

Economic Development and Clean Energy: Impacts on Carbon Intensity and Renewable Electricity Adoption

Aesha Gandhi Yiyun (Leo) Yao Junhao Xu Sebine Scaria

2025-12-03

1 Data Science Memo

Stakeholder (Target Audience). This memo is written for a **senior energy policy analyst at a multilateral development bank** (for example, the World Bank). The analyst advises on financing decisions for electricity-sector projects across low-, middle-, and high-income countries. They are familiar with basic statistics and regression output but are not a specialist in econometrics or energy systems modeling.

1.1 Executive Summary / TL;DR

Rapid decarbonization of electricity is essential to meeting global climate goals, yet fossil fuels still provide most of the world's power. Using a global country-year dataset on energy and economic indicators from 1965–2023, we address two questions: **(1)** How does a country's economic development relate to the carbon intensity of its electricity generation, and does this relationship differ by income level? **(2)** What factors are associated with a country achieving a **>50% renewable** electricity share?

For **Research Question 1**, we find that **economic growth has opposite associations with power-sector emissions in poor vs. rich countries**. In **low-income countries**, higher GDP per capita is associated with **more carbon-intensive electricity**: a one-unit increase in log GDP per capita corresponds to approximately **+105 grams of CO₂ per kWh**. In **high-income countries**, the same increase in log GDP per capita corresponds to about **-81 grams of CO₂ per kWh**. In other words, as the richest countries grow, their electricity tends to become cleaner, while in the poorest countries, growth comes with dirtier electricity.

For **Research Question 2**, we define a binary outcome indicating whether a country in a given year generates **more than 50% of its electricity from renewable sources**. Only

about 7% of country-year observations in our data meet this threshold. A logistic regression shows that the **share of fossil fuels in the power mix** is by far the dominant factor: each 1 percentage-point increase in fossil fuel share multiplies the odds of a renewables-majority by about **0.87** (an 80% reduction in odds for a 10-point increase). **Although GDP per capita and total electricity generation are statistically significant predictors, their effect sizes are modest; in contrast, fossil fuel share is both statistically and practically significant, showing the largest influence on whether a country exceeds 50% renewable electricity.**

Implications. For low-income countries, economic growth tends to **increase** the carbon intensity of electricity, unless there is proactive investment in clean power. For high-income countries, growth can coincide with **decreasing** carbon intensity if it is accompanied by policies that push the grid toward renewables. Combining the two insights together, we see that achieving a majority-renewable electricity system is more primarily a matter of **reducing fossil fuel dependence**, rather than simply being low- or high-income. Targeted financing for renewables, policies that discourage use of coal and gas, and support for grid flexibility are more effective than relying on income growth alone.

1.2 Key Decisions for the Stakeholder

1. **How aggressively to finance clean power in low-income countries.** Since growth in low-income economies is currently associated with more fossil-based generation, the bank must decide whether to scale up funding for renewable generation and grid expansion to help these countries avoid a fossil-heavy development path.
2. **How to prioritize policy and technical assistance.** The strongest predictor of a majority-renewable grid is low fossil fuel dependence. The stakeholder must decide whether to prioritize support for policies that directly reduce fossil generation (e.g., coal retirement programs, carbon pricing, etc.) over more general economic development interventions.
3. **Where to focus grid and storage investments.** Moving from a minority to majority-renewable grid requires high levels of grid flexibility and reliability. The stakeholder must decide which countries need support for storage, regional grid integration, and modern grid management tools to enable higher renewable shares without compromising reliability.
4. **How to allocate analytic resources.** Finally, the stakeholder must decide whether to conduct deeper country-specific researches (e.g., resource assessments, policy diagnostics, etc.) for countries near the 50% threshold, to better understand the obstacles and opportunities for decarbonizing their power systems.

1.3 Background and Problem Motivation

Electricity generation is one of the largest contributors to global greenhouse gas emissions. Many countries have committed to decarbonizing their power sectors, but progress varies widely. Some nations already generate most of their electricity from renewables, often leveraging hydro or geothermal resources, while others still rely heavily on coal, oil, and gas.

