Proceedings of Business and Economic Studies, 2026, Volume 9, Issue 10
https://bbwpublisher.com/journal/PBES
Online ISSN: 2209-265X
Print ISSN: 2209-2641

Market Response to ICVCM Methodology Decisions in the Voluntary Carbon Market

Yiming Zhang

School of Advanced International Studies, Johns Hopkins University, Washington, DC, United States

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: The Integrity Council for the Voluntary Carbon Market was created to identify carbon credits that meet a common threshold for integrity, but a governance determination matters economically only if buyers alter their behavior in response. This article tests whether the Integrity Council's 2024 methodology decisions changed revealed demand, measured by project-level credit retirements. It combines the Integrity Council's dated decisions with the Berkeley Voluntary Registry Offsets Database, which covers 11,468 projects across six registries and reports annual issuance and retirement histories. The principal analysis examines 1,448 renewable-energy projects and 9,833 project-years from 2018 through 2025. Projects using one of eight methodologies that the Integrity Council rejected in August 2024 are compared with other renewable-energy projects in the same overlapping project types. Models include project fixed effects, project-type-by-year effects, registry-by-year effects, available credit supply, and methodology-clustered standard errors. Contrary to the predicted demand penalty, rejected methodologies show no detectable decline in 2025 retirement intensity or in the probability of recording a retirement. The preferred estimates are 0.019 (standard error 0.288) for inverse-hyperbolic-sine retirements and 0.048 (standard error 0.041) for the retirement probability. Both estimates are imprecise but inconsistent with an immediate, clearly visible negative market response. Results remain similar in a balanced panel, a single-methodology sample, a shorter event window, and specifications without the supply control. A separate comparison of approved ozone-depleting-substance methodologies is too small and has nonparallel pre-trends, so it is reported descriptively rather than interpreted causally. The findings do not show that integrity labels lack value. They show that external governance judgments did not automatically translate into near-term retirement reallocation. This result provides a first large-sample test of buyer demand as a governance condition and identifies the public data needed to evaluate integrity reforms more conclusively.

Keywords: Voluntary carbon market; Carbon credits; ICVCM; Certification; Credit retirements; Private governance; Market integrity

Online publication: November 16, 2026

1. Introduction

Carbon-credit governance has recently shifted from asking whether projects disclose enough information to

asking whether the rules producing the credits are credible. The Integrity Council for the Voluntary Carbon Market, or ICVCM, embodies that shift. It assesses carbon-crediting programs and methodology categories against the Core Carbon Principles and permits eligible programs to label credits only when both the program and the methodology have passed its assessment. In June 2024, the Integrity Council approved the first methodologies for ozone-depleting-substance destruction and landfill-gas projects. In August, it determined that eight widely used renewable-energy methodologies did not meet its threshold. In November, it approved three newer REDD+ approaches while making clear that no credits had yet been issued under them [1–3].

These decisions created a rare test of private governance. The determinations were public, methodology-specific, and economically consequential in principle: credits generated under a rejected methodology could not receive the CCP label, while credits meeting both parts of the two-tick process could. The central empirical question is whether buyers responded. If external integrity judgments help markets distinguish credible from weak credits, projects affected by an adverse decision should experience a relative decline in demand, while projects covered by favorable decisions should benefit. If retirements do not move, the label may be informative without being behaviorally binding.

This question follows directly from Zhang's 2026 analysis of digital technology and governance in the voluntary carbon market [4]. That article argues that trustworthy market signals require three conditions to hold together: data governance independent of commercial interests, enforceable liability for misreporting, and genuine buyer demand for quality. The first two conditions concern how quality is produced and enforced. The third concerns whether quality affects market behavior. The original article offered comparative and process-tracing evidence but could not directly observe this third condition after the Integrity Council began issuing methodology decisions. The present article supplies that test.

The analysis uses the Berkeley Carbon Trading Project's Voluntary Registry Offsets Database, version 2026-06. The database consolidates project, issuance, and retirement information from the American Carbon Registry, Architecture for REDD+ Transactions, Climate Action Reserve, Gold Standard, Isometric, and Verra. The uploaded workbook contains 11,468 uniquely identified projects, of which 6,016 have issued credits, and 4,948 have recorded at least one retirement. Annual issuance and retirement totals reconcile to the project's reported totals. The database therefore permits construction of a project-year panel without converting aggregate market reports into fictitious transactions.

The primary design focuses on the August 2024 renewable-energy decision. Eight rejected methodology families appear extensively in the Berkeley data, while other methodologies remain available within the same renewable project types. The analysis restricts the sample to Gold Standard and Verra, the programs named in the decision, retains project types containing both rejected and comparison methodologies, and includes a project only after it has issued credits. The resulting panel contains 1,448 projects, 1,293 treated and 155 comparison projects, contributing 9,833 at-risk project-years between 2018 and 2025. The main outcomes are the inverse hyperbolic sine of annual retirements and an indicator for any retirement. Project fixed effects absorb stable differences in project quality and scale. Project-type-by-year and registry-by-year effects absorb category-specific and program-specific demand shocks. A time-varying measure of available supply accounts for the basic fact that credits cannot be retired unless they exist.

The evidence does not support the predicted demand penalty. In the preferred specification, the 2025 interaction for rejected methodologies is 0.019 with a methodology-clustered standard error of 0.288 for retirement intensity. For the probability of any retirement, the estimate is 0.048 with a standard error of 0.041.

The corresponding normal-approximation 95 percent intervals are -0.545 to 0.583 and -0.033 to 0.128. Thus, the analysis cannot rule out moderate changes in retirement intensity, but it finds neither a negative point estimate nor a statistically detectable negative extensive-margin effect. The result persists in a balanced panel of projects already at risk by 2018, after excluding projects listing multiple methodologies, in a 2020–2025 window, and without the supply control. Pre-decision interactions are individually indistinguishable from zero in the preferred renewable-energy specifications.

