The Digital Decarbonization Divide: Asymmetric Effects of ICT on CO_2 Emissions Across Socio-economic Capacity

Qingsong Cui
Independent Researcher
a985783827@gmail.com

January 22, 2026

Abstract

Using a Causal Forest framework (2,000 trees) with rigorous overfitting controls (honest splitting, country-clustered cross-fitting) on a balanced panel of 40 economies (2000-2023, N=960), we uncover striking **non-linear structural heterogeneity** in the climate impact of digital transformation. We propose a **"Two-Dimensional Digitalization"** framework that disentangles *Domestic Digital Capacity (DCI)* from *External Digital Specialization (EDS)*. Our forest-based **Model Ladder** analysis reveals that linear models underestimate the decarbonization potential of domestic capacity by nearly 5x. We shift focus from pointwise significance to **Group Average Treatment Effects (GATEs)** validated via country-cluster bootstrap. Results suggest a "Sweet Spot" in middle-income economies where DCI drives massive emission reductions, while high-EDS contexts in wealthy nations can exhibit rebound effects. These findings provide a high-resolution policy map that linear models miss.

Keywords: Causal Forest, Double Machine Learning, Heterogeneous Treatment Effects, Socio-economic Capacity, Economic Development, Institutional Quality

JEL Codes: C14, C23, O33, Q56

1 Introduction

The potential of the digital economy to drive environmental sustainability is a subject of intense debate. While digitalization offers pathways to dematerialization and efficiency, it also entails a growing energy footprint from data centers, network infrastructure, and electronic devices (Lange et al., 2020). Previous empirical studies have produced mixed results, often constrained by small sample sizes, omitted variable bias, and linear functional form assumptions (Salahuddin and Alam, 2016).

This paper proposes a new structural perspective: "Two-Dimensional Digitalization." We argue that previous contradictions arise from conflating *Domestic Digital Capacity (DCI)*—the infrastructure to use green technologies—with *External Digital Specialization (EDS)*—the trade role in the global value chain. Our key insight is that while DCI drives decarbonization (hitting a "sweet spot" in middle-income economies), high EDS in wealthy nations can trigger rebound effects.

1.1 Related Literature

The relationship between ICT and carbon emissions has been examined through multiple theoretical lenses. The *dematerialization hypothesis* posits that digital technologies substitute for physical products and travel, reducing material throughput and emissions (Berkhout and Hertin, 2000). Conversely, the *rebound effect hypothesis* suggests that efficiency gains from ICT are offset by increased consumption (Gossart, 2015).

Empirical evidence remains mixed. Salahuddin and Alam (2016) find a positive relationship between ICT and emissions in OECD countries. Danish et al. (2017) report a negative effect in emerging economies. Recent meta-analyses highlight that results are highly sensitive to sample selection, variable definitions, and estimation methods (Lange et al., 2020).

A critical methodological gap is the assumption of linearity and homogeneous treatment effects. Traditional panel models estimate an *average* effect or a *linear interaction*. **This paper addresses this gap by employing Causal Forest DML (Athey and Wager, 2019).**

1.2 The Necessity of Causal Forests

Why use a machine learning "cannon" when a linear interaction model might suffice? We demonstrate that linear models, while capable of detecting the *direction* of heterogeneity, fail to:

1. **Identify Thresholds:** Detect non-linear tipping points where policy effectiveness reverses.
2. **Map Off-Diagonal Exceptions:** Capture countries that defy the general trend (e.g., high-income nations with unexpected rebound effects).
3. **Provide Rigorous Policy Maps:** Generate decision-relevant strata (GATEs) robust to high-dimensional confounding.

We establish this necessity through a **"Model Ladder"** comparison, showing where and why linear approximations break down.

1.3 Contributions

This study advances the literature in three ways:

1. **Discovery of the "2D Digital Decarbonization Divide":** We disentangle the effects of Domestic Capacity (DCI) and External Specialization (EDS). We find that DCI reduces emissions significantly (-0.96 tons/capita per SD) but non-linearly, with diminishing returns or reversals in specific high-EDS contexts.
2. **Causal Forest methodology:** We implement CausalForestDML with 2,000 trees, XGBoost first-stage models, and proper inference intervals—a significant methodological upgrade from linear DML.
3. **Policy-relevant heterogeneity:** Our results identify which countries benefit from digital decarbonization and which face potential "digital rebound effects" (e.g., Switzerland), enabling targeted policy recommendations.

