves that second objective, estimating nuisance functions with flexible methods while the parameter of interest retains a valid standard error through orthogonal scores and cross-fitting (Chernozhukov et al., 2018), which adds to the predictive results an estimate of the effect of each finance dimension under the assumption that the controls capture what confounds the relation.

In climate science, machine learning is established for physical modeling, but in the political economy of climate its adoption is incipient, with allocation and ambition studies relying mostly on parametric regressions with specified functional forms. Work on implementation risk has separated distances often conflated: Roelfsema et al. (2020) estimate a median implementation gap of 22.4 to 28.2 GtCO₂eq in 2030, and Peterson and van Asselt (2025) identify five risks and group 21 countries and the European Union into four profiles, three facing threats to full compliance, with institutional capacity as a structural constraint. Both pose the question taken up here yet classify few countries on judgement or scenarios, without out-of-sample validation and without a measure recalculable every year for the whole set with a target.

There is no reason to expect the distance to the target to be a linear, additive function of a country's characteristics, and the aim here is to forecast who will fall short rather than to estimate a parameter under an imposed form, the purpose for which machine learning was designed (Athey & Imbens, 2019). The study reverses the position the literature assigns to finance, testing predictability, concessionality and fragmentation as predictors of non-compliance risk. Combining model comparison, SHAP decomposition and double machine learning answers which countries are at risk, which finance dimensions are associated with that risk and whether any dimension affects it; Online Resource 1 gathers the studies on climate finance and NDC implementation reviewed here.

3. Methodology

This section presents the study design in three parts, the first of which delimits the sample, justifies the filters applied and describes the construction of the NCGI together with the sources for emissions and targets. The second decomposes climate finance into predictability, concessionality and fragmentation, and presents the controls that make up the country-year panel. The third describes the identification strategy, the evaluation criteria and the robustness checks.

3.1 Data, sample and the compliance indicator

The study rests on a country-year panel with annual observations from 2013 to 2024, covering the economies outside Annex I of the UNFCCC, the recipients under the Convention: among them 85% declare a target conditional on external support, against 5% among Annex I members, so that testing whether finance design explains compliance makes sense only for countries that can condition ambition on support received. Three operational filters keep the indicator computable and comparable: a submitted and quantifiable NDC; an emissions series in PRIMAP-hist with at least ten consecutive years; and finance recorded in the OECD CRS in at least five years of the period. With a population cutoff of 500,000 inhabitants, 104 countries remain in the reference specification.

Emissions come from PRIMAP-hist (version 2.7), as national totals of the Kyoto gases in CO₂eq with AR4 factors, excluding land use, land-use change and forestry (LULUCF), the core reported most comparably across countries. The 2030 target is converted to MtCO₂eq following the quantification of Günther et al. (2021), which codes the conditional and unconditional target of each country separately. Where a target is worded ambiguously, the more ambitious bound is adopted, making the indicator conservative against the country.

The quantification is a snapshot of the NDCs in force in April 2021 and is held constant across the panel, since the source is published as a single version rather than a continuously updated series, so annual variation

of the indicator reflects the emissions trajectory, not revisions of ambition. Because revisions raised ambition in most cases, measuring compliance against the 2021 targets tends to understate risk, and section 4.4 gauges that effect against a revised target from a second source. Emissions use AR4 potentials to match the target metric, a choice that barely affects the diagnosis (rank correlation 0.999 between metrics; Online Resource 2 details the accounting choices).

The conditional target is what a country would reach with the external support declared as necessary, and the unconditional target what it undertakes with its own resources, so the conditional is by construction at least as ambitious. Comparing the projection with both separates non-compliance from failed domestic execution from non-compliance responding to the absence of promised support, a distinction often merged into a single number (Roelfsema et al., 2020) and a sensitive point in the Paris Agreement (Pauw et al., 2020). The reference specification adopts the conditional target and the unconditional reading enters as a check, over the subset declaring both.

The *PBL Climate Pledge NDC tool* (den Elzen et al., 2024) serves as a robustness source: it reports the target of the NDC in force and of the previous one, correlates at 0.98 with the reference, and isolates the effect of the target version on the diagnosis.

The trajectory projection starts from the average annual logarithmic growth rate of emissions in a window of k years ending in t , extended to 2030:

$$g_{i,t}^{(k)} = \frac{1}{k} \ln \left(\frac{E_{i,t}}{E_{i,t-k}} \right) \quad (1)$$

$$E_{i,2030|t}^{proj} = E_{i,t} \exp \left(g_{i,t}^{(k)} (2030 - t) \right) \quad (2)$$

The projection window parameter k is fixed at five years, with $k = 3$ and $k = 7$ as checks, and the emissions growth rate g has its highest 5% and lowest 5% in each year replaced by the cutoff values, so that atypical falls caused by crises, conflicts or inventory revisions do not dominate the variation. The continuous indicator and its binary risk version are:

$$NCGI_{i,t} = 100 \times \frac{E_{i,2030|t}^{proj} - E_{i,2030}^{NDC}}{E_{i,2030}^{NDC}} \quad (3)$$

$$Risk_{i,t} = 1(NCGI_{i,t} > \tau) \quad (4)$$

with $\tau = 5\%$ in the reference and $\tau = 0$ and $\tau = 10\%$ as checks. A value of $NCGI_{i,t} = 20$ indicates that the 2030 projection stands 20% above the target, and a negative value indicates a projection below the target. For the regression exercise, the highest 1% and the lowest 1% of the continuous indicator are replaced by the cutoff values, to contain extreme tails produced by targets close to zero; the binary version uses the indicator without that treatment.

The study builds a second version of the outcome that includes land use, $NCGI_{i,t}^{LU}$, because for countries in which the land sector accounts for a relevant share of the balance, excluding LULUCF leaves out most of the emissions story. Section 4.1 shows that this holds for about a third of the sample, whose risk classification reverses depending on whether the sector enters the accounting. In that version, the national total including land use comes from *Climate Watch*, is projected to 2030 by the same rule, and is compared with the NDC target in the coverage that includes LULUCF, $E_{i,2030}^{NDC,LU}$:

$$NCGI_{i,t}^{LU} = 100 \times \frac{E_{i,2030|t}^{proj,LU} - E_{i,2030}^{NDC,LU}}{E_{i,2030}^{NDC,LU}} \quad (5)$$

The versions with and without LULUCF are not interchangeable, since each rests on distinct sources and on targets with different accounting coverage, and the difference between them is itself informative: when a

country's NCGI changes with the counting rule, what this reveals is the fragility of the metric rather than a change in the trajectory, a fragility that stems from uncertain land-use inventories and from varying accounting conventions across NDCs (Grassi et al., 2021).