The paper also examines favorable decisions. Ozone-depleting-substance methodologies provide the cleanest conceptual comparison because the June 2024 announcement named specific protocols with already-issued credits. Yet the usable voluntary sample contains only 49 projects and five methodology clusters, and its pre-decision paths are not parallel. The positive-looking post-2024 coefficients from a saturated model therefore cannot support causal interpretation. The REDD+ approvals are even less suitable for a conventional event study because the Integrity Council stated at the decision date that no credits had yet been issued under the approved methodologies. These cases are analytically informative precisely because they reveal the boundary between a public decision and an estimable treatment.

The contribution is threefold. First, the paper provides a project-level empirical test of whether private integrity governance changes revealed carbon-credit demand. Existing work documents over-crediting, price dispersion, certification arrangements, or aggregate market contraction, but rarely observes a common external determination attached to identifiable methodologies. Second, it treats a null result as evidence with boundaries rather than converting absence of significance into proof of no effect. The estimates show no detectable near-term retirement penalty, while the intervals and data structure make clear what remains unresolved. Third, the paper connects the result to a governance mechanism. If adverse quality determinations do not reduce retirements, then buyer demand cannot be assumed merely because credible information has been supplied. It must be created through claim-eligibility rules, disclosure obligations, or other institutions that make quality consequential.

The rest of the paper proceeds as follows. Section 2 connects certification, information economics, and carbon-market demand. Section 3 describes the Integrity Council's 2024 decisions. Section 4 presents the data, treatment coding, outcomes, and empirical design. Section 5 reports descriptive, event-study, and robustness results. Section 6 interprets the bounded null finding and relates it to the governance conditions developed in the companion article. Section 7 draws implications for market design and data disclosure. Section 8 states the limitations, and Section 9 concludes.

2. Certification quality and buyer demand

2.1. Signals require consequences

The economic function of certification is to reduce quality uncertainty. In Akerlof's 1970 account, markets can unravel when buyers cannot distinguish high-quality from low-quality goods and therefore will not pay sellers for unobservable quality [5]. Spence's 1973 signaling model adds a condition: a signal separates types only when its cost differs systematically with the underlying quality being signaled [6]. A label that is equally easy for high- and low-quality suppliers to obtain cannot support a separating equilibrium. The certifying institution must therefore do more than publish information. It must make the signal difficult to obtain without the relevant underlying quality.

Certification agencies also face their own incentives. Bolton, Freixas, and Shapiro showed how issuer-paid ratings can be inflated when agencies compete for business, and investors rely mechanically on ratings [7]. Their analysis is directly relevant to carbon markets, where project developers or project owners generally commission validation and verification. The existence of a certifier does not remove information asymmetry; it moves the problem to the rules governing the certifier, the liability attached to error, and the behavior of information users.

Research on transnational private governance reaches a compatible conclusion from an institutional direction. Auld and Gulbrandsen distinguished transparency about governance procedures from transparency about regulated outcomes [8]. Schleifer, Fiorini, and Auld showed that voluntary sustainability programs often disclose formal procedures without disclosing the processes that produce certification decisions [9]. Bartley argued that private certification remains constrained by its dependence on firms, auditors, and consumers [10]. Together, this literature implies that more disclosure can coexist with weak market discipline.

Zhang applied these ideas to digital monitoring, reporting, and verification in carbon markets [4]. Digital tools can improve measurement of observable conditions, but they cannot directly observe the counterfactual baseline against which many credits are issued. The article identifies three governance conditions: independent data governance, enforceable liability, and genuine buyer demand for quality. These conditions are interdependent. Independence without liability may reveal problems without changing payoffs. Liability without buyer demand may preserve a niche of high-integrity credits while leaving a large low-integrity segment intact. Buyer demand without credible production and enforcement rules may reward labels rather than actual integrity.

The present paper isolates a limited but observable component of the third condition. A retirement removes a credit from circulation and records its use or final claim. Retirement is therefore a revealed-demand outcome, though not a complete measure of demand: a credit may be purchased well before it is retired, held in inventory, transferred between intermediaries, or retired under a preexisting contract. The test is accordingly whether integrity decisions redirect realized retirements in the near term, not whether they change every form of buyer valuation.

2.2. Carbon-credit demand and market differentiation

Carbon credits are heterogeneous assets. Project type, location, vintage, co-benefits, methodology, registry, buyer purpose, and transaction size may all affect value. Hedonic analysis decomposes price into implicit values for product attributes [11]. Conte and Kotchen applied this approach to voluntary offsets and showed that observable project and certification attributes help explain price variation [12]. Ecosystem Marketplace reports have continued to document large differences in prices and volumes across project categories and standards [13,14].

Price is not the only margin on which buyers can respond. A buyer can stop purchasing a category, reduce quantity, delay use, substitute into another methodology, or continue buying but demand a discount. Public registry data observe issuance and retirement but usually not transaction prices or contracting dates. That distinction matters. A rejection can affect price without changing retirement quantity, especially if demand is inelastic or contracts were signed earlier. Conversely, a fall in retirements can reflect category-wide demand or supply changes unrelated to the governance decision. A credible design must therefore compare affected projects with projects exposed to similar market conditions and control for available credit supply.

The literature documenting carbon-credit integrity makes a demand response plausible. West *et al.*, Probst *et al.*, and Swinfield *et al.* reported substantial over-crediting or sensitivity to baseline choices in important

project categories [15–17]. Delacote *et al.* argued that transparent and systematic ex-post evaluation is needed to restore credibility [18,19]. Christiansen showed that digital monitoring is also used to manage buyer reputational risk [20]. Together, these studies imply that buyers should value credible differentiation after the market's 2022–2023 integrity crisis. Whether they actually reallocate demand when a governance body differentiates methodologies remains an empirical question.