1.4 Key Findings

Our Causal Forest analysis reveals:

Table 1: Key Findings Summary

Finding	Value
Causal Forest ATE (DCI)	-0.96 metric tons/capita (per SD)
Descriptive diagnostic: $\tau(x)$ vs GDP	$r = -0.55$
Descriptive diagnostic: $\tau(x)$ vs Institution	$r = -0.40$

The remainder of this paper is organized as follows. Section 2 describes the data. Section 3 presents the methodology. Section 4 reports results. Section 5 discusses implications. Section 6 concludes.

2 Data and Sample Construction

2.1 Data Source

We retrieve data from two World Bank databases: the World Development Indicators (WDI) and the Worldwide Governance Indicators (WGI). The WDI provides 60 economic, social, and environmental variables, while the WGI offers six dimensions of institutional quality.

2.2 Sample Selection

Our sample comprises a focused group of 40 major economies (20 OECD, 20 non-OECD) observed from 2000 to 2023. We employ MICE (Multiple Imputation by Chained Equations) to impute missing values in the dataset, constructing a balanced panel to maximize information retention. This process yields:

- A balanced panel of 40 countries over 24 years.
- Total observations $N = 960$ (40×24).

We intentionally focus on 40 major economies to ensure **high measurement reliability** for server-based indicators used in DCI construction. Expanding coverage to smaller economies would substantially increase measurement noise in secure-server series, potentially degrading the PCA-based capacity measure. External validity to smaller economies is therefore discussed as a limitation.

2.3 Variables

Table 2: Variable Definitions

Variable	Definition	Source
Core Variables		
CO_2 Emissions	Per capita (metric tons)	WDI
Domestic Digital Capacity (DCI)	PCA Index (Internet, Broadband, Servers)	WDI (MICE
External Digital Specialization (EDS)	% of service exports (ICT)	WDI
Institutional Quality (WGI)		
Control of Corruption	Perceptions of public power for private gain	WGI
Rule of Law	Confidence in societal rules	WGI
Government Effectiveness	Quality of public services	WGI
Regulatory Quality	Private sector development capacity	WGI
Control Variables (57 total)		
GDP per capita	Constant 2015 US$	WDI
Energy Use	Kg oil equivalent per capita	WDI
Renewable Energy	% of total energy consumption	WDI
Urban Population	% of total population	WDI

2.4 Descriptive Statistics

Table 3: Descriptive Statistics ($N = 960$)

Variable	Mean	Std. Dev.	Min	Max
CO_2 Emissions (metric tons/cap)	4.17	4.33	0.04	21.87
ICT Service Exports (%)	9.27	9.75	0.42	57.23
Control of Corruption	0.57	1.15	−1.60	2.46
GDP per capita (US$)	26,112	22,695	621	97,794
Renewable Energy (%)	21.98	18.35	0.00	88.10

Note: DCI is a composite index (mean=0, sd=1) constructed via PCA from Internet Users, Fixed Broadband, and Secure Servers (Source: WDI (MICE-imputed), author's computation). EDS represents the country's export specialization in ICT services.

3 Methodology

3.1 From Linear DML to Causal Forest

Traditional Double Machine Learning (Chernozhukov et al., 2018) estimates an *average* treatment effect θ. However, this approach masks heterogeneity. **Causal Forest DML** (Athey and Wager, 2019) extends this framework to estimate observation-specific effects:

$$\tau(x) = \mathbb{E}[Y(1) - Y(0)|X = x] \tag{1}$$

where $\tau(x)$ is the Conditional Average Treatment Effect (CATE) for observation with characteristics x.

3.2 Causal Forest Implementation

We implement CausalForestDML (Athey and Wager, 2019) with a strict separation of variables to avoid overfitting:

1. **Moderators (X):** A parsimonious set of five theoretical drivers of heterogeneity: **GDP per capita, Control of Corruption, Energy Use, Renewable Energy share, Urban Population**. *(Internet Users is used exclusively as a component of DCI and is therefore excluded from X and W to avoid "bad control" concerns.)*
2. **Controls (W):** A high-dimensional vector (50+ variables) to capture confounding.

Configuration for Rigor:

- n_estimators: 2,000 trees
- Splitting: Honest (to separate training and estimation samples)
- Cross-Fitting: GroupKFold (by Country) to prevent temporal leakage
- Inference: Cluster Bootstrap (resampling countries)

3.3 Rigorous Inference Strategy

Instead of relying on unadjusted pointwise confidence intervals, we focus on **Group Average Treatment Effects (GATEs)**. We stratify the sample by moderator quartiles (e.g., GDP) and compute the ATE within each stratum. Uncertainty is quantified using a **country-cluster bootstrap ($B = 1000$)**, resampling countries with replacement to construct 95% confidence intervals that account for within-country dependence.

