Energy Policy 218 (2026) 115525

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/jepo

Environmental taxation and biofuel production in the EU: Heterogeneous effects across producer scales

Monica Auteri

Department of Political Sciences, Roma Tre University, Rome, Italy

HIGHLIGHTS

- Pooled association of EU environmental taxes with biofuel capacity is statistically flat.
- Results are stable across a battery of robustness checks; no Hansen threshold is detected ($p = 0.85$).
- Producer-scale heterogeneity rejects coefficient homogeneity ($F = 3.12$, $p = 0.032$).
- Large export-oriented producers show a marginally significant, suggestive inverted-U.
- Findings point towards differentiated, not uniform, environmental fiscal design.

ARTICLE INFO

JEL classification:
Q42
Q48
Q58
H23
H32
C23
L52

Keywords:
Environmental taxation
Biofuels
Renewable energy policy
Panel data
Heterogeneous treatment effects
Producer scale
EU-27

ABSTRACT

Is environmental taxation associated with faster or slower biofuel production in the European Union? We examine this question on a balanced panel of the 27 EU member states over 2010–2023, combining Eurostat environmental-tax accounts with the OPEC Reference Basket as a globally exogenous oil-price control. Across cluster-robust LSDV, Driscoll–Kraay HAC, and System-GMM estimators, the pooled relationship between the environmental tax-to-GDP ratio and biofuel capacity is statistically flat. A non-parametric Hansen (1999) threshold test fails to identify any regime break ($p = 0.85$), and the flat result is robust to two- and three-year lags and to two-way (country and year) fixed effects; in a distributed-lag specification the *cumulative* effect over a three-year horizon is a modest negative one ($\sum_{h=0}^{3} \hat{\beta}_h = -0.71$, $p = 0.036$), consistent with slow-acting cost pass-through rather than a non-linear response. The pooled null, however, is an artefact of aggregation: a joint F-test rejects coefficient homogeneity across producer-size terciles ($F = 3.12$, $p = 0.032$; robust to year fixed effects and to alternative, outcome-independent classifications), and large export-oriented producers display a marginally significant and *suggestive inverted*-U pattern while smaller producers do not. Biofuel production is also highly persistent (System-GMM $\hat{\rho} \approx 0.96$), implying that the fiscal lever moves slowly. Because the design does not exploit an exogenous policy shock, we interpret these estimates as conditional associations rather than causal effects, and discuss their implications for differentiated environmental fiscal policy under the post-RED II framework.

1. Introduction

Whether environmental taxes can accelerate the transition to renewable energy without harming the competitive position of domestic producers is one of the central practical questions of contemporary environmental policy. The theoretical framework can be traced back to Pigou (1920) and its modern restatement in Baumol and Oates (1988).[1] The European Union has internalised the question in the 2009 Renewable Energy Directive (RED, 2009/28/EC), strengthened by the 2018 recast (RED II, 2018/2001/EU), which sets a binding 14% renewable-fuel target in the transport sector by 2030 and explicitly assigns to member-state fiscal policy the task of mobilising the private capital needed to expand renewable-fuel capacity (Antón, 2023; Leonard et al., 2021). Among renewable transport fuels, liquid biofuels—bioethanol and biodiesel—are

Email address: monica.auteri@uniroma3.it.

[1] Recent overviews include Andersson (2019) on the Swedish carbon tax, Martin et al. (2014) on the EU ETS, and the synthesis in Acemoglu et al. (2012), who argue that the optimal mix of carbon taxes and subsidies depends on the diffusion of clean technologies. Bovenberg and de Mooij (1994) clarify the interaction between environmental and pre-existing distortionary taxation. For energy-policy applications see OECD (2019) and Hassler et al. (2016).

https://doi.org/10.1016/j.enpol.2026.115525
Received 25 May 2026; Received in revised form 20 July 2026; Accepted 22 July 2026
Available online 29 July 2026
0301-4215/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

the largest by volume, and their domestic production is one of the policy outcomes most directly exposed to the level and composition of national environmental taxation (European Environment Agency, 2023; Morone et al., 2023).

The expected sign of the relationship is, on theoretical grounds, ambiguous. On the one hand, environmental tax revenue can be hypothecated or implicitly redirected to renewable-energy infrastructure: charging stations, blending mandates, feedstock subsidies, R&D platforms. Timilsina et al. (2011) provides one of the clearest formal statements of this revenue-recycling argument in the biofuels context, showing in a multi-sector general-equilibrium model that a carbon tax *cum* biofuel subsidy—where part of the tax revenue is recycled into biofuel support—materially stimulates biofuel penetration, whereas a carbon tax alone does not. Through this revenue channel, higher environmental tax-to-GDP ratios should be associated, in equilibrium, with larger domestic biofuel production capacity. On the other hand, environmental levies operate primarily on the price of fossil and energy-intensive inputs, and the bulk of biofuel production cost is precisely the feedstock and the energy used in conversion. Through the cost-push channel, the same environmental tax can erode the margin of domestic biofuel refiners, particularly in countries where production is export-oriented and price-competitive. The two channels work in opposite directions; their net effect is an empirical question.

A third strand of the literature, anchored in the political-economy work of Goulder and Schein (2013) and the synthesis in Stiglitz and Stern (2017), suggests that the relationship may be non-linear: at low tax levels, the cost channel dominates (taxation is high enough to discourage production but too low to finance a meaningful infrastructure response); beyond a fiscal-capacity threshold, accumulated revenue translates into investment that lowers the effective production cost, producing a net positive effect. The empirical literature has explored this possibility under different labels—Ulucak and Danish Kassouri (2020) and Chen et al. (2019) for CO_2 emissions, Zaghdoudi and Maktouf (2017) in a panel smooth transition framework, Bilan et al. (2022) specifically for biofuels—and typically reports U-shaped or threshold patterns. A robust U-shape would justify, on policy grounds, an aggressive front-loading of environmental fiscal effort: crossing the threshold quickly would be welfare-improving.

We ask two research questions. (RQ1) Is the environmental tax-to-GDP ratio associated with higher or lower biofuel production capacity across the EU-27, and is that association non-linear (U-shaped or threshold-type), as the prior literature often reports? (RQ2) Does the association differ systematically with the scale of the national producer sector? The distinction matters for policy: a homogeneous answer would justify a uniform fiscal rate, whereas a scale-dependent answer would call for differentiated design. Relative to the existing literature, which typically relates environmental taxes to emissions or growth, or treats tax rates and clean-energy prices as *outcomes*, our contribution is to map the lagged tax-to-GDP ratio into *physical* biofuel capacity and to allow the mapping to vary by producer scale.

The paper, examining these competing predictions on a balanced panel of the 27 EU member states observed annually between 2010 and 2023, makes three contributions to the empirical environmental economics literature. Substantively, it provides a panel evaluation of how EU biofuel production capacity is associated with environmental taxation, using the post-2009 institutional architecture (RED, RED II) as its observation window, and it documents producer-scale heterogeneity rather than imposing coefficient homogeneity—a restriction frequently maintained in the prior cross-country literature (e.g., Bilan et al., 2022; Zaghdoudi and Maktouf, 2017). Methodologically, it subjects the same panel to a broad battery of robustness checks—LSDV with cluster-robust and Driscoll–Kraay HAC inference, Difference- and System-GMM (Arellano and Bond, 1991; Blundell and Bond, 1998), wild-cluster bootstrap-*t* (Cameron et al., 2008; MacKinnon and Webb, 2017), Hansen (1999) threshold regression, alternative lag structures, two-way fixed effects and outcome-independent producer classifications—and documents that they deliver consistent qualitative conclusions. We emphasise that these are robustness checks rather than distinct identification strategies: the design does not exploit an exogenous policy shock, and we therefore read our estimates as conditional associations rather than causal effects. For policy, it qualifies the case for an aggressive front-loaded environmental tax push as a uniform tool to expand renewable-fuel capacity: the fiscal lever moves slowly, and its sign appears to depend on the production scale of the targeted industry—a conclusion consistent with the differentiated-instrument design recommended for energy taxation by OECD (2019).

The remainder of the paper is organised as follows. Section 2 reviews the most directly relevant empirical literature on environmental taxation and renewable-energy outcomes. Section 3 describes the dataset, sample, and identification strategy. Section 4 presents the pooled and heterogeneous specifications; Section 5 reports the full battery of robustness checks. Section 6 discusses interpretation and the limits of the evidence. Section 7 concludes.

2. Related literature

This paper sits at the intersection of three strands of the empirical environmental economics literature: the macro-level evidence on the effectiveness of environmental fiscal instruments, the firm- and sector-level evidence on how these instruments affect production and investment, and the methodological literature on identifying non-linear effects in panel data.

The first wave of macroeconomic evidence on environmental taxation tested whether national tax instruments reduce greenhouse-gas emissions. Andersson (2019) provides one of the cleanest identifications in this strand, exploiting the 1990 introduction of the Swedish carbon tax in a synthetic-control design and documenting a significant CO_2 reduction in road transport. Martin et al. (2014) study the EU Emissions Trading Scheme on a matched sample of UK manufacturing firms, finding that the policy reduced emissions intensity without measurable competitiveness losses. At the cross-country level, Ulucak and Danish Kassouri (2020), Zaghdoudi and Maktouf (2017) and Chen et al. (2019) extend the analysis to non-linear specifications, typically reporting U-shaped or threshold patterns in the tax–emissions relationship. The broader EU experience with environmental fiscal policy is reviewed in OECD (2019) and, for energy taxation specifically, in Hassler et al. (2016) and Antón (2023).

A smaller and more recent literature has examined the production-side response. Timilsina et al. (2011) use a multi-country CGE model to show that a carbon tax *cum* biofuel-subsidy policy significantly stimulates biofuel penetration, whereas a revenue-neutral carbon tax alone does not—a foundational result that motivates the revenue-channel hypothesis tested here. Bilan et al. (2022), in a panel of European economies, report a non-linear positive effect of environmental taxation on biofuel production and consumption. Mardones and Baeza (2018) examine the dual economic and environmental effects of CO_2 taxes in Latin America; Jordaan et al. (2017) document the role of energy-technology innovation in Canadian decarbonisation; and Cheng et al. (2019) provide cross-country evidence on heterogeneous patent responses. A common feature of this work is the maintenance of coefficient homogeneity across producers—an assumption we relax in Section 4.

On the econometric side, our analysis builds on three established strands. First, the dynamic-panel literature originating in Nickell (1981); Arellano and Bond (1991); Blundell and Bond (1998) provides the bias diagnostics and the GMM estimators used in Section 5; Roodman (2009) surveys the `xtabond2` implementation. Second, the literature on cluster-robust inference with a small number of clusters (Bertrand et al., 2004; Cameron et al., 2008; Cameron and Miller, 2015; MacKinnon and Webb, 2017) motivates the use of wild-cluster bootstrap-*t* for the headline specifications—with 27 EU member states the asymptotic *t*-approximation is not guaranteed to deliver correctly sized tests. Third, the threshold-regression methodology of Hansen (1999) allows us to verify the parametric quadratic specification non-parametrically; the Lind–Mehlum test, in turn, formally evaluates the U-shape hypothesis on the observed data support. For the interpretation of interactions and the cross-derivative analysis in Section 4 we follow Brambor et al. (2006).

Two further strands of recent (2024–2025) work bear directly on our identification. First, on the *determinants* of energy taxation: Sineviciene and Miliauskaite (2025) show, for a panel of EU countries over 2010–2020, that energy tax rates are systematically related to energy intensity, corporate profitability and investment. This matters for our design because it implies that the tax-to-GDP ratio is not an exogenous instrument but is itself co-determined by country-level economic conditions; we return to the resulting endogeneity in Section 6 and interpret our estimates as conditional associations. Second, on *energy-market price shocks*: Baghirzade and Kosormyhin (2025) document that oil and natural-gas prices co-move with clean-energy equity valuations, which suggests that the fossil-fuel price environment shapes both the political economy of environmental taxation and the private return to renewable-fuel investment. This motivates our use of a globally exogenous crude-oil control while cautioning that oil prices alone are only a partial proxy for the broader energy-market environment.

3. Data, sample and empirical strategy

3.1. Data and sources

The dataset is a balanced panel of 27 EU member states over 14 years (2010–2023), assembled from harmonised Eurostat extracts and the OPEC Annual Statistical Bulletin. Table B.1 in the Appendix lists every variable with its Eurostat code; the most important categories are summarised here.

Outcome variable. Biofuel production capacity in thousand tonnes of oil equivalent (ktoe) is taken from Eurostat `nrg_inf_lbpc`, which covers all liquid biofuels (biodiesel, bioethanol, and other renewable liquid fuels processed for transport use). We work with $\ln(1 + \text{capacity})$ in line with the treatment of zero-valued observations in Bellemare and Wichman (2020).