At the same time, countries are at very different stages of economic development. High-income countries increasingly propose “decoupling” economic growth from emissions, while low-income countries often argue that they should not be required to forego fossil-fueled growth without financial and technological support. Understanding how **GDP per capita**, **income level**, and **energy mix** relate to both the **carbon intensity of electricity** and the **likelihood of a renewables-majority grid** is therefore essential to designing effective development and climate strategies.

Our analysis is intended to inform these decisions using historically observed relationships across nearly six decades of global data.

1.4 Data and Approach (High Level)

We use a dataset compiled by *Our World in Data* that combines multiple international sources and reports annual statistics for most countries worldwide from 1965 onward. Each observation represents a **country–year** with indicators on:

- **Electricity generation (TWh)**
- **Carbon intensity of electricity** (grams of CO₂ per kWh)
- **Share of electricity from fossil fuels** (coal, oil, gas)
- **Share of electricity from renewable sources** (including hydro, wind, solar, and others)
- **Gross domestic product (GDP)** and **population**, from which we compute **GDP per capita**

We augment this with **World Bank income group classifications** (Low, Lower middle, Upper middle, High income) to capture development stage. After dropping observations with missing key fields, we retain roughly nine thousand country–year observations, spanning over 180 countries between 1965 and 2023.

Analytically:

- For **RQ1**, we fit a **linear regression** of carbon intensity on **log GDP per capita**, income group, and their interaction, allowing the effect of income to differ across development stages.

- For **RQ2**, we define a binary outcome indicating whether **renewables provide more than half of electricity**, and fit a **logistic regression** where the predictors are **log GDP per capita**, **log electricity generation**, and **fossil fuel share of electricity**.

Technical details, such as data cleaning steps and diagnostic plots, are presented in the technical appendix.

1.5 Results: Research Question 1

1.5.1 Does GDP per capita affect carbon intensity of electricity, and does this differ by income group?

1.5.2 Descriptive Patterns

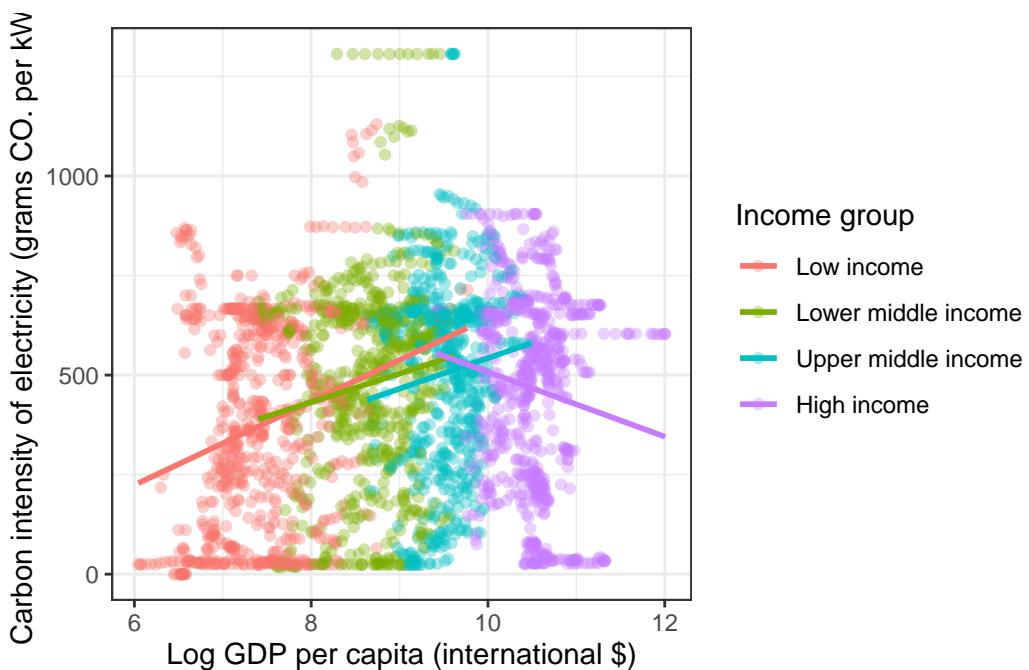


Figure 1: Carbon intensity vs. GDP per capita by income group.