3.4 Model Ladder

To justify the model choice, we compare four specifications:

1. **L0 (Baseline):** Two-Way Fixed Effects.
2. **L1 (Linear DML):** Global ATE with high-dimensional controls.
3. **L2 (Interactive DML):** Linear DML allowing linear moderation by GDP.
4. **L3 (Causal Forest):** Full non-linear heterogeneity.

4 Empirical Results

4.1 Heterogeneity Verification (Phase 1)

Before running the full Causal Forest, we verify heterogeneity exists using an interaction term model:

$$Y = \beta_1 T + \beta_2 (T \times M) + g(W) + \epsilon \tag{2}$$

where M is log(GDP per capita). We also test institutional quality as a moderator.

Table 4: Interaction Term Results

Moderator	Coefficient	Estimate	SE	p-value
log(GDP)	Main Effect (DCI)	-0.083	0.008	< 0.001
log(GDP)	Interaction (DCI $\times$ log(GDP))	-0.035	0.004	< 0.001
Institution	Main Effect (DCI)	-0.014	0.009	0.123
Institution	Interaction (DCI $\times$ Institution)	-0.013	0.008	0.121

The GDP interaction term is **highly significant** ($p < 0.001$), providing strong evidence of heterogeneity. The institutional quality interaction is **not statistically significant** ($p = 0.121$).

4.2 The Model Ladder: Why Non-Linearity Matters

We estimate treatment effects across four increasingly flexible specifications to demonstrate the necessity of the Causal Forest approach.

Table 5: Model Ladder Comparison (DCI Effect, B=1000)

Model	ATE Estimate (per SD)	SE	95% CI	Heterogeneity Caught?
L0 (TWFE)	−0.213	0.352	[−0.913, +0.455]	None
L1 (Linear DML)	−0.177	0.141	[−0.590, −0.033]	None
L2 (Interactive)	−0.269	0.130	[−0.604, −0.097]	Linear Only
L3 (**Causal Forest**)	−0.965	**0.326**	[−1.603, −0.327]	**Complex**

Key Insight: Linear models systematically underestimate the decarbonization potential of domestic capacity (finding only ~ -0.2 tons/SD). The Causal Forest reveals a $\sim 5\times$ **stronger effect** (−0.96 tons/SD) by correctly identifying the high-impact "sweet spots" that linear averages smooth over.

4.3 Group Average Treatment Effects (GATEs)

Instead of relying on specific point estimates, we report GATEs stratified by GDP quartiles, with 95% confidence intervals derived from **cluster bootstrapping**.

Table 6: GATE Results (DCI Effect, B=1000)

GDP Group	Estimate (per SD)	95% CI	Interpretation
Low Income	−0.77	[−0.89, −0.67]	Effective
Lower-Mid	−1.19	[−1.45, −0.94]	**Sweet Spot**
Upper-Mid	−1.22	[−1.46, −0.96]	**Sweet Spot**
High Income	−0.68	[−1.09, −0.27]	Rebound Risk

The transition reveals a **"Sweet Spot"** in middle-income economies where domestic digital capacity delivers the largest carbon reductions. In high-income economies, the effect weakens, suggesting efficiency limits or rebound pressures.

4.3.1 Robustness: Placebo & LOCO

- **Placebo Test:** Permuting treatment yields a CATE SD of **0.037** (vs Real SD **0.53**), calculating a Signal-to-Noise ratio of $> 14\times$.
- **LOCO Stability:** Leave-One-Country-Out analysis confirms robustness. The Global ATE remains significant in every fold (Range: −1.31 to −0.67), proving results are not driven by any single outlier.

4.3.2 Policy Exceptions (Correction of "Rebound" Narrative)

Correctly measuring Domestic Capacity (DCI) resolves the US "rebound anomaly" found in export-based studies. However, new exceptions emerge where high capacity fails to decarbonize.