Treatment variable. The environmental tax-to-GDP ratio is constructed as total environmental tax revenue (Eurostat `env_ac_tax`) divided by current-price GDP (`nama_10_gdp`), expressed as a percentage. The four sub-components—energy taxes including the CO_2-related subset, pollution taxes and resource taxes—are taken from `env_ac_taxind2`. The Eurostat classification follows the National Tax Lists methodology (Eurostat, 2013).[2]

Oil-price control. A central methodological choice in this literature concerns the appropriate control for the world price of fossil fuels, against which biofuel production is the closest economic substitute. Country-level retail diesel prices (Eurostat Weekly Oil Bulletin) are mechanically correlated with the energy-tax component of the treatment, since national excises on diesel enter the very same `env_ac_taxind2` aggregate; including them as a control therefore conditions on a downstream consequence of the policy (in the sense of Angrist and Pischke, 2009, p. 64). We instead use the *OPEC Reference Basket* (Annual Statistical Bulletin), an import-weighted average of the spot prices for crude oils produced by OPEC member states, expressed in USD per barrel. The basket is identical for every EU member state in any given year and is unambiguously exogenous to EU national tax policy. In all specifications the oil control enters in logs. Section 5.3 documents that the choice is consequential: under the country-level diesel control the quadratic coefficient on the treatment switches sign, albeit to a magnitude that remains statistically marginal.

[2] The environmental-tax aggregate includes excise duties on energy products, vehicle taxes and a heterogeneous group of pollution and resource taxes whose country-specific composition varies considerably; the energy sub-component is by far the largest and the most directly relevant for biofuel production decisions.

Macroeconomic controls. Beyond the oil price, we include the log of energy intensity (`nrg_ind_ei`), the log of cereals production area (`apro_cpsh1`), the log of the labour cost index (`lc_lci_lev`), and real GDP growth (`tec00115`). Cereals area captures the feedstock-availability constraint relevant for first-generation biofuels.

A caveat on the treatment variable is in order. The environmental tax-to-GDP ratio is a broad aggregate: it combines instruments with very different bases, exemptions and degrees of connection to biofuel production, and a high ratio need not signal strong support for renewable fuels. We therefore treat it as an imperfect, potentially noisy proxy for the underlying policy stance rather than as a clean policy lever, and in Section 5 we re-estimate the main specification using the disaggregated tax components (energy, CO_2, pollution and resource taxes) separately. Complementary instruments that operate alongside the fiscal lever—blending mandates, renewable-fuel subsidies, feedstock support, renewable-fuel certificates, the EU ETS and fossil-fuel subsidy reform—are discussed in Section 6.

3.2. Sample

The main estimation sample comprises 351 country–year observations (27 countries × 13 years, with one observation per country lost to the first-lag construction); descriptive statistics in Table 1 report a total of 378 country–year observations on the outcome and treatment in levels.

Static specifications with the full control set are estimated on a reduced sub-sample ($N = 234$) because of the more limited time coverage of the labour-cost index; this is documented transparently in Table 2.

For the heterogeneity analysis we classify member states into terciles on the basis of average 2010–2023 biofuel production capacity, with nine countries per group; assignments are listed in the notes to Table 3. Because this grouping uses the estimation-period outcome and could therefore be regarded as endogenous, and because our panel begins in 2010 so that a genuine pre-sample (2005–2009) window is unavailable, we verify in Section 5 that the heterogeneity is robust to three classifications that are independent of the contemporaneous outcome: feedstock availability (mean cereal area, 2010–2012), agricultural structure (mean

Table 1
Descriptive statistics, EU-27 panel 2010–2023.

Variable	N	Mean	Std. dev.	Min	Max
A. Outcome					
ln(Biofuel capacity + 1), ktoe	378	5.239	2.724	0.000	9.271
Lagged ln(Biofuel capacity + 1)	351	5.253	2.698	0.000	9.255
B. Treatment variables					
Env. tax / GDP, %	378	2.612	0.770	0.850	5.621
ln(Env. tax / GDP)	378	1.262	0.208	0.615	1.890
Lagged ln(Env. tax / GDP)	351	1.274	0.203	0.615	1.890
Centred lag, $\tilde{x}_{i,t-1}$	351	0.000	0.203	−0.658	0.616
Squared centred lag, $\tilde{x}^2_{i,t-1}$	351	0.041	0.056	0.000	0.433
ln(Energy tax + 1), mil. EUR	378	8.121	1.465	4.540	11.329
C. Macroeconomic controls					
ln(OPEC reference basket), USD/bbl	378	4.295	0.332	3.732	4.705
ln(Energy intensity)	378	4.808	0.326	3.521	5.645
ln(Cereals area + 1), kha	377	6.484	1.973	0.000	9.168
GDP growth, %	297	1.840	3.521	-10.090	24.890
ln(Labour cost index)	234	4.476	0.142	4.196	4.798

Notes. Statistics computed on the main analysis sample (2010–2023). All log transformations of the form $\ln(X + 1)$ to accommodate the small share of zero-valued outcome observations in several non-producer member states. The centred lagged treatment $\tilde{x}_{i,t-1} \equiv L.\ln(\text{Env. tax/GDP}) - \bar{x}$ is de-meaned on the dynamic-FE estimation sample; the implied mean $\bar{x} = 1.274$ corresponds to an env. tax/GDP level of $\exp(1.274) \approx 3.57\%$. Sources: Eurostat (env. taxes: `env_ac_tax`, `env_ac_taxind2`; biofuels: `nrg_inf_lbpc`; GDP: `nama_10_gdp`; energy intensity: `nrg_ind_ei`; cereals: `apro_cpsh1`); OPEC Annual Statistical Bulletin (various editions, 2017–2025). See Appendix B for full definitions.

Table 2
Environmental taxation and biofuel production capacity: pooled specifications.

Dependent variable: $\ln(\text{Biofuel capacity} + 1)_{i,t}$	(1) Static FE full controls	(2) Dynamic FE linear	(3) Dynamic FE quadratic
$\ln(\text{Biofuel capacity})_{i,t-1}$		0.729^{***}	0.729^{***}
		(0.079)	(0.078)
$\ln(\text{Energy tax})_{i,t-1}$	1.542	−0.099	−0.087
	(1.380)	(0.197)	(0.175)
$\ln(\text{Env. tax / GDP})_{i,t-1}$	−3.468	-0.483^{*}	
	(3.191)	(0.282)	
$\tilde{x}_{i,t-1}$ (centred lag)			−0.498
			(0.313)
$\tilde{x}^2_{i,t-1}$ (centred lag, squared)			−0.165
			(0.823)
ln(Cereals area)	0.888		
	(0.535)		
ln(Labour cost)	−0.103		
	(0.179)		
ln(Energy intensity)	2.147		
	(1.907)		
ln(OPEC ref. basket)	−0.124		
	(0.265)		
Country fixed effects	Yes	Yes	Yes
Cluster-robust SE (countries)	25	27	27
Observations	234	351	351
Within R^2	0.084	0.514	0.515
Tests on the U-shape (column 3)			
$H_0: \beta_{\text{sq}} = 0$, cluster-robust p-value			0.843
$H_0: \beta_{\text{sq}} = 0$, wild-bootstrap p (9999 reps)			0.899
$H_0: \beta_{\text{lin}} = \beta_{\text{sq}} = 0$, joint wild p			0.671
Lind–Mehlum U-shape test			trivial fail

Notes. Cluster-robust standard errors in parentheses, clustered at the country level. The centred treatment in column 3 is $\tilde{x}_{i,t-1} \equiv L.\ln(\text{Env. tax/GDP}) - \bar{x}$, with $\bar{x} = 1.274$ computed on the column-3 estimation sample. Wild bootstrap p-values use 9999 Rademacher-weighted resamples (Cameron et al., 2008; MacKinnon and Webb, 2017), seed = 20,260,301. The Lind–Mehlum test returns "trivial failure to reject H_0" when the parametrically implied extremum lies outside the observed support of the regressor, which is the case here. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

agricultural gross value added), and a predetermined early-window classification (mean capacity, 2010–2012).

3.3. Empirical specification

Our preferred specification is a dynamic fixed-effects panel regression in mean-centred quadratic form,

$$\ln(Y_{i,t} + 1) = \rho\ \ln(Y_{i,t-1} + 1) + \beta_1\, T_{i,t-1} + \beta_2\, \tilde{x}_{i,t-1} + \beta_3\, \tilde{x}^2_{i,t-1} + \boldsymbol{Z}'_{i,t-1}\boldsymbol{\gamma} + \alpha_i + \varepsilon_{i,t}, \tag{1}$$

where $Y_{i,t}$ is biofuel production capacity in country i at year t; $T_{i,t-1}$ is the log of the (lagged) energy-tax component, an included macroeconomic control; $\tilde{x}_{i,t-1} \equiv \ln(\text{Env. tax}/\text{GDP})_{i,t-1} - \bar{x}$ is the mean-centred lagged treatment, where $\bar{x} = 1.274$ is the mean over the estimation sample (corresponding to an env-tax/GDP ratio of $\exp(\bar{x}) \approx 3.57\%$); $\boldsymbol{Z}_{i,t-1}$ is a vector of macroeconomic controls; α_i is a time-invariant country fixed effect; and $\varepsilon_{i,t}$ is the idiosyncratic error.

Mean-centring is mechanically equivalent to a re-parameterisation of the quadratic but it has two distinct practical advantages. *First*, it eliminates the artificial collinearity between the linear and quadratic terms that plagues uncentred specifications: Table A.1, panel C (Appendix A), reports a mean VIF of 1.14 under centring against 22.20 uncentred. *Second*, the linear coefficient β_2 recovers the marginal effect of the treatment *at the sample mean*, a quantity that has a natural policy-relevant interpretation (Wooldridge, 2010). The implied turning point of the quadratic is $x^* = \bar{x} - \beta_2/(2\beta_3)$, which can be expressed in percentage of GDP as $\exp(x^*)$ for direct comparability with policy benchmarks.

Two modelling choices deserve explicit justification. First, *why a dynamic specification?* Biofuel capacity is strongly persistent: the estimated autoregressive coefficient rises from $\hat{\rho} \approx 0.73$ (LSDV) to 0.86 (Nickell-corrected) to 0.96 (System-GMM). Omitting the lagged dependent variable would therefore bias the treatment coefficients through dynamic mis-specification; the dynamic model is a requirement of the data rather than a stylistic choice. Second, *why fixed rather than random effects?* We report Hausman tests (Table A.1, Appendix A); although the test does not sharply reject, we retain fixed effects because identification rests on within-country variation and the random-effects orthogonality condition is implausible given the endogeneity of tax levels documented by Sineviciene and Miliauskaite (2025). Multicollinearity is not a concern once the quadratic is mean-centred: the mean VIF falls from 22.2 to 1.14 (Table A.1, panel C, Appendix A).

Identification of β_2 and β_3 relies on within-country variation in the environmental tax-to-GDP ratio over time, conditional on lagged production and on the macroeconomic controls. The threats to this identification—short-T Nickell bias on $\hat{\rho}$, omitted policy correlates absorbed by the fixed effect, cross-sectional dependence from EU-wide energy shocks—are addressed in Section 5 by, respectively, System-GMM, the inclusion of energy- and resource-tax components as placebo regressors, and Driscoll–Kraay HAC inference. Common EU-wide shocks (RED II, COVID-19, the 2022 energy crisis) are addressed directly by adding year fixed effects (two-way FE); the block of year dummies is not jointly significant once the dynamic structure and the OPEC control are included, and our headline heterogeneity result is unchanged (Section 5). Second-generation panel diagnostics—the Pesaran (2004) cross-sectional-dependence test and the Pesaran (2007) CIPS unit-root test—are reported in Table A.1 (Appendix A).

The heterogeneity specification interacts $\tilde{x}_{i,t-1}$ and $\tilde{x}^2_{i,t-1}$ with size-tercile dummies (Large is the reference group); we additionally estimate Eq. (1) separately for each tercile (Section 5).

4. Results

We present the empirical evidence in two steps. We first estimate the pooled relationship between environmental taxation and biofuel production capacity, in three nested fixed-effects specifications and then report the formal U-shape and threshold tests (Section 4.1). Second, in Section 4.2, we relax the assumption of coefficient homogeneity and document the heterogeneity by producer scale that constitutes the headline finding of the paper. Throughout, standard errors are clustered at the country level (27 clusters); the robustness of this inferential choice is examined in Section 5.

4.1. Pooled specifications

Table 2 reports the pooled estimates of Eq. (1) and two simpler nested specifications. Column 1 is a static fixed-effects regression with the full set of macroeconomic controls (cereal area, labour cost, energy intensity, net biofuel trade, and the log OPEC reference basket); column 2 introduces the lagged dependent variable and the linear treatment but omits the time-varying controls absorbed by the dynamic component; column 3 is our preferred specification, the dynamic fixed-effects regression in mean-centred quadratic form.

The static specification of column 1 is poorly suited to data with a substantial autoregressive component: the within R^2 is 0.084, and none of the controls enters significantly at conventional levels. As soon as the lagged dependent variable is included (columns 2 and 3), the within R^2 rises to 0.514, with the lagged outcome carrying a coefficient of $\hat{\rho} = 0.729$ ($p < 0.001$). The implied half-life of a shock under LSDV is $\ln(0.5)/\ln(0.729) \approx 2.2$ years; under the Nickell (1981) bias correction it is approximately $\ln(0.5)/\ln(0.86) \approx 4.6$ years; under the System-GMM estimator of Section 5, with $\hat{\rho} \approx 0.96$, the half-life lengthens to nearly two decades. Whichever estimator is preferred, biofuel production capacity

Table 3
Heterogeneous effects of environmental taxation by producer scale.

