An Autopsy of the Voluntary Carbon Market

Ugo Panizza*, Francesco Tripoli†, and Beatrice Weder di Mauro*

*Geneva Graduate Institute & CEPR

†Harvard Business School

January 2026

Abstract

The voluntary carbon market (VCM) is built on the premise that one offset credit compensates one tonne of emissions. While a growing body of evidence has raised concerns about the environmental integrity of many credits, much less is known about how access to voluntary offsets affects firms' incentives to reduce their own emissions. Using a buyer-linked transaction dataset covering the near-universe of voluntary carbon offset retirements, we exploit the 2023 market integrity scandals as a quasi-natural experiment to study this question. We find that firms that reduce or exit voluntary offsetting following the shock subsequently achieve larger reductions in operational emissions intensity than firms that continue to rely on offsets. This pattern persists under conservative assumptions about credit effectiveness and is consistent with a moral-licensing mechanism in which access to offsets weakens incentives for internal abatement. Beyond this behavioral channel, we document persistent oversupply, intermediary opacity, and sharp price and volume responses to integrity shocks in the VCM.

Keywords: Voluntary Carbon Market; Carbon credits; Integrity controversy; Hedonic pricing and market sentiment.

JEL: G14, L11, L51, Q54

1 Introduction

In academic tradition, an *anatomy* offers a comprehensive and systematic exploration of a phenomenon or body of literature. Despite our original intention to follow this path and provide a detailed study of the voluntary carbon market (VCM), we were quickly confronted with a stark reality: we stand in front of a dead body. The market displays clear signs of systemic failure: disrupted price trends, stagnating credit retirements, and a persistent lack of transparency among market actors. Faced with this reality, we soon realized that this article would report the findings of an *autopsy* of the VCM rather than those of an anatomy. Under the title *An Autopsy of the Voluntary Carbon Market*, the paper offers a diagnostic investigation into the market's demise, which we trace to oversupply, integrity controversies, intermediary opacity, and, crucially, a moral hazard mechanism whereby firms use offsets as a license to delay or dilute genuine decarbonization.

The voluntary carbon market emerged in the late 1990s and early 2000s as a complement to regulated carbon markets created under the Kyoto Protocol (Kollmuss et al., 2008; Newell et al., 2013). Built on the premise that one credit compensates one tonne of emissions, the market expanded rapidly and became a prominent component of corporate net-zero strategies, especially for sectors with hard-to-abate emissions such as aviation, steel, and cement (Tzachor et al., 2023; Black et al., 2024). Yet, a growing empirical literature now challenges the market's core equivalence. Across major crediting categories such as tropical forest conservation, improved forest management (IFM), and household energy enhancement, independent studies have documented systematic over-crediting, weak safeguards against non-permanence, and methodological choices that bias credits upward (Haya et al., 2020; Badgley et al., 2022; West et al., 2020; Gill-Wiehl et al., 2024; Hanna et al., 2016; Song et al., 2025; Calel et al., 2025). If a large share of credits does not represent additional, durable tonnes, the environmental ledger that justifies their use vanishes.

Although integrity concerns had circulated among specialists for years, they entered the public domain forcefully between August 2022 and January 2023. In August 2022, the HBO programme *Last Week Tonight with John Oliver* devoted a full episode to carbon offsets, arguing that many credits represent "phantom" climate benefits and prompting unusually direct responses from major standard-setters.[1] The wave of media attention culminated in January 2023 with a joint investigation by *The Guardian–Die Zeit*–SourceMaterial on rainforest offsets,[2] which amplified academic critiques of leading REDD+ methodologies (West et al., 2020, 2023; Guizar-Coutiño et al., 2022). These revelations were followed by intensified NGO campaigns and greenwashing litigation against major buyers, one of the most known being the challenge to Delta's claim to be the "world's first carbon neutral airline."[3] As buyers confronted reputational risk and confidence eroded, prices fell and retirements plateaued, as we show in Section 5.1.

However, in this paper we take the autopsy metaphor seriously and therefore do not restrict our analysis to the causes behind the VCM's failure to deliver what it promised. We also examine what this failure reveals about the interaction between offsetting and firms' own decarbonization choices. Policy debates and NGO reports have long raised the concern that offsets may create a form of moral hazard or "moral licensing," whereby firms that purchase credits feel justified in emitting more or in postponing investments in emissions reductions (Broekhoff et al., 2019; Carton et al., 2020). In principle, well-designed guidance mitigates this risk by restricting the use of offsets to unavoidable, hard-to-abate residual emissions after deep value-chain decarbonization (SBTi, 2021). In practice, however, firms face strong economic incentives to rely on relatively inexpensive offsets rather than on more expensive decarbonization plans.

To date, this moral-hazard hypothesis has remained largely conjectural. We still lack systematic evidence on whether firms that rely on offsets actually emit more than otherwise similar firms that do not. This gap partly arises from limited availability of granular, firm-level data on offset purchases and partly from an inherent reverse-causality issue: increases in firms' emissions are likely to trigger increased offset purchases, making it challenging to determine whether offsets lead to higher emissions or merely respond to them.

[1]See "John Oliver puts punchlines over truth".

[2]See "Revealed: more than 90% of rainforest carbon offsets by biggest certifier are worthless, analysis shows".

[3]See *Berrin v. Delta Air Lines Inc.*, Columbia Law School.

We address this gap by exploiting a large, buyer-linked dataset of voluntary offset transactions, combined with firm-level environmental accounts and financial data. To address reverse causality, our empirical strategy leverages the 2022–2023 integrity shocks as a quasi-natural experiment. As we show in Section 5.1, this sequence of investigations constitutes a sharp negative shock to the perceived credibility of major crediting methodologies and to the reputational value of offsetting. In the wake of the shock, some firms exited the market while others continued offsetting on essentially the same terms as before. A salient example is EasyJet, a leading European low-cost airline, which in late 2022 announced that it would no longer rely on offset purchases and instead reallocate resources towards fleet renewal, operational efficiencies, and sustainable aviation fuels.[4] We treat the divergence between "exiters" and "stayers" as an endogenous but highly informative response to an exogenous integrity shock, and we compare their subsequent emissions trajectories.

We find that firms that stop purchasing offsets in the wake of the scandals reduce their operational greenhouse-gas emissions intensity by about 12.7% more than observably similar firms that continue to buy offsets over the same horizon. When we subtract all purchased offsets from reported emissions, under the conservative assumption that every credited tonne perfectly compensates one tonne of CO_2e, the estimated difference remains economically meaningful, at about 7.4%. In other words, even under a best-case assumption about offset quality, firms that continue to rely on offsets emit substantially more than firms that abandon the practice. To the best of our knowledge, this is the first empirical evidence that voluntary offsetting is associated with subsequent higher emissions, consistent with a moral-licensing mechanism. We further show that these effects are concentrated in operational (Scope 1) emissions intensity, with no robust differential adjustment in Scope 2 or total emissions, consistent with the idea that the primary adjustment margin is firms' own abatement effort rather than upstream or value-chain emissions.

Literature: Our analysis contributes to several strands of literature. First, we add to work documenting structural weaknesses in offset programmes and crediting methodologies

[4]See "EasyJet to stop offsetting CO_2 emissions from December".

by showing how these design flaws play out at the level of firms' realised mitigation behaviour, rather than solely in project-level additionality or generous accounting (Gill-Wiehl et al., 2024; West et al., 2020; Cabiyo et al., 2025; OECD, 2024; Stolz and Benedict, 2025; NewClimate Institute and Carbon Market Watch, 2024; Günther et al., 2020).

Second, we contribute to the literature on corporate net-zero strategies and climate governance by linking transaction-level data from the VCM to firm-level emissions and financial accounts, thereby opening the "black box" of how offsets are used within corporate mitigation strategies (Foerster, 2023; Tilsted et al., 2023; Boiral et al., 2024; Burger et al., 2022; Stolz and Benedict, 2025). One closely related paper is Kim et al. (2024), who study a sample of approximately 1,400 offset projects purchased by nearly 900 publicly listed firms. They show that high-emitting firms respond to an exogenous downgrade in ESG ratings by reducing emissions, while lower-emitting firms increasingly rely on cheaper, lower-quality offsets—particularly firms with weaker climate governance. Our work differs along three dimensions. First, we study a fundamentally broader slice of the market: our dataset covers the near-universe of voluntary carbon market projects (more than 36,000) and links them to up to about 17,000 firms. Second, we leverage a different source of variation: we focus on high-profile integrity scandals that triggered sharp shifts in beliefs about additionality and credit quality, and we study how these confidence shocks propagate through prices, volumes, and firms' offsetting choices. Third, we address a different research question: whereas Kim et al. (2024) primarily study which firms purchase offsets and how offset use responds to changes in ESG scrutiny, we use the integrity shock to examine how reliance on offsets shapes firms' subsequent decarbonization effort, providing evidence consistent with a moral-licensing channel.

Third, we bring fresh evidence to debates on rebound effects and moral licensing in environmental policy (Greening et al., 2000; Sorrell et al., 2009; Gillingham et al., 2016; Tiefenbeck et al., 2013; Burger et al., 2022). By exploiting shocks that reduce reliance on routine offsetting, we provide market-based evidence on the magnitude of emissions reductions associated with a shift away from offset use.

Layout: The remainder of this paper is structured in three main sections. First, in Section 2 we provide a primer on the VCM for readers unfamiliar with its structure and key participants. In Section 3 we characterize the market through stylized facts that document its architecture, persistent oversupply, and the role of intermediation in shaping prices and mark-ups. In Section 3, we show how the integrity shock associated with the 2022–2023 scandals manifested in collapsing prices and plateauing retirements. We then exploit this shock in Section 5.2 to estimate the effect of exiting the VCM on firm-level emissions and interpret the resulting gaps as evidence of moral hazard in the use of offsets. Section 6 concludes by discussing implications for corporate net-zero frameworks, the regulation of environmental claims, and the future role (if any) of voluntary offsets in global climate governance.

2 A Primer on the Voluntary Carbon Market

The voluntary carbon market (VCM) emerged in the late 1990s and early 2000s as a complement to compliance carbon markets created under the Kyoto Protocol (Kollmuss et al., 2008; Newell et al., 2013). At its core is the carbon credit: a transferable claim that one metric tonne of carbon dioxide equivalent (CO_2e) has been reduced or removed relative to a counterfactual baseline. Credits are generated by project developers, whose main activity is the design and implementation of projects intended to prevent, reduce, or remove greenhouse-gas emissions, and whose primary source of revenue is the sale of the credits they generate. Purchasers can retire these credits against their own emissions, with the stated objective of compensating residual emissions by financing mitigation activities undertaken by another actor.