The plot shows that, at low income levels, carbon intensity tends to increase with GDP per capita, while at high income levels it decreases with GDP per capita.

1.5.3 Main Model Results

Table 1: Linear regression of carbon intensity of electricity on log GDP per capita, income group, and their interaction.

Characteristic	Estimate	Std. Error	95% CI	p-value
Log GDP per capita	104.75	17.02	71.38, 138.12	<0.001
Lower-middle income (vs. low)	262.54	217.00	-162.96, 688.04	0.226
Upper-middle income (vs. low)	178.96	310.01	-428.92, 786.84	0.564
High income (vs. low)	1722.88	273.46	1186.66, 2259.09	<0.001
Log GDP × lower-middle income	-33.00	26.49	-84.94, 18.95	0.213
Log GDP × upper-middle income	-27.91	34.29	-95.14, 39.32	0.416
Log GDP × high income	-185.81	28.72	-242.12, -129.51	<0.001

Interpretation:

- For **low-income countries** (reference group), a one-unit increase in log GDP per capita is associated with an increase of about **105 grams CO per kWh** in carbon intensity.
- For **high-income countries**, the interaction term reverses the slope: the net effect of log GDP per capita is about **81 grams CO per kWh lower** per log unit of GDP compared to low-income countries. This implies that, among high-income countries, higher GDP per capita is associated with **lower** carbon intensity.
- For **lower-middle and upper-middle income countries**, the interaction terms are negative but not statistically significant at the 5% level, suggesting that their GDP–carbon intensity relationship is intermediate between low- and high-income countries.

1.6 Results: Research Question 2

1.6.1 What factors are associated with having >50% renewable electricity?

1.6.2 Descriptive Patterns

Country-years with a renewables majority tend to have very low fossil shares, while those without such a majority rely heavily on fossil fuels.