The Integrity Council decisions permit two directional hypotheses. The adverse-decision hypothesis is the main confirmatory test:

H1: Projects using methodologies that the Integrity Council determined did not meet the CCP threshold experience a relative decline in retirement intensity and in the probability of retirement after the decision.

The favorable-decision hypothesis is conceptually symmetric:

H2: Projects using approved methodologies experience a relative increase in retirement intensity and in the probability of retirement after the decision.

The available data support a substantially stronger test of H1 than H2. Rejected renewable methodologies had long histories and comparison projects within the same broad project types. Approved ODS methodologies offer only a small number of voluntary projects and methodology clusters. Approved REDD+ methodologies had not issued credits at the event date. The empirical hierarchy therefore places H1 in the main analysis and treats H2 as descriptive and exploratory.

3. The 2024 Integrity Council decisions

The Integrity Council uses a two-tick structure. A carbon-crediting program must first be CCP-Eligible. A credit can then receive the CCP label only when it is issued under a methodology category that the Integrity Council has also approved. The distinction is important: approval of a program does not imply approval of all methodologies used by that program, and the Integrity Council is not itself a registry.

On 6 June 2024, the Integrity Council announced its first approved methodology categories. The decision covered three ODS protocols: ACR's Destruction of ODS from International Sources version 1.0, CAR's Article 5 Ozone Depleting Substances Project Protocol versions 1-2, and CAR's US Ozone Depleting Substances Project Protocol versions 1-2. It also covered four landfill-gas approaches: ACM0001 versions 15-19, AMS-III. G version 10, ACR's Landfill Gas Destruction and Beneficial Use Projects versions 1-2, and CAR's US Landfill Protocol version 6. Approval of ACM0001 and AMS-III.G was conditional for electricity-generating projects, applying to projects with capacity no greater than 10 MWe, while non-electricity projects were covered. The announcement estimated that the approved groups included 27 million credits, although it cautioned that methodology versions were sometimes unknown [1].

On 6 August 2024, the Integrity Council announced that eight renewable-energy methodologies used by Gold Standard and Verra did not meet its requirements. The affected families were ACM0002, ACM0006, ACM0018, AM0036, AM0072, AMS-I.D, AMS-I.L, and AMS-I.A, subject to the version ceilings stated in the decision. The Board also rejected AM0065 for sulfur-hexafluoride avoidance. The decision explained that the renewable methodologies did not sufficiently establish additionality in settings where renewable energy had become commercially competitive. It also stated that when a project used a combination of methodologies, use of any rejected methodology made the generated credits ineligible for the CCP label [2].

On 15 November 2024, the Integrity Council approved ART-TREES version 2.0 at the relevant crediting level, Verra's VM0048, and Verra's Jurisdictional and Nested REDD+ Framework version 4.1. The

announcement emphasized changes in baseline construction under VM0048. It also stated that no credits had yet been issued under the approved approaches and that eligible credits were expected to begin appearing in 2025 [3]. This timing prevents those decisions from supporting a standard comparison of already-trading treated credits before and after approval.

The decisions are not randomly assigned. Methodologies were assessed because they existed in particular categories, and approval or rejection depended on substantive integrity criteria. The design therefore does not treat the process as a randomized experiment. Its leverage comes from timing, within-project changes, overlapping renewable project types, and rich fixed effects. The resulting estimates answer a narrower question: after accounting for stable project differences, category-year shocks, registry-year shocks, and available supply, did retirement behavior change differentially for projects covered by an adverse decision?

4. Data and empirical strategy

4.1. Berkeley Voluntary Registry Offsets Database

The primary source is the Berkeley Carbon Trading Project's Voluntary Registry Offsets Database, version 2026-06 [21]. The workbook reports that it contains projects, issuances, and retirements through 30 June 2026. It consolidates six registries: ACR, ART, CAR, Gold Standard, Isometric, and Verra. The PROJECTS worksheet contains registry project identifiers, project names, voluntary status, harmonized scope and project type, reduction or removal classification, methodology or protocol, methodology version when available, region, country, developer, verifier, annual issuance by issuance year, annual retirements by calendar year, annual credits by vintage, and cumulative totals.

The audit identified 11,468 project rows and no duplicate registry-project keys. The database contains 5,002 Verra projects, 4,115 Gold Standard projects, 1,272 CAR projects, 986 ACR projects, 65 Isometric projects, and 28 ART projects. A methodology string is present for 9,859 projects, with 391 distinct strings. Annual issuance sums to 2.655 billion credits and annual retirements sum to 1.562 billion, matching the cumulative columns apart from floating-point rounding. These reconciliations provide a basic check that the year columns were located and reshaped correctly.

The raw workbook is preserved unchanged and identified by a SHA-256 checksum. Analysis scripts reshape only the populated research columns; blank formatting extending beyond the table is ignored. Each project is expanded into annual observations from 2018 through 2025. A project enters the risk set in the first year its cumulative issuance becomes positive and remains at risk thereafter. This rule avoids treating years in which a project had no issued credits as observations of zero demand.

4.2. Outcomes and available supply

The first outcome is annual retired credits transformed using the inverse hyperbolic sine. This transformation is defined at zero, reduces the influence of very large projects, and behaves similarly to a logarithm for large values. Because many active projects have no retirement in a given year, the second outcome is an indicator equal to one when annual retirements are positive. The first captures retirement intensity; the second captures the extensive margin of whether a project records any realized use.