Historically, the voluntary carbon market is closely linked to the Clean Development Mechanism (CDM). This project-based carbon offset system, established under the Kyoto Protocol, allowed emission-reduction projects in emerging and developing economies to generate Certified Emission Reductions (CERs), which advanced economies could use to meet their climate commitments. The CDM introduced many of the institutional features that later migrated into the VCM and, during the late 2000s, generated a large stock of CERs. Demand for these credits collapsed after the end of the first Kyoto commitment period and the tightening of CER eligibility in major compliance markets such as the EU Emissions Trading System. As a result, many CDM methodologies, project developers, and legacy credits transitioned—formally or informally—into the voluntary market. This legacy has shaped both the scale and structure of today's VCM, while also carrying over longstanding concerns about integrity.

Unlike compliance systems—such as the EU Emissions Trading System, China's national ETS, or California's cap-and-trade program—which impose legally binding emission caps and trade a fixed supply of allowances, the VCM is not driven by regulatory mandates. Instead, it allows firms and individual consumers to voluntarily offset or "compensate" emissions by purchasing credits. Lacking a centralized authority that defines uniform rules for project eligibility, credit issuance, or credit use, the VCM relies on third-party registries that develop methodologies, oversee project validation and verification, issue credits, and record their retirement. Each registry sets its own rules for eligibility, baselines, monitoring, and crediting, and maintains publicly accessible registries of issued and retired credits. Major registries include the American Carbon Registry, Gold Standard, Climate Action Reserve, and Verra.

In addition to registries, a large ecosystem of intermediaries has emerged around the VCM. Brokers, exchanges, and consultants facilitate transactions and market access, while specialized rating agencies—such as Calyx, Sylvera, and BeZero—assess credit quality. These assessments typically examine whether projects generate emissions reductions or removals that would not have occurred in the absence of credit revenues (additionality), whether emissions reductions in one location are accompanied by increases in emissions elsewhere (leakage), the durability of emissions reductions or removals over time (permanence), and the soundness of the methodologies used to quantify and verify outcomes. The coexistence of multiple registries, methodologies, and rating systems contributes to heterogeneity in credit quality and to opacity in market pricing.

Credits are generated by a wide range of project types that differ markedly in technologies, costs, and environmental risks. A useful organizing distinction is between emissions avoidance or reduction projects, which claim to reduce emissions relative to a baseline, and carbon removal projects, which aim to physically remove CO_2 from the atmosphere and store it over long horizons. Historically, the VCM has been dominated by avoidance-based credits, with removal credits representing a relatively small share of issuance and retirements.

Forestry and land-use projects constitute one of the largest categories by volume. These include afforestation and reforestation, improved forest management (IFM), and avoided deforestation projects, often grouped under REDD+. Such projects generate credits by increasing or preserving carbon stocks in biomass and soils relative to a counterfactual scenario. While they can produce large volumes of low-cost credits and offer biodiversity or development co-benefits, they are also associated with significant concerns regarding additionality, permanence, and leakage. Most forestry credits are best characterized as avoided-loss credits rather than permanent removals, as they rely on continued protection of carbon stocks over time.

Renewable energy projects—such as wind, solar, and small hydropower—were among the earliest and most prevalent sources of voluntary offsets. These projects generate credits by displacing fossil-fuel electricity generation, but their importance in the VCM has declined as renewable technologies have become cost-competitive and increasingly supported by public policy. Energy efficiency projects, which reduce emissions by lowering energy consumption through more efficient equipment or processes, share similar characteristics and challenges, relying heavily on modeled baselines and usage assumptions.

Household and community-level projects, including improved cookstoves, clean heating, and water purification, have historically generated a substantial share of credits, particularly in developing countries. These projects are typically classified as avoidance-based and are often justified by health and development co-benefits, but they face scrutiny over baseline inflation, monitoring difficulties, and the persistence of behavioral change. Methane abatement projects—such as landfill gas capture, wastewater treatment, and agricultural methane destruction—represent another major category by volume, reflecting methane's high global warming potential. Industrial gas destruction projects, which eliminate highly potent greenhouse gases from industrial processes, can also generate large quantities of credits, though they have long raised concerns about perverse incentives.

By contrast, carbon removal projects remain a small segment of the VCM. These include nature-based removals such as biochar and some soil carbon projects, as well as technological approaches like direct air capture with storage (DACCS) and bioenergy with carbon capture and storage (BECCS). Removal credits are typically more expensive and scarcer but are often viewed as conceptually closer to the "one tonne out" promise of offsetting, particularly when storage is durable and verifiable.

Grounded in these objectives, the VCM expanded rapidly, reaching a peak of roughly $2 billion in market value and approximately 500 $MtCO_2$ in transacted volume (Swinfield and Scott, 2025). Yet integrity concerns have intensified, and high-profile investigations and litigation have increased scrutiny of widely used methodologies and of the credibility of offset-based claims. As confidence eroded, both prices and transaction volumes fell sharply, setting the stage for the market integrity shock analyzed in Section 5.1.

3 Market Landscape: Supply, Demand, and Prices

3.1 Data

Our analysis builds on a comprehensive dataset from AlliedOffsets, which aggregates information from all major public registries and project developers and links credit retirements to identifiable buyers. The database covers 36,444 offsetting and removal projects worldwide and contains detailed project-level information, including registry and standard, sectoral classification, host country and geographic coordinates, crediting period, credit type, and additional attributes such as methodology and accreditation program. For each issuance, transfer, and retirement event, the dataset records the project identifier, the volume of credits, the transaction date, buyer information, and—when available—an estimated unit price derived from broker quotes, trading platforms, legacy and crypto exchanges, and other secondary sources.

These data allow us to reconstruct both the evolution of aggregate supply and demand and the composition of project portfolios at the level of individual buyers. Throughout the paper, we focus on the period from 2006—when retirement volumes become non-negligible and registry coverage relatively complete—to 2025. Transactions are classified as *issuances*, *retirements*, or *other* (including registry cancellations and offtakes), and we construct project-level and buyer-level aggregates by summing the corresponding volumes. On the buyer side, AlliedOffsets links retirements to corporate entities using a combination of registry account names, buyer metadata, and external corporate identifiers, yielding a highly granular transaction-level dataset that tracks the credit purchases of individual firms over time.

We link these data to two additional sources of firm-level information. First, we obtain greenhouse-gas (GHG) accounts for listed and large private companies from the S&P Trucost environmental dataset. Trucost reports scope 1 (direct), scope 2 (purchased energy), and scope 3 (value-chain) emissions, both in absolute terms and as emissions intensities measured per unit of revenue. The database combines firm-reported emissions with model-based estimates for companies that do not disclose complete inventories and has been widely used in the academic literature (see, among others, Bolton and Kacperczyk, 2021, Bolton and Kacperczyk, 2022, and Ilhan et al., 2021). We focus on emissions intensities (tCO_2e per unit of revenue) for scopes 1, 2, and 3, which we match to AlliedOffsets buyers using firm names and identifiers. Second, we characterize the financial and sectoral attributes of buyers using Compustat IQ, which provides balance-sheet and income-statement variables and industry classifications.

To begin exploiting this rich data structure, in the remainder of this Section, we move beyond aggregate volumes and characterize, in turn, the supply of credits (§3.2), the evolution of registered issuances and retirements, the composition and geography of demand (§3.3), and the credit pricing (§3.4).

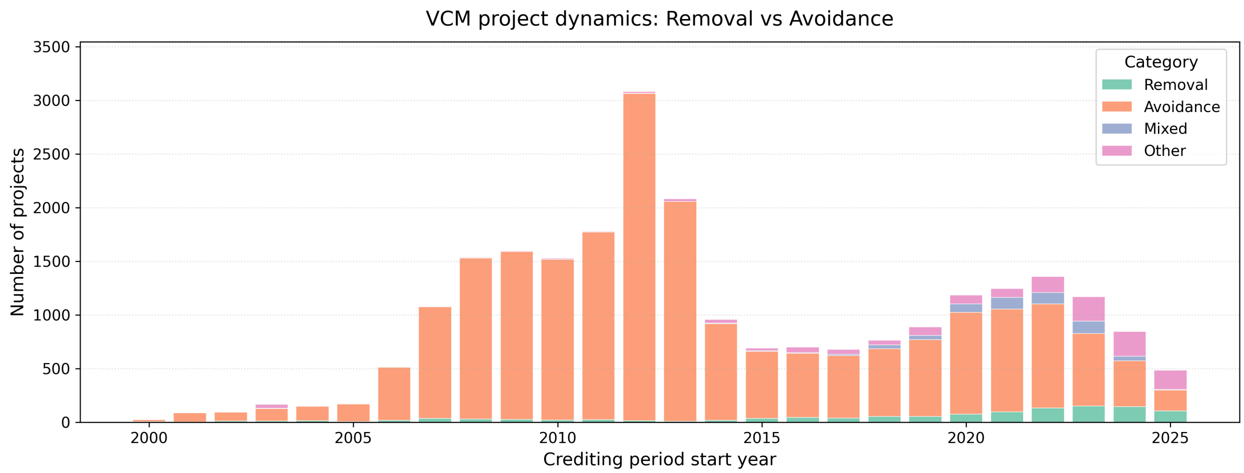


Figure 1: **VCM project dynamics by crediting-period start year.** The stacked bars show two waves of project origination and the dominance of avoidance over removal projects in the supply pipeline.

3.2 Supply-side patterns

As practice, we start the victim's autopsy from a comprehensive description of its body. Figure 1 traces the evolution of the project pipeline by crediting-period start year and project type. Two features are immediate. First, project origination expanded rapidly from the mid-2000s to the early 2010s and then experienced a second, smaller wave after 2018. Second, throughout both waves, the pipeline is dominated by avoidance projects, while removal activities account for only a marginal part of total supply.

This composition matters for what carbon "offsetting" can achieve. A high-integrity removal credit corresponds to a physical withdrawal of CO_2 (or equivalent GHGs) from the atmosphere, for example through afforestation or soil carbon sequestration. Retiring one such credit for each tonne emitted could, in principle, restore the atmospheric stock to its counterfactual level. Avoidance credits, by contrast, are issued when a project is deemed to emit less than a baseline, for example when a renewable plant substitutes for fossil fuel generation or an improved cookstove replaces traditional wood burning. Although retiring an avoidance credit prevents an additional tonne from being emitted elsewhere, it does not "undo" the tonne emitted by the buyer. If we think of this in terms of a global carbon "balance sheet", the use of avoidance credits moves the bottom line from two tonnes to one, not from one tonne to zero. The dominance of avoidance projects in Figure 1 therefore implies that, even under optimistic assumptions about additionality and permanence, the VCM has primarily financed the avoidance of incremental emissions rather than the removal of carbon already in the atmosphere.