1.6.3 Main Model Results

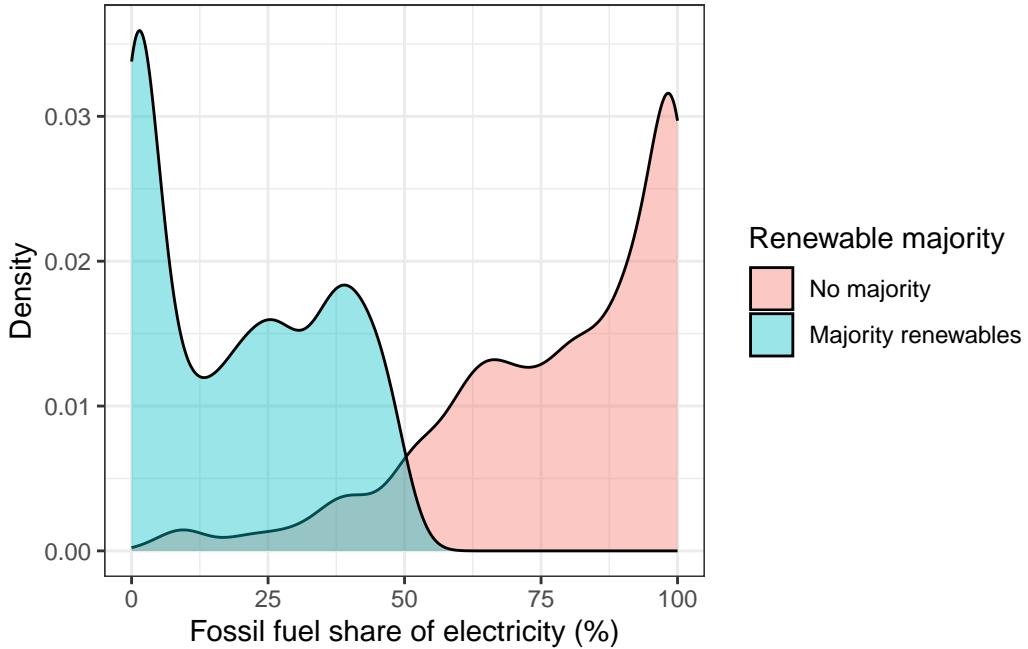


Figure 2: Fossil fuel share for years with and without $>50\%$ renewable electricity.

Table 2: Logistic regression for probability of having $>50\%$ renewable electricity (odds ratios).

Characteristic	OR	95% CI	p-value
GDP per capita (log)	0.521	0.442, 0.612	<0.001
Electricity generation (log, TWh)	0.700	0.637, 0.768	<0.001
Fossil fuel share of electricity (%)	0.874	0.865, 0.882	<0.001

Key findings:

- **Fossil fuel share** has an odds ratio well below 1, meaning each additional percentage point of fossil generation sharply reduces the odds of having a renewables majority.
- **GDP per capita** and **total electricity generation** are statistically significant predictors, but their odds ratios are closer to 1 and their effect sizes are modest compared with fossil fuel share. In other words, the composition of the power mix matters much more than income level or grid size.

1.6.4 ROC Analysis (Model Performance)

The ROC curve and AUC summarize how well the logistic model distinguishes between country-years with and without a renewables majority.

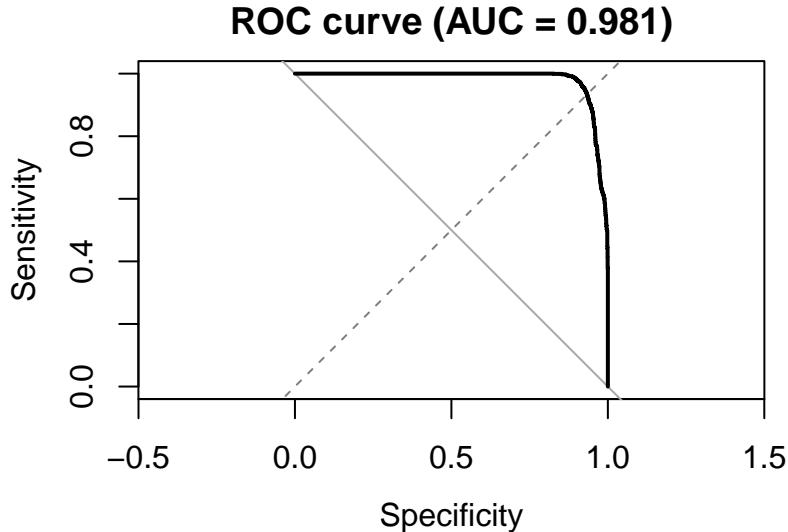


Figure 3: ROC curve for renewable majority model (AUC 0.98).

The ROC curve lies well above the diagonal reference line, and the AUC (area under the curve) is about **0.98**, indicating excellent ability to distinguish between years with and without a renewables majority. Such a strong performance is driven mainly by the fossil fuel share predictor.

1.7 Conclusion and Recommendations

Across nearly six decades of global data, we find:

- Economic development and carbon intensity of electricity are linked in **opposite ways** for low- and high-income countries.
- Achieving a **majority-renewable electricity mix** is primarily a function of **limiting fossil fuel share**, not simply of income level or grid size.

Recommendations:

1. **Prioritize clean power investments in low-income countries.** Financing solar, wind, hydro, and supporting grid infrastructure can help these countries grow without sharply increasing power-sector emissions.