3.2 Finance design and controls

The treatment variables describe, besides design, the size of the finance received. The OECD CRS records each operation by donor, recipient, instrument and Rio marker, from which we obtain the projects marked for mitigation or adaptation with a score of one or two. Coverage is that of international support of public origin, and does not include private flows, domestic finance or national instruments that also count as climate finance under broader definitions, such as subsidized credit lines and rural insurance. To smooth the annual oscillation of flows and approximate the execution horizon, each variable is aggregated in a three-year rolling window ending in t , and all variables enter lagged by one year relative to the outcome, so that finance observed through $t - 1$ predicts the gap in t . Both choices are subjected to sensitivity analysis in section 4.4, with two-year and five-year windows and with the contemporaneous measure in place of the lagged one. The lag breaks reverse simultaneity, whereby a country already far from the target attracts more resources for being in a worse position, which would invert the reading of the coefficient by making finance appear associated with larger gaps when finance in fact responds to the observed distance, and measuring finance before the outcome cuts that loop. The phenomenon, known as aid endogeneity, is documented in work on development aid (Burnside & Dollar, 2000; Rajan & Subramanian, 2008) and on climate finance (Weiler et al., 2018).

Fragmentation measures how dispersed support is across donors. Let $s_{d,i,t}$ be the share of donor d in climate disbursements to country i in the three-year rolling window, so that fragmentation is the complement of the Herfindahl-Hirschman index (HHI):

$$Frag_{i,t} = 1 - \sum_{d=1}^D s_{d,i,t}^2 \quad (6)$$

rising when many donors split small shares, following Aldasoro et al. (2010). Concessionality is the fraction of finance that arrives as an actual grant, measured by the ratio between the *grant equivalent* and the amount committed in the window:

$$Conc_{i,t} = \frac{\sum_{\text{window}} \text{GrantEquiv}_i}{\sum_{\text{window}} \text{Committed}_i} \quad (7)$$

Predictability is the ratio between what was disbursed and what was committed in the window:

$$Prev_{i,t} = \frac{\sum_{\text{janela}} \text{Disbursed}_i}{\sum_{\text{janela}} \text{Committed}_i} \quad (8)$$

Values close to 1 indicate that what was promised arrives, and low values signal delay or non-materialization. Volume enters as a scale control, through the logarithm of disbursement in the window, and the number of donors complements fragmentation, so that the effect of design is read net of the size of the flow.

The control vector gathers variables that affect the emissions trajectory and the receipt of resources, so the model does not attribute to finance design the variation arising from differences across countries: income and population in logarithms, the shares of industry and agriculture in output, trade openness, forest area and agricultural land as proportions of the territory, energy per capita, energy intensity and the renewables share, climate vulnerability and readiness, and the six governance dimensions. Table 1 lists every variable, with its definition and source.

Table 1. Panel variables, definitions and sources

Block	Variable	Description	Source
Outcome	$NCGI_{i,t}$	Percentage difference between the 2030 projection and the NDC target, excluding LULUCF	Authors' elaboration
Outcome	$NCGI_{i,t}^{LU}$	Same difference in the accounting that includes land use	Authors' elaboration
Outcome	$Risk_{i,t}$	Binary, equal to 1 if $NCGI_{i,t} > 5\%$	Authors' elaboration
Emissions	$E_{i,t}$	Total GHG emissions, Kyoto, GWP100 AR4, MtCO ₂ e, excluding LULUCF	PRIMAP-hist v2.7 (Gütschow et al., 2025)
Emissions	$E_{i,t}^{LU}$	National total including land use, MtCO ₂ e	Climate Watch (WRI)
Targets	$E_{i,2030}^{NDC}$	2030 NDC target, converted to MtCO ₂ e, conditional in the reference	NDCmitiQ (Günther et al., 2021)
Finance (treatment)	$Frag_{i,t}$	Fragmentation, one minus the donor concentration index, 3 year window, lagged 1 year	OECD CRS
Finance (treatment)	$Conc_{i,t}$	Concessionality, grant equivalent over committed, in the window	OECD CRS
Finance (treatment)	$Prev_{i,t}$	Predictability, disbursed over committed, in the window	OECD CRS
Finance (control)	$Vol_{i,t}$	Volume, logarithm of climate disbursement in the window	OECD CRS
Finance (control)	$NumDonors_{i,t}$	Number of distinct donors in the window	OECD CRS
Finance (control)	$HHI_{i,t-1}$	Donor concentration, Herfindahl-Hirschman index of climate disbursements, 3 year window, lagged 1 year. Enters only in the construction of fragmentation, of which the index is the complement, and is not part of the predictor set	OECD CRS
Control	$\ln PIBpc_{i,t}$	Log of GDP per capita, constant dollars	WDI (World Bank)
Control	$\ln Pop_{i,t}$	Log of population	WDI (World Bank)
Control	$Ind_{i,t}, Agro_{i,t}$	Industry and agriculture as a proportion of GDP	WDI (World Bank)
Control	$Com_{i,t}$	Trade as a proportion of GDP	WDI (World Bank)
Control	$Flor_{i,t}, TerAgr_{i,t}$	Forest area and agricultural land as a proportion of the territory	WDI (World Bank)
Control	$EnerPC_{i,t}, IntEner_{i,t}$	Energy per capita and energy intensity	Our World in Data
Control	$Renov_{i,t}$	Share of renewables in final consumption	WDI (World Bank)
Control	$Vuln_{i,t}, Pront_{i,t}$	Climate vulnerability and readiness	ND-GAIN (Chen et al., 2015)
Control	$Governance_{i,t}$	Six dimensions, estimate	WGI (Kaufmann et al., 2023)

Authors' elaboration. Note: finance variables in a three-year rolling window, from $t - 2$ to t , lagged one year relative to the outcome. Emissions in MtCO₂e with AR4 GWP100. CRS recipient codes are converted to ISO3, and regional aggregates are dropped from the sample.