Supply is also highly uneven in geographic space. Figure 2 maps registered projects by sector. Project activity is concentrated in China, India, Central America and the eastern United States. Forestry and land-use projects cluster in tropical forest regions, notably the Amazon basin, Central America, the Congo basin and Southeast Asia, while renewable-energy projects concentrate in fast-growing emerging economies such as India, China, Brazil and South Africa. Europe and North America host a mix of renewable, forestry and industrial projects, but on a smaller scale relative to their economic size, reflecting the availability of alternative regulatory instruments and the historical role of the Clean Development Mechanism (CDM).

Not all registered projects translate into retired credits. Between 2006 and 2011, developers issued carbon credits amounting to roughly 1 GtCO2e, yet almost none had been retired. By 2020, cumulative issuances had increased to about 3 GtCO2e, while cumulative retirements amounted to only around 0.6 GtCO2e. By 2025, cumulative issuances had reached approximately 5 GtCO2e, leaving a gap of nearly 3.5 GtCO2e between issued and retired credits (Figure 3).

If all projects were of uniformly high quality and integrity, such oversupply could be beneficial, allowing environmentally motivated consumers and firms to access a large market at relatively low cost. In the presence of substantial heterogeneity in offset quality, however, oversupply is likely to give rise to a dynamic akin to Gresham's law, whereby lower-quality credits drive higher-quality ones out of the market. Buyers can claim large volumes of offsetting at low cost by purchasing cheaper, lower-integrity credits, weakening demand for more credible and costly projects. As a result, firms may engage in greenwashing by selecting projects that best fit budget constraints or marketing narratives rather than those that maximize genuine climate impact.

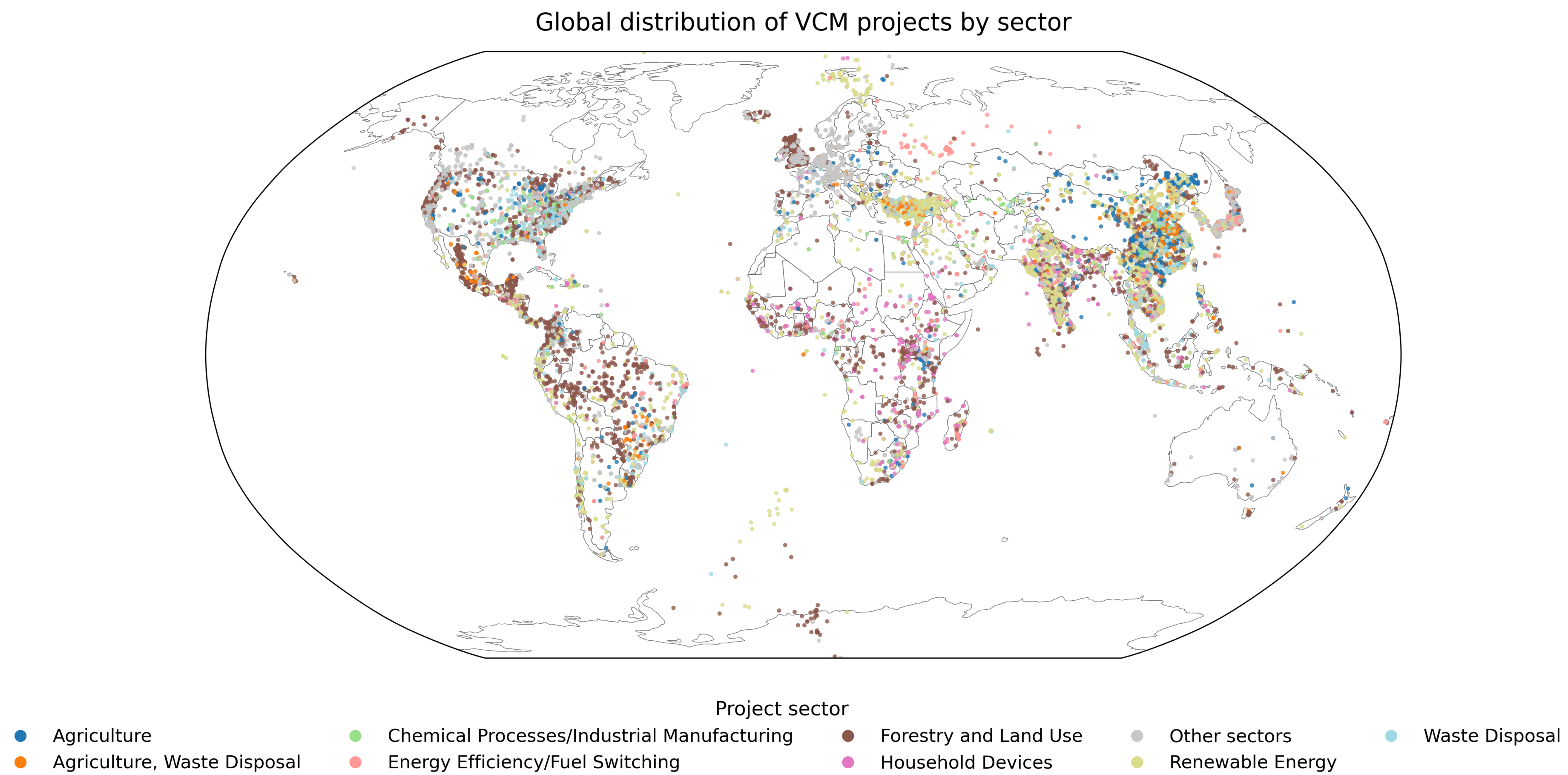


Figure 2: **Global distribution of VCM projects by sector.** Each point denotes a registered project, colored by sector. Supply is concentrated in a limited set of host countries, especially in emerging economies and tropical forest regions.

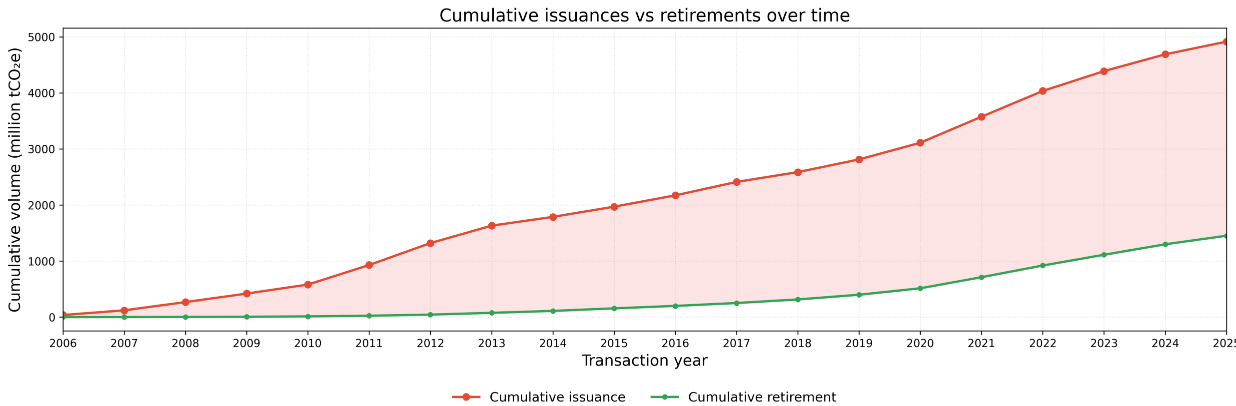


Figure 3: **Cumulative issuances and retirements in the VCM, 2006–2025.** Cumulative issuances (upper line) far outpace cumulative retirements (lower line), generating a stock of more than 3 $GtCO_2e$ of outstanding, unsold credits.

3.3 Demand-side patterns

Having characterized what is supplied, we now turn to who buys credits and which projects ultimately supply retired volumes. The main buyers are firms in carbon-intensive and consumer-facing sectors, such as airlines, energy utilities, materials and heavy manufacturing, as well as financial institutions and large technology firms with notable climate commitments. Even among these buyers, removal credits account for only a small fraction of portfolios; most volumes come from renewable-energy, forest-conservation, waste-management and household-energy projects, mirroring the supply composition in Figure 1. The geographic distribution differs markedly: some buyers purchase almost exclusively from projects located on the same continent as their headquarters, while others, especially European firms, source from projects across all major regions.

Taking the complementary perspective of projects, a small number of large, often flagship projects account for a disproportionate share of total retirements. Many of these are renewable-energy or industrial installations projects registered under the CDM or Verra, with long crediting periods and large installed capacities. Others are large-scale forestry and land-use projects that attracted substantial corporate demand in the late 2010s and early 2020s, when biodiversity co-benefits rose to prominence.

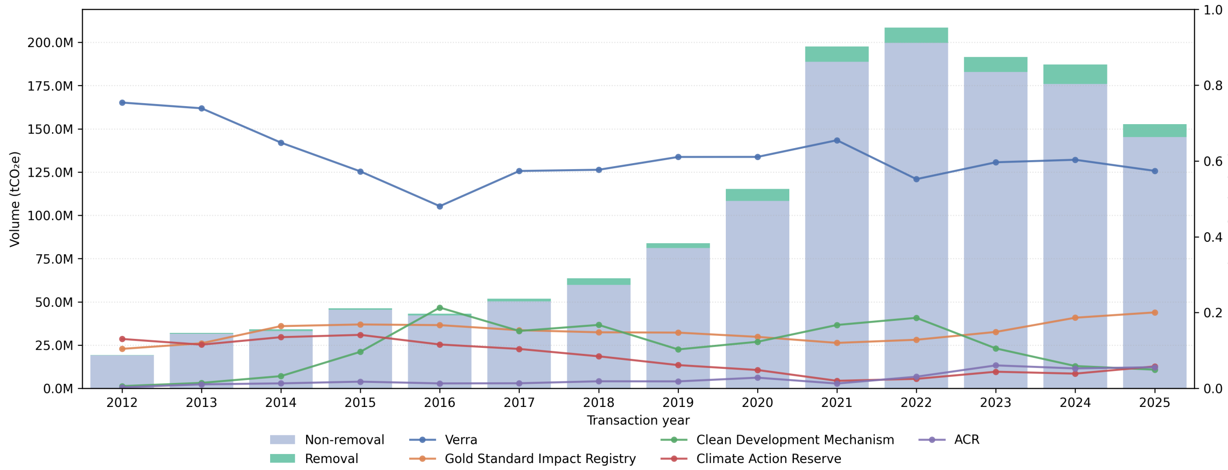


Figure 4: **Retirements and registry market shares over time.** Bars show annual retired volumes split into removal and non-removal credits; lines report the share of each major registry. Demand booms in 2019–2021 and then plateaus as integrity scandals erode confidence in the market.