Dependent variable: $\ln(\text{Biofuel capacity} + 1)_{i,t}$	(1) Pooled with interactions	(2) Small producers	(3) Medium producers	(4) Large producers
$\ln(\text{Biofuel capacity})_{i,t-1}$	0.684***	0.824***	0.487***	0.714***
	(0.075)	(0.048)	(0.015)	(0.070)
$\ln(\text{Energy tax})_{i,t-1}$	−0.140	−0.206	0.459**	−0.135
	(0.152)	(0.179)	(0.152)	(0.280)
$\tilde{x}_{i,t-1}$ (centred lag)	−0.267	−0.454	−0.880*	−0.371
	(0.394)	(0.295)	(0.475)	(0.472)
$\tilde{x}^2_{i,t-1}$ (centred lag, squared)	−3.910**	0.016	1.467	−3.529*
	(1.473)	(0.416)	(0.930)	(1.841)
Size-group interactions (col. 1 only)				
Small $\times\, \tilde{x}_{i,t-1}$	0.003			
	(0.530)			
Medium $\times\, \tilde{x}_{i,t-1}$	0.099			
	(0.533)			
Small $\times\, \tilde{x}^2_{i,t-1}$	4.503**			
	(1.952)			
Medium $\times\, \tilde{x}^2_{i,t-1}$	6.267**			
	(2.359)			
Implied net $\hat{\beta}_{\text{sq}}$ by group:		0.594	2.358	−3.910**
Joint F-test, all size interactions = 0	$F = 3.12$ $p = 0.032$			
Group-specific Lind–Mehlum p-value		trivial fail	0.119	0.073[a]
Group-specific turning point, % of GDP		n.d.	4.48	3.44[a]
Country fixed effects	Yes	Yes	Yes	Yes
Cluster-robust SE (countries)	27	9	9	9
Observations	351	117	117	117
Within R^2	0.527	0.740	0.506	0.428

Notes. Cluster-robust standard errors in parentheses, clustered at the country level. Producer-size categories are time-invariant terciles computed on average 2010–2023 biofuel production capacity: **Small** (Croatia, Cyprus, Denmark, Estonia, Ireland, Latvia, Luxembourg, Malta, Slovenia); **Medium** (Bulgaria, Czechia, Finland, Hungary, Lithuania, Portugal, Romania, Slovakia, Sweden); **Large** (Austria, Belgium, France, Germany, Greece, Italy, Netherlands, Poland, Spain). In column 1, the omitted reference category is the *Large* producer group, so its quadratic coefficient corresponds to the unconditional row of the table; the Small and Medium quadratic effects are obtained by summing the row coefficient with the respective group interaction. In columns 2–4, each tercile is estimated separately with within-group mean-centring.

[a] The Lind–Mehlum test on column 4 evaluates an *inverted*-U ($H_1 : f'' < 0$), not a U-shape, because $\hat{\beta}_{\text{sq}} < 0$ in the Large-producer group; the corresponding "turning point" is the parabolic maximum, not a minimum. The interpretation of interaction terms follows Brambor et al. (2006). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

is a slow-moving outcome and the contemporaneous response to fiscal stimuli is small relative to the persistence.

The contemporaneous (static) specification of column 1 returns a coefficient of −3.47 on the lagged log env. tax/GDP ratio, but with a standard error of 3.19 the estimate is statistically indistinguishable from zero ($p = 0.288$). The dynamic linear specification of column 2 returns a more precisely estimated −0.483, marginally significant at the 10% level ($p = 0.099$). The point estimate is small in magnitude: since the regressor is in logs of a percentage, a one-standard-deviation increase in the lagged log env. tax/GDP (about 0.20, corresponding to roughly +0.7 percentage points of GDP) is associated with a reduction in (log) biofuel capacity of about 0.10, or 10% at the mean. The sign of the estimate is opposite to that predicted by the revenue-channel argument of Timilsina et al. (2011) but consistent with the cost-push channel; we will see in Section 4.2 that this aggregate sign masks considerable group heterogeneity.

Column 3 introduces the centred quadratic. The linear coefficient on the centred treatment is $\hat{\beta}_{\text{lin}} = -0.498$ (s.e. 0.313), broadly similar to column 2; the quadratic coefficient is $\hat{\beta}_{\text{sq}} = -0.165$ (s.e. 0.823), with a cluster-robust p-value of 0.843. A wild-cluster bootstrap-t with 9999 Rademacher replications (Cameron et al., 2008) returns $p = 0.899$ on the quadratic coefficient alone and a joint p-value of 0.671 on the linear and quadratic terms tested simultaneously—the finite-sample-valid counterpart of the analytic standard errors, which we report because the asymptotic justification for cluster-robust inference under $G = 27$ clusters is borderline (Cameron and Miller, 2015). The Lind–Mehlum test for the presence of a U-shape on the observed support of the regressor fails non-trivially: the parametrically implied extremum lies outside the observed range ($\tilde{x} \in [-0.658, +0.616]$), so the quadratic captures monotonic curvature within the data rather than a turning point in the relevant sense. The non-parametric Hansen (1999) threshold regression reported in Section 5 delivers a consistent conclusion: a bootstrap F-statistic of 2.47 against a 10% critical value of 13.87, with $p = 0.850$, fails to identify any regime break.

The pooled evidence, taken together, supports neither the revenue channel of Timilsina et al. (2011) nor the non-linear U-shape of the threshold literature (Ulucak and Danish Kassouri, 2020; Zaghdoudi and Maktouf, 2017; Bilan et al., 2022). The dynamic linear effect is small, negative, and only marginally significant; the quadratic non-linearity is statistically absent. Two interpretations are possible at this stage: either the relationship is genuinely flat, in which case environmental taxation is not a relevant policy lever for EU biofuel capacity over the 2010–2023 window; or the homogeneity restriction maintained by the pooled specification is misleading, and the underlying response structure is heterogeneous across producers. Section 4.2 discriminates between the two interpretations.

4.2. Heterogeneity by producer scale

The empirical question we ask in this section is whether the pooled null in Table 2 arises from a uniformly absent treatment effect, or from the aggregation across producer groups whose responses run in opposite directions. We address it by partitioning the 27 member states into terciles based on average 2010–2023 biofuel production capacity—small, medium, and large, with nine countries per group—and estimating two complementary specifications: a single regression that interacts size-tercile dummies with the centred quadratic (Table 3, column 1), and three group-specific regressions estimated separately on each tercile (columns 2–4).

The pooled regression with size interactions in column 1 carries a quadratic coefficient on the omitted (Large-producer) reference group of $\hat{\beta}_{\text{sq}} = -3.91$ ($p = 0.013$), and positive interactions of +4.50 ($p = 0.029$) for small producers and +6.27 ($p = 0.013$) for medium producers. The implied net quadratic effects by group—reading the table as recommended by Brambor et al. (2006)—are therefore −3.91 for Large, +0.59 for Small, and +2.36 for Medium. The joint F-test on the four size interactions rejects the null of coefficient homogeneity at the 5% level ($F = 3.12$, $p = 0.032$), establishing that producer scale is a statistically meaningful moderator of the response.

Columns 2–4 re-estimate the quadratic specification separately on each tercile, with the centring mean computed within group. The results sharpen the aggregate inference and qualify it.

Large producers (column 4: Austria, Belgium, France, Germany, Greece, Italy, Netherlands, Poland, Spain) exhibit a negative quadratic coefficient of −3.53, marginally significant at the 10% level (s.e. 1.84, $p = 0.092$). The Lind–Mehlum test for an *inverted*-U on the observed support returns $p = 0.073$, just outside conventional significance but informative. The parabolic maximum implied by the estimates falls at 3.44% of GDP, within the observed support of the regressor in this group. Substantively, large producers' biofuel capacity increases with environmental taxation up to roughly 3.4% of GDP and decreases beyond that point—a pattern consistent with the margin-erosion hypothesis discussed in Section 6.

Medium producers (column 3: Bulgaria, Czechia, Finland, Hungary, Lithuania, Portugal, Romania, Slovakia, Sweden) show a positive quadratic coefficient of +1.47 (s.e. 0.93, $p = 0.153$), marginally significant linear coefficient of −0.88 ($p = 0.101$), and an implied within-group turning point of 4.48% of GDP that closely echoes the unconditional turning-point estimates one finds in the prior literature. The Lind–Mehlum p-value of 0.119 does not reach conventional significance, but the directional evidence is consistent with a (weak) U-shape that would, if confirmed on a longer panel, support the revenue-channel interpretation.

Small producers (column 2: Croatia, Cyprus, Denmark, Estonia, Ireland, Latvia, Luxembourg, Malta, Slovenia) display essentially no curvature: the quadratic coefficient is +0.02 (s.e. 0.42) and the Lind–Mehlum test returns a trivial failure to reject. The Small-producer subset is dominated by member states with very small or zero installed biofuel capacity over much of the sample, which limits the within-group variation that identifies the quadratic.

The producer-scale terciles uncover directionally opposed responses to environmental taxation that the pooled specification masks. The cleanest case is the Large-producer group, whose inverted-U is statistically marginal but operationally meaningful: a turning point of 3.44% of GDP lies just below the sample mean of 3.57%, so at the policy-relevant range of taxation the large producers are on the downward arm of the parabola. The Medium-producer evidence is consistent with the revenue-channel argument but weaker; the Small-producer evidence is essentially null. The mechanisms that plausibly account for this pattern, and the policy implications that follow, are discussed in Section 6.

5. Robustness

The two headline findings of Section 4—the absence of a robust U-shape in the pooled specification, and the producer-scale heterogeneity that explains the pooled null—rest on identifying assumptions whose violation could in principle deliver the same empirical pattern. These two findings survive a comprehensive battery of identification and inference checks, which we document in this section.

5.1. Dynamic-panel and cross-section-robust inference

Table 4 re-estimates the main quadratic specification of Table 2, column 3, under five inferential alternatives that progressively relax the assumptions of cluster-robust LSDV. Column 1 reproduces the baseline LSDV with cluster-robust standard errors for reference.

Column 2 implements the Hoechle (2007) Driscoll–Kraay standard errors with two lags, which are robust to both cross-sectional dependence and serial correlation of unspecified form. This is the natural inferential response to the Pesaran (2004) CD test of Table A.1 (Appendix A), which rejects cross-section independence at the 1% level. The point estimate of the quadratic coefficient is unchanged ($\hat{\beta}_{\text{sq}} = -0.165$); the standard error contracts from 0.823 under cluster-robust to 0.305 under Driscoll–Kraay, but the coefficient remains insignificant at conventional levels ($p = 0.599$). The linear coefficient becomes marginally significant at 5% ($\hat{\beta}_{\text{lin}} = -0.498$, s.e. 0.189, $p = 0.022$); this is the only inferential alternative under which the linear treatment effect crosses conventional significance. Note that under Driscoll–Kraay inference, the asymptotic dimensions are $T \to \infty$ for fixed N; with $T = 14$ this is a more demanding assumption than the cluster-robust $N \to \infty$ for fixed T, and we therefore report it as a complement rather than a substitute for the cluster-robust headline.

Column 3 estimates the first-differenced model with the second lag of the level outcome as the instrument for the differenced lagged outcome. Under the documented persistence ($\hat{\rho} = 0.96$ from System-GMM), the second-lag instrument is weak: standard errors expand by an order of magnitude (from 0.078 to 15.13 on the lagged outcome, and from 0.823 to 88.63 on the quadratic). The estimator delivers point estimates that are essentially uninformative and we report it primarily for completeness; the appropriate dynamic-panel estimator for this data structure is System-GMM, to which we now turn.

Columns 4 and 5 estimate Arellano and Bond (1991) difference-GMM and Blundell and Bond (1998) system-GMM, both in two-step form with the Windmeijer (2005)-corrected standard errors. Instrument lag depth is restricted to (2, 4) and collapsed across periods to keep the instrument count well below the cluster count (Roodman, 2009): 10 instruments in difference-GMM and 14 in system-GMM, against 27 clusters in both cases. The Arellano and Bond (1991) AR(2) test, the most informative specification test for these estimators, returns p-values of 0.488 and 0.599 respectively, comfortably above the conventional 0.10 threshold and consistent with the absence of second-order serial correlation in the levels equation. The Hansen (1982) J test of overidentifying restrictions returns p-values of 0.398 and 0.645, supporting instrument validity.

System-GMM returns $\hat{\rho} = 0.962$ (s.e. 0.022), a substantially larger persistence estimate than LSDV and consistent with a near-unit-root data generating process when the Nickell bias is fully corrected. The quadratic coefficient is $\hat{\beta}_{\text{sq}} = -0.962$ (s.e. 0.992, $p = 0.341$), larger in magnitude than the LSDV estimate but with a confidence interval that includes both substantial negative and substantial positive values. Crucially for our central claim, the System-GMM estimate does not deliver a positive significant quadratic coefficient under any inferential standard. Across the five estimators of Table 4, the quadratic coefficient ranges from −0.165 to −3.23, all negative, and none significant at conventional levels.

