The Paper Cliff: Subsidy Clawback Thresholds and the Limits of Generator Incentives in Renewable Electricity Markets

APEP Autonomous Research* @ai1scl

March 30, 2026

Abstract

Germany's renewable energy law creates a sharp incentive cliff: when wholesale electricity prices go negative for more than a threshold number of consecutive hours, generators lose their entire subsidy for the episode. This threshold was progressively tightened from six hours (pre-2021) to four (2021–2023) to three (2024). Using 15-minute generation data from Fraunhofer ISE covering 288 negative-price episodes across 2019–2024, I test whether generators strategically curtail output to avoid triggering the clawback. I find no evidence of curtailment. Generation near the threshold is weakly *higher*, not lower, and cross-country placebos show that the episode duration distribution in Germany is indistinguishable from countries without clawback rules. The clawback creates a collective-action problem: individual price-taking generators cannot end a system-wide negative-price episode by curtailing alone, rendering the duration-based penalty ineffective as a behavioral instrument.

JEL Codes: Q48, L94, H25

Keywords: renewable energy subsidies, negative electricity prices, bunching, clawback thresholds, curtailment

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 30m).

1. Introduction

In 2024, German wholesale electricity prices went negative for 457 hours—roughly one hour in every nineteen. During these episodes, renewable generators were effectively paying the grid to accept their electricity. Yet Germany's renewable energy law (Erneuerbare-Energien-Gesetz, or EEG) imposes a sharp penalty for persistent negative prices: once an episode exceeds a threshold number of consecutive hours, generators lose their market premium for the entire episode. The threshold has been progressively tightened—from six hours before 2021, to four in 2021–2023, to three from 2024—reflecting policymakers' belief that sharpening the incentive will induce generators to curtail output during oversupply.

This paper tests whether it works. The prediction is straightforward: if the clawback creates effective incentives, generators should curtail output as episodes approach the threshold, producing a "curtailment cliff" in within-episode generation profiles and bunching in the duration distribution just below the cutoff. I test both predictions using 15-minute generation data from the Fraunhofer ISE Energy-Charts API, which provides generation by 21 fuel types across 2019–2024—a level of temporal granularity that prior work on negative electricity prices has not exploited.

The answer is a clear null. Within-episode generation profiles show no decline as episodes approach the threshold. If anything, renewable output is modestly *higher* in the hours immediately before the clawback would trigger (+287 MW, $p = 0.035$), driven by solar generation tracking its diurnal cycle. The episode duration distribution shows some visual bunching at the 4-hour threshold under the 2021–2023 regime, but this pattern is not statistically distinguishable from zero (bunching coefficient $\hat{b} = 10.1$, SE $= 40.5$) and is replicated at the same duration in France and the Netherlands, neither of which has an equivalent clawback rule. A permutation test randomizing the near-threshold indicator across episode-hours yields a p-value of 0.028 for the generation result—but in the "wrong" direction, confirming that what the regression picks up is the natural rise of solar output during daytime hours, not strategic curtailment.

These findings contribute to three literatures. First, I add to the bunching literature pioneered by Saez (2010) and Kleven (2016) by documenting a setting where a sharp notch in incentives produces no detectable behavioral response. Chetty et al. (2011) showed that bunching responses depend on adjustment costs and salience; the electricity market offers a particularly clean test because the threshold is well-known and the affected units are sophisticated firms with real-time market access. The absence of bunching thus reveals something fundamental about market structure, not inattention.

Second, I contribute to the literature on negative electricity prices and renewable energy policy. Nicolosi (2010) documented the emergence of negative prices in Germany under the EEG's priority dispatch rules. Hirth (2013) and Ketterer (2014) analyzed how variable renewables depress wholesale prices, and Paraschiv et al. (2014) traced the merit-order effect. Sensfuß et al. (2008) provided early evidence of the merit-order effect in Germany, while Kyritsis et al. (2017) and Antweiler and Muesgens (2021) examined its long-run evolution. More recently, de Lagarde and Lantz (2022) studied how feed-in tariffs interact with negative prices and Brait and Müller (2020) analyzed implications for renewable support schemes. This paper provides the first micro-level test of whether the EEG's duration-based clawback actually changes generator behavior—the entire rationale for the policy.

