A Functional Carbon Emissions Analysis

Sam Lee, Everett Andrew

Brigham Young University

April 2025

Outline

Motivation: The Climate Imperative

"It's time to stop burning our planet, and start investing in the abundant renewable energy all around us."

— **António Guterres, United Nations Secretary-General**

Motivation: Emissions in 2023

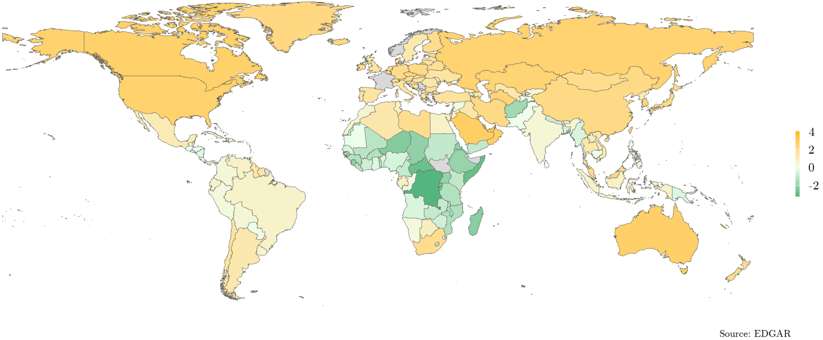
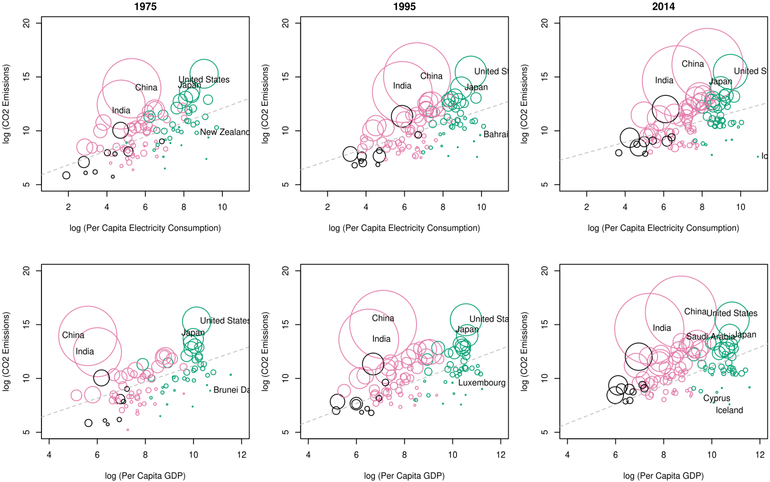


Figure: Log per capita CO_2 emissions in 2023

- Emissions per capita remain highest in the Global North
- Heterogeneous distribution
- What policies are effective?
- When are the effects realized?

Literature Review: Using FDA to Model Carbon Emissions



- FLM to model log CO_2 emissions using electricity consumption and GDP (Elayouty and Abou-Ali, 2023)

Research Question

Main Question

How does a country's share of renewable energy consumption affect its carbon emissions over time, holding other macroeconomic and political covariates constant?

Data: Sources and Scope

- **Emissions Database for Global Atmospheric Research (EDGAR)**: Annual country-level CO_2 emissions per capita (1970-2023)

Data: Sources and Scope

- **Emissions Database for Global Atmospheric Research (EDGAR)**: Annual country-level CO_2 emissions per capita (1970-2023)
- **World Bank**: Functional Macroeconomic Controls (1990-2020)
  - Percent renewable energy consumption (out of total energy consumption)
  - GDP per capita, Human Development Index, Inflation

Data: Sources and Scope

- **Emissions Database for Global Atmospheric Research (EDGAR)**: Annual country-level CO_2 emissions per capita (1970-2023)
- **World Bank**: Functional Macroeconomic Controls (1990-2020)
  - Percent renewable energy consumption (out of total energy consumption)
  - GDP per capita, Human Development Index, Inflation
- **World Governance Indicators**:
  - Measures of corruption, government effectiveness, rule of law, etc. (1996-2023)