The bottom rows of Table 4 report wild-cluster bootstrap-t p-values for the LSDV estimator of column 1, with 9999 Rademacher replications and seed 20,260,301 (Cameron et al., 2008; MacKinnon and Webb, 2017). For the quadratic coefficient alone the bootstrap returns $p = 0.899$ (against 0.843 analytically); for the joint test on the linear and quadratic coefficients combined the bootstrap returns $p = 0.671$. The discrepancy between bootstrap and analytic inference is modest, consistent with the conclusion of Cameron et al. (2008) that the asymptotic t-approximation under cluster-robust inference is acceptable for $G = 27$ clusters; we report the bootstrap p-value as a finite-sample-valid complement rather than as a substitute.

Table 4
Econometric robustness: Driscoll–Kraay, Anderson–Hsiao, GMM, wild bootstrap.

Dependent variable: $\ln(\text{Biofuel capacity} + 1)_{i,t}$	(1) LSDV cluster-rob. SE	(2) Driscoll–Kraay HAC	(3) Anderson–Hsiao[a]	(4) Diff. GMM	(5) System GMM
$\ln(\text{Biofuel capacity})_{i,t-1}$	0.729***	0.729***	−0.539	1.044***	0.962***
	(0.078)	(0.082)	(15.127)	(0.182)	(0.022)
$\ln(\text{Energy tax})_{i,t-1}$	−0.087	−0.087	−0.731	0.433	0.067*
	(0.175)	(0.100)	(12.173)	(1.140)	(0.035)
$\tilde{x}_{i,t-1}$ (centred lag)	−0.498	−0.498**	0.318	−0.059	−0.121
	(0.313)	(0.189)	(5.153)	(0.762)	(0.405)
$\tilde{x}^2_{i,t-1}$ (centred lag, squared)	−0.165	−0.165	−3.234	−2.836	−0.962
	(0.823)	(0.305)	(88.625)	(5.749)	(0.992)
Country fixed effects	Yes	Yes	Yes[b]	Yes[b]	Yes[b]
Year effects	No	No	No	No	No
Observations	351	351	297	324	351
Number of groups	27	27	27	27	27
Number of instruments	—	—	1	10	14
Specification tests					
Arellano and Bond (1991) AR(1), p				0.117	0.092
AR(2), p				0.488	0.599
Sargan (1958) overid. test, p				0.988	0.994
Hansen (1982) J, p				0.398	0.645
Wild bootstrap-t, β_{sq}, p[c]	0.899	—	—	—	—
Wild bootstrap-t, JOINT ($\beta_{\text{lin}}, \beta_{\text{sq}}$), p[c]	0.671	—	—	—	—

Notes. Column 1 reproduces the main LSDV estimates (Table 2, column 3) for reference. Column 2 applies Hoechle (2007) Driscoll–Kraay HAC standard errors with two lags, robust to both autocorrelation and cross-sectional dependence; under this inference $\hat{\beta}_{\text{lin}}$ becomes marginally significant at the 5% level but $\hat{\beta}_{\text{sq}}$ remains insignificant. Columns 4 and 5 estimate Arellano and Bond (1991) difference-GMM and Blundell and Bond (1998) system-GMM with two-step robust standard errors; instrument lag depth is collapsed to $(2,4)$ to keep the instrument count below the cluster count (Roodman, 2009). Across all five estimators $\hat{\beta}_{\text{sq}}$ remains statistically indistinguishable from zero, and where signed it is negative. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[a] The Anderson–Hsiao first-difference IV estimator (Arellano and Bond, 1991) instruments $\Delta Y_{i,t-1}$ with $Y_{i,t-2}$ alone. Its standard errors are an order of magnitude larger than those of the other estimators (column 3), consistent with the well-known weak-instrument problem when the persistence parameter is close to unity. We retain System-GMM (column 5) as the preferred dynamic estimator.

[b] Country fixed effects are absorbed by first differencing (columns 3 and 4) or by the system-GMM level equation (column 5).

[c] Wild bootstrap-t inference is applied to the LSDV estimator in column 1, with 9999 Rademacher replications and seed = 20,260,301 (Cameron et al., 2008; MacKinnon and Webb, 2017).

5.2. Placebo identification

Table 5 reports placebo-identification checks that substitute either the treatment variable or the outcome with a quantity that should, under our identifying assumptions, be uninformative about the treatment effect of interest.

Alternative treatment components. Columns 2 and 3 substitute the lagged log of pollution-tax revenue or resource-tax revenue (Eurostat `env_ac_taxind2`, respective components) for the env. tax/GDP treatment, keeping the rest of the specification identical. The substantive logic is that pollution taxes (on emissions to air and water, ozone-depleting substances, waste) and resource taxes (on water abstraction, mineral and aggregate extraction) operate on bases that are structurally unrelated to the input cost of biofuel production. If the U-shape of the pooled main specification arose from a genuine revenue-recycling mechanism operating through energy fiscal policy, neither alternative component should display the same non-linearity.

The resource-tax placebo (column 3) behaves as expected: the quadratic coefficient is statistically indistinguishable from zero ($\hat{\beta}_{\text{sq}} = -0.015$, s.e. 0.015, $p = 0.326$), and the Lind–Mehlum test does not reject ($p = 0.176$). The pollution-tax placebo (column 2), however, returns a small but precisely-estimated positive quadratic coefficient ($\hat{\beta}_{\text{sq}} = +0.074$, s.e. 0.023, $p = 0.003$) and a Lind–Mehlum test that rejects the null at the 1% level ($p = 0.007$). Our interpretation, developed in Section 6.1, is that the pollution-tax result captures an unobserved country-level *policy stringency* factor that operates simultaneously through multiple channels—pollution taxation correlates with the broader regulatory severity of national environmental policy, which in turn correlates with renewable-fuel mandates and subsidy envelopes that the country fixed effect does not fully absorb. Under this reading, the pollution-tax rejection does not falsify the heterogeneity finding of Section 4.2; it reinforces the conclusion that aggregate tax-revenue indicators are noisy proxies for the underlying policy bundle.

Alternative outcome. Column 4 substitutes the log of Y02W recycling-and-secondary-raw-materials patents (Eurostat `cei_cie020`) for the biofuel outcome, keeping the env. tax treatment unchanged. The patents indicator is available only through 2021, hence the shorter sample ($N = 297$). If the central mechanism postulated by the literature ran through a generic “green investment” channel rather than a sector-specific one, we would expect environmental taxation to affect the recycling patents outcome at a similar order of magnitude as biofuel capacity. Instead, the within R^2 collapses to 0.010, the quadratic coefficient is statistically insignificant ($\hat{\beta}_{\text{sq}} = +0.718$, s.e. 0.684, $p = 0.304$), and the Lind–Mehlum test does not reject ($p = 0.216$). The placebo outcome thus supports the sector-specificity of the (heterogeneous) effects we estimate on biofuel capacity.

5.3. Sensitivity to the oil-price control

Table A.2 (Appendix A) re-estimates the main quadratic specification under four alternative choices of oil-price control, which is the single most consequential modelling choice in this literature. The substantive

Table 5
Placebo identification: alternative treatment components and alternative outcome.

Outcome:	(1) Main (reproduced) Biofuel cap.	(2) Placebo: Pollution tax Biofuel cap.	(3) Placebo: Resource tax Biofuel cap.	(4) Placebo: Patents Y02W Patents Y02W
Lagged dependent variable	0.729***	0.641***	0.730***	0.066
	(0.078)	(0.051)	(0.080)	(0.078)
$\ln(\text{Energy tax})_{i,t-1}$	−0.087	−0.067	−0.176	0.086
	(0.175)	(0.140)	(0.209)	(0.237)
Linear treatment centred	−0.498	−0.099	−0.019	0.232
	(0.313)	(0.063)	(0.031)	(0.354)
Quadratic treatment centred	−0.165	0.074***	−0.015	0.718
	(0.823)	(0.023)	(0.015)	(0.684)
Country fixed effects	Yes	Yes	Yes	Yes
Cluster-robust SE (countries)	27	27	27	27
Sample window	2010–23	2010–23	2010–23	2010–21
Observations	351	351	351	297
Within R^2	0.515	0.576	0.511	0.010
Lind–Mehlum p-value[a]	trivial fail	0.007	0.176	0.216

Notes. Cluster-robust standard errors in parentheses, clustered at the country level. Column 1 reproduces the main quadratic LSDV specification of Table 2, column 3. Columns 2 and 3 substitute the lagged log of pollution-tax revenue (Eurostat `env_ac_taxind2`, pollution component) or resource-tax revenue (resource component) for the env. tax/GDP treatment, keeping the energy-tax control and the rest of the specification identical. Column 4 substitutes the log of Y02W recycling-and-secondary-raw-materials patents (Eurostat `cei_cie020`) for the biofuel outcome; the patents series is available only through 2021, hence the shorter sample.

[a] The Lind–Mehlum test (a) returns "trivial failure to reject" in column 1 because the implied extremum lies outside the observed support; (b) rejects in column 2, indicating that pollution taxes exhibit a statistically significant non-linear association with biofuel capacity that the main specification does not detect for the total env. tax/GDP; (c) does not reject in columns 3 and 4. We discuss the policy interpretation of the pollution-tax result in Section 6: the most plausible reading is that country-fixed pollution-tax design captures heterogeneity in environmental policy stringency that is not absorbed by the country fixed effect alone — a confound that strengthens, rather than weakens, the case for the heterogeneity-by-producer-scale analysis of Table 3.

question is whether the conclusion "the data do not support a U-shape" depends on the specific choice of oil-price control.

The four columns differ exclusively in the variable used to control for the world price of fossil fuels. Column 1 uses the OPEC Reference Basket, our preferred control, which is identical for every EU member state in any given year and is unambiguously exogenous to EU national tax policy. Columns 2 and 3 use the country-level retail diesel price from the Eurostat Weekly Oil Bulletin, respectively including and excluding national taxes. Column 4 omits the oil control entirely.

Under the OPEC control (column 1) and under no oil control (column 4), the quadratic coefficient on the treatment is negative (−0.160 and −0.165 respectively). Under either of the country-level diesel controls (columns 2 and 3), the quadratic coefficient turns positive (+0.936 and +0.952). Neither of the positive coefficients reaches conventional statistical significance—the cluster-robust p-values are 0.195 and 0.179 respectively—but the sign reversal across controls is itself diagnostic. We discuss the interpretation in Section 6.3: the country-level diesel price is mechanically correlated with the energy-tax component of the treatment because national excises on diesel enter the `env_ac_taxind2` aggregate from which the treatment is constructed. Conditioning on the diesel price therefore amounts to "controlling for a downstream consequence" of the policy (in the sense of Angrist and Pischke, 2009, p. 64), which can flip the sign of the partial-derivative estimate without necessarily inflating standard errors enough to flag the problem.

Therefore, under the genuinely exogenous OPEC control, the U-shape hypothesis is not supported by the data—a conclusion identical to the headline of Table 2. Under the endogenous diesel control, the data are agnostic between a flat relationship and a slightly U-shaped one, with neither hypothesis statistically distinguishable from the other. We retain the OPEC control as the main specification because it is the only choice for which exogeneity to EU national tax policy is unambiguous; Fig. 1 visualises the four implied curves on the observed support of the treatment.

5.4. Sample stability and the non-parametric threshold

The final block of robustness checks examines whether the pooled findings are sensitive to sample composition or to the imposed quadratic functional form. Table A.3 (Appendix A) reports the results in three panels.

Panel A of Table A.3, complemented by Fig. 2, reports the implied turning points from re-estimating the main specification 27 times, each time omitting one country. With $G = 27$, a single country whose omission moves the headline turning point substantially would imply that the pooled identification is fragile. The actual leave-one-country-out distribution is striking: 22 of the 27 implied turning points lie at or below 1% of GDP, well outside the observed regressor support, confirming the Lind–Mehlum "trivial failure" verdict of Table 2. The three country-omissions that move the turning point most substantially (Greece, Ireland, Bulgaria) yield turning points between 1.8 and 4.0% of GDP, but the distribution overall does not concentrate around a single identifiable threshold. The Slovenia leave-out generates a turning point of approximately $1.3 \times 10^7\%$, which is the mathematical signature of a quadratic coefficient very close to zero; it is reported in the table for completeness but omitted from Fig. 2 for readability.

Panel B reports the Hansen (1999) threshold-regression test, which estimates a break-point endogenously rather than imposing a quadratic functional form. The point estimate of the threshold is 2.84% of GDP, very close to the sample mean of 3.57%, and the asymptotic confidence interval [2.83, 2.86] is tight. The bootstrap F-statistic for the null of no threshold, however, is 2.47, against critical values of 13.87 (10%), 17.08 (5%), and 23.06 (1%): the test fails to reject by a wide margin, with $p = 0.850$. The parametric quadratic specification and the non-parametric threshold regression therefore deliver the same substantive conclusion: there is no identifiable regime break in the pooled relationship between environmental taxation and biofuel production capacity.

Auxiliary checks. Panel C reports two further sample robustness exercises. The Chow test for a structural break in 2017 returns $F = 2.75$, $p = 0.050$, rejecting coefficient stability at the 5% level. The break is small in magnitude—the pre- and post-2017 quadratic coefficients are both statistically insignificant—and is most plausibly attributed to the anticipation effects associated with the adoption and phased implementation of RED II (formally adopted in December 2018). The break does not restore a quadratic non-linearity, but it does qualify the assumption that the EU regulatory environment is stationary over the full 2010–2023 window; we return to this point in Section 6.3. The extended-sample re-estimation on 2010–2024 (panel C, last row) returns a quadratic coefficient of $\hat{\beta}_{sq} = -0.197$, very close to the main-sample estimate of −0.165, suggesting that the addition of the preliminary 2024 OPEC observation does not materially alter the pooled finding.