The time profile of demand is summarized in Figure 4, which plots annual retirement volumes and registry market shares. We immediately notice that retirements grew rapidly between 2019–2021 before plateauing after the 2022–2023 integrity scandals. On the registry side, although early retirements were dominated by CDM credits, reflecting the carry-over of Kyoto-era Certified Emission Reductions into voluntary portfolios, after the end of the first Kyoto commitment period, CDM retirements decline and Verra's Verified Carbon Standard expanded rapidly, becoming the main registry. Gold Standard, the American Carbon Registry and domestic schemes such as the Australian Carbon Credit Units play smaller but non-negligible roles.

Finally, Figure 5 links buyers to project geography. Each point corresponds to a project location and is annotated with the most frequent (mode) ISO country code of buyers' headquarters. Buyers are predominantly headquartered in Europe, North America, East Asia and Oceania, and they tend to purchase from projects either in their own region or in a limited set of preferred host regions. North and South American firms primarily source credits from projects within their respective continents; African and Oceanian buyers display a similar regional preference, though less pronounced. European buyers, by contrast, show a highly diversified sourcing pattern, with credits originating from projects on all major continents.

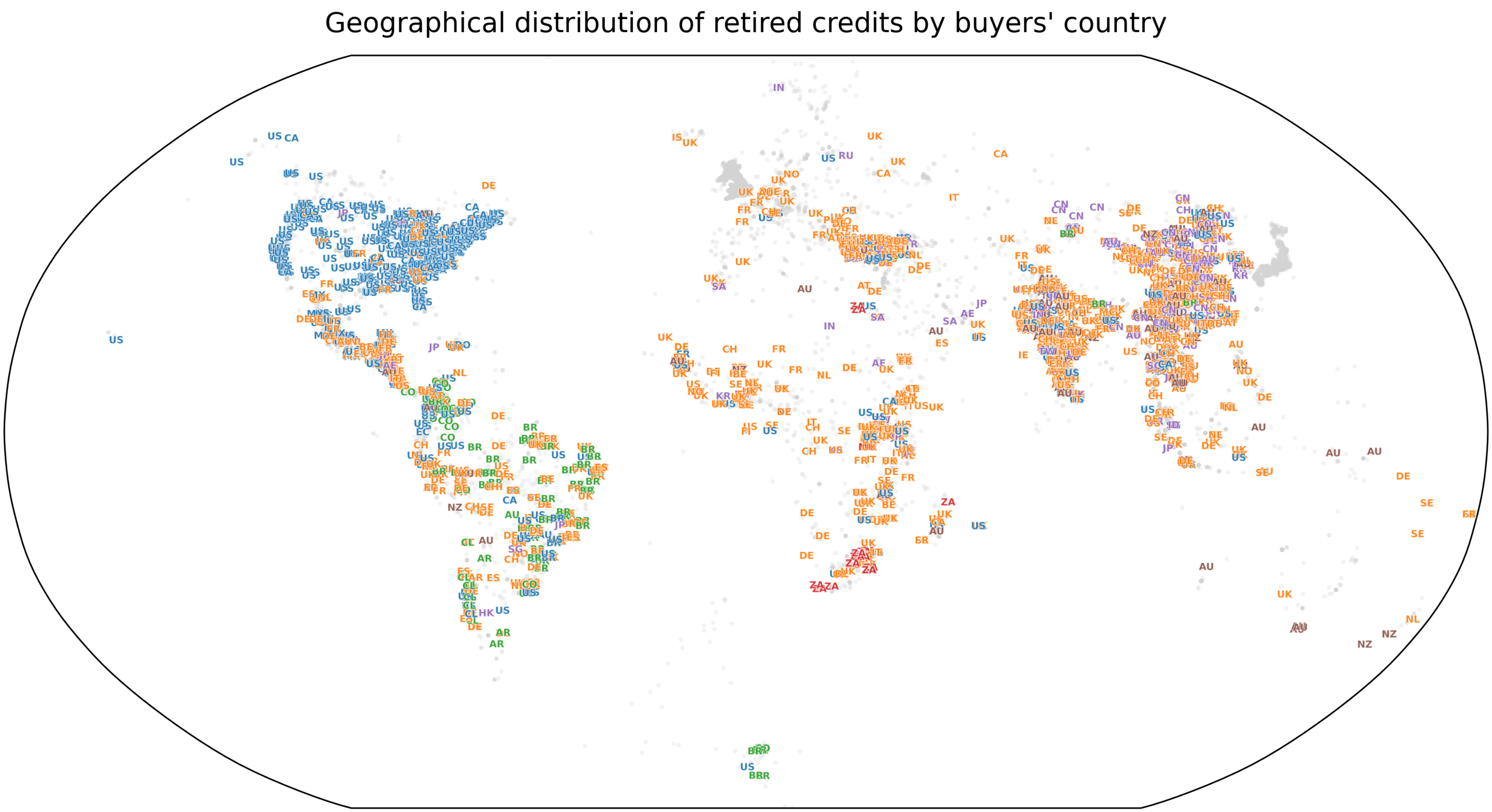


Figure 5: **Geographical distribution of retired credits by buyers' headquarter country.** Each label marks a project and indicates the modal headquarter country code among buyers retiring credits from that project. Clusters reveal strong regional sourcing for most buyers, with European firms displaying the most globally diversified portfolios.

Buyers headquartered in East Asia exhibit an intermediate behavior, combining strong regional sourcing from China and Southeast Asia with selected purchases elsewhere. PCA results shown in the online Appendix confirm these patterns: buyers from the Americas, Africa and Oceania form distinct clusters in the space of project characteristics, while European buyers are dispersed throughout, consistent with more diversified portfolios.

Taken together, the supply- and demand-side facts portray a market currently characterized by a considerable oversupply of credits, and cross-regional flows that are shaped by registry choices and buyer geography. We now move to analyzing the pricing implications of such a market structure.

3.4 Credit pricing patterns

We now turn from quantities to prices and analyze how the market values different types of credits. To do so, we exploit AlliedOffsets' estimates of credit prices, obtained by combining hundreds of weekly samples of bid, ask and transaction prices from brokers, exchanges and developers with project–level characteristics in a machine–learning model that "nowcasts" the price of credits. AlliedOffsets reports that the pricing model explains about 97% of the variation in observed sample prices ($R^2 \approx 0.97$) and directly observes prices for about 50% of retirements in the last three years. Thus, we restrict the analysis on prices to the 2022–2025 period since estimates for this range are more reliable due to the observed prices. Nevertheless, we still treat the AlliedOffsets prices as high–quality but model–based estimates rather than direct transaction records. As a result, we interpret the findings in this subsection as descriptive and do not attempt to recover structural demand elasticities. To further mitigate measurement error, we exclude observations that the provider flags as having low confidence.

As a first step, Figure 6 provides an high–level view of how prices evolved across project types. We rely on price indexes constructed as volumes-weighted price averages of the largest projects in a given segment of the market. The overall *AlliedOffsets 500* index is

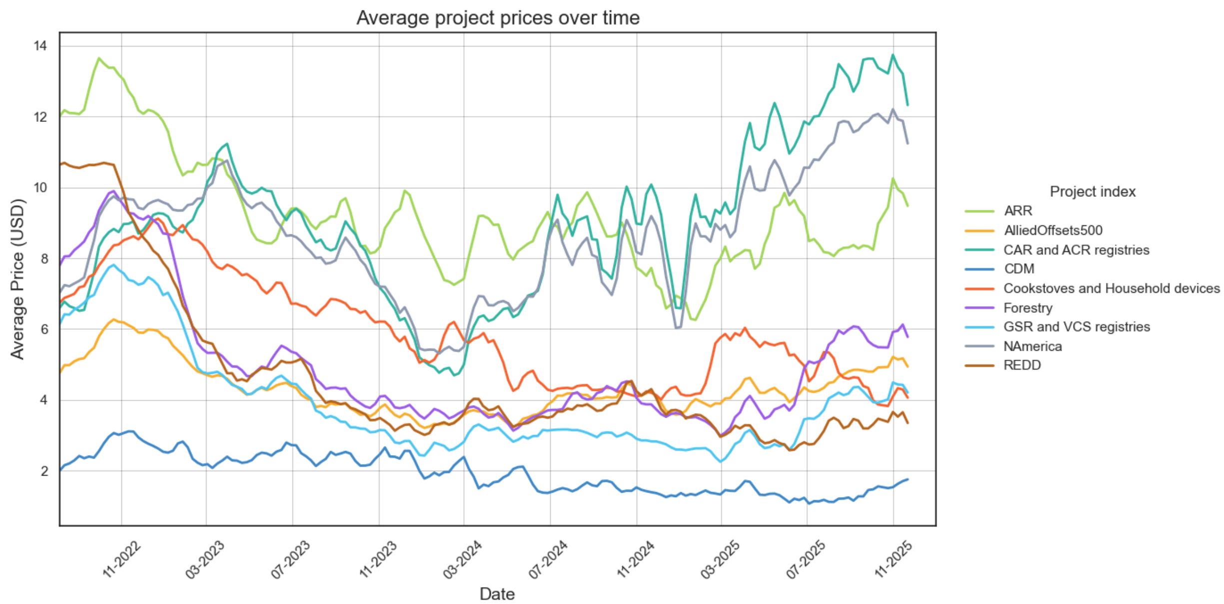


Figure 6: **Estimated price indexes by project type, 2022–2025.** Each line shows the average weekly prices of project indexes constructed from the largest projects by retirements in a given category.

defined as the weighted average of the top 500 projects, serving as a market–wide benchmark. The remaining series track specific groups such as North American projects, afforestation–reforestation (ARR), forestry credits, etc. Three patterns emerge. First, the market is clearly segmented into two broad price tiers. Credits on CAR and ACR, North American projects and nature-based removal (ARR) projects consistently traded at higher levels than the overall benchmark, while legacy CDM projects and many land–use removal credits remain in a lower price band. Second, CDM prices are persistently depressed and decline slowly over the sample, consistent with their legacy status and limited role in contemporary offsetting. Third, across most non–CDM segments prices display a pronounced common downturn in late 2022 and early 2023. The fall is particularly sharp for forestry and Gold Standard- or Verra-certified credits, whose prices drop abruptly around January 2023, immediately after the joint *Guardian*–*Die Zeit*–SourceMaterial investigation into rainforest offsets. This common movement anticipates the consequences of the integrity shock that we analyze in the next section: the price indexes already suggest that the scandals were absorbed as a market–wide negative revaluation of the most abundant credit types.