Data: Sources and Scope

- **Emissions Database for Global Atmospheric Research (EDGAR)**: Annual country-level CO_2 emissions per capita (1970-2023)
- **World Bank**: Functional Macroeconomic Controls (1990-2020)
  - Percent renewable energy consumption (out of total energy consumption)
  - GDP per capita, Human Development Index, Inflation
- **World Governance Indicators**:
  - Measures of corruption, government effectiveness, rule of law, etc. (1996-2023)
- Final dataset includes **178 countries** over **34 years** (1990–2023).

Data: Carbon Emissions Across Countries

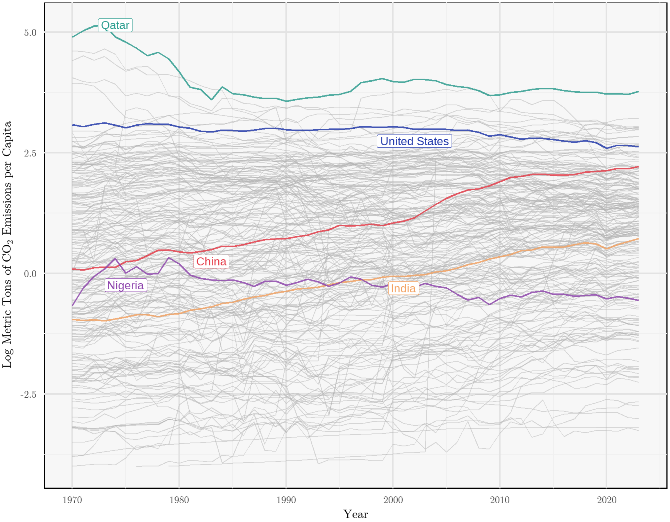


Figure: Log-transformed CO_2 emissions per capita across 178 countries

Data: Renewable Energy Consumption Over Time

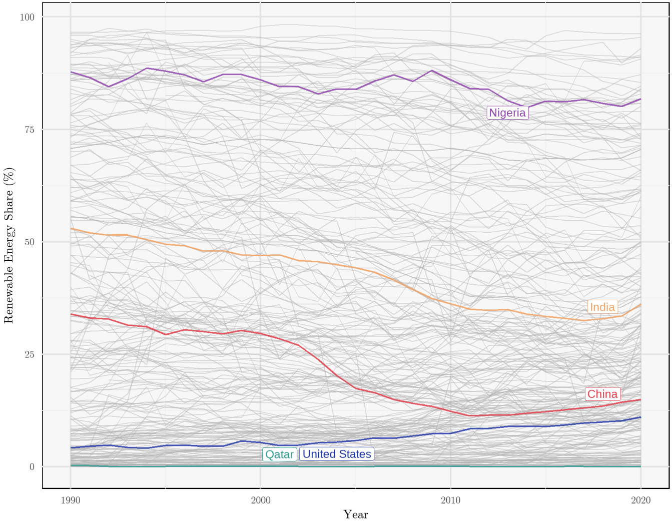


Figure: Percent of total energy consumption from renewable sources

Data: Summary Statistics

Variable	Min	Median	Mean	Max	%Missing
Key Variables					
log CO_2 Emissions (Metric Tons per Capita)	−6.0992	0.7331	0.4728	5.1559	0.00
GDP per Capita	0.4942	13.1340	21.4102	179.2007	0.00
Renewable Share (%)	0.00	21.62	31.44	98.30	0.36
Human Development Index (HDI)	0.2120	0.6960	0.6703	0.9670	7.27
Inflation (%)	−31.57	4.59	35.46	26 762.02	1.32
Interest Rate (%)	−97.69	5.88	5.97	139.96	39.07
Unemployment Rate (%)	0.04	6.66	7.98	38.80	41.70
Governance Indicators					
Corruption Index	−1.85	−0.26	−0.03	2.46	0.76
Government Effectiveness	−2.44	−0.15	−0.01	2.47	1.06
Political Stability	−3.31	0.04	−0.04	1.96	0.36
Rule of Law	−2.59	−0.19	−0.04	2.12	0.00
Regulation Quality	−2.55	−0.14	−0.01	2.31	0.99
Voice and Accountability	−2.31	−0.02	−0.05	1.80	0.00