Third, the results speak to market design in electricity. Fabra and Reguant (2014) and Reguant (2019) study how firms respond to cost shocks in electricity markets. Borenstein and Bushnell (2015) review the restructured U.S. electricity industry, emphasizing that market power depends on the ability to move prices. My null result is consistent with Ito (2014)'s insight that agents respond to the price schedule they perceive, not necessarily the one policymakers design. In a uniform-price electricity market, individual generators are price-takers during negative-price episodes; curtailing one wind farm does not end the episode for the system. The clawback creates a collective-action problem that individual incentives cannot solve.

The rest of the paper proceeds as follows. Section 2 describes the EEG clawback mechanism and the institutional setting. Section 3 presents the data. Section 4 lays out the empirical strategy. Section 5 presents results, and Section 6 discusses implications.

2. Institutional Background

Germany's Renewable Energy Sources Act (EEG) has undergone multiple revisions since its introduction in 2000, but the core incentive structure for most renewable generators above 400 kW has converged on the *market premium* model: generators sell electricity on the wholesale market and receive a technology-specific premium above the market price. When prices are positive, the premium supplements market revenue. When prices are negative, generators face a choice: continue producing and absorb the negative price (offset by the premium), or curtail.

The clawback rule adds a discontinuity. Under § 51 of the EEG (as amended), if the day-ahead price is negative for N or more *consecutive* hours, generators in certain categories lose their market premium for the *entire episode*—not just the hours beyond the threshold. The premium forfeited depends on the generator's capacity, technology, and contracted premium rate, but for a typical 3 MW onshore wind turbine receiving approximately €60/MWh in market premium, a 6-hour clawback episode implies a loss of roughly €1,080.

The threshold N has been progressively tightened:

- **Pre-2021:** 6 consecutive hours
- **2021–2023 (EEG 2021):** 4 consecutive hours
- **2024–2025 (EEG 2023):** 3 consecutive hours
- **2026:** 2 consecutive hours (scheduled)
- **2027+:** 1 hour (scheduled)

The policy rationale is that tightening the threshold strengthens generators' incentive to curtail during oversupply, thereby reducing the frequency and depth of negative prices. The implicit behavioral model is that generators monitor episode duration in real time and, as the threshold approaches, choose to shut down or reduce output to prevent the episode from reaching the critical duration.

Why curtailment may be difficult. Several features of the electricity market limit individual generators' ability to respond. First, negative-price episodes reflect *system-level* oversupply: total generation exceeds total demand plus export capacity. An individual wind farm curtailing 3 MW does not appreciably change the market-clearing price when system-wide oversupply may exceed 10 GW. Second, wind turbines face physical and contractual constraints on curtailment—grid operators may require continued feed-in for system stability, and curtailment may void warranty or grid-connection agreements. Third, solar generators have limited real-time curtailment capability; inverters can be programmed to respond to price signals, but many smaller installations lack the control infrastructure. Fourth, the market-clearing price is determined in the day-ahead auction, not in real time. By the time an episode is unfolding, generators have already committed their bids. Intra-day adjustments are possible but face transaction costs and liquidity constraints.

3. Data

I use the Fraunhofer ISE Energy-Charts public API, which provides two datasets at high temporal resolution for Germany and 42 European countries.

Prices. Hourly day-ahead electricity prices for the DE-LU bidding zone (Germany-Luxembourg) and four placebo countries (France, Austria, Netherlands, Spain) for 2019–2024. The data originate from the Bundesnetzagentur via SMARD.de and are licensed CC BY 4.0. I observe 52,608 price-hours per country, for a total of 263,040 observations.