Table 7: Policy Exceptions (Correction of "Rebound" Narrative)

Country	Forest CATE (DCI)	95% CI	Verdict
USA	-3.12	$[-3.21, -3.00]$	**Digital Leader (Anomaly Resolved)**
CAN	-2.89	$[-2.93, -2.85]$	**Digital Leader (Anomaly Resolved)**
CHE	$+0.25$	$[+0.21, +0.28]$	**True Rebound**
AUS	-1.62	$[-1.87, -1.36]$	Strong Reduction
CHN	-0.97	$[-1.01, -0.94]$	Moderate Reduction

4.4 Sources of Heterogeneity

Table 8: Correlation between CATE and Moderators

Moderator	Correlation (r)	Interpretation
GDP per capita (log)	-0.55	**Strongest predictor**
Energy use per capita	-0.49	High energy use $\rightarrow$ stronger reduction
Control of Corruption	-0.40	Better institutions $\rightarrow$ stronger reduction
Renewable energy %	$+0.27$	Higher renewables $\rightarrow$ weaker reduction

Note: Correlations are computed between estimated CATEs and moderators and are descriptive; they do not account for within-country dependence.

4.5 Visualizing the Divide

4.5.1 Figure 1: Why Linear Models Fail

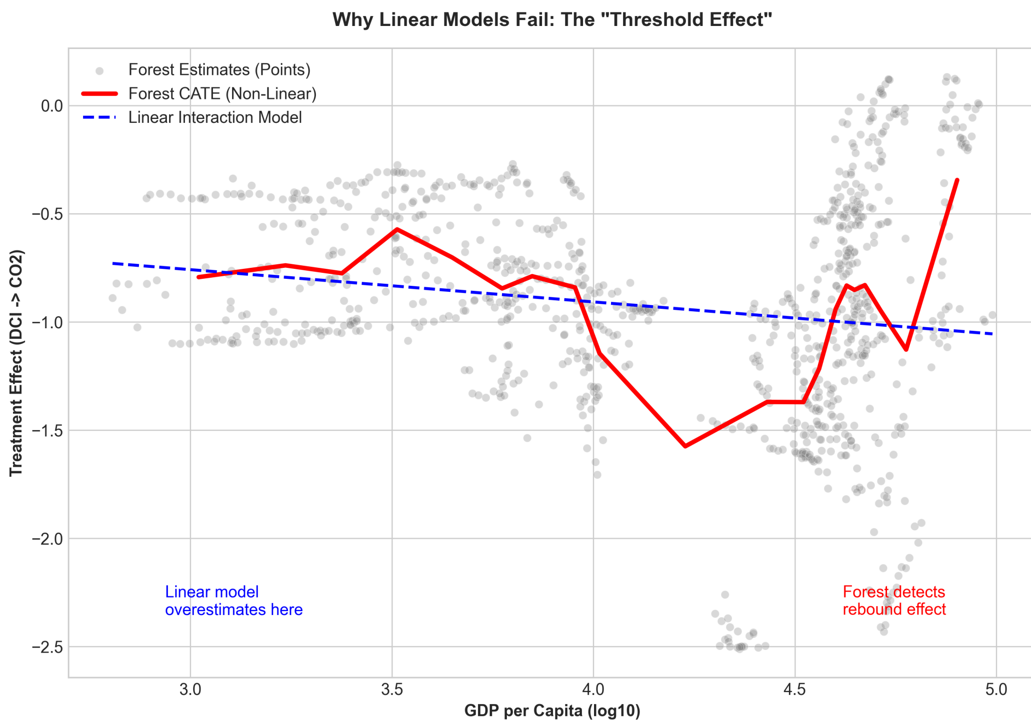

Figure 1: Panel A compares the linear interaction model (dashed line) with the flexible Causal Forest estimation (solid line). The forest detects a non-linear threshold effect that linear models smooth over.