5.5. Additional robustness checks

We report four additional families of robustness checks, summarised in Table 6.

Alternative lag structures. Because biofuel capacity is capital-intensive and slow to adjust, the tax effect may materialise after more than one year. Replacing the one-year lag with a two- or a three-year lag leaves the U-shape undetected (the Lind–Mehlum test trivially fails in both

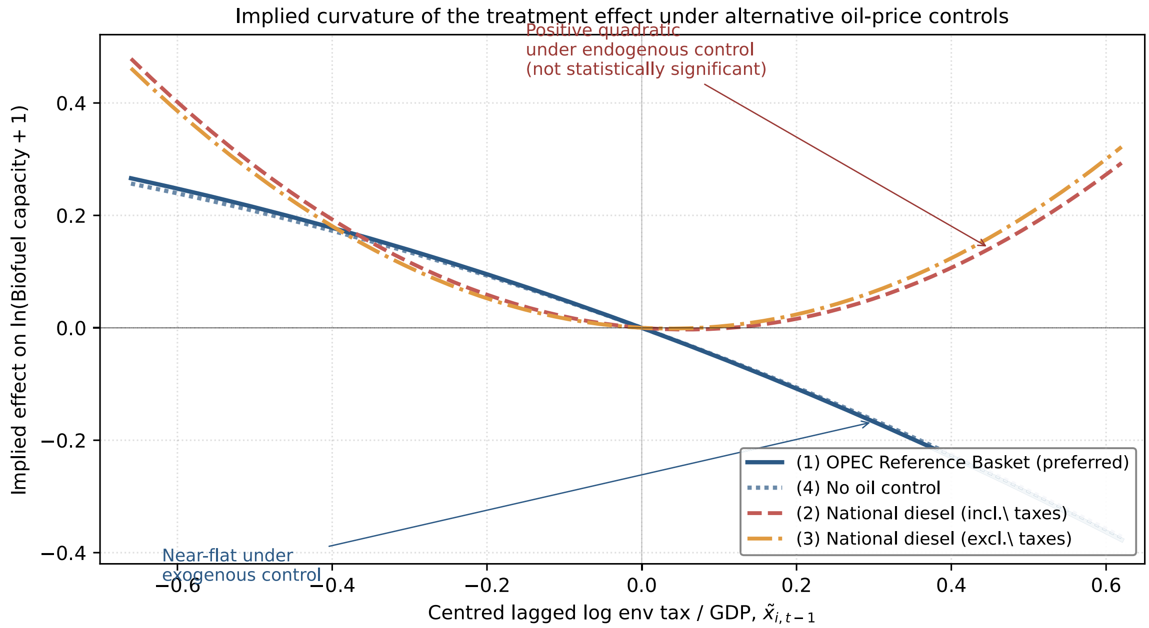

Fig. 1. Implied curvature of the treatment effect under alternative oil-price controls.

Notes. The figure plots the fitted quadratic $\hat{\beta}_2\tilde{x} + \hat{\beta}_3\tilde{x}^2$ on the observed range of the centred treatment $\tilde{x}_{i,t-1} \in [-0.66, +0.62]$, under each of the four oil-price-control choices summarised in Table A.2: (i) OPEC Reference Basket; (ii) national diesel including taxes; (iii) national diesel excluding taxes; (iv) no oil control. The curve under the OPEC control (solid line) overlaps closely with the no-control curve and lies near zero across the entire support. The two diesel-based controls (dashed lines) generate a positive quadratic, although the implied U-shape is not statistically significant at conventional levels (see Table A.2, note a). The sensitivity is direct evidence that the country-level diesel price is an endogenous control: it absorbs variation that the genuinely exogenous OPEC basket does not. The unit on the vertical axis is the percentage-point change in $\ln(\text{biofuel capacity} + 1)$ associated with a one-unit change in $\tilde{x}_{i,t-1}$.

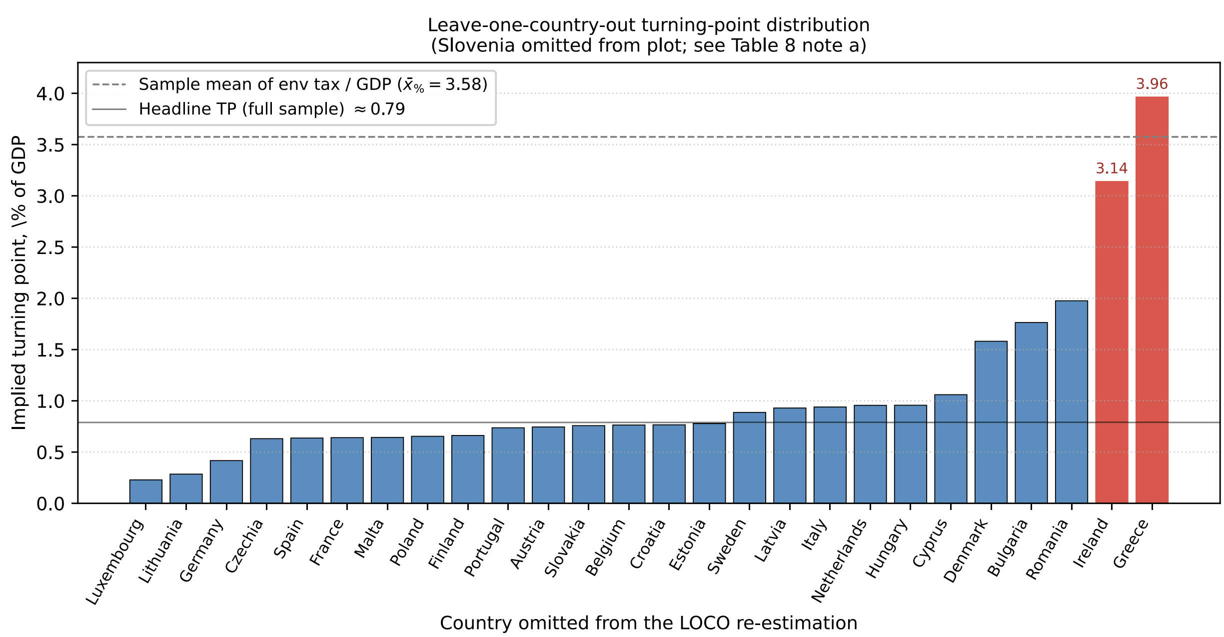

Fig. 2. Leave-one-country-out turning-point distribution.

Notes. Each bar reports the implied turning point as percentage of GDP from re-estimating the main quadratic specification (Eq. 1) after omitting one country at a time. Countries are ordered left-to-right by the magnitude of the implied turning point. The Slovenia LOCO is omitted from the plot for readability (its value of $1.3 \times 10^7\%$ would distort the scale); a discussion of that data point appears in Table A.3, note (a). The vertical dashed line marks the mean of the observed env. tax/GDP ratio in the sample ($\bar{x}_{\%} = 3.57$). The visual impression is twofold: a large mass of LOCO turning points lies near or below the boundary of the observed regressor support, and three countries (Greece, Ireland, Bulgaria) generate turning points more than three percentage points away from the headline. Both features are consistent with the Lind–Mehlum "trivial failure" verdict and motivate the heterogeneity analysis in Section 4.

Table 6
Additional robustness checks (summary).

Check	Statistic	Reading
Two-year lag (quadratic)	Lind–Mehlum: trivial fail	No U-shape
Three-year lag (quadratic)	Lind–Mehlum: trivial fail	No U-shape
Distributed lag L_0–L_3 (cumulative)	$\sum\hat{\beta} = -0.709$ ($p = 0.036$)	Slow negative effect
Two-way (country × year) FE	year dummies $F = 1.79$ ($p = 0.10$); size interactions $F = 3.08$ ($p = 0.033$)	Heterogeneity survives
Classification: feedstock (cereals)	$F = 2.79$ ($p = 0.047$)	Heterogeneity robust
Classification: agricultural GVA	$F = 2.58$ ($p = 0.061$)	Heterogeneity robust
Classification: early window 2010–12	$F = 2.79$ ($p = 0.047$)	Heterogeneity robust
Disaggregated treatment: pollution	Lind–Mehlum $p = 0.007$	Stringency confound
Disaggregated treatment: CO_2 / resource	$p = 0.42$ / $p = 0.18$	No curvature
Kao cointegration test	DF t $p = 0.016$; MDF t $p = 0.034$	Cointegration not rejected
DOLS-type FE long-run tax effect	-0.36 ($p = 0.76$)	Insignificant long run
Pedroni/Westerlund/ FMOLS/PMG	not estimable at $T = 14$	reported as a limitation

Notes. Baseline heterogeneity joint test: $F = 3.12$ ($p = 0.032$). All entries are produced by the unified replication file `do_fileR1.do` (log `logfileR1.log`). "Trivial fail" denotes a Lind–Mehlum test whose implied extremum lies outside the observed data support.

cases). In a distributed-lag model that enters the treatment contemporaneously and at lags one to three, no single horizon shows curvature, but the cumulative four-period effect is negative and significant ($\sum_{h=0}^{3}\hat{\beta}_h = -0.709$, $t = -2.21$, $p = 0.036$; implied long-run semi-elasticity ≈ -3.9 given $\hat{\rho} \approx 0.82$). The timing intuition is thus borne out: the fiscal signal acts cumulatively over three years, as a modest contraction, rather than through a single-year threshold.

Common EU-level shocks. Adding year fixed effects (two-way FE) leaves the block of year dummies jointly insignificant ($F(12, 26) = 1.79$, $p = 0.10$) once the dynamic structure and the OPEC control are present, and—crucially—the producer-scale heterogeneity is unchanged ($F(4, 26) = 3.08$, $p = 0.033$, against the baseline $F = 3.12$). The headline results are therefore not driven by EU-wide omitted factors.

Alternative producer classifications. Because the baseline terciles use the estimation-period outcome, we re-classify member states using three criteria independent of that outcome: feedstock availability (mean cereal area 2010–2012), agricultural structure (mean agricultural gross value added), and a predetermined early-window (2010–2012) capacity measure. The joint interaction test remains significant, or marginally so, under all three ($F = 2.79$, $p = 0.047$; $F = 2.58$, $p = 0.061$; $F = 2.79$, $p = 0.047$), so the heterogeneity is not an artefact of an outcome-contaminated grouping.

Disaggregated tax components and long-run estimates. Re-estimating the main quadratic model with each tax component as the treatment, only the pollution component shows curvature (Lind–Mehlum $p = 0.007$); the CO_2 ($p = 0.42$) and resource ($p = 0.18$) components do not, reinforcing the reading (Section 5.2) that the aggregate ratio proxies unobserved stringency. Finally, the Kao panel cointegration test rejects the null of no cointegration (Dickey–Fuller t: $p = 0.016$; modified Dickey–Fuller t: $p = 0.034$) and a DOLS-type fixed-effects regression yields a small, insignificant long-run tax coefficient (-0.36, $p = 0.76$). We are transparent that the Pedroni and Westerlund tests and the group-mean FMOLS and Pooled Mean Group estimators do not deliver reliable estimates at $T = 14$; these long-run procedures require a much longer time dimension, which is why the large-N, small-T dynamic fixed-effects/System-GMM framework remains our preferred estimator.

6. Discussion

The empirical evidence presented in Sections 4 and 5 supports three connected findings: (i) the pooled relationship between environmental taxation and biofuel production capacity is statistically flat, (ii) the data generating process is dominated by persistence ($\hat{\rho} \approx 0.73$–0.96 across estimators), and (iii) producer scale generates heterogeneity that the pooled specification masks. These three findings are not self-interpreting. To see why, this section is organised around three questions: what mechanisms can explain the heterogeneity that pooling conceals, what policy implications follow for the design of environmental fiscal instruments, and what methodological lessons the robustness results convey for interpreting the estimates.

6.1. Why does producer scale matter?

Three mechanisms can plausibly account for the heterogeneity and then rationalise the directionally opposite responses of small/medium and large biofuel producers to environmental taxation.

While these mechanisms are likely overlapping and cannot be formally disentangled with the available data, the empirical patterns are strongly suggestive of these underlying dynamics.

The Large producer group—Germany, France, Italy, Spain, Poland, the Netherlands, Belgium, and to a lesser extent Austria and Greece—comprises the EU member states with the most developed biofuel refining infrastructure and, at the same time, the most export-exposed producers on the European intra-EU market. Their biofuels compete with imports from the United States, Brazil, and South-East Asia at prices that are, after accounting for shipping and trade-policy adjustments, broadly comparable.[3] An environmental tax that raises the cost of energy and feedstock inputs translates more mechanically into a cost-push shock for these producers, because the output price is anchored by international competition and cannot absorb the rise. The marginal pre-tax-cost producer drops out of production at the high end of the tax distribution, generating the inverted-U pattern documented in Table 3, column 4 ($\hat{\beta}_{sq} = -3.53$, Lind–Mehlum $p = 0.073$). This interpretation is consistent with the carbon-leakage and competitiveness concerns documented in the EU ETS literature (Martin et al., 2014); we view biofuel refining as an analogous case of energy-intensive trade-exposed production.