Table 1: Summary Statistics: Negative-Price Episodes in Germany

	Pre-2021 (6h)	2021–2023 (4h)	2024 (3h)
Number of episodes	109	89	90
Mean duration (hours)	4.7	5.7	5.1
SD duration	4.5	4.8	3.3
Median duration (hours)	3.0	5.0	5.0
Mean price (EUR/MWh)	-9.5	-8.3	-7.5
Min price (EUR/MWh)	-90.0	-500.0	-135.4
% above clawback threshold	27.5	66.3	76.7
% ending just below threshold	8.3	18.0	12.2

Notes: A negative-price episode is a maximal consecutive run of hours with day-ahead price below zero in the German (DE-LU) bidding zone. The clawback threshold is the minimum episode duration triggering loss of EEG market premium: 6 hours before 2021, 4 hours in 2021–2023, and 3 hours from 2024. "Just below threshold" means duration equals threshold minus one. Data from Fraunhofer ISE Energy-Charts, 2019–2024.

Generation. Electricity generation at 15-minute resolution by 21 fuel types for Germany. The generation data, also from Fraunhofer ISE under CC BY 4.0, provide approximately 35,000 observations per year per fuel type. I aggregate the four 15-minute observations within each hour to mean MW (average power output), then classify renewable generation into three categories: wind (onshore + offshore), solar, and total renewable (adding biomass, hydro run-of-river, and geothermal). All generation figures in the paper are reported in MW (instantaneous power), not MWh (energy).

Episode construction. I define a negative-price episode as a maximal consecutive run of hours with day-ahead price strictly below zero. This yields 288 episodes in Germany across 2019–2024, with a mean duration of 5.1 hours and a maximum of 36 hours. The number of episodes varies by year—from 14 in 2022 (when the energy crisis suppressed negative prices) to 90 in 2024—and by regime, with 109 episodes under the 6-hour threshold (2019–2020), 89 under the 4-hour threshold (2021–2023), and 90 under the 3-hour threshold (2024).

For the placebo analysis, I apply the same episode construction to France (167 episodes), Austria (158), the Netherlands (238), and Spain (46).

4. Empirical Strategy

I pursue two complementary approaches: bunching estimation on the episode duration distribution, and within-episode regression analysis of generation profiles.

4.1 Bunching Analysis

If generators curtail to avoid the clawback, episodes should pile up at durations just below the threshold. Following Kleven (2016), I estimate the counterfactual distribution of episode durations using a polynomial of order k fitted to duration bins outside an exclusion window of ± 2 hours around the threshold. The bunching estimand is:

$$\hat{b} = \frac{\sum_{j=h^*-2}^{h^*}(c_j - c_j^0)}{\bar{c}^0} \tag{1}$$

where c_j is the observed count at duration j, c_j^0 is the polynomial counterfactual, h^* is the threshold, and $\bar{c}^0$ is the average counterfactual count in the bunching region. I estimate $\hat{b}$ separately for each regime and compute standard errors via 500 bootstrap replications of the duration distribution.

4.2 Within-Episode Curtailment Test

The bunching test asks whether episodes end at the "right" duration. The curtailment test asks whether generation drops as episodes approach the threshold. I estimate:

$$\text{Gen}_{eh} = \alpha_e + \gamma_h + \beta \cdot \mathbb{I}[\text{NearThreshold}_{eh}] + \varepsilon_{eh} \tag{2}$$

where e indexes episodes, h indexes hour-of-day, α_e are episode fixed effects (absorbing all episode-level variation including season, weather, and regime), and γ_h are hour-of-day fixed effects (absorbing the diurnal cycle of generation). The key variable $\mathbb{I}[\text{NearThreshold}_{eh}]$ equals one if the episode-hour is zero or one hours from the clawback threshold.

Under the curtailment hypothesis, $\hat{\beta} < 0$: generation should drop as episodes approach the cliff. I cluster standard errors by date to account for correlation across episodes occurring on the same day.