Available supply is constructed from the same accounting records. For project i in year t, beginning available supply equals cumulative issuance through t-1 minus cumulative retirements through t-1, bounded below at zero. Issuance during year t is then added to obtain supply available for retirement during that year.

The preferred models include the inverse hyperbolic sine of this quantity. The control is not interpreted as exogenous. It prevents a mechanical shortage of credits from being mistaken for a demand response. Specifications omitting supply are reported because contemporaneous issuance can itself respond to market conditions.

Retirements are an imperfect proxy for purchases. They may occur after the transaction date and can reflect inventory management or claims calendars. They nonetheless have two advantages: they record final use rather than an asking price, and they are consistently observed across the covered registries. The outcome should therefore be understood as revealed retirement demand, not transaction demand at the exact date of purchase.

4.3. Treatment and comparison groups

The primary treatment indicator equals one when the methodology string contains any of the eight rejected renewable families: ACM0002, ACM0006, ACM0018, AM0036, AM0072, AMS-I.D, AMS-I.L, or AMS-I.A. Projects listing multiple methodologies are treated if any listed method was rejected, following the decision's stated eligibility rule. The main sample is restricted to the Renewable Energy scope and to Gold Standard and Verra. It excludes missing methodology strings and retains only project types containing both treated and comparison projects.

The overlapping types are biomass, geothermal, hydropower, bundled renewable energy, centralized solar, distributed solar, solar lighting, solar water heaters, and wind. As shown in **Table 1**, of the 11,468 database projects, 2,501 fall within the harmonized Renewable Energy scope. Gold Standard and Verra account for 2,496 of those projects, and 2,223 have a non-missing methodology. After requiring project-type overlap and positive cumulative issuance during the study period, the analytical panel contains 1,448 projects and 9,833 project-year observations. There are 1,293 treated projects and 155 comparison projects across 35 observed methodology strings.

Table 1. Data and analytical sample

Sample stage	Count
All unique projects in workbook	11,468
Projects in harmonized Renewable Energy scope	2,501
Gold Standard and Verra renewable projects	2,496
Renewable projects with methodology recorded	2,223
Projects in final at-risk panel	1,448
Project-year observations, 2018–2025	9,833
Rejected-methodology projects	1,293
Comparison projects	155
Methodology-string clusters	35

Note: The at-risk panel includes only project-years after positive cumulative issuance. Counts are calculated from Berkeley Carbon Trading Project data, release v2026-06.

This imbalance reflects the market rather than a sampling choice: the rejected CDM-derived methodologies were widely used. It also limits common support. Project-type-by-year effects ensure that treatment is not identified simply by comparing wind projects with solar-water-heater projects, but treated projects can still differ from comparison projects in unobserved time-varying ways. The analysis therefore uses the language of

differential response and avoids presenting the treatment coefficient as a fully identified causal effect.

The favorable-decision sample is narrower. It contains voluntary ODS projects using one of the three approved protocols and comparison projects using ACR's High-GWP Foam methodology or Verra VM0016. Compliance-designated projects are excluded. After applying the risk-set rule, this sample contains 49 projects, 28 approved and 21 comparison projects, contributing 153 project-years and only five methodology clusters. It is unsuitable for strong clustered inference and displays nonparallel pre-decision coefficients.

Landfill-gas approvals are not used for the primary favorable test. The Berkeley workbook often lacks methodology-version information and does not consistently report the generating-capacity threshold needed to apply the Integrity Council's conditional approval. The Integrity Council itself cautioned that version ambiguity affected its landfill estimates. Coding all projects under a base methodology as approved would therefore create treatment misclassification.

4.4. Event-study specification

The main model estimates annual interactions between rejected-methodology status and year, with 2023 omitted as the reference year. It includes project fixed effects, common year effects, project-type-by-year effects, registry-by-year effects, and available supply. In notation, retirement outcome Y for project i in year t is modeled as a function of project effects, year effects, type-specific year shocks, registry-specific year shocks, a series of rejected-methodology-by-year indicators, and the inverse hyperbolic sine of available supply. Errors are clustered by the complete methodology string recorded in the workbook.

Project fixed effects remove all stable project attributes, including location, baseline scale, developer characteristics, and time-invariant differences in perceived quality. Project-type-by-year effects allow wind, hydropower, solar, biomass, and other categories to follow different annual demand paths. Registry-by-year effects allow Gold Standard and Verra to experience different shocks. The treatment coefficients are therefore identified from differential changes between treated and comparison projects within these controls.

The annual frequency determines interpretation. The August 2024 decision occurred during the retirement year, so the 2024 coefficient mixes roughly seven pre-decision months with five post-decision months and is treated as transitional. Calendar year 2025 is the first complete post-decision period. The June 2026 release provides only a partial 2026 total, which is not included in the confirmatory regression because retirement seasonality is unknown. It is retained for later descriptive updates.

The standard-error calculation uses a cluster-robust sandwich estimator with a finite-cluster correction and 35 methodology clusters in the primary sample. Reported *P*-values and 95 percent intervals use a normal approximation. This choice is adequate for transparent replication but should not be mistaken for exact small-sample inference. The treatment is assigned by methodology family, while some cluster strings contain combinations of methods. The robustness analysis therefore excludes multiple-method strings in one specification.

Four robustness checks address the most immediate design choices. First, a balanced sample retains only projects already in the risk set in 2018. Second, a single-methodology sample excludes strings containing multiple methodologies. Third, a shorter window begins in 2020. Fourth, the supply control is removed. Every robustness model retains project effects, project-type-by-year effects, and registry-by-year effects.