2. **Support policies that reduce fossil fuel dependence.** Instruments such as carbon pricing, renewable portfolio standards, and the removal of fossil fuel subsidies can shift the economics toward renewables and lower fossil shares.
3. **Invest in enabling technologies.** To sustain high renewable shares, countries need storage, flexible grids, and sometimes regional power integration. Supporting these technologies will make it easier for countries to move beyond the 50% renewables threshold.
4. **Encourage knowledge and technology transfer.** High-income countries that have begun to decarbonize can share policy experience and technology with lower-income countries, helping them avoid high-carbon development paths.

2 Technical Appendix

2.1 Data

The main dataset used in this project is a global energy dataset compiled by *Our World in Data*. It reports annual values for:

- Total electricity generation (TWh)
- Carbon intensity of electricity (grams of CO₂ per kWh)
- Shares of electricity from fossil fuels and renewables
- GDP and population

We restrict attention to **country-level** records between **1965 and 2023**, which is the period during which electricity statistics are relatively complete. Regional aggregates and historical entities are excluded.

Transformations and filters:

- GDP per capita was computed as GDP divided by population.
- For **RQ1**, we kept country-years with non-missing carbon intensity, GDP per capita, and income group.
- For **RQ2**, we kept country-years with non-missing renewable share, fossil share, GDP per capita, and electricity generation.
- Log transformations of GDP per capita and electricity generation were used to handle strong right skewness.
- We merged World Bank income group classifications (low, lower middle, upper middle, high income) by country and year.

We did not impute missing entries in key variables; instead, we excluded those country-years from modeling.

2.2 Models

2.2.1 RQ1: Carbon intensity and GDP per capita

We modeled **carbon intensity of electricity** as a continuous outcome in a linear regression:

$$\text{CarbonIntensity}_{it} = \beta_0 + \beta_1 \log(\text{GDPpc}_{it}) + \beta_2 \text{IncomeGroup}_{it} + \beta_3 [\log(\text{GDPpc}_{it}) \times \text{IncomeGroup}_{it}] + \varepsilon_{it},$$

where i indexes **countries** and t indexes **years**. Low-income countries serve as the reference group.

The log transformation of GDP per capita addresses extreme skew and makes proportional changes in income more interpretable. The interaction term allows the slope of GDP per capita to vary across income groups, which is essential for testing whether development operates differently in poorer vs. richer countries.

2.2.1.1 Model selection and transformation

As an initial step, we fit a model using **untransformed GDP per capita**:

$$\text{CarbonIntensity}_{it} = \beta_0 + \beta_1 \log(\text{GDPpc}_{it}) + \beta_2 \text{IncomeGroup}_{it} + \beta_3 [\log(\text{GDPpc}_{it}) \times \text{IncomeGroup}_{it}] + \varepsilon_{it},$$

Diagnostic plots for this raw model showed strong curvature in the residuals versus fitted values and strong heteroskedasticity, consistent with our exploratory plots in which GDP per capita was extremely right-skewed. After replacing GDP per capita with **log(GDP per capita)**, the residual plots became more linear and the spread of residuals became more stable across fitted values. We therefore retained this log-transformed model.

2.2.1.2 Full coefficient table

Table 3: Full linear regression output for carbon intensity model.

term	Estimate	Std. Error	Conf. Low	Conf. High	p-value
(Intercept)	-404.65	126.40	-652.49	-156.80	0.001
log_gdp	104.75	17.02	71.38	138.12	<0.001
income_groupLower middle income	262.54	217.00	-162.96	688.04	0.226
income_groupUpper middle income	178.96	310.01	-428.92	786.84	0.564
income_groupHigh income	1722.88	273.46	1186.66	2259.09	<0.001

term	Estimate	Std. Error	Conf. Low	Conf. High	p-value
log_gdp:income_groupLower middle income	-33.00	26.49	-84.94	18.95	0.213
log_gdp:income_groupUpper middle income	-27.91	34.29	-95.14	39.32	0.416
log_gdp:income_groupHigh income	-185.81	28.72	-242.12	-129.51	<0.001