3.3 Predictive modeling and effect estimation

The model explains the indicator and the risk observed in each year from finance aggregated in the three-year rolling window ending in $t - 1$ and from controls contemporaneous to the diagnosis:

$$NCGI_{i,t} = f(X_{i,t-1}^{fin}, X_{i,t}^{ctrl}) + \varepsilon_{i,t} \quad (9)$$

$$P(Risk_{i,t} = 1) = g(X_{i,t-1}^{fin}, X_{i,t}^{ctrl}) \quad (10)$$

The two outcomes are estimated independently, each with its own tuning, algorithm selection and SHAP decomposition, sharing only the regressors and the temporal split. Since the data are tabular panel rather than a time series, gradient boosting is the best-suited family, because tree models are robust to uninformative variables and learn thresholds and interactions without difficulty on medium-sized samples (Grinsztajn et al., 2022). Under the same split, the study compares a linear baseline (ridge in the continuous version, regularized logit in the binary one), a random forest, four gradient-boosting implementations (histogram-based, XGBoost, LightGBM and CatBoost) (Chen & Guestrin, 2016; Ke et al., 2017; Prokhorenkova et al., 2018) and a multilayer perceptron, with validation that trains on early years and tests on later ones to avoid leakage of the future (Bergmeir et al., 2018). Performance is measured out of sample by the coefficient of determination and, for classification, by the ROC and precision-recall AUC, the latter appropriate under class imbalance (Saito & Rehmsmeier, 2015); Online Resource 2 gives the metric definitions, the tuning grid and the treatment of missing values.

Interpretation uses SHAP, with the implementation for trees, which writes each forecast as the sum of a reference value and the contributions of each variable:

$$\hat{y}_{i,t} = \phi_0 + \sum_{j=1}^P \phi_j^{(i,t)} \quad (11)$$

read at the global scale through the mean absolute importance, ranking variables by how much each one informs the forecast, complemented by the *interaction values* between the three finance dimensions and the six governance dimensions plus institutional readiness, which verify whether support operates conditionally on the institutional capacity of the recipient (Lundberg et al., 2020).

SHAP measures predictive association, so, to understand the effect of each finance dimension on compliance, the study adds a layer of *double machine learning* in the *partially linear regression* specification (Chernozhukov et al., 2018). For each treatment D among predictability, concessionality and fragmentation, we write

$$NCGI_{i,t} = \theta D_{i,t-1} + \ell(Z_{i,t}) + U_{i,t} \quad (12)$$

$$D_{i,t-1} = m(Z_{i,t}) + V_{i,t} \quad (13)$$

Z gathers the 18 controls, the two remaining finance dimensions, the volume disbursed, the number of donors and the year. Volume and the number of donors enter so that the effect of each dimension is examined net of the size of the flow and of the number of partners, and the year captures shocks common to all countries, which tree models accommodate without imposing a functional form. Conditioning on volume carries an implication for interpretation: since predictability and concessionality are ratios whose denominator is the amount committed, and the amount disbursed is held fixed in Z , the variation that identifies the coefficients comes from what was committed, so that the answer is what happens to the gap when the committed portfolio changes, holding constant what actually arrives. The nuisance functions are estimated by *gradient boosting* and the coefficient is obtained through *cross-fitting* repeated over five folds and twenty repetitions, with aggregation by the median. Two procedures respect the panel structure: standard errors admit dependence among observations of the same country, and the *cross-fitting* folds keep each country entirely on the same side of the split, because observations of the same country in different years are nearly identical and, if placed in opposite folds, would let the nuisance functions learn the country identity in training and use that identity in estimation, which would contaminate the effect.

The estimated coefficient can be interpreted as the effect of finance on the gap only if four conditions hold, and each condition has a corresponding test. The condition that donors do not choose countries for reasons invisible to the model requires that finance allocation, once the controls in Z , the country's own finance history and the previous year's gap are netted out, does not depend on unobserved characteristics that also determine the future gap, and the dynamic specification described below incorporates the lagged gap into Z precisely to meet that requirement. Identification also requires that something remain to be compared, so that, even after conditioning on Z , variation in the treatment still exists, and such variation does exist: the residual dispersion of the treatment, measured out of fold, ranges from 80% to 95% of the original dispersion depending on the dimension and the coverage. Nothing in the future may explain the present, which requires that finance not yet received cannot cause today's gap, and the future treatment placebo reported in section 4.3 tests exactly that in two versions, with a one-year and a three-year lead, and in the three-year version the finance window lies entirely ahead of the gap, so that any coefficient appearing in that placebo measures the tendency of donors to choose certain countries rather than an effect of finance. Countries must also not interfere with one another, so that support received by one country does not affect the gap of another, a requirement placed under pressure because the donor budget is fixed in the short run, so that more support for one country means less for another; should that condition fail, the treatments become negatively correlated across countries, which makes finding an effect harder and therefore reinforces a null result instead of manufacturing one. Since conditioning assumptions are never fully verifiable, the sensitivity analysis of Chernozhukov et al. (2026) translates the uncertainty into a number: how strong an omitted factor would need to be, measured by the partial R^2 with the outcome and with the treatment, to conceal an effect capable of moving the indicator by ten percentage points across the interquartile range of the treatment.

Besides the reference specification, three variants test the robustness of the estimate: the specification without country clustering and without volume, number of donors and year in the conditioning set, retained for comparison with the earlier version; the dynamic one, which adds to Z the outcome lagged by one year;

and the correlated effects one, which adds the country means of the treatment and of the covariates (Mundlak, 1978) and approximates a panel estimator with fixed effects.

4. Results

4.1 Which countries are on track to miss their own target?

In 2024, of the 104 countries with a quantifiable target and an emissions series, 70 follow a trajectory that exceeds the NDC, so two thirds of the sample are not on track to meet their own commitment. The median distance is 26% above the target, with the first quartile at -6% and the third at 100%, so the typical country in the upper half projects emissions that double what was undertaken. The reference specification measures distance to the conditional target; the unconditional reading, declared by 77 countries, appears in section 4.4, and Online Resource 3 presents the full ranking.

Table 2 organizes the result for the ten largest emitters, selected by 2024 emissions volume rather than by distance to the target. Iran (71% above target), Saudi Arabia (52%) and Mexico (29%) combine volume with a large gap, alongside Kazakhstan (17%) and South Africa (13%), while Pakistan (-66%), Malaysia (-53%), Vietnam (-19%), Indonesia (-17%) and Thailand (-6%) project below target. Brazil, China and India are absent because none declared a quantifiable conditional target, appearing instead in the unconditional reading of section 4.4, with Brazil 161% above, China 21% above and India 12% below.