5. Results

5.1. Descriptive patterns

The raw data show a broad contraction in renewable-credit retirements rather than a clean divergence caused by the August decision. Among at-risk treated projects, retired volume was 46.75 million credits in 2023, 42.76 million in 2024, and 34.77 million in 2025. Mean project retirements fell from 38,348 in 2023 to 26,889 in 2025, while the share of projects with any retirement fell from 0.513 to 0.456.

Comparison projects also weakened. Their retired volume was 1.44 million in 2023, 0.76 million in 2024, and 0.92 million in 2025. Mean retirements fell from 10,800 in 2023 to 5,965 in 2025. The retirement probability declined from 0.489 in 2023 to 0.379 in 2024 before recovering to 0.445 in 2025. These movements show why comparing the treated group's 2025 level only with its own 2023 level would be misleading. Both groups experienced substantial changes, and the comparison group had an especially sharp transitional-year decline.

Supply differed considerably. Treated projects were older and larger on average and held substantially more available credits. The preferred model therefore uses within-project changes and includes available supply. The treatment coefficient should be read as the differential movement after these adjustments, not as a comparison of average project size.

5.2. Event-study estimates

Figure 1 presents the preferred dynamic estimates. As shown in **Table 2**, for retirement intensity, none of the pre-2024 treatment interactions is individually distinguishable from zero. Relative to 2023, the estimates are 0.153 in 2018, 0.203 in 2019, -0.575 in 2020, 0.040 in 2021, and -0.144 in 2022. Their clustered standard errors range from 0.430 to 0.802. The estimates fluctuate around zero without a monotonic differential trend.

The 2024 retirement-intensity coefficient is 0.471 with a standard error of 0.237. It is positive rather than negative and reaches the conventional 5 percent threshold under the normal approximation. Because most of 2024 preceded the August announcement and the annual data cannot date retirements within the year, this coefficient is not interpreted as an effect of rejection. It does, however, rule out a large immediate annual decline visible in the aggregate 2024 record.

The first full post-decision coefficient in 2025 is 0.019 with a standard error of 0.288 and a normal-approximation *P*-value of 0.947. The 95 percent interval extends from -0.545 to 0.583 inverse-hyperbolic-sine units. The estimate is nearly zero. The interval remains wide enough to include an economically meaningful negative change in retirement intensity, so the result is the absence of a detectable penalty rather than evidence of exact equivalence.

The extensive-margin model tells the same story. Pre-decision interactions are small relative to their clustered uncertainty. The 2024 coefficient is 0.067 with a standard error of 0.032, again positive. The 2025 estimate is 0.048 with a standard error of 0.041 and a *P*-value of 0.247. Its 95 percent interval runs from -0.033 to 0.128. In this specification, the data rule out a decline larger than approximately 3.3 percentage points at the lower end of the normal interval, but they do not rule out smaller declines or positive changes.

The predicted signs in H1 are negative for both outcomes. Neither preferred estimate is negative. Rejection therefore did not produce the observable near-term retirement penalty hypothesized by the buyer-demand mechanism. This conclusion is narrower than saying the decision had no market effect. Prices, forward contracts, inventories, and buyer composition are unobserved.

Table 2. Event-study estimates for rejected renewable methodologies

Year	asinh retirements	Any retirement
2018	0.153 (0.666)	0.052 (0.063)
2019	0.203 (0.773)	0.027 (0.076)
2020	−0.575 (0.802)	−0.046 (0.086)
2021	0.040 (0.430)	0.002 (0.052)
2022	−0.144 (0.449)	−0.003 (0.034)
2023	Reference	Reference
2024	0.471** (0.237)	0.067** (0.032)
2025	0.019 (0.288)	0.048 (0.041)

Note: Coefficients are rejected-methodology-by-year interactions relative to 2023; cluster-robust standard errors are in parentheses. Models include project fixed effects, year effects, project-type-by-year effects, registry-by-year effects, and asinh available supply. Standard errors are clustered by methodology string (35 clusters). $^{**}P < 0.05$ using a normal approximation. The 2024 coefficient is transitional because the decision occurred in August.

