Systematic review of the actual emissions reductions of carbon offset projects across all major sectors

Benedict Probst (bprobst@ethz.ch)
ETH Zurich https://orcid.org/0000-0002-1149-8938
Malte Toetzke
ETH Zurich https://orcid.org/0000-0002-1153-2702
Laura Diaz Anadon
University of Cambridge https://orcid.org/0000-0002-2688-118X
Andreas Kontoleon
University of Cambridge https://orcid.org/0000-0003-4769-898X
Volker Hoffmann
ETH Zurich

Article

Keywords:

Posted Date: July 27th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3149652/v1

Abstract

Net-zero targets have significantly increased carbon offset demand. Carbon offsets are issued based on ex-ante estimates of project emissions reductions, though systematic evidence on ex-post evaluations of achieved emissions reductions is missing. We synthesized existing rigorous empirical studies evaluating more than 2,000 offset projects across all major offset sectors. Our analysis shows that offset projects achieved considerably lower emissions reductions than officially claimed. We estimate that only 12% of the total volume of existing credits constitute real emissions reductions, with 0% for renewable energy, 0.4% for cookstoves, 25.0% for forestry and 27.5% for chemical processes. Our results thus indicate that 88% of the total credit volume across these four sectors in the voluntary carbon market does not constitute real emissions reductions. This offset achievement gap corresponds to almost twice the annual German CO_2 emissions. We complement evidence from offset projects with 51 additional studies conducting ex-post evaluations of field interventions with settings comparable to offset projects. For cookstoves and forestry projects, these field interventions were more effective at reducing emissions than the voluntary offset projects, likely due to more careful intervention targeting, stricter monitoring and enforcement of intervention protocols.

Introduction

The United States and the European Union want to reach net-zero emissions by 2050, China by 2060, and India by 2070. By December 2022, 138 countries had already made net-zero pledges covering more than 80% of global emissions[1]. Similarly, many large corporations – including Amazon and Volkswagen – have promised to reach carbon neutrality by mid-century or earlier [1], and various firms claim that they are already ‘carbon neutral’ today. Yet, many of these claims entail the purchase of carbon offsets. These are “reduction, avoidance or removal of a unit of greenhouse gas (GHG) emissions by one entity, purchased by another entity to counterbalance a unit of GHG emissions by that other entity.”[2] Offsets in the voluntary carbon market today almost exclusively rely on reducing or avoiding emissions through, for instance, more fuel-efficient cookstoves or improved forest protection. While offsets based on carbon removal are growing, they only constitute a minor share of current voluntary carbon markets and are not the focus of this study.

What explains the major role that offsets are playing in corporate strategies is the implicit assumption that carbon offsets are economically efficient, as emissions reductions are achieved where they are cheapest [3]. However, for an offset project to contribute to emissions reductions, offsets need to conform to environmental integrity criteria, such as additionality[4,5] (i.e., reduction/removals would not have occurred without the project), durability (i.e., reduction/removals are not subject to near-term reversal or renewed at fixed intervals), and not leading to leakage (i.e., merely displacing emissions elsewhere). In addition, many carbon offset projects aim to create additional positive environmental and socioeconomic co-benefits, such as enhanced biodiversity or poverty alleviation.[2]

Carbon offsets have come under considerable criticism, however, as the underlying projects may not lead to actual emissions reductions[6]. Carbon offsets are commonly issued by comparing the actual carbon reductions of a project to a hypothetical baseline scenario if the project had not been implemented. This counterfactual baseline scenario is typically based on extrapolating historical emission trends. Yet, historical baselines are commonly an imperfect guide to future emissions. It is, therefore, critical to contrast the ex-ante estimated emissions reductions to the ex-post achieved emissions reductions by offset projects. This allows us to gauge what the offset achieved relative to what has been claimed ex-ante. We call this the 'offset achievement ratio' (see Methods for detailed explanation), which is the share of achieved emissions reductions based on credible academic studies relative to the claims made by project developers ex-ante.

While several studies have assessed the actual emissions reductions that were realized in individual offset projects relative to the expected reductions claimed by the verifiers[6,7], systematic and large-scale evidence of the actual reductions covering the full range of offset sectors is missing (for definition and a full list of sectors, see Table 1). In line with conventional systematic review methodology[8] and based on a Context-Intervention-Mechanism(s)-Outcome(s) logic (CIMO) [8], the central question of this analysis is therefore: 'What is known in the scientific literature about the differences between ex-ante estimates and ex-post outcomes of individual carbon offsetting projects adopted to enable the transition towards a net-zero emission economy across multiple sectors?'

We proceed in four steps. First, we define keywords to identify potentially relevant scientific studies across all major carbon offset sectors. As many offsetting sub-sectors only constitute a fraction of a per cent, we focus on the largest sub-sectors which, combined, make up more than 90% of credits issued in the voluntary offset market (Figure 1). Second, we use the artificial-intelligence-supported systematic review tool AS Review[9] to filter for relevant studies (e.g., using experimental or rigorous observational research methodologies) from 64,993 potentially relevant studies identified in the first step (Supplementary Figure 1 and Supplementary Table 1 & 2 for search terms). Third, we download the full text of the studies identified using AS Review and manually check for relevance (see Supplementary Table 3 for criteria). Fourth, two researchers independently extract the ex-post computed emissions reductions from individual projects and other relevant aspects of the study detailed in our Codebook. Lastly, for each project, we compute an offset achievement ratio. For field interventions that did not officially issue offsets, we compute a 'synthetic' offset achievement ratio (i.e., the ratio of achieved emissions reductions if these projects had used assumptions of similar, real-world projects to issue offsets; see Methods for detailed approach). In total, our final sample comprises more than 2,000 offset projects, and 130 effect sizes from 61 studies (see Extended Data 1).