4.3 Cross-Regime Comparison

To test whether the 2021 threshold tightening (6h $\rightarrow$ 4h) amplified curtailment, I compare generation in near-threshold hours across regimes:

$$\text{Gen}_{eh} = \gamma_h + \delta_m + \beta_1 \cdot \mathbb{I}[\text{Near}] + \beta_2 \cdot \mathbb{I}[\text{Post}] + \beta_3 \cdot \mathbb{I}[\text{Near}] \times \mathbb{I}[\text{Post}] + \varepsilon_{eh} \tag{3}$$

where δ_m are month fixed effects, $\mathbb{I}[\text{Post}]$ indicates the post-2021 period, and β_3 captures the differential curtailment response under the tighter threshold.

5. Results

5.1 No Evidence of Duration Bunching

Bunching estimates for each regime show no statistically significant excess mass below the clawback threshold. The 2021–2023 regime (4-hour threshold) produces the largest bunching coefficient ($\hat{b} = 10.1$), but with a standard error of 40.5, the estimate is far from statistically significant. The pre-2021 regime ($\hat{b} = -0.5$, SE = 8.1) and 2024 regime ($\hat{b} = -6.9$, SE = 31.8) show negative point estimates, inconsistent with strategic bunching. The simple share of episodes ending just below the threshold—18.0% at duration 3 under the 4-hour regime, compared to 8.3% at duration 5 under the 6-hour regime (Table 1)—is suggestive, but the cross-country placebo test (Table 4) undermines a causal interpretation: France ($\hat{b} = 182.0$) and the Netherlands ($\hat{b} = 214.0$) show larger concentrations at the same duration, despite having no equivalent clawback rule. The 4-hour duration appears to be a natural clustering point for negative-price episodes driven by the solar cycle—roughly half a daylight period—rather than strategic behavior.

5.2 Generation Rises, Not Falls, Near the Threshold

Table 2 presents mean generation by hours to the threshold. Far from declining, renewable generation is flat or slightly increasing as episodes approach the clawback. At $h - 4$ (four hours before the threshold), mean renewable generation is 44,343 MW; at h (the threshold hour itself), it is 47,069 MW.

Table 3 reports the formal test. Column (1) shows that renewable generation is 287 MW *higher* in near-threshold hours ($p = 0.035$). Column (2) decomposes by fuel type: wind shows a small, insignificant decline (−50 MW), while solar drives the positive coefficient (+339 MW, $p = 0.010$). This pattern is consistent with the diurnal cycle: negative-price episodes that begin overnight and extend into daylight hours naturally see rising solar output as the sun rises, and these rising-solar hours mechanically coincide with the near-threshold period.

Column (4) tests whether the 2021 threshold tightening changed behavior. The interaction term (Near × Post-2021) is 152 MW with a standard error of 680, showing no detectable intensification of curtailment under the tighter regime.

Selection and power. Two caveats merit attention. First, the within-episode regression conditions on episodes reaching at least 3 hours, which excludes episodes that may have been "successfully" shortened by early curtailment. If curtailment operates primarily by preventing episodes from forming rather than by reducing generation within long episodes, the regression sample is selected on non-response. The bunching estimator, which uses

Table 2: Renewable Generation by Hours to Clawback Threshold

Hours to Threshold	Renewable (MW)	Wind (MW)	Solar (MW)	N
$h-4$	4.434e+04	2.800e+04	9,929	65
$h-3$	4.577e+04	2.618e+04	1.326e+04	140
$h-2$	4.656e+04	2.324e+04	1.706e+04	193
$h-1$	4.707e+04	2.318e+04	1.769e+04	183
h (threshold)	4.707e+04	2.327e+04	1.762e+04	158
$h+1$ (post-threshold)	4.731e+04	2.424e+04	1.691e+04	126
$h+2$ (post-threshold)	4.614e+04	2.355e+04	1.647e+04	103

Notes: Mean generation (MW) by fuel type, indexed by hours to the clawback threshold (h). $h-k$ denotes k hours before the threshold would trigger; $h+k$ denotes k hours after. Sample restricted to episodes with duration ≥ 3 hours. Renewable includes wind (onshore + offshore), solar, biomass, hydro run-of-river, and geothermal.

all episodes regardless of duration, does not suffer from this selection. Second, the 95% confidence interval for the curtailment coefficient spans approximately $[+22, +553]$ MW, meaning the data can rule out reductions larger than 22 MW—less than 0.05% of mean renewable generation during negative-price hours. While the study has adequate power to detect large curtailment responses, smaller but economically meaningful reductions (e.g., 1–2% of generation) cannot be excluded.