- Variables with 0–10% missingness were imputed using functional mean imputation
- Covariates with > 10% missingness excluded

Methodology: Why not Standard Functional Linear Regression?

In a traditional function-on-function model framework...

Methodology: Why not Standard Functional Linear Regression?

In a traditional function-on-function model framework...

- Covariate values from **any time** can influence outcomes at t
  - Violates basic temporal causality
  - Future policy shouldn't predict past emissions

Methodology: Why not Standard Functional Linear Regression?

In a traditional function-on-function model framework...

- Covariate values from **any time** can influence outcomes at t
  - Violates basic temporal causality
  - Future policy shouldn't predict past emissions
- Overfit by "fitting noise" from future information

Methodology: Why not Standard Functional Linear Regression?

In a traditional function-on-function model framework...

- Covariate values from **any time** can influence outcomes at t
  - Violates basic temporal causality
  - Future policy shouldn't predict past emissions
- Overfit by "fitting noise" from future information
- Solution?

Methodology: Historical Functional Linear Model (HFLM)

Canonical HFLM

$$y_i(t) = \beta_0(t) + \int_{s_0(t)}^{t} x_i(s)\,\beta(s,t)\,ds + \varepsilon_i(t)$$

- Enforce temporal causality: predictor values must precede the outcome ($s \leq t$) (Malfait & Ramsay, 2003)
- $s_0(t) = \max(0, t - \delta)$ ensures a lag window
  - Fix δ or estimate through cross-validation

Methodology: Triangular Support

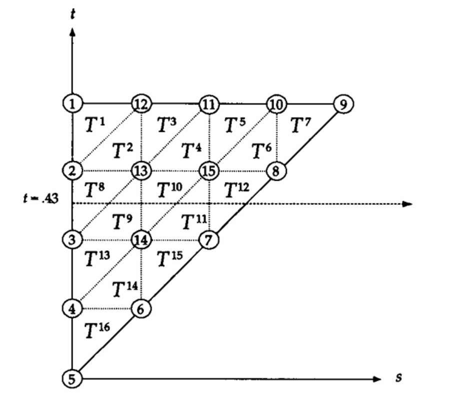


Figure: Triangular Support (Malfait & Ramsay, 2003)

- Covariate values influence the response only within the triangle $s \in [t - \delta, t], s \leq t$

Methodology: Application

HFLM with Penalized Tensor B-Splines

$$\log y_i(t) = \beta_0(t) + \sum_{k=1}^{10} \sum_{m=1}^{M} b_{km} \int_{s_0(t)}^{T_k} x_{ki}(s)\phi_m(s,t)\, ds + \varepsilon_i(t)$$

- Nomenclature:
  - Log per capita CO_2 Emissions in country i at time t: $\log y_i(t)$
  - 10 covariates (*GDP, HDI, WGI, etc.*) with different supports, T_k
  - Propose 20 year lag: $\delta = 20$
  - $\beta_k(s,t)$ smoothed using a penalized B-spline basis

- Riemann sum approximation

Methodology: Implementation in R

- `pffr()` from the `refund` package
- `limits` function enforces triangular support: $s \in [t - \delta, t]$

```r
model <- pffr(
  carbon ~ ff(energy, xind = s, yind = co2.s,
              limits = function(s, t){
                s >= pmax(1990, t - delta)
              }) + ...
)
```