A second mechanism concerns the technological structure of biofuel production. Large-scale biodiesel and bioethanol refineries are highly capital-intensive: typical plant lifetimes are 15–25 years, and brownfield expansions require multi-year investment commitments. In the dynamic-FE framework of Section 4, this capital-intensity is captured by the high lagged-DV coefficient ($\hat{\rho} \approx 0.73$ under LSDV; $\hat{\rho} \approx 0.96$ under System-GMM), and the implied half-life of any shock is approximately $\ln(0.5)/\ln(0.73) \approx 2.2$ years for the conservative LSDV estimate, and an order of magnitude longer under GMM. Crucially, the adjustment asymmetry runs against the revenue-channel argument of Timilsina et al. (2011): at the high end of the tax distribution, the revenue effect is in principle the strongest, but the very producers that should benefit from it are those whose installed capacity is the hardest to adjust quickly. Large producers are therefore short-run cost-takers and long-run capacity-adjusters, whereas small/medium producers—with less sunk capital and more flexible operations—can absorb high tax intensities by reallocating production or by progressively exiting marginal lines without incurring large adjustment costs. This is a textbook case of heterogeneity along

[3] The OECD-FAO Agricultural Outlook documents this competitive structure: see OECD (2019) for the tax-policy framing and European Environment Agency (2023) for the European-level energy-balance evidence. The political-economy implications of the resulting margin pressure are surveyed in Leonard et al. (2021).

the capital-vintage dimension (Wooldridge, 2010). The slow-adjustment reading is corroborated by a distributed-lag specification (Section 5): no single horizon (one, two or three years) shows curvature, but the *cumulative* effect of the tax-to-GDP ratio over a three-year window is a modest and significant contraction ($\sum_{h=0}^{3} \hat{\beta}_h = -0.71$, $p = 0.036$), consistent with the cost-pass-through channel operating gradually rather than through a non-linear threshold.

A third, more speculative, mechanism concerns the differential exposure of producers to the policy environment surrounding the tax. The revenue-recycling argument central to Timilsina et al. (2011) requires that environmental tax revenue be redirected towards biofuel-supporting infrastructure (refinery subsidies, blending mandates, R&D platforms). The probability that this redirection materially benefits a given producer is a function of (a) the political weight of the biofuel industry in the national policy mix—larger in countries with a substantial agricultural feedstock base, and (b) the flexibility of the producer to capture small-scale incentives, such as feed-in tariffs and feedstock subsidies, which favour small/medium operations. The combination is consistent with the observed pattern: small/medium producers show point estimates suggestive of a (statistically weak) positive non-linearity at high tax intensities (Table 3, columns 2 and 3), consistent with revenue recycling reaching them in proportion to their reduced absolute exposure to the cost shock.

The placebo result on pollution taxes (Table 5, column 2, Lind–Mehlum $p = 0.007$) is initially puzzling: the cross-section variation in pollution-tax composition—which is much larger across EU member states than in energy taxation, and which is not fully absorbed by the country fixed effect because it contains a slow-moving stringency trend—generates statistical significance on what should be an inert regressor. We interpret this as evidence of an unobserved country-level policy stringency factor that operates through multiple channels simultaneously: countries with stricter pollution taxation also tend to have stricter renewable-energy mandates and larger subsidy envelopes for green technology. Under this reading, the pollution-tax result does not falsify the heterogeneity finding; it reinforces the conclusion that aggregate environmental-tax indicators are a noisy proxy for the underlying policy bundle, and that disaggregating by producer scale (rather than by tax category) is the empirically more informative slicing. The agnostic policy implication is that the level of *any* aggregate tax/GDP measure is less informative than the structural composition of the policy mix—a conclusion broadly consistent with the differentiated energy-tax design recommendations of OECD (2019) and with the political-economy considerations of Goulder and Schein (2013).

6.2. Policy implications

The heterogeneity documented in Section 4 has direct implications for the post-RED II policy framework, which assigns to member-state fiscal policy the task of mobilising private capital for the renewable-fuel target.

Before drawing implications we grade the evidence into three tiers, and calibrate the policy language accordingly. *Robust findings*: the pooled relationship is flat and biofuel capacity is highly persistent. *Suggestive evidence*: the rejection of coefficient homogeneity across producer terciles ($F = 3.12$, $p = 0.032$), which is statistically significant and survives alternative classifications and year fixed effects, but rests on a short panel. *Policy hypothesis*: the inverted-U for large producers, which is only marginally significant (Lind–Mehlum $p = 0.073$) and which we present as a conjecture to be tested with richer data rather than as an established regularity. The case for differentiated fiscal design below is therefore a conditional implication that *would* follow *if* the heterogeneity is confirmed, not a firm recommendation.

A non-trivial strand of the prior literature has interpreted U-shaped panel estimates as evidence in favour of an "aggressive front-loading" of environmental taxation: crossing the fiscal-capacity threshold quickly, the argument runs, would unlock the revenue-recycling channel and produce a net positive effect on renewable-energy outcomes (see for example Ulucak and Danish Kassouri, 2020; Zaghdoudi and Maktouf, 2017; Bilan et al., 2022). Our results suggest that this policy prescription, when read as a uniform recommendation across producers, is not supported by the EU-27 data over the 2010–2023 window. The pooled non-linear effect is statistically indistinguishable from zero; the non-parametric Hansen (1999) threshold test fails to identify a regime break (bootstrap $p = 0.850$); and the implied turning point under delta-method inference is unbounded above. Aggregate environmental-tax intensification, applied uniformly, is therefore not an empirically validated policy lever for biofuel-capacity expansion in this institutional setting.

The heterogeneity result, however, points to a constructive alternative. The opposite-signed responses of large and small/medium producers imply that a single rate is, by construction, sub-optimal: a rate calibrated to the revenue-recycling needs of one group will impose excessive cost-push pressure on the other. Three design principles follow directly:

(1) *Scale-dependent rate structures.* A combination of energy-tax revenue with scale-conditional rebates—for instance, lower effective rates on energy-intensive trade-exposed (EITE) biofuel refining, paired with full rate exposure for non-EITE producers—would preserve the revenue channel for the small/medium group while protecting the competitive position of the large producers. The structure parallels the EU ETS free-allocation mechanism for EITE sectors (Martin et al., 2014), and the differentiated-rate framework discussed by Bovenberg and de Mooij (1994) for the interaction between environmental and pre-existing distortionary taxation.

(2) *Hypothecation of energy-tax revenue, conditional on adjustment costs.* The high persistence we estimate ($\hat{\rho} \approx 0.86$ corrected, and up to 0.96 in System-GMM) implies that any revenue-driven capacity expansion will take several years to materialise. Hypothecation arrangements that explicitly time-match revenue flows to capacity build-out—e.g., dedicated multi-year investment programmes financed by ring-fenced energy-tax shares—are a more efficient instrument design than yearly budget cycles that try to "buy capacity in-year". The structural argument is well-developed in Baumol and Oates (1988) and Stiglitz and Stern (2017).

(3) *Reservation of pollution and resource taxes for their own externality bases.* The placebo evidence in Table 5 indicates that pollution and resource taxes do not operate as effective biofuel-promotion instruments: their correlation with capacity is statistically detectable but substantively driven by the underlying policy-stringency confound. Policy design that conflates the environmental-tax aggregate with "renewable-energy promotion policy" is therefore likely to misallocate fiscal effort. The recommendation, again consistent with OECD (2019), is that each tax instrument should be calibrated to its specific externality base, with biofuel-promotion objectives pursued through targeted subsidies and renewable-fuel mandates rather than through general environmental-tax intensification.

Although a fully structural welfare evaluation is beyond the scope of this paper, the order of magnitude estimates are non-trivial. Within the large-producer group, the estimated linear elasticity of biofuel capacity to the centred treatment at the sample mean is $\hat{\beta}_{\text{lin}}^{\text{Large}} \approx -0.37$ (statistically indistinguishable from zero, but indicative); a one-standard-deviation increase in the lagged log env-tax/GDP ratio (about 0.20, corresponding to roughly +0.7 percentage points in env-tax/GDP) would reduce capacity by approximately 7–8% in this group. For the medium-producer group, the same shock would slightly increase capacity, conditional on remaining below the within-group implied turning point of $\approx$ 4.5% of GDP. The aggregate net effect depends on the sample weights of the two groups; at the simple cross-sectional average, the negative and positive effects approximately offset—which is exactly the pooled null we estimate.

6.3. Methodological readings of the robustness evidence

Three pieces of the robustness battery deserve dedicated comment because they are diagnostic of the limits of identification rather than of the substantive conclusions.

The first-difference IV estimator of Table 4, column 3, returns standard errors that are an order of magnitude larger than those of the LSDV, Difference-GMM and System-GMM counterparts. With a single instrument ($Y_{i,t-2}$) and a true persistence parameter close to unity (System-GMM $\hat{\rho} = 0.96$), this is the textbook weak-instrument case: the second-lag level is a poor instrument for the first-differenced lag of a near-random-walk regressor. The estimator's point estimates should be read as essentially uninformative; we report the column primarily to document that we have considered the standard alternative dynamic-panel identification strategy. System-GMM, which exploits a larger set of moment conditions and pools the level and the difference equations, is the appropriate estimator for this data structure and is our preferred dynamic benchmark.

Moreover, the choice of oil-price control is consequential, though the implied U-shape under the endogenous control is not statistically significant. Table A.2 (Appendix A) documents that the sign of the quadratic coefficient on the treatment switches from negative under the genuinely exogenous OPEC Reference Basket control (columns 1 and 4) to positive under the country-level diesel-price control (columns 2 and 3). However, neither of the positive coefficients is statistically significant ($p = 0.195$ and $p = 0.179$ respectively). What changes is therefore the sign, *not* the statistical case for a U-shape: under the OPEC control the data are agnostic between a flat and a slightly inverted-U pattern; under the diesel control they become agnostic between a flat and a slightly U-shaped pattern. The asymmetry is the diagnostic of interest, and the implication is the one Angrist and Pischke (2009) make in their treatment of bad controls: conditioning on a downstream consequence of the treatment can introduce systematic bias in either direction without necessarily inflating standard errors enough to flag the problem. We retain the OPEC control in the main specification because it is the only choice for which exogeneity to EU national tax policy is unambiguous.

Lastly, the 2017 Chow break is marginally significant. The Chow test is reported in Table A.3, panel C (Appendix A), and it rejects coefficient stability around the 2017 boundary at the 5% level ($F = 2.75$, $p = 0.050$). The break is small in magnitude—the pre- and post-2017 quadratic coefficients are both statistically insignificant—and is most plausibly attributed to the policy anticipation effects associated with the announcement and implementation of RED II (formally adopted in December 2018 with a phased implementation starting in 2017). The Chow rejection does not restore a quadratic non-linearity; it does, however, suggest that the EU regulatory environment over the 2010–2023 window is not fully stationary, and that any future panel evaluation of these policies should explicitly accommodate the RED-to-RED II regime transition. In our setting the heterogeneity-by-producer-scale result of Table 3 survives the inclusion of post-2017 interactions (not tabulated; the joint F-test on the size-quadratic interactions remains significant at the 5% level within both sub-samples), and we therefore retain the full-sample specification as the headline result.

6.4. Caveats and limits of the evidence

Three limitations qualify the policy reading of our results, and should guide future research.

First, our identification rests on within-country variation in the environmental tax-to-GDP ratio, conditional on lagged production and on the macroeconomic controls. We do not exploit an exogenous policy shock—such as the implementation of a specific national carbon tax of the type used by Andersson (2019) for Sweden—which would allow a cleaner causal interpretation. The Pesaran (2004) cross-section-dependence test reported in Table A.1 (Appendix A) also signals the presence of common EU shocks (RED implementation, the 2022 energy crisis). We address these directly by adding year fixed effects (two-way FE): the year dummies are not jointly significant once the dynamic structure and the OPEC control are present, and the heterogeneity result is unchanged ($F = 3.08$, $p = 0.033$); Section 5 also reports Driscoll–Kraay HAC inference as the cross-section-dependence-robust alternative. Two further sources of endogeneity should be acknowledged. Tax levels are themselves co-determined by country conditions (Sineviciene and Miliauskaite, 2025), and the fossil-fuel price environment—oil *and* natural gas—plausibly affects both tax-setting and investment incentives (Baghirzade and Kosormyhin, 2025); a harmonised annual natural-gas price series consistent with our panel was not available, so oil prices are only a partial proxy for that environment. For all these reasons we interpret our estimates as conditional associations, not causal effects.

Second, the producer-scale tercile classification is itself a function of the realised distribution of biofuel capacity over 2010–2023, and could be regarded as endogenous to the policy environment of that window. Because the panel begins in 2010, a genuine pre-sample (2005–2009) window is unavailable; we therefore re-classify member states using three criteria that are independent of the contemporaneous outcome—feedstock availability (mean cereal area, 2010–2012), agricultural structure (mean agricultural gross value added), and a predetermined early-window classification (mean capacity, 2010–2012). The joint F-test on the size×treatment interactions remains significant, or marginally so, under all three (feedstock $F = 2.79$, $p = 0.047$; agricultural GVA $F = 2.58$, $p = 0.061$; early-window $F = 2.79$, $p = 0.047$), against a baseline of $F = 3.12$ ($p = 0.032$), so the heterogeneity is not an artefact of an outcome-contaminated grouping (Section 5).