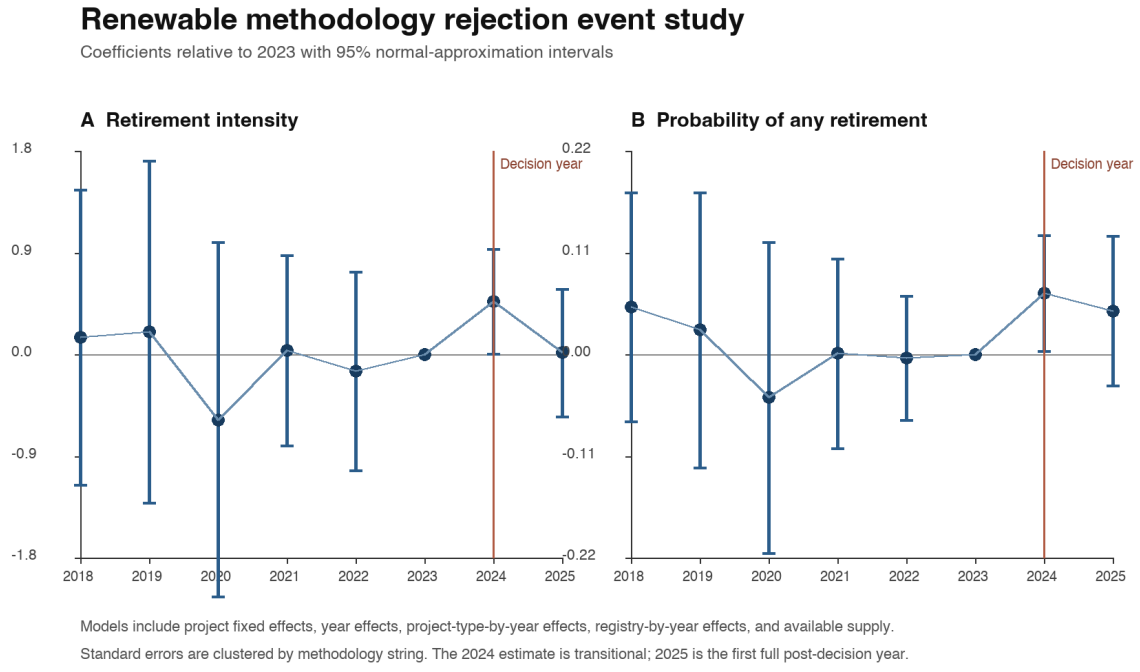


Figure 1. Dynamic estimates around the August 2024 rejection decision (Note: Points are rejected-methodology-by-year interaction estimates relative to 2023; bars show 95 percent normal-approximation intervals using standard errors clustered by methodology string. The 2024 estimate is transitional; 2025 is the first full post-decision year.)

5.3. Robustness

The main null finding is stable across four design variants. As shown in **Table 3**, without the supply control, the 2025 retirement-intensity coefficient is -0.090 with a standard error of 0.433. The retirement-probability coefficient is 0.038 with a standard error of 0.046. The sign of intensity becomes negative, but the estimate remains small relative to its uncertainty.

The balanced panel contains 951 projects and 26 methodology clusters. Its supply-adjusted 2025 coefficient is 0.009 (standard error 0.338) for intensity and 0.031 (standard error 0.040) for the retirement probability.

Restricting the analysis to 1,426 projects with a single methodology produces estimates of 0.075 (standard error 0.300) and 0.050 (standard error 0.044). Beginning the panel in 2020 yields 0.022 (standard error 0.285) and 0.048 (standard error 0.041). None provides evidence of a negative 2025 response.

Table 3. Robustness of the 2025 post-decision coefficient

Specification	asinh retirements	Any retirement
Preferred specification	0.019 (0.288)	0.048 (0.041)
Without supply control	−0.090 (0.433)	0.038 (0.046)
Balanced from 2018	0.009 (0.338)	0.031 (0.040)
Single-methodology projects	0.075 (0.300)	0.050 (0.044)
2020–2025 window	0.022 (0.285)	0.048 (0.041)

Note: Cluster-robust standard errors are in parentheses. All models retain project fixed effects, year effects, project-type-by-year effects, and registry-by-year effects. The first, third, fourth, and fifth specifications include available supply.

The consistency of the point estimates reduces concern that entry into the risk set, combination methodologies, the early part of the event window, or the supply control alone generates the conclusion. It does not resolve treatment misclassification. Methodology versions are frequently absent, and the 2026 workbook reports the methodology currently associated with each project rather than a complete annual history of methodology assignments. Projects that changed methodology could therefore be classified by a later status. The single-methodology restriction removes combinations but cannot reconstruct historical version changes.

5.4. Favorable decisions

The ODS comparison illustrates why statistical significance cannot substitute for design credibility. The sample contains 28 at-risk projects using an approved protocol and 21 projects using the two selected comparison methodologies. Approved projects recorded 103,773 retirements in 2023, 485,009 in 2024, and 460,048 in 2025. Comparison projects recorded 41,013, 160,332, and 149,629. In 2025, the retirement probability was 0.536 for approved projects and 0.571 for comparison projects.

A saturated event model can generate positive post-2024 coefficients for approved projects, including a 2025 probability estimate of 0.282. That number is not treated as evidence for H2. Only five methodology clusters are available, the risk-set composition changes substantially over time, and the pre-decision interactions vary sharply. For example, approved projects have negative relative coefficients in 2019 and 2020 and a positive relative coefficient in 2022. The parallel-path assumption is not credible.

The REDD+ case is even clearer. The November announcement stated that no credits had yet been issued under the three approved methodologies. Later observations can describe the creation and uptake of a new credit class, but they cannot supply a conventional before-and-after outcome for the same stock of treated credits. A study that labeled older REDD+ projects as treated by the VM0048 approval would conflate transition, requantification, and entry with a label effect.

The favorable evidence is therefore descriptive. It shows that some approved protocols subsequently recorded retirements, but it does not isolate approval as the cause. H2 remains unresolved in the public annual data.

6. Interpretation

6.1. A governance decision without an observable quantity penalty

The central finding is not that the Integrity Council's decision was uninformative. The Board made a clear adverse determination, explained the additionality concerns, and rendered affected credits ineligible for the CCP label. The finding is that this information did not create a detectable negative differential in project-level retirements during the first complete year after the decision.

This distinction matters for signaling theory. Information can be credible yet economically weak if recipients do not bear a cost for ignoring it. The rejected credits remained valid units within their issuing registries. Existing buyers could retire them. Contracts signed before the decision could still be performed. Firms whose claims did not require CCP-approved credits could continue using them. The decision changed eligibility for a voluntary quality label, not the legal status of the underlying credits.

In the three-condition framework, the Integrity Council strengthened information about governance quality but did not by itself create C3, genuine buyer demand. Demand becomes a governance condition when quality differences change the choices or liabilities of buyers. The near-zero retirement estimates suggest that, at least over the observed horizon, publication of an adverse assessment was insufficient to produce a visible reallocation in final use.

The result qualifies a statement in the companion article. That article cited aggregate market evidence suggesting buyers paid more or traded more in categories where CCP approval had attached. The project-level evidence here shows that favorable aggregate movements cannot be generalized into a symmetric penalty for rejected renewable methodologies. Aggregate prices or values can rise because of category composition, supply, or registry-specific changes. Directly observing retirements produces a more cautious conclusion: the market response was selective, delayed, or expressed on an unobserved margin.

