Quantifying the Effects of Climate Policy Stringency on Verified Emissions and Satellite-Derived NOx

Master's Thesis

Arnav Agrawal

Columbia University
MA in Quantitative Methods in the Social Sciences

December 2025

Research Question & Motivation

Core Question: How does EU ETS policy stringency affect emissions at the facility level?

Dual-Outcome Approach:

- **Verified CO_2**: Administrative data from EU ETS registry (gold standard)
- **Satellite NOx**: Independent physical measurement from TROPOMI

Why Both?

- CO_2: Accurate compliance trajectories, global climate impact
- NOx: Local air quality co-benefits, health impacts (respiratory, cardiovascular)
- Cross-validation: Agreement between independent systems strengthens causal claims

Sample: 521 large combustion plants across 82 NUTS2 regions (2018–2023)

EU ETS Background

Cap-and-Trade System (established 2005)

- World's largest carbon market, covers ∼40% of EU GHG emissions
- Large combustion plants (≥20 MW) must hold allowances = verified emissions

Treatment Variable: Allocation Ratio

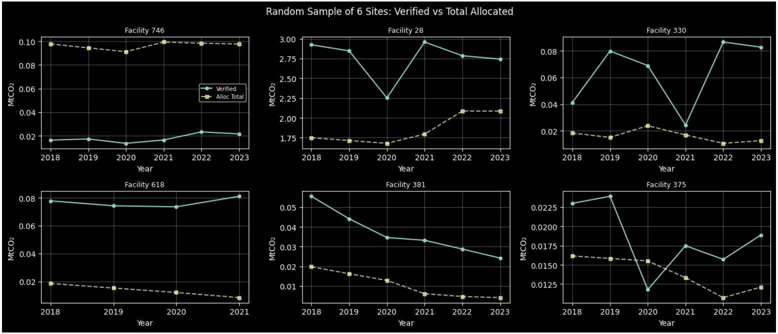
$$R_{it} = \frac{\text{Free Allocation}_{it}}{\text{Verified Emissions}_{it}}$$

- $R < 1$: Shortfall → must purchase allowances (policy stringency)
- $R > 1$: Surplus → can sell excess allowances

Key Policy Evolution (Phase III/IV):

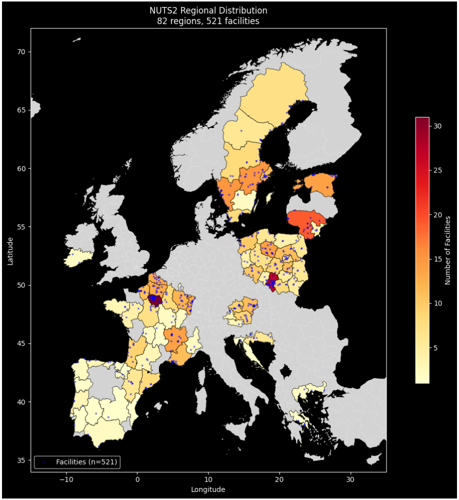
- Power sector lost most free allocation → full marginal carbon cost
- Electricity: mean $R = 0.52$ vs. Other sectors: mean $R = 1.03$

Policy Stringency: Verified vs. Allocated Emissions



- **Teal**: Verified emissions; **Yellow dashed**: Free allocation
- Gap above allocation line = shortfall requiring market purchases
- Within-facility variation over time provides identification

Geographic Distribution of Sample



- 521 facilities across 82 NUTS2 regions
- Concentration in Germany, Poland, Spain, France
- NUTS2 regions used for clustering and Region×Year FE

Data Sources & Sample Construction

Three Linked Data Sources:

1. **EEA Large Combustion Plant Registry**: Coordinates, capacity, fuel mix
2. **EU ETS (EUTL)**: Verified emissions, free allocations
3. **TROPOMI Satellite**: Daily NO_2 tropospheric columns (3.5×5.5 km)

Sample Attrition:

- 3,405 plants $\rightarrow$ 932 facilities (500m spatial clustering)
- $\rightarrow$ 521 facilities with ETS linkage & $\geq$3 years data
- $\rightarrow$ 291 facilities for satellite NOx ($\geq$20 valid days/year)

Key Statistics: Mean emissions 580 $ktCO_2$/yr, 80.8% electricity sector

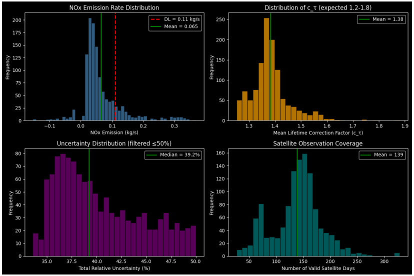
Satellite NOx: Beirle Flux-Divergence Method

Physical Basis: Continuity equation

$$\underbrace{\mathbf{w} \cdot \nabla V}_{\text{Advection}} \approx E - S$$

Implementation:

1. Spatial gradient of NO_2 (TROPOMI)
2. $\times$ wind velocity (ERA5-Land)
3. Integrate over 15 km disc
4. Lifetime & topographic corrections



Key Thresholds:

- Detection: 0.11 kg/s (conservative)
- Uncertainty: ∼35–45%

AlphaEarth Embeddings: High-Dimensional Controls

Google AlphaEarth Foundations (2025): 64-dim geospatial representations learned from:

- Multi-source satellite imagery (Sentinel-1/2, Landsat)
- Climate reanalysis (ERA5-Land), topography (GLO-30)
- Geotagged text (Wikipedia, GBIF)

What Embeddings Encode:

- Land use context (urban density, industrial areas)
- Infrastructure (roads, built environment)
- Vegetation, climate patterns, terrain

Why Use Them?