2.2.1.3 Diagnostics

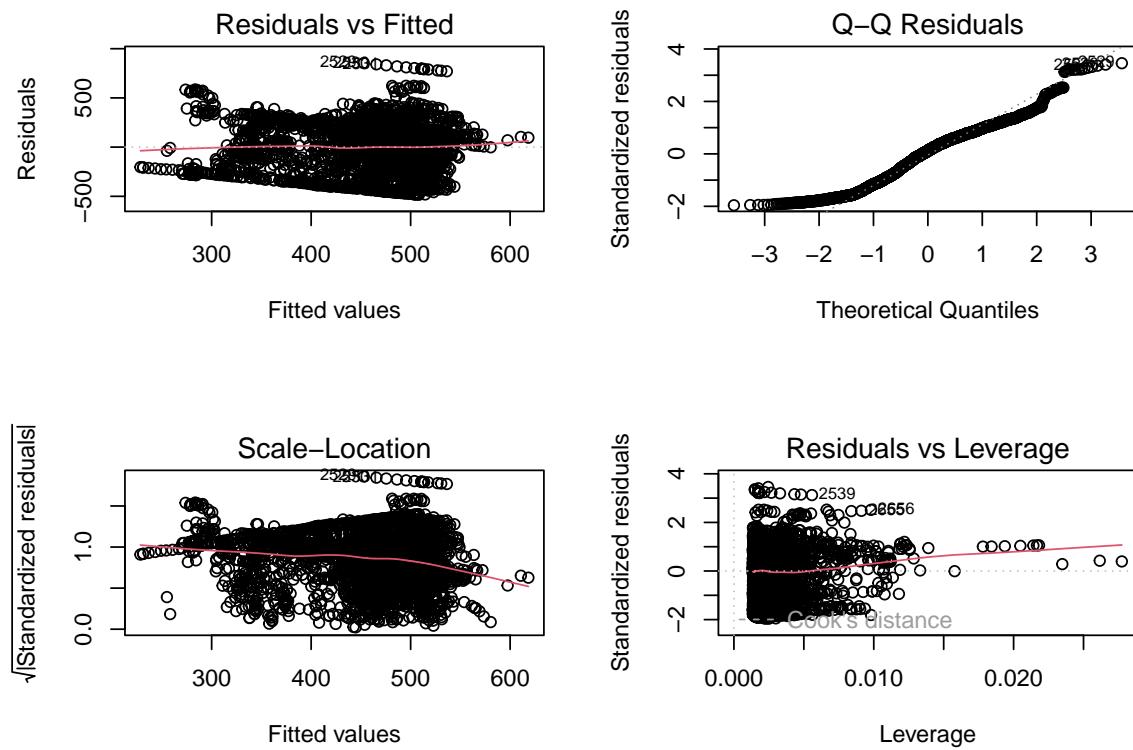


Figure 4: Regression diagnostics for carbon intensity model.

- Residual-fitted plots do not show strong curvature after log-transforming GDP per capita.
- Q-Q plots show mild tail deviations but are broadly consistent with normal residuals.
- Scale-location plots show some heteroskedasticity (fan shape), common in cross-country data.

- Cook's distances are well below 1 for all observations, so no single influential point.
- Variance Inflation Factors for predictors were all below 2, indicating no severe multi-collinearity.

Potential confounders. We do not control for some variables that may confound the association between GDP per capita and carbon intensity, including energy policy stringency, resource endowment (e.g., hydro or fossil reserves), and industrial structure. As a result, the estimated GDP effects may partly capture differences in these omitted structural and policy factors.

2.2.2 RQ2: Predicting >50% renewable electricity

We modeled the probability that a country–year has **more than 50% of its electricity from renewables** using a logistic regression:

$$\Pr(Y_{it} = 1) = \text{logit}^{-1} (\alpha_0 + \alpha_1 \log(\text{GDPpc}_{it}) + \alpha_2 \log(\text{Generation}_{it}) + \alpha_3 \text{FossilShare}_{it}),$$

where $Y_{it} = 1$ if the renewable share exceeds 50% and 0 otherwise.