6.2. Why a response may not appear in retirements

At least five mechanisms can reconcile a meaningful integrity decision with a null retirement result.

First, retirements lag purchases. A company may buy credits months or years before retiring them against an emissions claim. The 2025 retirement record can therefore contain credits contracted before August 2024. Without transaction or account-level transfer dates, the analysis cannot separate new demand from execution of existing commitments.

Second, the response may occur in price rather than quantity. Buyers may demand discounts for ineligible credits while continuing to retire similar volumes. This paper deliberately does not infer prices from annual reports. Transaction data would be required to test that margin.

Third, the CCP label may matter only to a subset of buyers. Firms following the VCMI Claims Code or facing external disclosure scrutiny may substitute toward approved credits, while other buyers continue purchasing inexpensive units. Aggregate project retirements combine these segments. A concentrated response among sophisticated buyers can disappear in an outcome without buyer identity.

Fourth, renewable credits were already viewed as a mature, relatively low-priced category. The Integrity Council's decision may have confirmed information that market participants had incorporated earlier rather than providing a new shock. In that case, the event date is institutionally important but economically anticipated.

Fifth, supply and category conditions changed simultaneously. The treated group contains many older wind and hydropower projects with large stocks, while comparison methodologies serve different niches. Project effects, category-year effects, registry-year effects, and supply controls address observable portions of this

problem but cannot recreate random assignment.

These mechanisms have different implications. Lagged contracts predict a delayed quantity response. Price adjustment predicts a greenium or discount without retirement changes. Buyer segmentation predicts heterogeneous effects linked to disclosure obligations. Anticipation predicts changes before August 2024. Future data can distinguish them, but the present annual registry panel cannot.

6.3. What the null result does and does not establish

The analysis rejects an unqualified claim that methodology rejection automatically redirects retirement demand. It does not reject the broader value of the Core Carbon Principles. Standards can improve practice by inducing programs to revise rules, by defining a common vocabulary for integrity, and by shaping future project development even when legacy retirements do not move immediately.

The intensity estimate is imprecise. Its 95 percent interval includes a sizable negative change, so a moderate demand penalty remains possible. The extensive-margin interval is tighter and excludes a large negative shift in the probability of any retirement, but it still permits small declines. The correct reading is a bounded null: no negative effect is detected, the point estimates are not negative, and the data are informative enough to reject some large extensive-margin penalties, while leaving important economic responses unresolved.

This interpretation also avoids a common error in null-result research. Failure to reject zero is not proof that the true effect is zero. Conversely, a theory should not be protected from evidence by treating every nonsignificant result as uninformative. Here the combination of near-zero point estimates, stable robustness results, and nonnegative 2025 coefficients weakens the prediction of an immediate retirement penalty. The width of the intervals limits how far that weakening can go.

7. Implications for governance and data design

The first implication concerns claim eligibility. A voluntary label changes behavior most reliably when buyers need it for a valued use. If firms may make equivalent offsetting claims with credits that failed the integrity assessment, rejection informs buyers but does not change the feasible set. Claiming frameworks, procurement standards, and disclosure rules can convert quality from a preference into an eligibility condition. This is the institutional mechanism by which C3 can be created rather than presumed.

The second implication concerns legacy credits. The two-tick system primarily determines which credits may carry the CCP label. It does not cancel rejected credits or impose buyer liability for using them. A transparent market should distinguish at least three statuses: credits eligible for the label, credits assessed and found not to meet the threshold, and credits not yet assessed. Buyers should disclose the status and methodology version of credits used in claims. That disclosure would make continued reliance on adversely assessed credits observable.

The third implication is data infrastructure. The Berkeley database makes cross-registry project analysis possible, but program records do not yet provide all fields required to evaluate the market response. Public records should preserve methodology-version histories, effective dates of transitions, credit-level CCP status, dates on which labels were applied, retirement transaction dates, and anonymized buyer categories. Landfill-gas approvals also require project capacity, which is not consistently available in the consolidated workbook. These are governance variables, not merely research conveniences: without them, the market cannot audit whether integrity determinations affect use.

The fourth implication is that price transparency and quantity transparency are complements. A retreat from transaction-level price analysis in this paper is not a reason to ignore price. It demonstrates why claims about a governance premium require accessible observations rather than figures reverse-engineered from aggregate charts. Registries and trading platforms could publish privacy-preserving transaction summaries linked to methodology, vintage, volume, and CCP status. Such data would permit direct tests of whether rejected credits receive discounts even when retirement quantities remain stable.

Finally, oversight bodies should evaluate their impact rather than infer it from label issuance. Useful indicators include the share of newly issued eligible credits actually carrying the label, changes in retirement shares after decisions, the buyer composition of labeled and rejected credits, and program corrections induced by assessment. Publishing these measures would turn private governance from a rule-making exercise into an auditable intervention.

8. Limitations

The study has a few principal limitations. First, annual retirements are temporally coarse. The August 2024 decision cannot be separated from the pre-decision portion of that calendar year, and 2025 provides only one complete post-decision observation. The June 2026 data are partial and cannot be annualized confidently without knowing retirement seasonality.

Second, retirement is a lagged measure of realized use, not a purchase date. Existing contracts and inventories can attenuate an immediate response. The design cannot determine whether transaction prices or new purchases changed.

Third, treatment is not random. Rejected methodologies differ substantively from comparison methodologies, and the treated group is much larger. Project fixed effects and flexible time effects address stable differences and broad category shocks but not unobserved project-specific changes correlated with assessment status.

These limitations narrow the conclusion but do not erase the contribution. The paper observes real projects, real issuance, and real retirements; it applies the official decision rules as far as public fields allow; and it reports the result that the data support rather than the positive premium anticipated in the original outline.