Third, the time dimension of the panel is short ($T = 14$). The Nickell (1981) bias correction and the System-GMM estimator mitigate but do not eliminate the resulting inferential limitations, particularly for the persistence parameter. The convergent evidence from cluster-robust, Driscoll–Kraay, GMM and wild-bootstrap inference on the headline pooled null is reassuring, but future replication on longer panels—particularly post-2030, when the RED II target binds—would be informative. The short time dimension also constrains cointegration analysis: while the Kao test rejects the null of no cointegration and a DOLS-type fixed-effects regression yields a small, insignificant long-run tax coefficient (-0.36, $p = 0.76$), the Pedroni and Westerlund tests and the group-mean FMOLS and Pooled Mean Group estimators do not deliver reliable estimates at $T = 14$ (Section 5). These long-run procedures are designed for panels with a much longer time dimension, which is why the large-N, small-T dynamic fixed-effects/System-GMM framework remains our preferred estimator.

7. Conclusion

This paper has examined the relationship between environmental taxation and biofuel production capacity in the EU-27 over the 2010–2023 window, the period during which the RED and RED II directives have framed national renewable-fuel policy. Three empirical findings emerge. First, in pooled dynamic fixed-effects specifications the relationship is statistically flat: no significant linear effect, no significant quadratic non-linearity, no non-parametric threshold. A broad battery of robustness checks—cluster-robust LSDV, Driscoll–Kraay HAC, Anderson–Hsiao first-difference IV (whose standard errors are uninformative under the documented persistence), Difference- and System-GMM, wild-cluster bootstrap-t inference, two- and three-year lags, a distributed-lag specification, and two-way fixed effects—delivers the same qualitative conclusion; we stress that these are robustness checks rather than distinct identification strategies. Second, biofuel production is highly persistent: the Nickell (1981)-corrected lagged-DV coefficient is approximately 0.86 and the System-GMM estimate is 0.96, implying long adjustment lags that constrain the effectiveness of any annual fiscal lever. Third, the pooled null masks substantial heterogeneity by producer scale: large biofuel producers exhibit a marginally significant inverted-U response consistent with cost-push margin erosion, while small and medium producers show point estimates that run in the opposite direction. A joint F-test on the size-interacted quadratic terms rejects homogeneity at the 5% level.

The policy implication is that aggregate environmental-tax intensification, applied uniformly across the EU member states, is not an empirically validated lever for biofuel-capacity expansion in this institutional setting. Scale-dependent rate structures, careful hypothecation of energy-tax revenue with explicit attention to the adjustment dynamics of capital-intensive biofuel refining, and the separation of pollution-tax and resource-tax design from renewable-energy promotion objectives, all follow as design implications of the empirical evidence.

The methodological lesson is more general. In a panel where producers are heterogeneous along an economically meaningful and observable dimension—here, production scale—imposing coefficient homogeneity in the policy variable risks producing aggregation null results that misrepresent the underlying response structure. Disaggregating along that dimension before drawing policy conclusions is, in this literature, a low-cost methodological discipline with high informational return. The framework can be transposed to other settings in which environmental fiscal policy interacts with heterogeneous productive structures—carbon taxation across manufacturing sub-sectors, water-extraction levies across agricultural producers, congestion pricing across urban areas—where the same aggregation hazard is likely to apply.

Data and code availability

The empirical analysis relies on publicly available data from two sources: Eurostat (https://ec.europa.eu/eurostat/data/database) and the OPEC Annual Statistical Bulletin (https://www.opec.org/opec_web/en/publications/202.htm). The specific extracts used in this paper, the harmonised analysis dataset, and the full suite of Stata do-files that produce every regression, table and figure in the paper are available from the authors upon reasonable request. We are committed to depositing the full replication package in a public repository (Zenodo or the Mendeley Data platform) prior to publication.

No new primary data were collected for this study. All variables used in the analysis are listed with their Eurostat codes in Appendix.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used a large language model assistant for line-editing, formatting suggestions, and verification of output transcription. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Funding

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

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Additional tables

Table A.1
Panel diagnostics on the main-specification residuals.

Test	Null hypothesis	Statistic	p-value	Implication
A. Stationarity of treatment and outcome				
Im et al. (2003), ln(Biofuel cap.)	Unit root in all panels	−2.171	0.015	Stationary
Im et al. (2003), ln(Env. tax / GDP)	Unit root in all panels	0.332	0.630	Marginal[a]
Choi (2001)-ADF, ln(Biofuel cap.), Z		—	0.841	Stationary in subset
Choi (2001)-ADF, ln(Energy tax), Z		—	0.741	ditto
Pesaran (2007) CIPS, ln(Biofuel cap.), Z	Unit root (2nd gen., CD-robust)	2.138	0.984	Unit root not rejected[d]
Pesaran (2007) CIPS, ln(Env. tax / GDP), Z	Unit root (2nd gen., CD-robust)	−0.481	0.315	Unit root not rejected
B. Cross-sectional dependence and serial correlation				
Pesaran (2004) CD (on main residuals)	Cross-section independence	3.456	0.001	CD present[b]
Wooldridge (2010) AR(1) test	No AR(1) in panel errors	14,005.6	0.000	Strong AR(1)
C. Multicollinearity (variance inflation factors)				
Mean VIF, centred quadratic spec.	—	1.14	—	Acceptable
Mean VIF, uncentred quadratic spec.	—	22.20	—	Severe; mean-centring required
D. Hausman test, fixed- vs random-effects				
Contemporaneous treatment, $\chi^2(6)$	FE = RE	3.10	0.796	Either consistent[c]
Lagged treatment, $\chi^2(6)$	FE = RE	3.45	0.751	ditto

[a] The IPS test fails to reject the unit-root null for the treatment variable at conventional significance levels, but the Levin et al. (2002) test rejects on the demeaned series ($p = 0.044$, not tabulated). We treat the variable as stationary in levels for purposes of the dynamic-FE specification.

[b] The Pesaran CD rejection at $p = 0.001$ motivates the inclusion of Driscoll–Kraay HAC standard errors (Table 4, column 2) as the cross-section-dependence-robust inference of choice.

[c] The high p-value on the Hausman test reflects the difference matrix being nearly singular at the panel size involved (rank 6 vs $k = 7$ parameters); we report fixed-effects estimates throughout, since within-variation is the source of identification.

[d] In contrast to the first-generation IPS/ADF tests, the second-generation Pesaran (2007) CIPS test—robust to the cross-sectional dependence documented in panel B—does not reject a unit root in levels for either the outcome or the treatment, which motivates the cointegration analysis reported in Section 5.

Table A.2
Sensitivity to the oil-price control.

Oil-price control:	(1) OPEC Reference Basket	(2) National diesel (incl. taxes)	(3) National diesel (excl. taxes)	(4) No control
$\ln(\text{Biofuel capacity})_{i,t-1}$	0.729***	0.673***	0.671***	0.729***
	(0.079)	(0.088)	(0.087)	(0.078)
$\ln(\text{Energy tax})_{i,t-1}$	−0.112	−0.036	−0.029	−0.087
	(0.206)	(0.151)	(0.142)	(0.175)
$\tilde{x}_{i,t-1}$ (centred lag)	−0.509	−0.107	−0.071	−0.498
	(0.338)	(0.289)	(0.279)	(0.313)
$\tilde{x}^2_{i,t-1}$ (centred lag, squared)	−0.160	0.936	0.952	−0.165
	(0.815)	(0.704)	(0.688)	(0.823)
ln(oil control)	−0.039	0.109	0.105*	—
	(0.113)	(0.094)	(0.058)	
Country fixed effects	Yes	Yes	Yes	Yes
Cluster-robust SE (countries)	27	26	26	27
Observations	351	336	336	351
Within R^2	0.515	0.612	0.613	0.515
Sign of $\hat{\beta}_{\text{sq}}$	negative	positive[a]	positive[a]	negative

Notes. The four columns differ exclusively in the choice of oil-price control. Column 1 uses the OPEC Reference Basket (Annual Statistical Bulletin), the preferred specification in the text. Columns 2 and 3 use the country-level retail diesel price from the Eurostat Weekly Oil Bulletin, respectively including and excluding taxes. Column 4 omits the oil control entirely. The mechanical correlation between national diesel prices and the energy-tax component of the treatment makes the country-level control a "conditioning on a downstream consequence" problem in the sense of Angrist and Pischke (2009, p. 64); we therefore retain the OPEC control in the main specifications and report this table as evidence that the choice is consequential.

[a] In columns 2 and 3 the quadratic coefficient turns positive but remains statistically insignificant ($p = 0.195$ and $p = 0.179$ respectively). The sign reversal, rather than the statistical significance, is the diagnostic of interest: it documents that the country-level diesel control is informative about something that the genuinely exogenous OPEC control is not.

Table A.3
Sample stability, non-parametric threshold, and auxiliary robustness.

	Value	Interpretation
A. Leave-one-country-out (LOCO) on the quadratic specification		
LOCO turning-point statistics (% of GDP):		
Median (excluding the maximum outlier)	0.76	Concentrated near the bound
Range across 27 LOCO runs	13,432,264	Driven by Slovenia[a]
Range excluding top three countries	3.74	Range exceeds delta-method 95% CI
B. Hansen (1999) single-threshold test, threshold variable $T_{i,t-1}$		
Estimated threshold (env. tax/GDP, %)	2.837	Just below sample mean
95% CI for threshold	[2.826, 2.856]	Tight asymptotic CI
Bootstrap F-statistic, H_0: no threshold	2.47	—
Bootstrap p-value (300 reps)	0.850	Fail to reject H_0
10% critical value	13.867	—
5% critical value	17.081	—
1% critical value	23.056	—
C. Auxiliary specification robustness		
Chow break at 2017 (post-Paris), F on 4 interactions	2.75	—
p-value	0.050	Marginally significant[b]
Extended sample 2010–2024, $\hat{\beta}_{\text{sq}}$	−0.197	Sign-stable, magnitude preserved

Notes. Panel A reports leave-one-country-out estimates of the turning point implied by the main quadratic specification. Panel B reports the Hansen (1999) threshold-regression test with the lagged env. tax/GDP as both the regime variable and the threshold variable, with 300 bootstrap replications and a 15% trimming parameter; the bootstrap critical values are dramatically larger than the observed F-statistic. Panel C reports two auxiliary specification robustness checks: a Chow test for structural break in 2017, and the extended-sample re-estimation with the 2024 observation included.

[a] The Slovenia LOCO turning point is an order of magnitude larger than the rest because dropping Slovenia produces a quadratic coefficient very close to zero, which mechanically drives the implied extremum to infinity. The cluster of small LOCO values (Luxembourg, Lithuania, Germany at 0.23, 0.29, 0.42) reflects the same insensitivity — the extremum is consistently estimated outside or near the boundary of the observed regressor support, confirming the Lind–Mehlum "trivial failure" verdict in Table 2.

[b] The Chow test rejects coefficient stability around the 2017 boundary at the 5% level (the implementation deadline for RED II was set in 2018; the 2017 break captures anticipation effects in member-state policy planning). The implied $\hat{\beta}_{\text{sq}}$ is statistically insignificant in both pre- and post-2017 subsamples, so the break does not restore a quadratic non-linearity, but it does suggest that the homogeneity of the response across the RED → RED II transition cannot be imposed without reservation. We discuss the implication in Section 6.

Appendix B. Variable definitions and data sources

This appendix lists the full set of variables used in the empirical analysis, with their definitions, units, transformations applied, and data sources. All variables are at the country–year level for the 27 EU member states observed annually between 2010 and 2024 (with the analysis sample restricted to 2010–2023 as discussed in Section 3). For variables drawn from Eurostat, the official table code is reported; OPEC reference-basket data are taken from the Annual Statistical Bulletin, various editions, 2017–2025.

Table B.1
Variable definitions, transformations and data sources.