5.3 Cross-Country Placebos

Table 4 presents bunching estimates at the 4-hour duration for Germany and four European countries. If the German clawback drives bunching at 4 hours, we should see excess mass in Germany but not in countries without the rule. Instead, France and the Netherlands show comparable or larger concentrations. Austria, which shares the German bidding zone for part of the sample period, shows a small negative estimate. These results suggest that the duration distribution is shaped by weather and demand patterns common to European electricity markets, not by the German subsidy clawback.

5.4 Robustness

The bunching estimates are sensitive to the polynomial order of the counterfactual distribution (Table 5, Panel A), ranging from $\hat{b} = 0.1$ (order 4) to 82.0 (orders 6–7). This instability reflects the small sample—89 episodes in the 2021–2023 regime—rather than a genuine behavioral response. Bandwidth sensitivity (Panel B) shows a similar pattern: narrower windows produce small estimates, wider windows produce large but noisy ones. The donut specification (excluding episodes exactly at duration 4) yields $\hat{b} = 6.3$ (SE = 26.8). A placebo

Table 3: Within-Episode Curtailment at the Clawback Threshold

	(1) Renewable	(2) Wind	(3) Solar	(4) Renewable
Near threshold	287.3**	-49.6	338.8**	169.7
	(135.5)	(132.9)	(130.9)	(583.9)
Post-2021				1116.6
				(774.1)
Near threshold × Post-2021				151.8
				(679.5)
Episode FE	Yes	Yes	Yes	No
Hour-of-day FE	Yes	Yes	Yes	Yes
Month FE	No	No	No	Yes
Observations	1,357	1,357	1,357	932

Notes: Dependent variable is generation (MW). "Near threshold" equals one if the episode-hour is zero or one hours from the clawback threshold. Columns (1)–(3) include episode fixed effects and hour-of-day fixed effects; sample restricted to episodes with duration ≥ 3 hours. Column (4) replaces episode FE with month FE and adds an interaction with the 2021 reform (threshold tightened from 6h to 4h), using only pre-2021 and 2021–2023 episodes. Standard errors clustered by date. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

test at duration 5—which should show no bunching under the 4-hour threshold—confirms the null ($\hat{b} = -0.4$).

For the curtailment regression, alternative windows for the near-threshold indicator produce consistent results: restricting to the threshold hour alone yields +343 MW ($p = 0.017$), while expanding to a 3-hour window yields −26 MW ($p = 0.881$). The daytime/nighttime split shows a larger positive coefficient for daytime episodes (+365 MW) than nighttime (+85 MW), consistent with the solar-cycle interpretation. A permutation test randomly shuffling the near-threshold indicator within episodes produces a p-value of 0.028, confirming that the positive coefficient is systematic but not evidence of curtailment.

6. Discussion

The absence of strategic curtailment has a simple economic explanation: the clawback threshold is a *collective-action problem masquerading as an individual incentive.* The market-clearing price in the day-ahead auction depends on aggregate supply and demand, not on the output of any single generator. When system-wide oversupply drives prices negative, an individual wind farm curtailing 3 MW has negligible impact on whether the episode persists for a fourth hour. The clawback penalizes generators for a market outcome they cannot individually prevent.