Table 2. Distance to the 2030 target in the ten largest emitters of the sample, 2024

Country	2024 emissions (MtCO _{2e})	NCGI (%)	Status
Iran	1,102	71	Above target
Saudi Arabia	852	52	Above target
Mexico	757	29	Above target
Kazakhstan	355	17	Above target
South Africa	484	13	Above target
Thailand	383	-6	Within target
Indonesia	1,137	-17	Within target
Vietnam	528	-19	Within target
Malaysia	367	-53	Within target
Pakistan	456	-66	Within target

Authors' elaboration. Note: conditional target, emissions excluding land use. A positive NCGI indicates a projection above the target. Selection is by emissions volume in 2024, and not by distance to the target. The colour scale runs from green, for the largest slack, to red, for the largest gap. Online Resource 3 presents the complete sample.

The indicator measures distance in percentage terms, how far each country exceeds its own target, which differs from distance in tonnes, how much each country adds to the global gap; the two correlate at 0.75. In percentage terms the top is occupied by small economies such as Laos, Comoros, Angola, Barbados and Niger, whereas in tonnes Iran, Saudi Arabia, Nigeria, the Philippines and Mexico alone account for 55% of the 2,765 MtCO_{2e} of projected excess for 2030, so the countries that depart most from their promise are not those that weigh most in the total.

The land-use sector carries both emissions from deforestation and removals by the forest sink, and that double nature makes its inclusion consequential. Across 90 countries with both versions in 2023, the share at risk falls from 66% without LULUCF to 58% with it, and 33 countries (37% of the sample) reverse classification, 20 moving from risk to slack and 13 the opposite way, following the geography of forests: Laos moves from 7,914% to 140%, Angola from 1,551% to 81% and Niger from 1,208% to -10%, while Chad moves from 133% to 2,136%. Part of the difference between versions could stem from the change of source, but in the 21 countries with under 10% forest cover the median difference is 10 percentage points, far below the displacements above.

The unconditional reading delimits how much risk stems from domestic execution. Among the 77 countries with an unconditional target, 42% exceed it against 67% in the conditional reference, and across the 58 declaring both the share at risk falls from 64% to 34%. The 17 countries that change classification all move from risk to slack, expected because the conditional target is at least as ambitious, and the ordering is preserved (rank correlation 0.90), so for almost a third of the sample the risk depends on whether the conditioned finance materializes.

4.2 What does the model predict and what explains the risk?

With public data and six years before 2030, the countries that will not reach their own target can be identified, and that lead time is what lets the diagnosis serve a decision. Table 3 submits the model to the test that matters, predicting the years not used in estimation, training through 2021 and evaluating on 2022 to 2024. In the continuous version, the best model explains 66% of the variation in the distance without land use and 83% with it, against 8% and 12% for regularized linear regression. In the binary version, the best model correctly ranks 87% and 90% of country pairs, against 54% and 62% for the regularized logit and the 50% of random classification; with classes imbalanced (63% and 57% at risk), the precision-recall AUC confirms the result, reaching 0.914 and 0.918 against prevalences of 0.627 and 0.572.

Table 3. Performance outside the training period, by outcome form and accounting coverage

Model	Continuous w/o LU	Continuous w/ LU	Binary w/o LU	Binary w/ LU	Binary w/o LU	Binary w/ LU
	(R ²)	(R ²)	(AUC)	(AUC)	(PR-AUC)	(PR-AUC)
Linear baseline	0.084	0.119	0.538	0.620	0.695	0.620
Random forest	0.528	0.820	0.872	0.900	0.914	0.910
HistGB	0.518	0.676	0.866	0.887	0.911	0.903
XGBoost	-3.794	0.777	0.873	0.841	0.909	0.870
LightGBM	0.580	0.558	0.847	0.833	0.887	0.858
CatBoost	0.248	0.825	0.856	0.822	0.884	0.839
Multilayer perceptron	0.659	0.756	0.778	0.902	0.837	0.918
Observations (training/test)	765/255	681/152	765/255	681/152	765/255	681/152

Authors' elaboration. Note: reference specification, with conditional target. Metrics computed out-of-sample, on the 2022 to 2024 test set. The linear baseline is ridge regression in the continuous columns and regularized logit in the binary ones, both with the penalty chosen by cross-validation. The out-of-sample R² admits negative values when the model predicts worse than the mean of the test set, and does not correspond to the fit R² of a regression. In the AUC, 0.5 corresponds to a random draw; in the PR-AUC, the benchmark is the prevalence of the at-risk class in the test set, 0.627 without land use and 0.572 with land use. The color scale runs from green, for the best performance, to red, for the worst, within each column. In the continuous outcome without land use, the RMSE and the MAE are, respectively, 769.0 and 500.0 for the linear baseline, 552.3 and 173.2 for the random forest, 557.8 and 195.5 for HistGB, 1,759.4 and 985.3 for XGBoost, 520.5 and 283.2 for LightGBM, 697.0 and 406.6 for CatBoost, and 469.1 and 274.1 for the multilayer perceptron.

Since all algorithms use the same variables, differences across them follow from the functional form each imposes: only the linear baseline is restricted to an additive, linear relation, while the others admit thresholds and interactions, and the same algorithm can behave in opposite ways across specifications, which supports comparing models under the same split rather than fixing one in advance.

What best anticipates compliance is the structure of the economy and the quality of public institutions, measured also by ND-GAIN institutional readiness, not the support received, with the caveat that the ranking measures predictive contribution and not causality. Land-use and energy variables dominate when the land sector stays out of the accounting, whereas institutional indicators take the first positions once land emissions enter (Table 4), consistent with forest emissions depending on enforcement more than on sectoral composition. The volume disbursed appears in fourth place only in the binary specification with land use, the sole presence of finance among the top positions.