Results: Functional $R^2(t)$

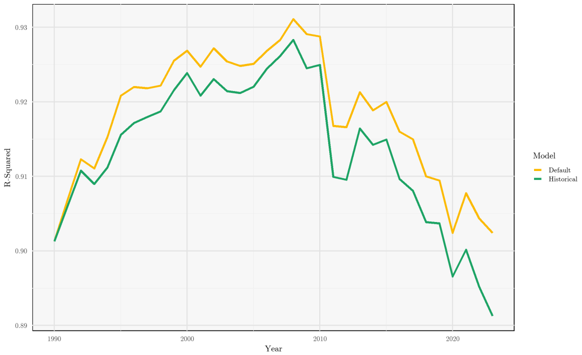


Figure: $R^2(t)$ comparing historical and traditional function-on-function models

- Both models fit extremely well
- Traditional model overfits by using post-t covariate values (non-causal)

Results: Effect of Renewable Energy Consumption

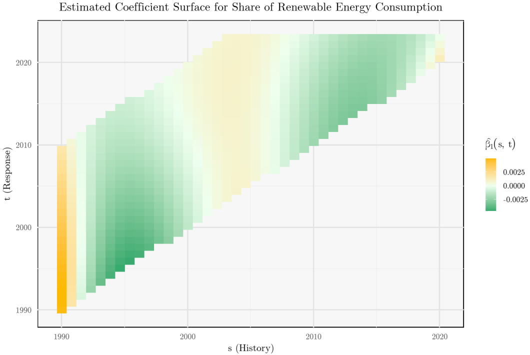
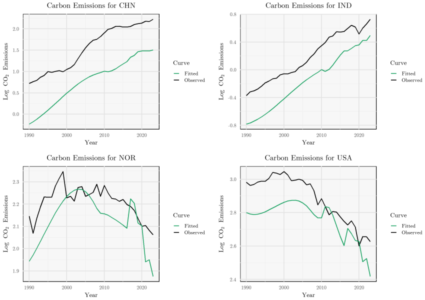


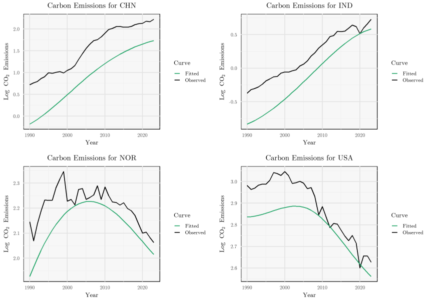
Figure: Estimated $\hat{\beta}_1(s,t)$ surface for the share of renewable energy consumption

- Strong negative region along diagonal: recent renewables $\implies$ lower emissions
- Effect greatest around 1995-2010 and 2005-2020
- Structural lags in renewable energy capital?

Results: Country-Level Comparison



HFLM



FLM

- Trends predicted well
- Under-fit intercept?
- HFLM more noisy?
  - Reimann sum integration
  - Less data $\implies$ higher variance

Conclusions & Policy Implications

- Renewable energy share has a persistent, long-run negative effect on CO_2 emissions
- Effects linger 5-10 years
  - Policy implications: Investments today shape emissions a decade from now
- Traditional function-on-function models inflate $R^2(t)$

Conclusion: Limitations & Future Work

Limitations

- Spatial correlation: Country-level errors not independent
- Sparse covariates: Some covariates inconsistently reported
- Renewable energy consumption is a ratio
  - A country can increase renewable share without reducing emissions if total energy demand rises
- Bias: Penalized intercept may distort level fits

Conclusion: Limitations & Future Work

Limitations

- Spatial correlation: Country-level errors not independent
- Sparse covariates: Some covariates inconsistently reported
- Renewable energy consumption is a ratio
  - A country can increase renewable share without reducing emissions if total energy demand rises
- Bias: Penalized intercept may distort level fits

Future Work

- Spatio-functional models
- Uncertainty quantification: Bootstrap over $\beta_k(s, t)$

Thank You!

Questions?

`https://github.com/SamLeeBYU/CO2-FDA`