- Control satellite retrieval confounders (terrain $\rightarrow$ AMF; urban $\rightarrow$ background NO_2)
- High-dimensional spatial confounders that would be impractical to specify manually
- Extends Veitch et al. (2019) text embeddings to geospatial domain

Embedding Dimensionality Reduction: PCA vs. PLS

Problem: 64 dimensions may cause overfitting with limited within-facility variation

PCA (Unsupervised):

- Projects onto directions maximizing variance in embedding space
- **Causally safe**—does not use outcome information

PLS (Supervised):

- Projects onto directions that predict NOx outcome
- **Risk:** Naive PLS on panel → regularization bias (outcome snooping)
- **Solution:** Train on *facility-level means* (cross-sectional), not panel obs. Then reproject every fac-year's embedding using the trained projector.

Key Design: PLS trained on $\bar{Y}_i^{\text{NOx}} = T^{-1} \sum_t Y_{it}$ (one obs/facility)

- Resulting embeddings are *time-invariant* within facility
- $\equiv$ pre-treatment covariates → no "bad controls" problem
- Both methods: 64-dim → 10 components

Empirical Strategy

Two-Way Fixed Effects (TWFE):

$$Y_{it} = \alpha_i + \gamma_{r(i),t} + \beta R_{it} + \mathbf{X}'_{it}\boldsymbol{\delta} + \varepsilon_{it}$$

- α_i: Facility FE (absorb time-invariant confounders)
- $\gamma_{r(i),t}$: NUTS2 Region $\times$ Year FE (absorb regional trends)
- $\mathbf{X}_{it}$: Capacity, fuel shares, AlphaEarth embeddings (NOx only)

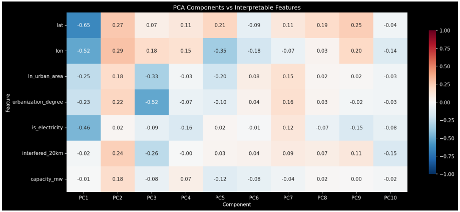
Identification: Within-facility variation in allocation ratios over time

Key Design Choices:

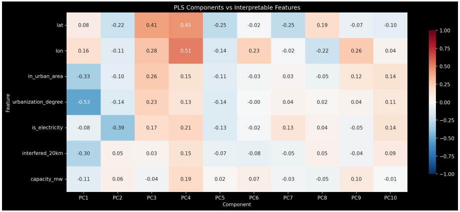
- Cluster SEs by NUTS2 region (82 clusters)
- AlphaEarth embeddings (64-dim $\rightarrow$ 10-dim via PCA/PLS) control satellite confounders
- Do *not* control for dispatch—mediator, not confounder

AlphaEarth Embeddings: Component Correlations

PCA (Unsupervised)

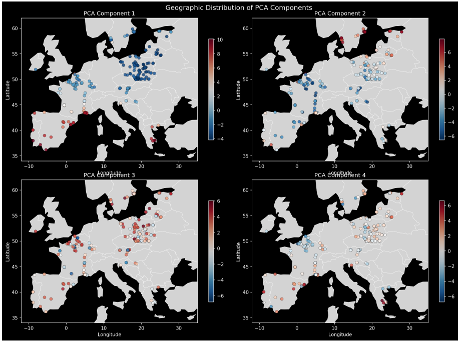


PLS (Supervised)

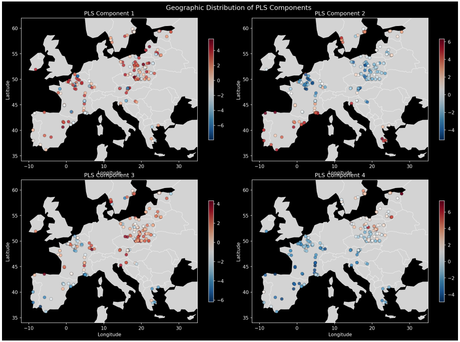


AlphaEarth Embeddings: Geographic Distribution

PCA Components 1–4



PLS Components 1–4

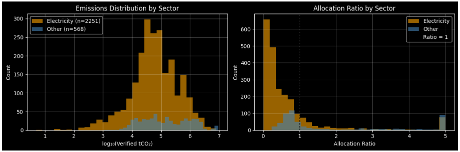


Main Results: Verified CO_2 Emissions

	Log CO_2
Allocation Ratio	-0.186^{***}
	(0.030)
95% CI	$[-0.25, -0.13]$
Observations	2,723
Facilities	521