Table 4. Ranking of the 23 variables by predictive contribution (SHAP), in the reference specification

Variable	Continuous	Continuous	Binary	Binary
	w/o LU	w/ LU	w/o LU	w/ LU
Energy intensity	4th	6th	1st	6th
Agriculture (% GDP)	6th	1st	5th	8th
Population (log)	3rd	11th	3rd	7th
Forest area (% territory)	2nd	2nd	7th	18th
Industry (% GDP)	8th	12th	8th	3rd
Institutional readiness	12th	3rd	16th	1st
Climate vulnerability	5th	7th	9th	12th
Government effectiveness	19th	4th	10th	2nd
Agricultural land (% territory)	1st	9th	11th	15th
Trade openness (% GDP)	14th	18th	2nd	5th
Control of corruption	17th	5th	13th	9th
Rule of law	9th	16th	4th	22nd
Renewables in final consumption	7th	19th	6th	19th
Volume disbursed (finance, control)	15th	15th	18th	4th
Voice and accountability	11th	20th	15th	10th
GDP per capita (log)	13th	17th	12th	16th
Political stability	10th	21st	19th	11th
Regulatory quality	16th	13th	21st	13th
Fragmentation (finance, treatment)	21st	14th	14th	14th
Energy per capita	20th	8th	17th	23rd
Concessionality (finance, treatment)	22nd	10th	20th	17th
Predictability (finance, treatment)	18th	22nd	22nd	20th
Number of donors (finance, control)	23rd	23rd	23rd	21st

Authors' elaboration. Note: conditional target. Position of each variable in the ranking by mean absolute SHAP contribution, from 1 (largest contribution) to 23. The columns cross the form of the outcome, continuous or binary, with the accounting coverage, without or with land use. Rows are ordered by average position, and the colour scale runs from green, for the first positions, to red, for the last. Variables coming from the CRS are identified in the label itself.

The ranking makes economic sense rather than aggregating correlations. Land and energy variables lead because they describe where emissions come from and how much reducing them costs: where emissions originate in power generation and heavy industry, abatement is decided in a few plants and responds to regulation and price, whereas emissions from land use and agriculture spread across many producers, raising the cost of monitoring and enforcement. Institutional indicators measure how much of that cost each country converts into policy. Compliance is decided by characteristics that no international transfer alters within the 2030 horizon. The result converges with Peterson and van Asselt (2025) and Fekete et al. (2021), the difference being that here the constraint emerges from a forecast validated out of sample.

4.3 Does climate finance explain compliance?

Table 4 places the three finance dimensions between 10th and 22nd of 23 in the predictive ranking, none reaching the upper third in more than one specification, so the support received does not help anticipate who will fall short. The volume disbursed, a scale control, ranges from 4th to 18th with no discernible pattern, so once the structural context enters, finance design adds little about who will miss the targets. Because the reference target is conditional on external support, section 4.4 repeats the exercise against the unconditional target and reaches the same conclusion.

Finance might still operate conditionally, mattering only where institutional capacity turns it into policy. The SHAP interactions between the three dimensions and the seven government-capacity variables address this: the largest single interaction reaches 1.1% and 2.0% of the leading variable in the continuous specifications and 10.5% and 11.6% in the binary ones,, and can exceed the finance dimension's own small contribution, so the little that finance informs passes partly through institutional quality. No interaction rivals the structural predictors, so conditional effectiveness softens the null reading without reversing it.

Contribution to prediction is not effect, since a variable may rank low either for bearing no relation to the outcome or for carrying information already brought by others. Double machine learning estimates the effect of each dimension on the distance to the target, with inference requiring care: because observations of the same country are nearly identical across years, folds keep each country entirely on one side of the split, standard errors admit within-country dependence, and cross-fitting is repeated twenty times with median aggregation (Online Resource 4 details the correction).

Table 5 reports the reference specification for the four combinations of accounting coverage and target family, in percentage points of the indicator per unit of treatment. Fragmentation is estimated with a standard error of the same order as the point estimate, so the 95% intervals contain zero in every coverage and the sign is not identified. Concessionality and predictability are negative and statistically significant at 5% in a single combination, the conditional target without land use, at -153.2 and -58.3, which over the interquartile range correspond to 83 and 24 percentage points against the median gap of 26%; in the other three combinations the same coefficients lie between -44.0 and 5.8 and are not statistically significant. The instability rules out a causal reading, since an effect of finance design would have no reason to appear under one target and vanish under the other for the same countries, nor to depend on the land-use rule.

Table 5. Effect of the finance dimensions on the indicator, reference double machine learning specification

Outcome	Dimension	Effect	Standard error	p-value	95% CI	n	Countries
Without land use, conditional	Fragmentation	-52.9	352.1	0.881	-844.3 to 637.3	592	74
Without land use, conditional	Concessionality	-153.2	78.3	0.050	-315.9 to 0.3	592	74
Without land use, conditional	Predictability	-58.3	20.7	0.005	-103.4 to -17.7	592	74
With land use, conditional	Fragmentation	89.8	123.5	0.467	-154.9 to 331.8	537	69
With land use, conditional	Concessionality	0.1	25.1	0.997	-51.8 to 49.3	537	69
With land use, conditional	Predictability	5.8	11.5	0.613	-18.4 to 28.3	537	69
Without land use, unconditional	Fragmentation	-24.8	54.9	0.652	-136.6 to 82.8	438	54
Without land use, unconditional	Concessionality	-0.4	16.7	0.981	-33.1 to 32.4	438	54
Without land use, unconditional	Predictability	-0.2	5.1	0.972	-11.0 to 9.8	438	54
With land use, unconditional	Fragmentation	-47.7	71.0	0.502	-161.9 to 91.6	391	50
With land use, unconditional	Concessionality	-44.0	51.8	0.396	-138.1 to 57.6	391	50
With land use, unconditional	Predictability	-6.7	14.3	0.640	-36.6 to 21.3	391	50

Authors' elaboration. Note: standard errors clustered by country and cross-fitting folds grouped by country; twenty repetitions over five folds, with aggregation by the median. Z gathers the eighteen controls, the remaining finance dimensions, the volume disbursed, the number of donors and the year.

Table 6 identifies the origin of the result by applying the anticipation placebo, in which the treatment is moved forward in time and compared with the present gap. Finance not yet received cannot cause the current gap, so that any coefficient estimated in that configuration measures donor selection, that is, allocation responding to the trajectory of the country. The exercise is run in two forms: in the first, with a one-year lead, the rolling window covers $t - 1$ to $t + 1$; in the second, with a three-year lead, the window covers $t + 1$ to $t + 3$ and lies entirely ahead of the gap, with no past finance inside the measure. In the combination where the lagged coefficients are statistically significant, future finance produces estimates of larger magnitude in both versions, and the largest appear in the three-year version. Concessionality moves from -153.2 in the lagged estimate to -256.2 with a one-year lead and to -301.5 with a three-year lead, whereas predictability moves from -58.3 to -116.9 and to -134.2, both significant at 5% in the three-year version. Since finance to be received in three years cannot alter the current gap, the coefficient captures the allocation rule: donors direct predictable and concessional support to countries whose gap is already falling, and the persistence of the indicator transfers that selection to the lagged estimates. In the remaining combinations, the placebo is not significant in either version, consistent with the corresponding null results, and fragmentation, whose coefficient falls to -1.3 in the three-year version, confirms the absence of a pattern. Estimates that attain significance only where the placebo attains significance as well constitute a signature of selection, and not of a causal effect.