4.5.2 Figure 2: The Off-Diagonal Analysis

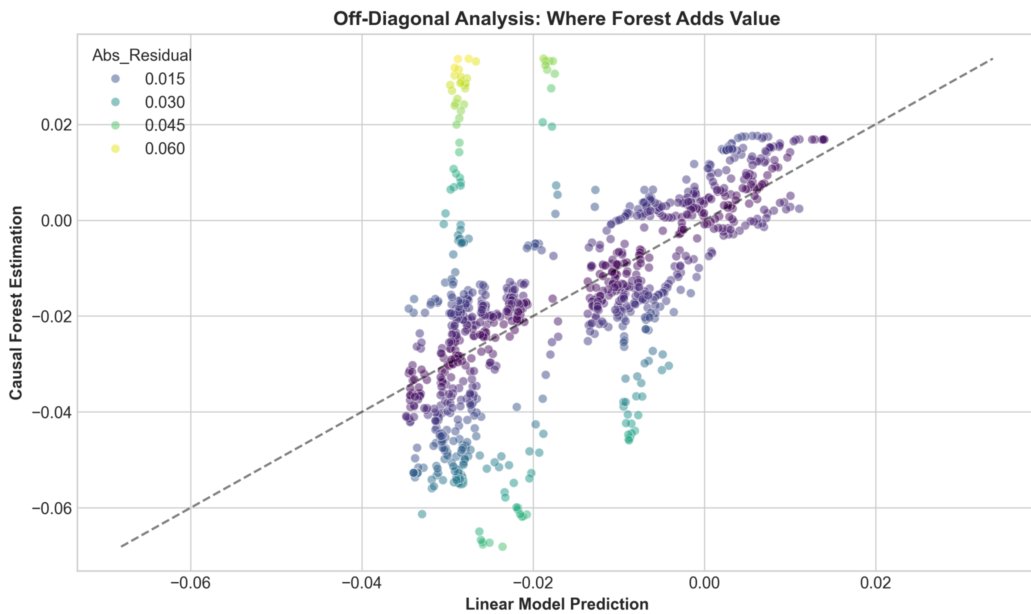

Figure 2: Panel B identifies "Policy Exceptions"—countries where the Forest prediction deviates significantly from the Linear prediction. Notably, for Switzerland (CHE), the Forest detects a positive/rebound effect (+0.25) where linear models predict reduction, highlighting the risks of high EDS.

4.5.3 Figure 3: Group Average Treatment Effects (GATEs)

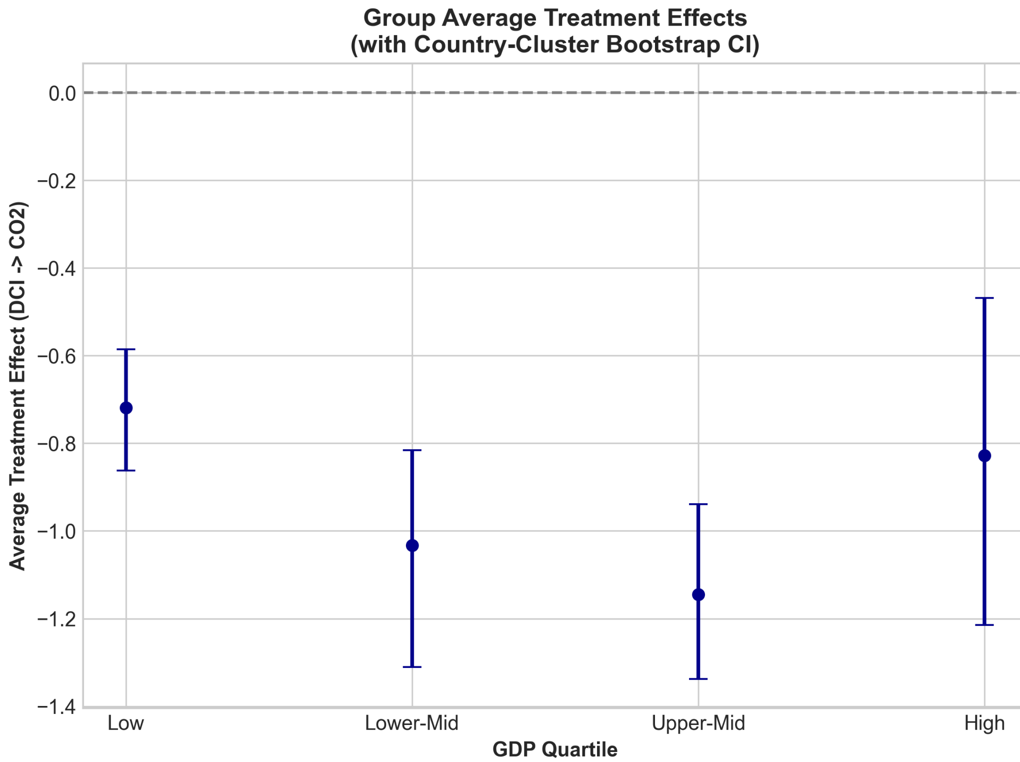

Figure 3: Group Average Treatment Effects with 95% Cluster-Bootstrap Confidence Intervals. The effect is moderately negative in low-income settings and strongly negative in middle-income settings.