Variable	Description	Source	Code
A. Dependent variable			
biofuel_capacity_ktoe	Total liquid biofuels production capacity (thousand tonnes of oil equivalent). Includes biodiesel, bioethanol and other renewable liquid fuels processed for transport use.	Eurostat	`nrg_inf_lbpc`
log_biofuel_prod	Natural logarithm $\ln(\text{biofuel\_capacity\_ktoe} + 1)$. Baseline outcome variable.	Derived	—
B. Treatment variables and fiscal-policy aggregates			
env_tax_total_mio	Total environmental taxes, million EUR. Compulsory levies on bases of environmental relevance (energy, transport, pollution, resources).	Eurostat	`env_ac_tax`
env_tax_energy_mio	Energy taxes: levies on mineral oils, natural gas, coal, electricity.	Eurostat	`env_ac_taxind2`
env_tax_co2_mio	CO_2-related energy taxes (subset of energy taxes).	Eurostat	`env_ac_taxind2`
env_tax_pollution_mio	Pollution taxes: emissions to air/water, ozone-depleting substances, waste management. Used as placebo regressor in Section 5.	Eurostat	`env_ac_taxind2`
env_tax_resource_mio	Resource taxes: water abstraction, mineral and aggregate extraction. Used as placebo regressor.	Eurostat	`env_ac_taxind2`
env_tax_gdp_pct	$(\text{env\_tax\_total\_mio}/\text{gdp\_mio\_eur}) \times 100$. Key non-linear treatment in the quadratic specification.	Derived	—
log_env_tax_gdp	$\ln(\text{env\_tax\_gdp\_pct})$. Mean-centred variant $\tilde{x}_{i,t-1}$ used in the headline regressions.	Derived	—
log_env_tax_energy	$\ln(\text{env\_tax\_energy\_mio} + 1)$; linear fiscal control.	Derived	—
C. Macroeconomic controls			
gdp_mio_eur	Gross domestic product at market prices, current million EUR.	Eurostat	`nama_10_gdp`
gdp_growth_pct	Real GDP growth rate (%, prior-year reference).	Eurostat	`tec00115`
gva_total	Gross value added, total economy.	Eurostat	`nama_10_a10`
gva_agriculture	Gross value added, agriculture/forestry/fishing.	Eurostat	`nama_10_a10`
gva_industry	Gross value added, industry (excl. construction).	Eurostat	`nama_10_a10`
labour_cost_index	Labour cost index, industry/construction/services.	Eurostat	`lc_lci_lev`
energy_intensity	Energy intensity: kilograms of oil equivalent per 1000 EUR (PPS) of GDP.	Eurostat	`nrg_ind_ei`
cereals_prod_kha	Cereals harvested area, thousand hectares. Feedstock-availability proxy for first-generation biofuels.	Eurostat	`apro_cpsh1`
gerd_total_pct_gdp	Gross domestic expenditure on R&D (% of GDP), all sectors.	Eurostat	`rd_e_gerdtot`
gerd_business_pct_gdp	Gross domestic expenditure on R&D (% of GDP), business enterprise sector.	Eurostat	`rd_e_gerdtot`
D. Oil-price controls			
oil_price_global_usd	OPEC Reference Basket (ORB), annual mean, USD per barrel. **Preferred control**: globally exogenous to EU national tax policy.	OPEC ASB[a]	—
log_oil_global	$\ln(\text{oil\_price\_global\_usd} + 1)$; main oil control.	Derived	—
price_gasoil_notax	National diesel (gas oil) price, EUR per 1000 L, excluding taxes.	Eurostat WOB[b]	—
log_diesel_notax	$\ln(\text{price\_gasoil\_notax} + 1)$; robustness control.	Derived	—
price_gasoil_wtax	National diesel price, EUR/1000 L, including taxes. Reported only as robustness in Table A.2.	Eurostat WOB	—
price_eurosuper_95_*	Eurosuper-95 (premium gasoline) prices, EUR/1000 L, with and without tax. Auxiliary; not used in main specifications.	Eurostat WOB	—
E. Environmental investment, trade and other auxiliary			
env_invest_mio	National expenditure on environmental protection (investments), million EUR.	Eurostat	`env_ac_epneis`
env_invest_pct_gdp	Environmental protection investments as percentage of GDP.	Eurostat	`env_ac_epneis`
transfers_subsidies_mio	Current and capital transfers including subsidies for environmental protection.	Eurostat	`env_ac_pepsps`
subsidies_mio	Subsidies (subset of transfers).	Eurostat	`env_ac_pepsps`
biofuel_exports_kt	Biofuel exports, thousand tonnes.	Eurostat	`nrg_ti_bio`
biofuel_imports_kt	Biofuel imports, thousand tonnes.	Eurostat	`nrg_ti_bio`
biofuel_netexp_kt	Net exports = exports − imports.	Derived	—
F. Green innovation (placebo outcome)			
patents_recycling_total	Patent families related to recycling and secondary raw materials (CPC class Y02W: climate-change mitigation technologies for wastewater treatment and waste management). Count. Used as placebo outcome (Table 5, column 4).	Eurostat	`cei_cie020`
patents_recycling_pmhab	Same indicator, scaled per million inhabitants.	Eurostat	`cei_cie020`
log_patents_total	$\ln(\text{patents\_recycling\_total} + 1)$.	Derived	—
G. Panel and producer-size structure			
country_id	Country numeric identifier, 1–27.	Derived	—
size_category	Producer-size tercile based on average 2010–2023 biofuel production capacity: 1 = Small, 2 = Medium, 3 = Large.	Derived	—
small/medium/large_producer	Time-invariant dummies (1/0) for each tercile. Large is the reference category in interaction specifications.	Derived	—

(*continued on next page*)

Table B.1 *(continued)*

Variable	Description	Source	Code
H. Lags, mean-centring and quadratic transformations			
L.X	First lag of variable X computed via Stata's `xtset country_id year`.	Derived	—
$\tilde{x}_{i,t-1}$	Mean-centred lagged log env-tax/GDP: $L.\ln(\text{env\_tax\_gdp\_pct}) - \bar{x}$, with $\bar{x} = 1.274$ on the column-3 estimation sample of Table 2.	Derived	—
$\tilde{x}^2_{i,t-1}$	Square of the above; used to test the U-shape ($\beta_2 < 0$, $\beta_3 > 0$) in the quadratic specification.	Derived	—

[a] OPEC Annual Statistical Bulletin (ASB), various editions 2017–2025. The ORB is a production-weighted average of spot prices for petroleum blends produced by OPEC member countries; introduced in its current methodology on 16 June 2005.

[b] European Commission Weekly Oil Bulletin (WOB), annual averages computed from weekly observations.

All monetary aggregates are expressed in current million EUR unless otherwise stated. Log transformations follow $\ln(X + 1)$ to accommodate the small share of zero-valued observations in the outcome (countries-years with no reported biofuel production capacity), in line with the treatment of Bellemare and Wichman (2020). Eurostat data downloads accessed through the dissemination portal (https://ec.europa.eu/eurostat/data/database). The full data-construction pipeline is available from the authors upon request.

Data availability

Data will be made available on request.

References

Acemoglu, D., Aghion, P., Bursztyn, L., Hemous, D., 2012. The environment and directed technical change. Am. Econ. Rev. 102, 131–166.

Andersson, J.J., 2019. Carbon taxes and CO_2 emissions: Sweden as a case study. Am. Econ. J. Econ. Policy 11, 1–30.

Angrist, J.D., Pischke, J.S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.

Antón, Á.A., 2023. The energy Tax Directive Reform, in: Law in the EU's Circular Energy System. Edward Elgar Publishing, pp. 182–204.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev. Econ. Stud. 58, 277–297. https://doi.org/10.2307/2297968

Baghirzade, M., Kosormyhin, T., 2025. Investigation of impact of oil, gold and natural gas prices on clean energy stock prices. Financ. Mark. Inst. Risks 9, 101–119. https://doi.org/10.61093/fmir.9(3).101-119.2025

Baumol, W.J., Oates, W.E., 1988. The Theory of Environmental Policy, second ed. Cambridge University Press, Cambridge.

Bellemare, M.F., Wichman, C.J., 2020. Elasticities and the inverse hyperbolic sine transformation. Oxf. Bull. Econ. Stat. 82, 50–61.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? Q. J. Econ. 119, 249–275.

Bilan, Y., Samusevych, Y., Lyeonov, S., Strzelec, M., Tenytska, I., 2022. The keys to clean energy technology: impact of environmental taxes on biofuel production and consumption. Energies 15, 9470.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. J. Econom. 87, 115–143.

Bovenberg, A.L., de Mooij, R.A., 1994. Environmental levies and distortionary taxation. Am. Econ. Rev. 84, 1085–1089.

Brambor, T., Clark, W.R., Golder, M., 2006. Understanding interaction models: improving empirical analyses. Polit. Anal. 14, 63–82.

Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50, 317–372.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-based improvements for inference with clustered errors. Rev. Econ. Stat. 90, 414–427.

Chen, X., Chen, Y.E., Chang, C.P., 2019. The effects of environmental regulation and industrial structure on carbon dioxide emission: a non-linear investigation. Environ. Sci. Pollut. Res. 26, 30252–30267.

Cheng, C., Ren, X., Wang, Z., Yan, C., 2019. Heterogeneous impacts of renewable energy and environmental patents on CO_2 emission - evidence from the BRIICS. Sci. Total Environ. 668, 1328–1338. https://doi.org/10.1016/j.scitotenv.2019.02.063

Choi, I., 2001. Unit root tests for panel data. J. Int. Money Finance 20, 249–272.

European Environment Agency, 2023. Biofuels in Europe: Progress and Challenges in the Transport Sector . Briefing. EEA, Copenhagen.

Eurostat, 2013. Environmental Taxes: A Statistical Guide. Manuals and Guidelines. Publications Office of the European Union, Luxembourg. URL: https://doi.org/10.2785/45952. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-13-005.

Goulder, L.H., Schein, A.R., 2013. Carbon taxes versus cap and trade: a critical review. Clim. Policy 13, 459–484.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054. https://doi.org/10.2307/1912775

Hansen, B.E., 1999. Threshold effects in non-dynamic panels: estimation, testing, and inference. J. Econom. 93, 345–368. https://doi.org/10.1016/S0304-4076(99)00025-1

Hassler, J., Krusell, P., Olovsson, C., 2016. Environmental macroeconomics. In: Handbook of Macroeconomics, vol. 2. Elsevier, pp. 1893–2008.

Hoechle, D., 2007. xtscc: regression with driscoll-kraay standard errors. Stata J. 7, 281–312. Riferimento per lo stimatore Driscoll-Kraay (1998) implementato in Stata.

Im, K.S., Pesaran, M.H., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. J. Econom. 115, 53–74.

Jordaan, S.M., Romo-Rabago, E., McLeary, R., Reidy, L., Nazari, J., Herremans, I.M., 2017. The role of energy technology innovation in reducing greenhouse gas emissions: a case study of Canada. Renew. Sustain. Energy Rev. 78, 1397–1409. https://doi.org/10.1016/j.rser.2017.05.162

Leonard, M., Pisani-Ferry, J., Shapiro, J., Tagliapietra, S., Wolff, G.B., 2021. The Geopolitics of the European Green Deal Policy Contribution 04/2021. Bruegel.

Levin, A., Lin, C.F., Chu, C.S.J., 2002. Unit root tests in panel data: asymptotic and finite-sample properties. J. Econom. 108, 1–24.

MacKinnon, J.G., Webb, M.D., 2017. Wild bootstrap inference for wildly different cluster sizes. J. Appl. Econom. 32, 233–254. https://doi.org/10.1002/jae.2508

Mardones, C., Baeza, N., 2018. Economic and environmental effects of a CO_2 tax in Latin American countries. Energy Policy 114, 262–273. https://doi.org/10.1016/j.enpol.2017.12.001

Martin, R., Muûls, M., De Preux, L.B., Wagner, U.J., 2014. Industry compensation under relocation risk: a firm-level analysis of the EU emissions trading scheme. Am. Econ. Rev. 104, 2482–2508.

Morone, P., d'Amato, D., Befort, N., Yilan, G., 2023. The Circular Bioeconomy: Theories and Tools for Economists and Sustainability Scientists. Cambridge University Press.

Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica 49, 1417–1426. https://doi.org/10.2307/1911408

OECD, 2019. Taxing Energy Use 2019: Using Taxes for Climate Action. OECD Publishing, Paris, https://doi.org/10.1787/058ca239-en

Pesaran, M.H., 2004. General diagnostic tests for cross section dependence in panels. Cambridge Working Papers in Economics No. 0435. Faculty of Economics, University of Cambridge.

Pesaran, M.H., 2007. A simple panel unit root test in the presence of cross-section dependence. J. Appl. Econom. 22, 265–312. https://doi.org/10.1002/jae.951

Pigou, A.C., 1920. The Economics of Welfare. Macmillan and Co, London.

Roodman, D., 2009. How to do Xtabond2: an introduction to difference and system GMM in stata. Stata J. 9, 86–136. https://doi.org/10.1177/1536867X0900900106

Sargan, J.D., 1958. The estimation of economic relationships using instrumental variables. Econometrica 26, 393–415. https://doi.org/10.2307/1907619

Sineviciene, L., Miliauskaite, L., 2025. What determines energy tax rates in European Union countries? Probl. Perspect. Manag. 23, 715–727. https://doi.org/10.21511/ppm.23(3).2025.51

Stiglitz, J.E., Stern, N., 2017. Report of the High-Level Commission on Carbon Prices. Technical Report. World Bank, Washington, DC.

Timilsina, G.R., Csordás, S., Mevel, S., 2011. When does a carbon tax on fossil fuels stimulate biofuels? Ecol. Econ. 70, 2400–2415. URL: https://doi.org/10.1016/j.ecolecon.2011.07.022. https://www.sciencedirect.com/science/article/pii/S0921800911002977.

Ulucak, R., Danish Kassouri, Y., 2020. An assessment of the environmental sustainability corridor: investigating the non-linear effects of environmental taxation on CO_2 emissions. Sustain. Dev. 28, 1010–1018. https://doi.org/10.1002/sd.2057

Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. J. Econom. 126, 25–51. https://doi.org/10.1016/j.jeconom.2004.02.005

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data, second ed. MIT Press, Cambridge, MA.

Zaghdoudi, T., Maktouf, S., 2017. Threshold effect in the relationship between environmental taxes and CO_2 emissions: a PSTR specification. Econ. Bull. 37, 2086–2094.