To move beyond aggregate indexes, we next examine how prices differ across projects with different characteristics. Table 1 reports ordinary least squares regressions of the logarithm of the estimated price on project attributes. The estimates confirm the visual patterns from Figure 6. Conditional on other covariates, projects on CAR and ACR have substantially higher prices than comparable CDM or VCS projects, while credits on some other registries trade at a discount. Nature–based removal projects, on the contrary, receive a premium. Furthermore, different project methodologies trade at distinct prices, with renewable energy credits being the cheapest option and Biochar the most expensive. More recent vintages are associated with higher prices, in line with buyer preferences for credits that are closer in time to retirement. Finally, offsets exhibit decreasing returns to scale since larger projects result in cheaper credits.

We then go one step further by using a Shapley–value decomposition to assess how much of the explained variation in log prices can be attributed to different groups of covariates. Using this method, the contribution of a given group (for example, project technology and geography versus intermediary identifiers such as developer, registry, or verifier) is defined as its average marginal impact on the model's fit when added across all possible orderings of covariate groups (Shapley, 1953). This results in a model–agnostic allocation of the regression R^2 across sources of variation. [5]

The results of the decomposition are reported in Table 2. Project–side attributes such as technology, vintage, and host country account for a large share of the explained variance, as one would expect if prices broadly reflect underlying mitigation opportunities and risk profiles. However, intermediary identifiers also explain a non–trivial fraction of the variation, amounting to an overall 47.43%. We interpret this as suggestive evidence that branding,

[5] Formally, let $\mathcal{G}$ denote the set of groups of regressors, and let $v(S)$ be the value function that maps any subset $S \subseteq \mathcal{G}$ to the regression fit when only variables in S are included (in our case, $v(S)$ is the R^2 from the corresponding specification). The Shapley value for group $g \in \mathcal{G}$ is

$$\phi_g = \sum_{S \subseteq \mathcal{G} \setminus \{g\}} \frac{|S|! \left(|\mathcal{G}| - |S| - 1\right)!}{|\mathcal{G}|!} \Big[v(S \cup \{g\}) - v(S) \Big],$$

which averages the marginal contribution of g to the fit over all possible subsets S that could precede it. By construction, the Shapley values satisfy $\sum_{g \in \mathcal{G}} \phi_g = v(\mathcal{G})$, so that the total explained variance is exactly decomposed into additive contributions from each group.

Table 1: **Determinants of Carbon Credit Prices**
This table reports a set of OLS regressions where the dependant variable is the log of the estimated offset price and the explanatory variable are project and issuer characteristics. Some levels of the factor variable have been omitted for space constraints.

	log(price)			
	(1)	(2)	(3)	(4)
Registry (ref: ACR)				
BioCarbon Standard	-0.906***	-0.565***	-0.545***	78.771***
	(0.100)	(0.107)	(0.107)	(5.249)
Clean Development Mechanism	-1.516***	-1.235***	-1.276***	78.018***
	(0.084)	(0.080)	(0.093)	(5.250)
Gold Standard	-0.381***	0.020	0.019	-0.018
	(0.071)	(0.072)	(0.072)	(0.065)
Verra	-0.847***	-0.485***	-0.469***	-0.416***
	(0.067)	(0.070)	(0.070)	(0.064)
Woodland Carbon Code	0.206**	0.238**	0.240**	
	(0.089)	(0.101)	(0.100)	
Project type (ref: avoidance/reduction)				
Mixed Avoidance/Removal	0.386***	0.389***	0.379***	0.301***
	(0.062)	(0.057)	(0.057)	(0.056)
Removal	0.626***	0.573***	0.577***	0.618***
	(0.079)	(0.080)	(0.079)	(0.081)
Project sector (ref: household devices)				
Biochar	3.387***	3.834***	3.821***	3.725***
	(0.092)	(0.100)	(0.099)	(0.102)
Renewable energy	-0.733***	-0.421***	-0.433***	-0.341***
	(0.022)	(0.043)	(0.043)	(0.040)
Utilization	2.973***	2.763***	2.752***	2.551***
	(0.088)	(0.101)	(0.100)	(0.100)
Waste disposal	-0.184**	-0.105*	-0.109**	-0.013
	(0.072)	(0.055)	(0.054)	(0.048)
Geography (ref: Asia)				
Africa		0.389***	0.390***	0.348***
		(0.043)	(0.044)	(0.041)
Europe		0.708***	0.717***	0.688***
		(0.040)	(0.038)	(0.052)
North America		0.684***	0.683***	0.667***
		(0.048)	(0.048)	(0.042)
South America		0.348***	0.343***	0.389***
		(0.045)	(0.045)	(0.042)
Governance / quality				
Article 6 registered (True)			0.079	0.087
			(0.079)	(0.075)
CORSIA eligible (True)			0.094***	0.077***
			(0.027)	(0.026)
Scale (ref: large)				
Micro scale				0.152***
				(0.025)
Small scale				0.159***
				(0.020)
Letter of Authorization (True)				0.291***
				(0.103)
Vintage (year)				0.039***
				(0.003)
Intercept	2.392***	1.634***	1.626***	-77.812***
	(0.071)	(0.083)	(0.083)	(5.247)
Observations	23005	23000	23000	22242
R^2	0.719	0.752	0.753	0.745
Month FE	Yes	Yes	Yes	Yes

Standard errors in parenthesis are clustered at the project level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: **Shapley Decomposition of Credit Price Variation.**
This table reports the results of a Shorrocks–Shapley decomposition of the results of Column 4 of Table 1. Each entry reports the Shapley value (contribution to R^2) and the corresponding percentage of the explained variance. Percentages do not sum exactly to 100 due to time FE. Project characteristics group includes project type, sector, CORSIA approval, CCP likelihood, methodology, country, vintage, size, and Article 6 status.

Group	Shapley value (R^2)	Share of explained variance (%)
Developer	0.2707	33.09
Registry	0.0914	11.18
Verifier	0.1122	13.71
Project Characteristics	0.2324	28.41
Total	0.8181	100.00

reputation and market power among intermediaries contribute materially to observed spreads in carbon–credit prices.

4 A Toy Model of Offsetting and Moral Licensing

Before turning to the firm-level evidence, we pause to clarify why a "license to emit" can arise from a purely economic mechanism. The preceding sections documented that the voluntary carbon market is characterized by heterogeneity in credit types and substantial dispersion in prices. A natural concern — and the central behavioral hypothesis tested in this paper — is that access to relatively inexpensive offsets may relax the effective constraint firms face when managing their climate footprint, thereby crowding out costly internal abatement. Importantly, this mechanism does not hinge on whether credits are of low integrity. Even if one assumes the benchmark case in which offsets perfectly compensate emissions, a moral-licensing channel can arise on the demand side whenever the metric that is salient for external scrutiny is a *net* footprint and firms allocate scarce resources across alternative compliance instruments. The purpose of this section is therefore not to provide a full structural model of firm decarbonization. Rather, it offers a deliberately stylized framework, a "toy model", that isolates the economic force behind moral licensing and yields a transparent, testable implication that motivates the empirical strategy in Section 5.2.

Consider a firm that produces a fixed level of output (or revenue) $Y > 0$ within a period (e.g., a fiscal year). In that period, baseline gross emissions are $\bar{e}Y$, where $\bar{e} > 0$ is baseline emissions intensity. The firm can reduce emissions through *internal abatement*, which captures operational and process changes that lower emissions at the source. Denote by $a \geq 0$ the abatement effort undertaken within the period, which reduces gross emissions contemporaneously according to

$$E(a) \ = \ \bar{e}Y - \eta a, \qquad \eta > 0. \tag{1}$$

In addition, the firm can purchase and retire offsets $O \geq 0$ at unit price $p_O > 0$. Offsets do not change gross emissions mechanically, but they enter the *net* footprint measure that is salient in voluntary climate claims, reputational assessments, and internal KPIs:

$$\tilde{E}(a, O) \ \equiv \ E(a) - \gamma O, \qquad \gamma \in (0, 1]. \tag{2}$$

The parameter γ captures the perceived (or claimed) effectiveness of offsets in netting out emissions. The benchmark scenario in which "a ton is a ton" corresponds to $\gamma = 1$.

We assume that firms typically allocate a *finite* amount of resources to decarbonization within a reporting period. Let $B > 0$ denote the budget allocated to actions that can be used to manage the salient net footprint. The firm can spend this budget on internal abatement and on offsets:

$$C(a) + p_O O \ \leq \ B, \tag{3}$$

where $C(\cdot)$ is increasing and convex, with $C'(a) > 0$ and $C''(a) > 0$.

Further suppose that the firm has pledged to a salient *net-emissions-intensity* commitment

$$\frac{\tilde{E}(a, O)}{Y} \ \leq \ \tilde{\bar{e}}, \tag{4}$$

where $\tilde{\bar{e}}$ is the maximum net intensity consistent with the firm's public target, stakeholder expectations, or internal performance benchmark. The firm chooses (a, O) to meet (4) at minimum cost:

$$\min_{a\geq 0,\, O\geq 0} \; C(a) + p_O O \quad \text{s.t.} \quad \bar{e} - \frac{\eta}{Y}a - \frac{\gamma}{Y}O \;\leq\; \bar{\tilde{e}}. \tag{5}$$

When both instruments are used (an interior solution), the first-order conditions imply that the Lagrange multiplier $\nu > 0$ on the constraint satisfies

$$C'(a^*) = \nu\frac{\eta}{Y}, \qquad p_O = \nu\frac{\gamma}{Y} \;\;\Rightarrow\;\; C'(a^*) = \frac{\eta}{\gamma}\,p_O. \tag{6}$$

Equation (6) formalizes substitution: since $C'(\cdot)$ is increasing, a lower offset price p_O (or a higher perceived effectiveness γ) implies lower optimal abatement a^*. Intuitively, when offsets offer a relatively inexpensive way to improve the salient net metric, the firm optimally relies more on offsets and less on internal abatement. The licensing implication follows from a simple accounting point: when offsets are available and cheap, the firm can meet the *net* target (4) with relatively little internal abatement, and therefore it chooses a higher level of net emissions than it would choose if it were forced to rely on abatement alone under the same budget. To see this, define the *achieved* net intensity:

$$\tilde{e}(a,O) \;\equiv\; \frac{\tilde{E}(a,O)}{Y} \;=\; \bar{e} - \frac{\eta}{Y}a - \frac{\gamma}{Y}O. \tag{7}$$

Under (5), the firm will typically choose a bundle that *just satisfies* the constraint (targets are met but not over-achieved), so that

$$\tilde{e}(a^*,O^*) = \bar{\tilde{e}}. \tag{8}$$

Now consider the counterfactual in which offsets are unavailable (or no longer accepted), i.e. $O \equiv 0$, but the budget B is unchanged. Then the firm must satisfy the same target using only abatement:

$$\min_{a\geq 0} \; C(a) \quad \text{s.t.} \quad \bar{e} - \frac{\eta}{Y}a \;\leq\; \bar{\tilde{e}}. \tag{9}$$

If the target is feasible, the no-offset solution requires abatement $a^0 \geq \frac{Y}{\eta}(\bar{e}-\bar{\tilde{e}})$, implying that the firm must undertake *more internal abatement* than under offsetting whenever $O^* > 0$.