Table 4: Full logistic regression output (odds ratios) for renewable majority model.

term	OR	Std. Error	z-stat	p-value	Conf.		
					Low	Conf.	High
(Intercept)	207503.317	0.805	15.213	2.90e-52	44327.357	1041265.587	
log_gdp	0.521	0.083	-7.880	3.28e-15	0.442		0.612
log_generation	0.700	0.048	-7.495	6.62e-14	0.637		0.768
fossil_share_elec	0.874	0.005	-26.283	2.97e-152	0.865		0.882

2.2.2.1 Diagnostics

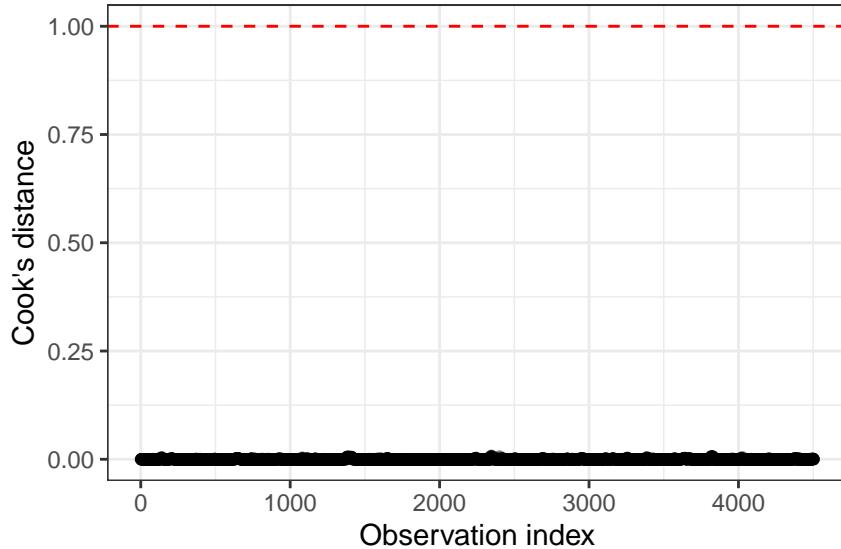


Figure 5: Influential points measured by Cook’s distance for the logistic regression model.

All Cook’s distances are essentially near zero and well below 1, which means no observation has unusually large influence on the logistic model—this matches what we observed in our plot.

We also examined the ROC curve and AUC (see the ROC figure in the memo section), which confirm that the model has very high discriminative ability, driven mainly by the fossil fuel share predictor.

Potential confounders. Key omitted variables that may confound the relationships in this model include renewable resource potential (e.g., solar irradiation, hydro capacity), policy support for renewables (feed-in tariffs, renewable portfolio standards), and grid flexibility constraints. Countries with strong renewable resources and supportive policies are both more likely to have low fossil shares and to reach majority-renewable status, so part of the fossil-share effect may reflect these underlying structural advantages.

2.3 Next Steps

Limitations and potential extensions include:

- **Causality.** Our models are associative, not causal. Future work could use fixed-effects panel models or instruments to isolate the causal effects of income and policy variables.
- **Omitted variables / confounding.** Resource endowments (solar, wind, hydro potential), policy strength, and technology costs were not included but may confound some of the associations we observe.

- **Threshold choice.** We focused on a 50% renewable threshold; examining other thresholds (e.g., 30%, 80%) could provide a more nuanced picture of transition stages.
- **Temporal dynamics.** Countries’ transitions are path-dependent. Dynamic models that incorporate lagged variables, investment cycles, or structural breaks could yield deeper insights.

3 References

- Our World in Data. *Energy Data*. Retrieved from the “Global primary energy” and “Electricity mix” datasets.
- World Bank. *World Bank Country and Lending Groups*. Classification by income level.
- International Energy Agency (IEA). *World Energy Outlook* (various years).
- Intergovernmental Panel on Climate Change (IPCC). *Sixth Assessment Report: Mitigation of Climate Change*.
- Ember Climate. *Global Electricity Review* (various years).