This diagnosis has broader implications for renewable energy policy design. Duration-

Table 4: Placebo: Bunching at 4-Hour Threshold Across Countries

Country	Episodes	Bunching $\hat{b}$	SE	EEG Clawback?
DE	89	10.050	(40.455)	Yes
FR	167	182.000**	(87.887)	No
AT	158	-0.196	(56.681)	No
NL	238	214.000**	(94.245)	No
ES	46	NA	(NA)	No

Notes: Bunching estimated at the 4-hour duration mark for all countries using the full 2019–2024 sample. Germany (DE) is the only country with an EEG-style clawback at this threshold. France, Austria, Netherlands, and Spain experience negative prices but have no equivalent subsidy clawback rule. Estimates use a 5th-order polynomial counterfactual with ±2 hour exclusion bandwidth. Standard errors from 500 bootstrap replications. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

based clawback rules assume that generators can influence the length of negative-price episodes, but this assumption fails in markets with many small, price-taking renewable generators. The policy may inadvertently penalize generators for weather outcomes rather than encouraging efficient dispatch. A more effective approach might target the system-level conditions that produce negative prices—transmission congestion, inflexible conventional generation, or insufficient demand response—rather than imposing penalties on individual renewable producers.

The null result also speaks to the bunching literature. Chetty et al. (2011) identified adjustment costs and optimization frictions as key determinants of bunching responses. In the electricity market, the relevant "adjustment cost" is not ignorance or inattention—generators are sophisticated firms with real-time market access—but rather the inability to affect the relevant price through individual action. This parallels findings in tax bunching where responses depend on the controllability of the running variable: taxpayers bunch at kink points they can manipulate (income reporting) but not at thresholds determined by external factors.

Several limitations deserve mention. First, I observe aggregate generation by fuel type, not individual generator output. If some generators curtail while others increase output—for instance, if utility-scale wind operators respond while distributed solar installations do not—the aggregate signal could mask offsetting behavioral responses. Individual-level dispatch data, which are not publicly available, would be needed to rule out heterogeneous micro-responses. Second, the 15-minute generation data capture realized output, not dispatch decisions; curtailment choices made in the day-ahead auction are not directly observed. Third, the relatively small number of episodes (288 total, 89 in the key 2021–2023 regime) limits statistical power, particularly for the bunching analysis, where the confidence intervals are wide enough to include economically meaningful bunching coefficients. Fourth, the most recent

Table 5: Robustness of Bunching Estimates (2021–2023, 4h Threshold)

Specification	$\hat{b}$	SE
Panel A: Polynomial Order		
Order 3	1.741	(1.840)
Order 4	0.118	(6.135)
Order 5	10.050	(41.565)
Order 6	82.000**	(35.235)
Order 7	82.000**	(33.933)
Panel B: Bandwidth		
± 1 hours	0.168	(0.709)
± 2 hours	10.050	(40.301)
± 3 hours	-10.881	(47.282)
Panel C: Additional Tests		
Donut (excl. duration = 4)	6.259	(27.312)
Placebo threshold (5h)	-0.397	(1.104)

Notes: All specifications use episodes from 2021–2023 and test bunching at the 4-hour clawback threshold. Panel A varies the polynomial order of the counterfactual distribution (baseline: 5th order). Panel B varies the exclusion bandwidth (baseline: ±2 hours). Panel C reports a donut specification and placebo test at 5h. Standard errors from 500 bootstrap replications. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

threshold tightenings (to 2 hours in 2026 and 1 hour in 2027) may yet produce detectable responses as the clawback becomes harder to avoid by chance—though the collective-action logic suggests that even very tight thresholds will fail to change behavior unless individual generators acquire market power over the system-level price.

7. Conclusion

Sharp incentive discontinuities do not always produce sharp behavioral responses. Germany's EEG clawback threshold creates a discrete subsidy loss at a well-defined duration—exactly the conditions under which bunching theory predicts a response. Yet 15-minute generation data reveal no curtailment cliff. The episode duration distribution is shaped by weather and demand, not by the subsidy schedule. The result exposes a design flaw in duration-based clawback rules: they impose individual penalties for collective outcomes that no single generator can control.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers
Contributors: @ai1scl
First Contributor: https://github.com/ai1scl

References

Antweiler, Werner and Felix Muesgens, "On the Long-Term Merit Order Effect of Renewable Energies," *Energy Economics*, 2021, *99*, 105275.

Borenstein, Severin and James Bushnell, "The U.S. Electricity Industry After 20 Years of Restructuring," *Annual Review of Economics*, 2015, *7*, 437–463.