This simple model yields a direct empirical implication. Holding fixed pre-shock exposure to offsetting, an exogenous disruption that induces some firms to cease purchasing offsets (a discrete drop from $O^* > 0$ to $O = 0$) forces those firms to substitute toward internal abatement in order to continue meeting the salient net-emissions benchmark (4). Because a affects gross emissions contemporaneously through (1), this substitution predicts a discrete improvement in operational emissions outcomes for firms that exit the market relative to similar firms that continue offsetting. This prediction is the basis of the design in Section 5.2. Importantly, it does not rely on integrity failures: the substitution channel operates even under $\gamma = 1$.

5 Results

5.1 Integrity shock: the effect of a media scandal

As discussed in the Section 1, two waves of media scrutiny brought integrity concerns into the public eye. The first, in August 2022, was triggered by the HBO program *Last Week Tonight with John Oliver*. This was followed by a second shock in January 2023, when a joint investigation by *The Guardian*, *Die Zeit* and SourceMaterial argued that the vast majority of rainforest offsets issued under leading REDD+ methodologies did not correspond to real emission reductions. Given the price reactions that the preceding exploratory analyses documented, in what follows, we focus on the January 2023 investigation as our main "integrity shock", while noting that all results are robust to centering the analysis on the August 2022 episode (see online Appendix).

To quantify its impact on market outcomes, we estimate a regression discontinuity design (RDD) in time that exploits the sharp timing of the January 2023 investigation release (Imbens and Lemieux (2008); Cattaneo et al. (2020)). Let t index months and normalise $t = 0$ at January 2023, so that $t < 0$ denotes pre-shock months and $t > 0$ post-shock months. For an outcome y_t (either the average price of offsets or the total volume of retirements in month t), we estimate a RDD of the form

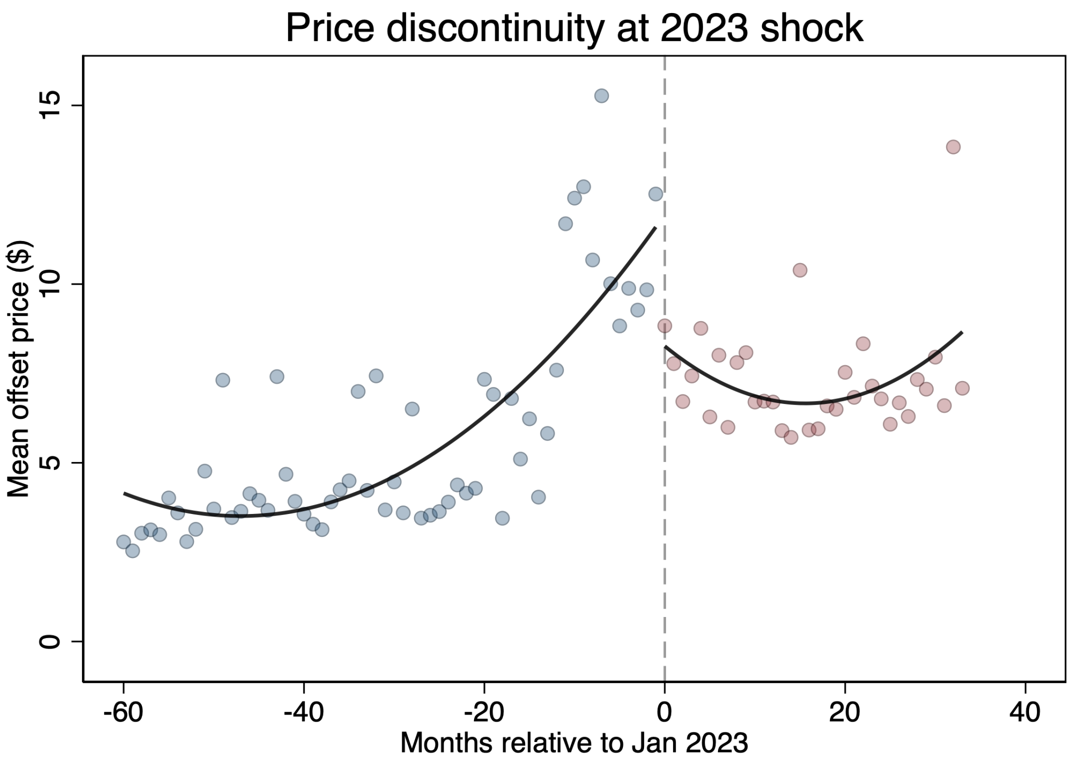
$$y_t = \alpha + \tau \mathbf{1}\{t \geq 0\} + f_-(t)\mathbf{1}\{t < 0\} + f_+(t)\mathbf{1}\{t \geq 0\} + \varepsilon_t, \tag{10}$$

where $\mathbf{1}\{\cdot\}$ is an indicator function, $f_-(\cdot)$ and $f_+(\cdot)$ are smooth functions capturing the pre- and post-shock time trends, and τ is the parameter of interest: the discrete jump in the conditional expectation of y_t at the discontinuity point.

Figure 7 summarizes the resulting patterns for both prices (upper panel) and retirement volumes (lower panel). Before the shock, average offset prices followed a rising trend. Immediately after January 2023, we observe a discrete downward shift: the fitted post-shock path lies well below the extrapolated pre-shock trend, and the point estimates imply an economically large reduction in average prices relative to the counterfactual. At the same time, the lower panel shows that total monthly retirements, which had been growing steadily during the 2019–early 2022 boom of the market, lose momentum after the shock. Although volumes did not collapse, they stopped increasing and instead plateaued. In other words, the integrity scandal appears to have been absorbed by the market both as a negative repricing of widely used credits and as a halt in the expansion of demand.

Interpreted through the lens of Section 3.4, these discontinuities are consistent with a revision of beliefs about credit integrity and heightened reputational risk for buyers. The timing closely matches the sharp decline in price indexes for forestry and Verra- or Gold Standard-certified projects. Importantly, the RDD framework allows us to separate the discrete impact of the integrity shock from the gradual trends in prices and retirements driven by broader macroeconomic and climate-policy conditions. Robustness checks centering the cutoff on the August 2022 John Oliver episode or using alternative polynomial specifications of the pre- and post-shock trend functions yield qualitatively similar conclusions, although the January 2023 cutoff delivers the clearest break in both prices and volumes.

These results provide the key to our identification strategy for firm-level behavior. The integrity scandal was not caused by changes in the emissions of individual companies, but by external investigations into credit methodologies and registry practices. Yet, the RDD



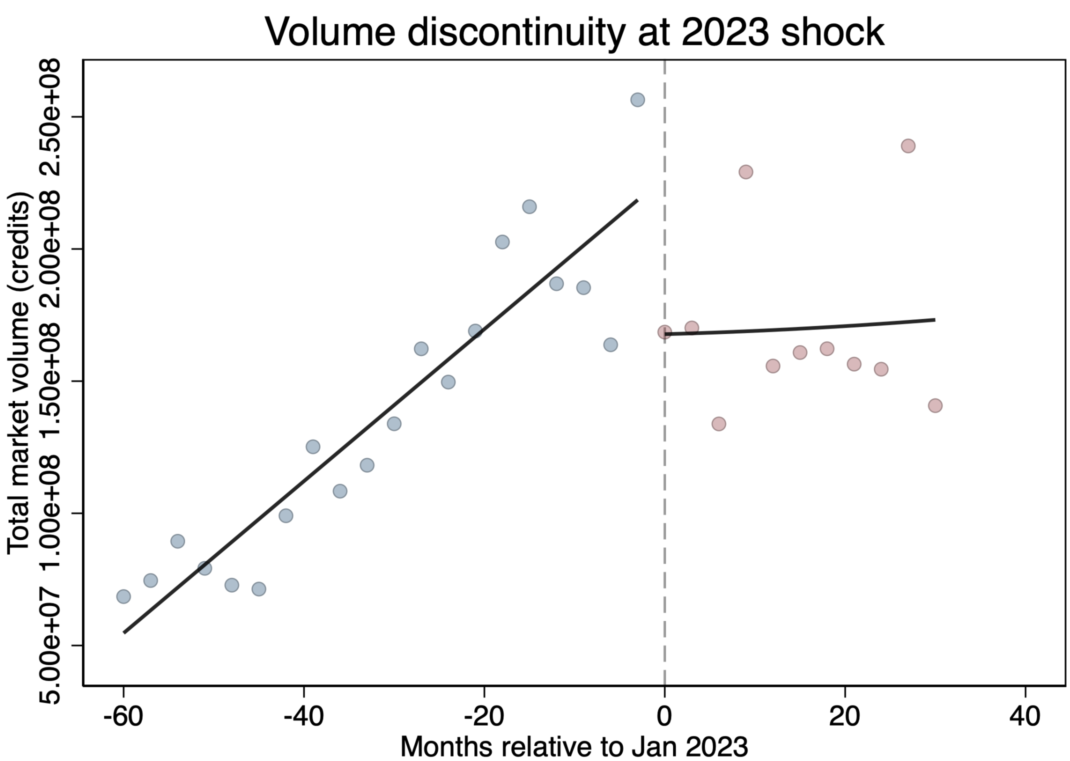


Figure 7: **Regression discontinuity in prices and retirements around the January 2023 integrity shock.** Dots show observed monthly outcomes; solid lines plot local–polynomial fits on either side of the cutoff.

evidence in this section establishes that the January 2023 event constitutes a sharp, plausibly exogenous shift in the perceived value of offsetting. Indeed, the shock led several firms to cease buying offsets altogether due to reputational concerns on the use of offsets. We can therefore use it as a quasi-natural experiment to classify firms into "exiters" and "stayers" and to compare their subsequent emissions trajectories. By conditioning on pre-shock emissions, sector, financial characteristics and offsetting histories, this strategy allows us to overcome the reverse-causality concern inherent in attempts to trace how access to, and reliance on, offsets affects firms GHG emissions.