Table 6. Anticipation placebo: future finance on the present gap

Outcome	Dimension	Lead	Effect	Standard error	p-value	95% CI	n
Without land use, conditional	Fragmentation	1 year	-126.3	400.0	0.752	-862.1 to 657.7	663
Without land use, conditional	Fragmentation	3 years	-1.3	403.7	0.997	-769.6 to 789.9	645
Without land use, conditional	Concessionality	1 year	-256.2	117.0	0.028	-453.1 to -27.0	663
Without land use, conditional	Concessionality	3 years	-301.5	125.8	0.017	-534.4 to -54.8	645
Without land use, conditional	Predictability	1 year	-116.9	30.8	0.000	-174.5 to -56.6	663
Without land use, conditional	Predictability	3 years	-134.2	38.7	0.001	-204.7 to -58.3	645

With land use, conditional	Fragmentation	1 year	33.4	121.4	0.783	-230.5 to 271.3	601
With land use, conditional	Fragmentation	3 years	16.9	120.5	0.888	-218.4 to 253.1	587
With land use, conditional	Concessionality	1 year	-18.4	17.0	0.281	-44.4 to 15.0	601
With land use, conditional	Concessionality	3 years	-1.9	19.9	0.923	-45.2 to 37.0	587
With land use, conditional	Predictability	1 year	0.2	8.0	0.983	-13.8 to 15.8	601
With land use, conditional	Predictability	3 years	-1.1	7.2	0.877	-16.0 to 13.0	587
Without land use, unconditional	Fragmentation	1 year	-11.6	46.7	0.804	-114.3 to 80.0	491
Without land use, unconditional	Fragmentation	3 years	-21.7	49.5	0.660	-104.4 to 75.2	481
Without land use, unconditional	Concessionality	1 year	-0.5	16.6	0.974	-32.3 to 31.9	491
Without land use, unconditional	Concessionality	3 years	14.2	11.2	0.204	-8.9 to 36.2	481
Without land use, unconditional	Predictability	1 year	0.7	5.1	0.895	-9.1 to 10.7	491
Without land use, unconditional	Predictability	3 years	5.2	6.2	0.401	-7.6 to 17.4	481
With land use, unconditional	Fragmentation	1 year	-38.9	59.5	0.513	-163.5 to 77.6	438
With land use, unconditional	Fragmentation	3 years	-19.7	51.7	0.703	-136.9 to 81.6	432
With land use, unconditional	Concessionality	1 year	-50.0	38.7	0.196	-124.4 to 25.8	438
With land use, unconditional	Concessionality	3 years	-26.1	25.9	0.315	-74.9 to 24.8	432
With land use, unconditional	Predictability	1 year	-10.3	10.6	0.331	-33.1 to 10.5	438
With land use, unconditional	Predictability	3 years	1.9	7.0	0.787	-12.0 to 15.5	432

Authors' elaboration. Note: under the one-year lead, the treatment is the three-year rolling window ending in $t+1$, which covers $t-1$ to $t+1$; under the three-year lead, the window covers $t+1$ to $t+3$ and is strictly future relative to the gap in t . All other choices are identical to Table 5. In the absence of anticipation, the coefficient measures donor selection, and not an effect of finance.

Outside the combination explained by the placebo, predictability shows the narrowest intervals, from -11.0 to 9.8 in the unconditional analysis, which rules out effects of appreciable magnitude, whereas fragmentation runs from -844.3 to 637.3 and lacks the precision to answer, requiring exogenous variation. The null result matches the allocation literature: mitigation finance is directed to countries with larger sinks, better governance and lower income (Halimanjaya, 2015; Weiler et al., 2018), the same variables this article finds at the top of the explanation of compliance, so conditional on forest, governance and income finance has little to add. This is what one would expect if finance were more a symptom of the implementation environment than a cause, compatible with Rashid et al. (2023) and with the 0.03% elasticity of Wei (2025), which would move the compliance of no country in this sample; the placebo in Table 6 documents that selection, since finance responds to the trajectory of the gap before it arrives.

4.4 Robustness

The results were submitted to eight lines of checking, summarized here, with Tables C1 to C5 in Online Resource 5 and the extended discussion in Online Resource 4; none alters the conclusions. Replacing the conditional target with the unconditional one lowers the share at risk from 67% to 42% and the median from 26% to -2%, an expected difference, while preserving the ordering (rank correlation 0.90 across the 58 countries declaring both), out-of-training performance and the low ranking of the finance dimensions, with no effect estimate reaching significance (Table 5). Replacing the target with its revised PBL version, which correlates at 0.98 with the reference, raises the share at risk from 38% to 64% across 55 countries, since 42 revised downward, so the unrevised target makes the diagnosis conservative. Varying the growth window between three and seven years leaves the share at risk between 62% and 65%, and moving the risk threshold from five to zero and ten percentage points moves it from 67% to 55% (Table C4). Rebuilding the finance windows with two and five years keeps performance, while the contemporaneous measure raises fragmentation to 5th position (Table C5). Removing the energy controls, which could mediate the finance channel, preserves the single significant pair and leaves fragmentation imprecise (Table C3), and the structural controls barely vary within country (within-variance shares of 0.1% to 7.7%), so no reconfiguration finance could have caused occurred within the horizon. On identification, the dynamic and correlated-effects variants keep the same pair while fragmentation oscillates, a sensitivity analysis shows that hiding an effect of predictability would require an omitted confounder with partial R^2 of 0.14 against at most 0.01 for government effectiveness, and a donor-side shift-share instrument (Nunn & Qian, 2014) has first-stage strength only for fragmentation and yields a negative, imprecise estimate (Tables C1 and C2).