Brait, Luisa and Simon Müller, "Negative Electricity Prices and Their Implications for Renewable Energy Support," *Energy Economics*, 2020, *86*, 104634.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Fabra, Natalia and Mar Reguant, "Pass-Through of Emissions Costs in Electricity Markets," *American Economic Review*, 2014, *104* (9), 2872–2899.

Hirth, Lion, "The Market Value of Variable Renewables: The Effect of Solar Wind Power Variability on Their Relative Price," *Energy Economics*, 2013, *38*, 218–236.

Ito, Koichiro, "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing," *American Economic Review*, 2014, *104* (2), 537–563.

Ketterer, Janina C., "The Impact of Wind Power Generation on the Electricity Price in Germany," *Energy Economics*, 2014, *44*, 270–280.

Kleven, Henrik, "Bunching," *Annual Review of Economics*, 2016, *8*, 435–464.

Kyritsis, Evangelos, Jonas Andersson, and Apostolos Serletis, "Electricity Prices, Large-Scale Renewable Integration, and Policy Implications," *Energy Policy*, 2017, *101*, 550–560.

Lagarde, Cyril Milhaud de and Frédéric Lantz, "How Renewable Energy Feed-In Tariffs Interact with Negative Electricity Prices," *Energy Policy*, 2022, *163*, 112829.

Nicolosi, Marco, "Wind Power Integration and Power System Flexibility — An Empirical Analysis of Extreme Events in Germany under the New Negative Price Regime," *Energy Policy*, 2010, *38* (11), 7257–7268.

Paraschiv, Florentina, David Erni, and Ralf Pietsch, "The Impact of Renewable Energies on EEX Day-Ahead Electricity Prices," *Energy Policy*, 2014, *73*, 196–210.

Reguant, Mar, “The Efficiency and Sectoral Distributional Impacts of Large-Scale Renewable Energy Policies,” *Journal of the Association of Environmental and Resource Economists*, 2019, *6* (S1), S129–S168.

Saez, Emmanuel, “Do Taxpayers Bunch at Kink Points?,” *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

Sensfuß, Frank, Mario Ragwitz, and Massimo Genoese, “The Merit-Order Effect: A Detailed Analysis of the Price Effect of Renewable Electricity Generation on Spot Market Prices in Germany,” *Energy Policy*, 2008, *36* (8), 3086–3094.

Table 6: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
Renewable generation	287.3	135.5	7,311	0.039	0.019	Small positive
Wind generation	-49.6	132.9	7,883	-0.006	0.017	Small negative
Solar generation	338.8	130.9	1.018e+04	0.033	0.013	Small positive
Panel B: Heterogeneous (Sample Splits)						
Renewable (daytime)	364.5	155.7	6,267	0.058	0.025	Moderate positive
Renewable (nighttime)	85.2	256.9	6,801	0.013	0.038	Small positive

Notes: **Country:** Germany. **Research question:** Do renewable generators curtail electricity output as negative-price episodes approach the EEG subsidy clawback threshold? **Policy mechanism:** Under the EEG market premium, renewable generators above 400 kW lose their subsidy for an entire negative-price episode once it reaches the clawback duration; we test whether generation drops in the hours immediately before this threshold triggers, which would indicate strategic curtailment to preserve subsidies. **Outcome definition:** Hourly electricity generation (MW) by fuel type during negative-price episodes, aggregated from 15-minute Fraunhofer ISE Energy-Charts data. **Treatment:** Binary; episode-hour is within zero or one hours of the clawback threshold versus farther from it. **Data:** Fraunhofer ISE Energy-Charts, 2019–2024, episode-hour level, 1,357 observations across 209 episodes. **Method:** OLS with episode and hour-of-day fixed effects; standard errors clustered by date. **Sample:** Negative-price episodes in Germany with duration ≥ 3 hours. SDE $= \hat{\beta}/\mathrm{SD}(Y)$ where SD(Y) is the pre-treatment (pre-2021) standard deviation. Classification refers to magnitude, not statistical significance: Large ($|\mathrm{SDE}| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).

A. Standardized Effect Sizes