5.2 Moral hazard evidence revealed by the integrity shock

The preceding subsection establishes that the January 2023 investigation constitutes a sharp and plausibly exogenous integrity shock to the voluntary carbon market: it triggered a discrete repricing of widely used credits and a plateau in aggregate retirements. We now leverage the same event to study firm-level behavior. Our goal is to test whether reliance on voluntary offsetting weakens incentives for internal decarbonization—a moral-licensing or "license to emit" mechanism—while addressing the key empirical challenge highlighted in the introduction: offset use is endogenous to emissions. Firms with higher emissions may mechanically buy more credits, so simple correlations between offsets and emissions confound behavioral responses with reverse causality.

Our identification strategy exploits the divergence in firms' offsetting choices following the January 2023 shock. We restrict attention to firms that were already active buyers prior to the shock (i.e., "VCM-exposed" firms), and we classify them based on their post-shock offsetting behavior after a two-month grace period. We define *exiters* as firms that were purchasing offsets before January 2023 and then stop retiring offsets after the grace period, and *stayers* as firms that were purchasing offsets before January 2023 and continue retiring offsets after the grace period. We then compare the evolution of emissions outcomes for exiters and stayers before versus after the shock. Because the shock originates in external scrutiny of credit methodologies and the reputational value of offsetting, rather than in changes in any single firm's emissions, it provides a quasi-experimental break that helps overcome the reverse-causality concern.

Formally, let i index firms and t index years. For each outcome y_{it}—measured as the log of an emissions-intensity metric from Trucost—we estimate the following two-way fixed-effects difference-in-differences specification:

$$y_{it} \;=\; \alpha_i \;+\; \lambda_t \;+\; \beta\left(\text{Post}_t^{2023} \times \text{Quit}_i\right) + \eta\;\text{Post}_t^{2023} \;+\; \varepsilon_{it}, \tag{11}$$

where α_i are firm fixed effects, λ_t are year fixed effects, Post_t^{2023} is an indicator equal to one in the post-shock period (year $\geq$ 2023), and Quit_i is an indicator for firms classified as exiters. The coefficient of interest is β, which captures the differential change in emissions intensity for exiters relative to stayers after the January 2023 integrity shock. Standard errors are clustered at the firm level to allow for serial correlation in firms' emissions outcomes over time (Bertrand et al., 2004).

Table 3 reports the main estimates. Column (1) uses log scope 1 emissions intensity as the outcome and yields $\beta = -0.1265$ (s.e. 0.0436), implying that exiters reduce operational emissions intensity by about 12.65% more than stayers after the shock.[6] This gap is difficult to reconcile with a best-case view in which voluntary offsetting is merely a neutral substitute for abatement: even if credit retirements fully and reliably "cancel" firms' own emissions one-for-one, then continuing to offset should not be associated with systematically higher realized operational emissions intensity than exiting, once we condition on firm fixed effects and common time shocks.

To probe this interpretation, column (2) considers an "adjusted" scope 1 measure in which we subtract firms' retired offsets from scope 1 emissions under the conservative assumption that each credited tonne perfectly compensates one tonne of CO_2e. Even under this best-case accounting, the estimated coefficient remains negative and economically meaningful ($\beta = -0.0736$, s.e. 0.0376), corresponding to an additional reduction of about 7.36% for

[6]Since the dependent variables are in logs, coefficients can conveniently be interpreted as approximate percentage changes.

Table 3: **Difference-in-Differences Estimates of GHG Emissions**

This table reports difference-in-differences estimates of greenhouse-gas (GHG) emissions intensities following the 2023 integrity shock. Columns (1), (2), (4)–(6) compare firms that *stopped* purchasing offsets after the shock to firms that continued offsetting. Column (3) focuses on firms those that *reduced* (rather than stopped) their offset purchases after the shock. The dependent variables are in all in logs. Robust standard errors, clustered at the firm level, are reported in parentheses.

	(1) Scope 1	(2) Adj. Scope 1	(3) Scope 1	(4) Scope 2	(5) Scope 3	(6) Total
Post shock × Quit	-0.1265^{***}	-0.0736^{*}		-0.0149	-0.0287	-0.0430
	(0.0436)	(0.0376)		(0.0386)	(0.0177)	(0.0275)
Post shock × Reduce			-0.0327			
			(0.0740)			
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	129,331	127,849	129,331	129,356	129,356	129,356

exiters. This is the key behavioral result: even when we grant offsets their maximum possible effectiveness in netting out operational emissions, firms that continue to rely on offsets display weaker improvements in operational emissions outcomes than firms that stop offsetting. Put differently, the "tonne is a tonne" assumption is not sufficient to eliminate the observed behavioral wedge between exiters and stayers, which is consistent with a moral-licensing mechanism in which access to offsets weakens incentives for internal abatement.

Column (3) sharpens this conclusion by distinguishing *exiting* from *reducing*. We define *reducers* as firms that, after the grace period, retire at least 70% fewer offsets than in the pre-shock period, and we estimate the same specification replacing Quit_i with an indicator Reduce_i. The estimated coefficient for reducers is statistically indistinguishable from zero ($\beta = -0.0327$, s.e. 0.0740). This pattern suggests that partial retrenchment in offsetting is not associated with a clear differential improvement in operational emissions intensity; instead, the economically meaningful shift occurs among firms that stop offsetting altogether. One interpretation is that firms that remain active buyers—even at reduced scale—continue to treat offsetting as part of their decarbonization portfolio, so the incentives to reallocate resources toward internal abatement are weaker than for firms that exit completely.

Consistent with our priors about controllability, columns (4) and (6) show no robust evidence of differential post-shock changes in scope 2 emissions intensity ($\beta = -0.0149$, s.e. 0.0386) or total GHG emissions intensity ($\beta = -0.0430$, s.e. 0.0275) between exiters and stayers. These outcomes reflect factors that are less directly under managerial control over short horizons (purchased-energy emissions for scope 2 and broader portfolio effects for total emissions), and the absence of a strong differential response is therefore not surprising. The effect is also not robust for scope 3 emissions intensity since the levers available to firms operate through complex value-chain relationships and are therefore outside of the firm direct control.

Finally, we examine heterogeneity patterns that are directly motivated by the mechanism emphasized above: if offsets are used to sustain climate claims rather than to complement internal abatement, then the post-shock reallocation toward abatement should be weaker among firms with a greater propensity to greenwash and among firms with fewer feasible abatement options. Table 4 reports two sets of cross-sectional splits using the same specification as equation (11). Panels A and B split firms by whether they ever experienced an environmental controversy during the sample window, using Refinitiv's controversy flag (a binary indicator recorded when a firm is impacted by at least one environmental controversy in a given year). Among firms with no recorded controversies (Panel A), the point estimates remain negative for both scope 1 and adjusted scope 1, but become imprecise—a loss of statistical power driven by the sharp reduction in sample size when matching to Refinitiv. Among firms with at least one environmental controversy (Panel B), the coefficient on adjusted scope 1 is positive and statistically significant ($\beta = 0.1930$, s.e. 0.1105). A natural interpretation is that these firms may have relied on offsets primarily to support claims about "net" performance; after the shock, they exit offsetting due to heightened scrutiny, but do not substitute toward internal abatement, so their adjusted (netted) operational measure deteriorates mechanically when offsets are no longer subtracted. Given the small number of observations in this split, we interpret this evidence with caution.

Panels C and D then split firms by sectoral abatement feasibility. Panel C (not hard-to-abate sectors) shows effects similar in sign and magnitude to the baseline for scope 1

Table 4: **Cross-Sectional Diff-in-Diff estimates across cross-sectional splits.**
This table reports difference-in-differences estimates of the impact of the 2023 integrity shock on firm-level emissions intensity, estimated separately across cross-sectional splits. Panel A restricts the sample to firms that never experienced an environmental controversy, while Panel B restricts the sample to firms that experienced at least one controversy. Panel C restricts the sample to firms in sectors that are *not* hard-to-abate, and Panel D restricts the sample to firms in hard-to-abate sectors. Dependent variables are in logs and robust standard errors, clustered at the firm level, are reported in parentheses.

	(1) Scope 1	(2) Adj. Scope 1	(3) Scope 2	(4) Total GHG
Panel A: No Environmental Controversies				
Post shock × Quit	−0.0989	−0.0354	0.0926	0.0661
	(0.1101)	(0.0617)	(0.0990)	(0.0491)
Observations	3,407	3,406	3,426	3,426
Panel B: Environmental Controversies				
Post shock × Quit	0.2052	0.1930*	0.1225	0.0941
	(0.1560)	(0.1105)	(0.2835)	(0.1021)
Observations	515	515	515	515
Panel C: Not hard-to-abate sectors				
Post shock × Quit	−0.1178**	−0.0590	−0.0107	−0.0351
	(0.0463)	(0.0393)	(0.0401)	(0.0291)
Observations	107,841	106,453	107,866	107,866
Panel D: Hard-to-abate sectors				
Post shock × Quit	−0.1993*	−0.1893*	−0.0491	−0.1022
	(0.1092)	(0.1031)	(0.1294)	(0.0644)
Observations	21,141	21,046	21,141	21,141

($\beta = -0.1178$, s.e. 0.0463) but weaker and statistically insignificant for adjusted scope 1 ($\beta = -0.0590$, s.e. 0.0393). Panel D (hard-to-abate sectors) exhibits larger negative point estimates for both scope 1 ($\beta = -0.1993$, s.e. 0.1092) and adjusted scope 1 ($\beta = -0.1893$, s.e. 0.1031), albeit with wider standard errors due to the smaller sample. While the hard-to-abate estimates are less precise, their magnitude is consistent with the idea that the integrity shock forced an especially consequential re-optimization for firms facing high abatement costs: when offsets became reputationally costly or less valuable for claims, firms that exited offsetting exhibited comparatively larger operational-intensity improvements than those that continued. At the same time, the lack of precision cautions against strong conclusions.

Taken together, the evidence supports three conclusions. First, following the January 2023 integrity shock, firms that exit voluntary offsetting subsequently achieve larger reductions in operational emissions intensity than comparable firms that continue offsetting. Second, this gap persists even under a conservative best-case adjustment that subtracts offsets one-for-one from scope 1 emissions, which is consistent with a moral-licensing mechanism rather than a pure accounting artifact. Third, the effects are concentrated in operational emissions (scope 1) and do not affect emissions that are not under direct control from firms.

6 Conclusion

This paper has offered an *autopsy* of the voluntary carbon market (VCM), combining transaction-level evidence on supply, demand, and prices with firm-level emissions accounts to study how integrity controversies reshaped both market outcomes and corporate decarbonization behavior. We document a market that, even before the 2022–2023 scandals, exhibited structural fragilities: persistent oversupply, fragmented intermediation, and limited transparency in pricing and buyer behavior. We then show that the January 2023 integrity scandal constitutes a sharp and plausibly exogenous revaluation of offsetting, with a discrete downward break in prices and a marked loss of momentum in retirements. Finally, leveraging this shock as a quasi-natural experiment, we provide evidence consistent with a moral-hazard mechanism: firms that *exit* offsetting after the scandal reduce scope 1 emissions intensity substantially more than otherwise similar firms that *stay* in the market, and this difference remains economically meaningful even under a conservative "ton-for-ton" adjustment that subtracts purchased offsets from scope 1 emissions.