9. Conclusion

The Integrity Council's 2024 methodology decisions created a public test of whether external integrity governance changes carbon-credit demand. Using 9,833 at-risk observations from 1,448 renewable-energy projects, this paper finds no detectable negative retirement response in the first complete year after eight methodology families failed the CCP assessment. The preferred 2025 coefficients are close to zero for retirement intensity and slightly positive for the probability of any retirement. Balanced-panel, single-methodology, shorter-window, and no-supply-control specifications lead to the same conclusion.

The result is best understood as a failure of automatic transmission, not as a failure of governance. The Integrity Council produced a differentiating judgment. The public registry record does not show that buyers promptly converted that judgment into a retirement penalty. Information about integrity and demand for integrity are separate institutional achievements.

This distinction strengthens the central claim of Zhang (2026). Credible measurement and external

assessment cannot by themselves create structural integrity. Buyer-side rules must make quality consequential. Until methodology status affects claim eligibility, disclosure obligations, liability, or another valued use, rejected credits can remain economically active even after their governance weaknesses become public.

Future releases will add post-decision years and may improve methodology-version and label fields. Transaction data could reveal a price response hidden from retirements, and buyer-linked records could identify the segments in which quality is demanded. For now, the evidence supports a restrained conclusion: the market received the signal, but the first full year of public retirement data does not show that it acted on that signal in the predicted direction.

Disclosure statement

The author declares no conflict of interest.

References

[1] Integrity Council for the Voluntary Carbon Market, 2024, Integrity Council Announces First High-Integrity CCP-Labelled Carbon Credits, as Assessments Continue. 6 June. https://icvcm.org/integrity-council-announces-first-high-integrity-ccp-labelled-carbon-credits-as-assessments-continue/

[2] Integrity Council for the Voluntary Carbon Market, 2024, Carbon Credits from Current Renewable Energy Methodologies Will Not Receive High-Integrity CCP Label. 6 August. https://icvcm.org/carbon-credits-from-current-renewable-energy-methodologies-will-not-receive-high-integrity-ccp-label/

[3] Integrity Council for the Voluntary Carbon Market, 2024, Integrity Council Approves Three REDD+ Methodologies. 15 November. https://icvcm.org/integrity-council-approves-three-redd-methodologies/

[4] Zhang Y, 2026, Beyond Performative Transparency: Digital Technology, Governance Conditions, and the Legitimacy Crisis of the Voluntary Carbon Market. Journal of Applied Economics and Policy Studies 19(10): 203–219. https://doi.org/10.54254/2977-5701/2026.37218

[5] Akerlof GA, 1970, The Market for Lemons: Quality Uncertainty and the Market Mechanism. Quarterly Journal of Economics 84(3): 488–500.

[6] Spence M, 1973, Job Market Signaling. Quarterly Journal of Economics 87(3): 355–374.

[7] Bolton P, Freixas X, Shapiro J, 2012, The Credit Ratings Game. Journal of Finance 67(1): 85–111.

[8] Auld G, Gulbrandsen LH, 2010, Transparency in Nonstate Certification: Consequences for Accountability and Legitimacy. Global Environmental Politics 10(3): 97–119.

[9] Schleifer P, Fiorini M, Auld G, 2019, Transparency in Transnational Governance: The Determinants of Information Disclosure of Voluntary Sustainability Programs. Regulation and Governance 13(4): 488–506.

[10] Bartley T, 2007, Institutional Emergence in an Era of Globalization: The Rise of Transnational Private Regulation of Labor and Environmental Conditions. American Journal of Sociology 113(2): 297–351.

[11] Rosen S, 1974, Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy 82(1): 34–55.

[12] Conte MN, Kotchen MJ, 2010. Explaining the Price of Voluntary Carbon Offsets. Climate Change Economics 1(2): 93–111.

[13] Ecosystem Marketplace, 2024, On the Path to Maturity: State of the Voluntary Carbon Market 2024. Washington, DC: Forest Trends Association.

[14] Ecosystem Marketplace, 2025, Meeting the Moment: Renewing Trust in Carbon Finance. State of the Voluntary Carbon Market 2025. Washington, DC: Forest Trends Association.

[15] West TAP, Wunder S, Sills EO, et al., 2023. Action Needed to Make Carbon Offsets from Forest Conservation Work for Climate Change Mitigation. Science 381(6660): 873–877.

[16] Probst BS, Toetzke M, Kontoleon A, et al., 2024. Systematic Assessment of the Achieved Emission Reductions of Carbon Crediting Projects. Nature Communications 15: 9562.

[17] Swinfield T, Williams A, Coomes D, et al., 2026. Learning Lessons from Over-Crediting to Ensure Additionality in Forest Carbon Credits. Nature Communications 17: 3944.

[18] Delacote P, L'Horty T, Kontoleon A, et al., 2024. Strong Transparency Required for Carbon Credit Mechanisms. Nature Sustainability 7(6): 706–713.

[19] Delacote P, Chabé-Ferret S, Creti A, et al., 2025. Restoring Credibility in Carbon Offsets through Systematic Ex-Post Evaluation. Nature Sustainability 8(7): 733–740.

[20] Christiansen KL, 2025, Relegitimising the Voluntary Carbon Market: Visions of Digital Monitoring, Reporting and Verification. Environment and Planning A: Economy and Space 57(8): 1190–1205.

[21] Quartson P, Haya BK, Bernard T, et al., 2026, Voluntary Registry Offsets Database v2026-06. Berkeley Carbon Trading Project, University of California, Berkeley. https://gspp.berkeley.edu/berkeley-carbon-trading-project/offsets-database

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.