Interpretation:

- 10% shortfall $\rightarrow$ **1.9%** lower emissions
- At mean: $\sim$11 ktCO_2 reduction
- $p < 0.001$



Electricity (orange) vs. Other sectors (blue)

Main Results: Satellite-Derived NOx

	DL $\geq$ 0.03 kg/s		DL $\geq$ 0.11 kg/s	
	PCA	PLS	PCA	PLS
Allocation Ratio	-0.000	-0.000	-0.003^{**}	-0.003^{***}
	(0.000)	(0.000)	(0.001)	(0.000)
p-value	0.84	0.65	0.017	0.001
N	577	577	140	140

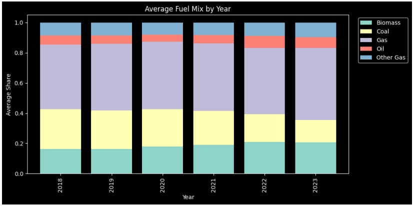
Key Pattern:

- **Null** at permissive threshold—noise dominates signal
- **Significant** at conservative threshold (DL $\geq$ 0.11 kg/s)
- 10% shortfall $\rightarrow$ $\sim$**1.7%** lower NOx (consistent with 1.9% CO_2!)

Cross-Validation: Directional agreement between administrative & satellite outcomes

Heterogeneity: ETS CO_2 by Fuel, Sector, Location

Dimension	Group	β	N
Fuel	Coal	-0.98^{***}	653
	Gas	-0.21^{***}	1,197
	Oil	-0.45	129
	Biomass	-0.11^{***}	438
Sector	Electricity	-0.21^{***}	2,173
	Other	-0.09^{***}	443
Location	Rural	-0.24^{***}	650
	Urban	-0.15^{***}	1,943



Coal declining, gas/biomass rising

Interpretation: Coal (∼95 tCO_2/TJ) faces highest carbon intensity → strongest response. Electricity lost free allocation → full marginal cost.

Heterogeneity: ETS CO_2 by Country & PyPSA Cluster

By Country:

Country	β	N
Spain	−1.25***	125
Poland	−0.33***	743
France	−0.25***	920
Austria	−0.24***	227
Sweden	−0.11***	421

By PyPSA-Eur Cluster:

Cluster	β	N
PL0 2 (Poland)	−1.46***	228
PL0 0 (Poland)	−1.09***	122
PL0 1 (Poland)	−0.38*	148
AT0 0 (Austria)	−0.26**	129
FR0 9 (France)	−0.25**	212

Interpretation: Polish clusters show strongest effects (5–7× pooled estimate)—coal-heavy generation mix. PyPSA clusters group facilities facing correlated prices & dispatch.

Heterogeneity: Satellite NOx (DL $\geq$ 0.11, PLS)

Dimension	Group	β	p	N
Sector	Electricity	-0.003***	.002	132
Interference	Yes (<20km)	-0.003**	.013	140
Stat Error	Low (<30%)	-0.003**	.013	140
Location	Urban	$+0.007$**	.045	128
Fuel	Gas	-0.003	.90	68
	Coal	$+0.051$***	.001	38
Country	France	-0.004***	.008	96
	Poland	$+0.079$**	.021	44

Findings: Electricity drives main effect; robustness to interference & stat error.

Anomalies: Urban/Coal/Poland show positive β—likely measurement noise (small N=140).

Robustness Summary

Cross-Outcome Consistency:

- Both CO_2 and NOx show negative effects (1.9% vs. 1.7%)
- Agreement across independent measurement systems

Embedding Method Stability:

- PCA vs. PLS nearly identical at DL ≥ 0.11 (-0.00282 vs. -0.00301)
- Not sensitive to dimensionality reduction approach

Measurement Quality:

- **Spatial interference**: Effects persist for interfered facilities ($\beta = -0.003^{**}$, $p = 0.013$)
- **Detection limit**: Null $\rightarrow$ significant pattern consistent with physics
- Inverse-variance weighting accounts for heteroskedasticity

Contributions & Conclusion

Three Methodological Contributions:

1. **AlphaEarth Embeddings**: Geospatial foundation model as high-dim controls (extends Veitch et al. 2019 to geospatial domain)
2. **NUTS2 + PyPSA-Eur**: Administrative regions for inference, network clusters for electricity heterogeneity
3. **Beirle Flux-Divergence**: Atmospheric physics adapted for panel econometrics

Key Empirical Findings:

- 10% allocation shortfall $\rightarrow$ **1.9% lower CO_2** ($p < 0.001$)
- Satellite NOx corroborates at conservative DL ($\sim$1.7%)
- Effects strongest for **coal-dominant** and **electricity-sector** facilities

Broader Impact: Framework for dual-outcome policy evaluation combining administrative records with satellite remote sensing