As with any quasi-experimental study of corporate environmental behavior, several limitations are important for interpreting these findings. First, both transaction-linked offsetting data and corporate emissions accounts can be measured with error and are reported at discrete time intervals; this may attenuate estimated effects, obscure heterogeneous adjustment paths, or leave some scope for residual confounding by contemporaneous disclosure and reporting changes. Second, although we observe outcomes at the firm level, we do not directly observe the underlying abatement technologies, investment decisions, or organizational changes that deliver the emissions-intensity reductions, so we cannot uniquely map the reduced-form responses into specific operational channels. Finally, our results pertain to firms active in the VCM and to the specific integrity shock studied here; extrapolating to other periods, regulatory regimes, or future market designs should therefore be done with caution.

Nevertheless, we want to make very clear that, despite what a first read of the paper might suggest, our main takeaway from this analysis is *not* that we should move away from the Voluntary Carbon Market, nor that offsetting is intrinsically illegitimate. Rather, our evidence motivates a shift in emphasis: recent scrutiny has focused overwhelmingly on the *supply side* of the market — methodologies, baselines, additionality, permanence, and certification practices — yet our results suggest there is also an important *demand-side* integrity problem. Even if every credit were of high quality, the ability to substitute offsets for abatement can slow real emissions-intensity reductions at the firm level by relaxing the perceived constraint that drives operational change. This demand-side channel has received comparatively little systematic scrutiny in the policy and academic debate, and it matters precisely because voluntary offsetting is frequently framed as a bridge to "net-zero" rather than as a last-resort instrument for residual emissions.

Why does this distinction matter for climate policy? First, the market's dominant composition implies that many credits function as imperfect substitutes for fossil emissions in the relevant time and risk dimensions. A large share of traded credits historically comes from *avoidance*-type activities rather than durable removals, and even when these projects deliver real mitigation, the mapping from "one credit" to "one tonne of permanent compensation" is conceptually and empirically contested. In that setting, using offsets to substitute for feasible abatement is not merely an accounting choice: it risks delaying the technological and organizational adjustments required to permanently reduce gross emissions intensities. Second, a forward-looking perspective highlights scarcity and prioritization. Today's oversupply reflects what we call *nominal* credit abundance: a large volume of units exists on registries and can be transacted at low prices, but this does not imply an equally abundant supply of high-integrity tonnes that can credibly serve as residual-emissions instruments in a tightening climate policy environment. In a future where hard-to-abate sectors must decarbonize rapidly, the relevant constraint may well be a shortage of credible, durable mitigation units, and using offsets as a substitute for abatement where abatement is viable risks misallocating what will become a scarce resource.

This trade-off is particularly salient for genuinely hard-to-abate activities. Aviation is a clear example: scalable substitutes such as sustainable aviation fuel (SAF) remain limited relative to sectoral demand, implying that near-term decarbonization pathways may still involve residual emissions for which offsets (or other mechanisms) are invoked. In that context, the appropriate response is not blanket abolition, but governance that prioritizes offset use for residual emissions and strengthens accountability for claims. The policy challenge, therefore, is to prevent offsetting from becoming a low-cost substitute for feasible operational abatement while preserving a role for credible instruments in sectors where technological options are not yet deployable at scale.

Our results point to three practical implications. First, corporate claims should be evaluated not only on the integrity of purchased credits, but also on whether offset use coincides with credible reductions in gross emissions intensity. Second, governance reforms should aim to compress the wedge between the private cost of offsetting and the social value of decarbonization. One promising direction is a more centralized (or tightly supervised) transaction architecture in which credits are priced in a way that better reflects benchmark abatement costs, while limiting intermediary mark-ups and reallocating any surplus revenues toward public purposes. Revenue recycling is feasible in practice: Switzerland's CO_2 levy, for instance, redistributes receipts to households via health-insurance premium reductions, illustrating one concrete channel for lump-sum recycling. In the EU, ETS auction revenues finance mechanisms such as the Innovation Fund and are embedded in a broader framework that links carbon-pricing proceeds to climate and energy objectives.

We conclude by calling for further research on how to design institutions that preserve any legitimate role of offsets while preventing substitution away from feasible abatement. A central agenda is to identify claim frameworks and market rules that (i) discipline buyer behavior, (ii) prioritize offsetting for residual emissions, and (iii) align private incentives with emissions-intensity reductions that are real, durable, and measurable. In that sense, the autopsy is not only diagnostic: it is also a starting point for designing a governance architecture in which voluntary climate finance, if it persists, supports rather than substitutes genuine decarbonization.

References

Badgley, G., Freeman, J., Hamman, J. J., and Haya, B. (2022). Systematic over-crediting of forest offsets. *Global Change Biology*.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Black, A. et al. (2024). State of the voluntary carbon market 2024. Ecosystem Marketplace Report.

Boiral, O., Heras-Saizarbitoria, I., and Brotherton, M. (2024). Achieving corporate carbon neutrality: A critical analysis of disclosures, offsetting practices, and claims. *Journal of Cleaner Production*, 434:139846.

Bolton, P. and Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, 142(2):517–549.

Bolton, P. and Kacperczyk, M. (2022). Global pricing of carbon-transition risk. *Management Science*, 68(5):3209–3231.

Broekhoff, D., Piggot, G., and Erickson, P. (2019). Securing climate benefit: A guide to using carbon offsets. Stockholm Environment Institute Report.

Burger, A. M., Schuler, J., and Eberling, E. (2022). Guilty pleasures: Moral licensing in climate-related behavior. *Global Environmental Change*, 72:102415.

Cabiyo, B. et al. (2025). Carbon offset markets must learn to mitigate the risk of overcrediting. *PNAS Nexus*. Discusses overcrediting as a structural risk in offset systems.

Calel, R., Colmer, J., Dechezleprêtre, A., and Glachant, M. (2025). Do carbon offsets offset carbon? *American Economic Journal: Applied Economics*, 17(1):1–40.

Carton, W., Asiyanbi, A., Beck, S., Buck, H. J., and Lund, J. F. (2020). Negative emissions and the long history of carbon removal. *WIREs Climate Change*, 11(6):e671.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge Elements: Quantitative and Computational Methods for Social Science. Cambridge University Press.

Foerster, A. (2023). Corporate net zero pledges: A triumph of private climate governance? *Australian Journal of International Affairs*, 77(5):587–606.

Gill-Wiehl, A., Kammen, D. M., Haya, B., et al. (2024). Carbon offsets for clean cookstoves are pervasively over-credited. *Nature Sustainability*.

Gillingham, K., Rapson, D. S., and Wagner, G. (2016). The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy*, 10(1):68–88.

Greening, L. A., Greene, D. L., and Difiglio, C. (2000). Energy efficiency and consumption — the rebound effect — a survey. *Energy Policy*, 28(6–7):389–401.

Guizar-Coutiño, A., Jones, J. P. G., Balmford, A., Carmenta, R., and Coomes, D. A. (2022). A global evaluation of the effectiveness of voluntary redd+ projects at reducing deforestation and degradation in the moist tropics. *Conservation Biology*, 36(6):e13970.

Günther, S. A., Staake, T., Schöb, S., and Tiefenbeck, V. (2020). The behavioral response to a corporate carbon offset program: A field experiment on adverse effects and mitigation strategies. *Global Environmental Change*, 64:102123.

Hanna, R., Duflo, E., and Greenstone, M. (2016). Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves. *American Economic Journal: Economic Policy*, 8(1):80–114.

Haya, B., Cullenward, D., Strong, A., et al. (2020). Managing uncertainty in carbon offsets: Improved forest management and the US forest offset market. *Climate Policy*, 20(9):1112–1131.

Ilhan, E., Sautner, Z., and Vilkov, G. (2021). Carbon tail risk. *Review of Financial Studies*, 34(3):1540–1571.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Kim, S., Li, T., and Wu, Y. (2024). Carbon offsets: Decarbonization or transition-washing? *European Corporate Governance Institute – Finance Working Paper*, (1122/2025).

Kollmuss, A., Lazarus, M., Lee, C., and Polycarp, C. (2008). A review of the voluntary carbon market. *Stockholm Environment Institute Working Paper*.

NewClimate Institute and Carbon Market Watch (2024). Corporate climate responsibility monitor 2024.

Newell, R. G., Pizer, W. A., and Raimi, D. (2013). Carbon markets 15 years after Kyoto: Lessons learned, new challenges. *Journal of Economic Perspectives*, 27(1):123–146.

OECD (2024). The interplay between voluntary and compliance carbon markets. Technical report, Organisation for Economic Co-operation and Development.

Science Based Targets initiative (2021). Sbti corporate net zero standard.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2:307–317.

Song, Y., Aldy, J. E., Holbrook, N. M., Gao, X., and Thompson, J. (2025). Selection and over-crediting in forest-based carbon offset projects. Working paper.

Sorrell, S., Dimitropoulos, J., and Sommerville, M. (2009). Empirical estimates of the direct rebound effect: A review. *Energy Policy*, 37(4):1356–1371.

Stolz, N. and Benedict, P. (2025). The negligible role of carbon offsetting in corporate climate strategies. *Nature Communications*, 16(7963).

Swinfield, T. and Scott, E. (2025). Scientific credibility for high-integrity voluntary carbon markets. Cambridge Centre for Carbon Credits, University of Cambridge. Cambridge Open Engage preprint.

Tiefenbeck, V., Staake, T., Roth, K., and Sachs, O. (2013). For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy*, 57:160–171.

Tilsted, J. P., Palm, E., Bjørn, A., and Lund, J. F. (2023). Corporate climate futures in the making: Why we need research on the politics of science-based targets. *Energy Research Social Science*, 103:103229.

Tzachor, A. et al. (2023). Voluntary carbon markets and net-zero corporate strategies. *Climate Policy*.

West, T. A. P., Börner, J., Sills, E. O., and Kontoleon, A. (2020). Overstated carbon emission reductions from voluntary redd+ projects in the brazilian amazon. *Proceedings of the National Academy of Sciences*, 117(39):24188–24194.

West, T. A. P., Wunder, S., Sills, E. O., Börner, J., Rifai, S. W., Neidermeier, A. N., Frey, G. P., and Kontoleon, A. (2023). Action needed to make carbon offsets from forest conservation work for climate change mitigation. *Science*, 381(6660):873–877.