5. Concluding remarks

This study asked which countries risk missing their own NDC by 2030 and whether the design of climate finance helps anticipate the answer. To the first, two thirds (67%) of the 104 countries follow a trajectory that exceeds the commitment, with a median gap of 26% and a third quartile at 100%, so the more distant half is heading toward doubling what was promised. To the second, the answer is negative: the three finance dimensions rank between 10th and 22nd of 23 in predictive contribution, and in the reference double machine learning specification two of twelve estimates cross the 5% threshold, both identified by the anticipation placebo as donor selection rather than effect. None of the checks displaces these answers.

What best anticipates the separation between compliant and non-compliant countries is productive structure, not the support received: agricultural land and forest as proportions of the territory, the weight of agriculture in output, population and energy intensity lead the ranking, and once land use enters the accounting institutional readiness and government effectiveness take the lead. Where emissions concentrate in a few fuel-burning sectors, abatement carries cost and governance distinct from those of economies whose emissions come from land and dispersed agriculture, and that difference already shows in compliance before any international transfer.

The accounting ambiguity of land use has the most immediate consequence for the climate architecture. Because the sector combines emissions from deforestation with removals by the forest sink, the choice to include or exclude it reclassifies 37% of countries and displaces weighty cases, with Laos moving from 7,914% to 140% and Chad the reverse way. This speaks to the literature describing climate-finance accounting as a tangle of competing definitions (Weikmans & Roberts, 2019) and shows the problem reaches the measurement of compliance itself: where the counting rule alters the outcome for a third of those assessed, the commitment ceases to be an observable price, and a regime combining self-declared targets, absence of sanction and freedom of accounting scope makes the promise cheap where counting is discretionary, at a cost to the credibility of the review mechanism.

The four distinguishing choices hold up unevenly. The null for finance does not invalidate the decomposition, which ruled out predictability as a lever within the estimated range and left fragmentation open owing to imprecision. Measuring compliance with and without land use proved the most consequential source of variation, determining even the sign of the causal coefficient. Separating association from effect blocked the conclusion that a low SHAP position rules an effect out, and coverage of 104 countries with annual recalculation delivers a continuous, auditable measure without loss of accuracy.

The diagnosis matters to three audiences. To national authorities and to the global stocktake, the indicator offers a measure recalculable every year, anticipating by six years what the 2030 verification would record too late, a gain that is informational, since targets remain self-determined and free of sanction. To finance providers, the return to reforming the design of support, harmonizing donors or raising the grant share, is low over the range observed, so attention moves to supporting the quality of public institutions and standardizing land-use accounting, the latter of low cost and high informational yield. To recipient governments, the comparison between conditional and unconditional targets separates the risk that responds to domestic constraints from the risk that depends on promised support, sustaining a negotiating position grounded in numbers.

Among the limitations, the projection assumes continuity of the recent trend and is a diagnosis of trajectory rather than a forecast conditional on future decisions; the causal sample of 391 to 592 observations produces wide intervals, so fragmentation signals a lack of precision rather than a demonstrated null; and the population cutoff removes the small island states, where per capita finance concentrates. Sectoral composition and the energy profile respond in part to the decarbonization effort itself, though within the horizon examined they barely move within country and their removal does not shift the coefficients in a definite direction, and the 2021 target is fixed across the panel, so time-varying targets remain a natural extension, as does an additionality approach comparing observed change with a counterfactual that public data do not offer.

References

Aldasoro, I., Nunnenkamp, P., & Thiele, R. (2010). Less aid proliferation and more donor coordination? The wide gap between words and deeds. *Journal of Development Studies*, 46(9), 1633-1652. <https://doi.org/10.1080/00220388.2010.492855>

- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685-725. <https://doi.org/10.1146/annurev-economics-080217-053433>
- Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120, 70-83. <https://doi.org/10.1016/j.csda.2017.10.001>
- Betzold, C., & Weiler, F. (2017). Allocation of aid for adaptation to climate change: Do vulnerable countries receive more support? *International Environmental Agreements: Politics, Law and Economics*, 17(1), 17-36. <https://doi.org/10.1007/s10784-016-9343-8>
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199-231. <https://doi.org/10.1214/ss/1009213726>
- Buchner, B., Naran, B., Padmanabhi, R., Stout, S., Strinati, C., Wignarajah, D., Miao, G., Connolly, J., & Marini, N. (2023). *Global landscape of climate finance 2023*. Climate Policy Initiative. <https://www.climatepolicyinitiative.org/publication/global-landscape-of-climate-finance-2023/>
- Burnside, C., & Dollar, D. (2000). Aid, policies, and growth. *American Economic Review*, 90(4), 847-868. <https://doi.org/10.1257/aer.90.4.847>
- Chen, C., Noble, I., Hellmann, J., Coffee, J., Murillo, M., & Chawla, N. (2015). *University of Notre Dame Global Adaptation Index: Country index technical report*. University of Notre Dame. <https://gain.nd.edu/our-work/country-index/>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). ACM. <https://doi.org/10.1145/2939672.2939785>
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), 1-68. <https://doi.org/10.1111/ectj.12097>
- Chernozhukov, V., Cinelli, C., Newey, W. K., Sharma, A., & Syrgkanis, V. (2026). Long story short: Omitted variable bias in causal machine learning. *The Review of Economics and Statistics*. <https://doi.org/10.1162/REST.a.1705>
- den Elzen, M. G. J., Dafnomilis, I., Hooijschuur, E., Nascimento, L., Kuramochi, T., Forsell, N., Gutiérrez, Z. A., & Gusti (2024). *Infographics PBL NDC tool March 2024* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.10875617>
- Fekete, H., Kuramochi, T., Roelfsema, M., den Elzen, M., Forsell, N., Höhne, N., Luna, L., Hans, F., Sterl, S., & Olivier, J. G. J. (2021). A review of successful climate change mitigation policies in major emitting economies and the links to nationally determined contributions. *Renewable and Sustainable Energy Reviews*, 137, 110602. <https://doi.org/10.1016/j.rser.2020.110602>
- Grassi, G., Stehfest, E., Rogelj, J., van Vuuren, D., Cescatti, A., House, J., Nabuurs, G.-J., Rossi, S., Alkama, R., Viñas, R. A., Calvin, K., Ceccherini, G., Federici, S., Fujimori, S., Gusti, M., Hasegawa, T., Havlik, P., Humpenöder, F., Korosuo, A., Perugini, L., Tubiello, F. N., & Popp, A. (2021). Critical adjustment of land mitigation pathways for assessing countries' climate progress. *Nature Climate Change*, 11(5), 425-434. <https://doi.org/10.1038/s41558-021-01033-6>
- Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? *Advances in Neural Information Processing Systems*, 35, 507-520. <https://doi.org/10.48550/arXiv.2207.08815>
- Günther, A., Gütschow, J., & Jeffery, M. L. (2021). NDCmitiQ v1.0.0: A tool to quantify and analyse greenhouse gas mitigation targets. *Geoscientific Model Development*, 14(9), 5695-5730. <https://doi.org/10.5194/gmd-14-5695-2021>
- Gütschow, J., Jeffery, L., Gieseke, R., & Günther, A. (2025). *PRIMAP-hist national historical emissions time series (v2.7)*. Zenodo. <https://doi.org/10.5281/zenodo.17090760>
- Halimanjaya, A. (2015). Climate mitigation finance across developing countries: What are the major determinants? *Climate Policy*, 15(2), 223-252. <https://doi.org/10.1080/14693062.2014.912978>
- Intergovernmental Panel on Climate Change. (2022). *Climate change 2022: Mitigation of climate change*. Cambridge University Press. <https://doi.org/10.1017/9781009157926>