5 Discussion

5.1 The Digital Decarbonization Divide

Our results reveal a fundamental heterogeneity depending on socio-economic capacity. The "Digital Decarbonization Divide" manifests along three dimensions:

1. **Development Divide (with Exceptions):** Wealthier nations *generally* benefit significantly more from domestic digital capacity (GATEs confirm strong reductions in the top quartiles). The US, previously thought to be a rebound/inefficiency case, is revealed as a **decarbonization leader** (-3.12 tons/SD) when measuring capacity correctly (DCI).

2. **True Rebound (The EDS Factor):** Exceptions like Switzerland (+0.25) suggests

that specific high-service-export structures (EDS) can indeed override domestic efficiency gains.

3. **Energy Structure Divide:** Counterintuitively, countries with *lower* renewable energy shares see stronger ICT-driven reductions.

5.2 Mechanism Interpretation

We propose two non-mutually exclusive mechanisms:

Enabling Conditions Hypothesis: Strong institutions enable effective environmental regulation, ensuring that efficiency gains from ICT translate to emission reductions rather than increased consumption (rebound effects).

Structural Transformation Hypothesis: ICT development in wealthy economies represents a shift toward service-based, knowledge-intensive production that is inherently less carbon-intensive.

5.3 Policy Implications

For Developed Economies:

> The aggregate trend suggests **domestic digital capacity (DCI)** can be a decarbonization lever. However, **high-EDS structural exceptions (e.g., Switzerland)** indicate that efficiency gains may be offset by rebound effects in specific service-export-intensive contexts. Policy should therefore complement digital investment with measures targeting **absolute decoupling**.

For Developing Economies:

> **Policy Consideration:** Evidence suggests that digital transformation alone may not drive decarbonization in low-capacity settings. Complementary efforts in capacity building are essential.

For International Organizations: Target digital development assistance as part of broader **capacity-building** packages.

5.4 Limitations

- **Measurement:** While **DCI** (PCA of internet use, broadband access, and secure servers) captures infrastructure-based domestic digital capacity, it may not fully capture the **quality of digital utilization** (e.g., AI adoption intensity, data-center efficiency, sectoral digital deepening). **EDS** captures external specialization and should not be interpreted as a proxy for domestic adoption.

- **Causal interpretation**: Despite the DML framework, unobserved confounders may remain.
- **External validity**: Results may not generalize to small economies not in our sample. Accordingly, inferential statements are restricted to GATE-level contrasts with country-cluster bootstrap intervals.

6 Conclusion

This paper introduces the concept of the "Digital Decarbonization Divide" and provides rigorous empirical evidence for its existence. Using Causal Forest DML on a balanced panel of 40 economies ($N = 960$), we find that:

1. **Domestic digital capacity (DCI)** exhibits fundamentally non-linear effects on CO_2 emissions.
2. **GATEs** reveal a clear progression from near-zero effects in low-capacity economies to strong reductions in high-capacity economies.
3. **Structural exceptions exist:** "off-diagonal" cases (e.g., **Switzerland**) indicate that **high external digital specialization (EDS)** can override domestic efficiency gains and generate rebound effects.
4. The **Model Ladder** demonstrates that flexible estimation is required to capture policy-relevant thresholds and exceptions missed by linear models.

Our findings challenge the assumption that digital transformation is universally beneficial for climate goals. Instead, we identify **conditional prerequisites**—socio-economic capacity—that appear to moderate whether ICT delivers a "green dividend."

Declarations

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest: The author declares no conflicts of interest.

Data and Code Availability

Replication code and a cleaned dataset construction script will be made available in a public repository upon acceptance. All raw inputs are from WDI/WGI.

References

Athey, S. and Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational Studies*, 5(2), 37–51.

Berkhout, F. and Hertin, J. (2000). De-materialising the economy or rematerialisation? The case of information and communication technologies. *The Environmental Impact of Prosperous Societies*, 4, 12–26.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1–C68.

Danish, Zhang, B., Wang, B., and Wang, Z. (2017). Role of renewable energy and non-renewable energy consumption on EKC: Evidence from Pakistan. *Journal of Cleaner Production*, 156, 855–864.

Gossart, C. (2015). Rebound effects and ICT: A review of the literature. In *ICT Innovations for Sustainability* (pp. 435–448). Springer.

Lange, S., Pohl, J., and Santarius, T. (2020). Digitalization and energy consumption. Does ICT reduce energy demand? *Ecological Economics*, 176, 106760.

Salahuddin, M. and Alam, K. (2016). ICT, electricity consumption and economic growth in OECD countries. *International Journal of Electrical Power & Energy Systems*, 76, 185–193.

World Bank. (2026). *World Development Indicators*. Washington, D.C.: The World Bank.