- International Energy Agency. (2024). *World energy balances*. IEA. <https://www.iea.org/data-and-statistics/data-product/world-energy-balances>
- International Monetary Fund. (2024). *World economic outlook, October 2024*. IMF. <https://www.imf.org/en/Publications/WEO>
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2023). *Worldwide governance indicators*. World Bank. <https://info.worldbank.org/governance/wgi/>
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 3146-3154. https://papers.nips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774. https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56-67. <https://doi.org/10.1038/s42256-019-0138-9>
- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46(1), 69-85. <https://doi.org/10.2307/1913646>
- Nor, M. I. (2025). Investigating the dynamics of climate finance disbursements: A panel data approach from 2003 to 2022. *PLOS ONE*, 20(3), e0318170. <https://doi.org/10.1371/journal.pone.0318170>
- Northrop, E., Biru, H., Lima, S., Bouye, M., & Song, R. (2016). *Examining the alignment between the intended nationally determined contributions and Sustainable Development Goals*. World Resources Institute. <https://www.wri.org/research/examining-alignment-between-intended-nationally-determined-contributions-and-sustainable>
- Nunn, N., & Qian, N. (2014). US food aid and civil conflict. *American Economic Review*, 104(6), 1630-1666. <https://doi.org/10.1257/aer.104.6.1630>
- OECD. (2024). *Climate finance provided and mobilised by developed countries in 2013-2022*. OECD Publishing. <https://doi.org/10.1787/19150727-en>
- Pauw, W. P., Castro, P., Pickering, J., & Bhasin, S. (2020). Conditional nationally determined contributions in the Paris Agreement: Foothold for equity or Achilles heel? *Climate Policy*, 20(4), 468-484. <https://doi.org/10.1080/14693062.2019.1635874>
- Peterson, L., & van Asselt, H. (2025). Assessing risks to the implementation of NDCs under the Paris Agreement. *Climate Policy*, 25(4), 462-476. <https://doi.org/10.1080/14693062.2025.2513023>
- Pickering, J., Skovgaard, J., Kim, S., Roberts, J. T., Rossati, D., Stadelmann, M., & Reich, H. (2015). Acting on climate finance pledges: Inter-agency dynamics and relationships with aid in contributor states. *World Development*, 68, 149-162. <https://doi.org/10.1016/j.worlddev.2014.10.033>
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31, 6638-6648. https://papers.nips.cc/paper_files/paper/2018/hash/14491b756b3a51daac41c24863285549-Abstract.html
- Rajan, R. G., & Subramanian, A. (2008). Aid and growth: What does the cross-country evidence really show? *Review of Economics and Statistics*, 90(4), 643-665. <https://doi.org/10.1162/rest.90.4.643>
- Rashid, S., Khan, M. R., & Haque, N. (2023). Does climate finance enhance mitigation ambitions of recipient countries? *Earth System Governance*, 17, 100188. <https://doi.org/10.1016/j.esg.2023.100188>
- Roelfsema, M., van Soest, H. L., Harmsen, M., van Vuuren, D. P., Bertram, C., den Elzen, M., Höhne, N., Iacobuta, G., Krey, V., Kriegler, E., Luderer, G., Riahi, K., Ueckerdt, F., Després, J., Drouet, L., Emmerling, J., Frank, S., Fricko, O., Gidden, M., ... Vishwanathan, S. S. (2020). Taking stock of national climate policies to evaluate implementation of the Paris Agreement. *Nature Communications*, 11, 2096. <https://doi.org/10.1038/s41467-020-15414-6>
- Rogelj, J., den Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K., & Meinshausen, M. (2016). Paris Agreement climate proposals need a boost to keep warming well below 2°C. *Nature*, 534(7609), 631-639. <https://doi.org/10.1038/nature18307>

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>

United Nations. (2024). *World population prospects 2024*. United Nations, Department of Economic and Social Affairs. <https://population.un.org/wpp/>

United Nations Framework Convention on Climate Change. (2015). *Adoption of the Paris Agreement (Decision 1/CP.21, FCCC/CP/2015/10/Add.1)*. UNFCCC. <https://unfccc.int/resource/docs/2015/cop21/eng/10a01.pdf>

Wei, B. (2025). Does climate finance achieve its goals in developing countries? An econometric assessment of mitigation and adaptation outcomes. *Discover Sustainability*, 6, 441. <https://doi.org/10.1007/s43621-025-01284-9>

Weikmans, R., & Roberts, J. T. (2019). The international climate finance accounting muddle: Is there hope on the horizon? *Climate and Development*, 11(2), 97-111. <https://doi.org/10.1080/17565529.2017.1410087>

Weiler, F., Klöck, C., & Dornan, M. (2018). Vulnerability, good governance, or donor interests? The allocation of aid for climate change adaptation. *World Development*, 104, 65-77. <https://doi.org/10.1016/j.worlddev.2017.11.001>

World Bank. (2024). *World development indicators*. World Bank. <https://databank.worldbank.org/source/world-development-indicators>

World Resources Institute. (2024). *Climate Watch NDC database*. WRI. <https://www.climatewatchdata.org/ndcs-explore>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformationanon.docx](#)