Our analysis extends the existing literature in two major ways: First, we provide the first cross-sectoral, quantitative assessment of the offset achievement ratio of carbon offset projects in the peer-reviewed literature[10] and highlight insights on durability, co-benefits, and other relevant factors from these studies (see Supplementary Table 4 for previous meta-analyses in the non-peer reviewed literature). Second, we complemented the evidence on offset projects with 51 ex-post evaluations from field interventions that tested interventions similar to offset projects and jointly comprise 1.2 million observations. For instance, less than half of projects that attempt to reduce deforestation from deforestation and forest degradation (REDD+) have issued offsets[11]. Yet, there is a large, high-quality literature that investigates the underlying effectiveness of such interventions[12,13] allowing us to assess whether assumptions made by project developers using these interventions for offset projects are realistic.

Offset projects and field interventions

We conducted a systematic review of the offset achievement ratio of offset projects. In total, our set of studies includes offset sectors that jointly issued around 90% of carbon offsets (Figure 1). These contain 7 main sectors and 14 sub-sectors as defined by the Berkeley Carbon Trading Project[14] (Table 1).

Table 1: Sectors, sub-sectors and descriptions of offset sectors. *Directly cited and text shortened from Berkeley Voluntary Registry Offsets database scope & types document (Version April 2021).*[12]

Forestry & Land Use	
REDD+	Reducing deforestation and forest degradation in the global south. Many REDD+ projects bundle several activities (e.g., improved forest management, afforestation/reforestation). The "+" in REDD+ refers to the many project co-benefits (e.g., biodiversity).
Improved Forest Management	Applying practices which increase above and below-ground carbon stocks including reducing timber harvest levels, extending timber harvest rotations, designating reserves, fuel load treatments, enrichment planting, and stand irrigation or fertilization
Afforestation & Reforestation	Planting trees and reducing barriers to natural regeneration in non-urban areas.
Renewable Energy	
Wind	Installing wind turbines for grid-connected electricity generation replacing traditional, fossil-fuel or natural gas combustion for electricity production
Hydropower	Installing large and small-scale hydroelectric power plant (HEPP) turbines to generate electricity through regular dam flow operations or additions to multipurpose reservoirs
Solar	Installing solar modules as electricity production for grid-connected energy use.
Biomass	Generating heat, electricity (grid-connected or direct use), and/or biogas from renewable biomass, commonly utilizing agricultural waste biomass.
Waste management	
Landfill/wastewater methane	Landfill: Reducing and combusting methane from landfills including municipal, industrial, and other solid waste facilities. Wastewater: Treating wastewater to capture and flare methane, process with anaerobic digesters, and/or dewater sludge by drying before disposal
Chemical processes	
Ozone-depleting substances	Collecting and destroying refrigerants that are ozone-depleting substances with high GWP from discarded equipment such as air conditioners, refrigerators, and foam. *We also include the recovery and destruction of SF6 and HFC-23 in this category.*
N2O destruction in nitric acid production	Installing abatement measures and catalytic reduction units to destroy N2O emissions from nitric acid factories and caprolactam production plants. Nitric acid (HNO3) and caprolactam are crucial components of fertilizer and synthetic fibre production.
Household and community	
Cookstoves	Building improved cookstoves to replace or minimize the use of dung or firewood for cooking. Carbon benefits are realized in the form of reduced emissions from burning biomass as well as reducing deforestation. Less smoke leads to improved health benefits.
Industrial manufacturing	

Mine methane capture	Capturing and destroying or using mine methane that would otherwise be released to the atmosphere from active and abandoned coal, trona, and precious and base metal mines.
Natural gas electricity production	Constructing new natural gas-fired grid-connected electricity generation plants replacing higher greenhouse gas intensity fuels like coal. The fuel sources for the plants are fossil fuel natural gas, not renewable natural gas harvested through decomposition processes.
Carbon capture and storage	
Carbon Capture and Enhanced Oil Recovery	Capturing carbon dioxide from industrial processes followed by compression, transport and injection for permanent storage underground while also enhancing oil recovery.

Projects from forestry and renewable energy projects dominate the voluntary carbon market and constitute 74% of issued credits (Figure 1a/b). Industrial manufacturing, waste management, chemical processes, and household and community jointly constitute 23% (we use household and cookstoves interchangeably throughout the text as there is only one category). Carbon Capture and Storage (CCS), agriculture, and transportation together account for around 3%. Each sector is composed of sub-sectors. For instance, the forestry sector contains projects related to REDD+, forest management, and afforestation (though for forestry there is substantial overlap between these categories as REDD+ is a broad term). In turn, the renewable energy sector contains projects from wind, solar, biomass, and hydropower.

Following the typology of the Berkeley Carbon Project, we classify each of the 61 studies in our review into one of seven sectors with 14 sub-sectors (Figure 1). We differentiate between studies that investigate projects that officially issued carbon offsets and those projects that used a similar field intervention but did not officially issue offsets. We found 10 studies investigating 2,244 offset projects across four sectors (Figure 1a) and 51 studies investigating field interventions without issued carbon credits with a total of 1.2m observations (Figure 1b). For the other 3 main sectors (waste management, industrial manufacturing, and carbon capture and storage), we could not find any ex-post studies using a credible control group.

We have the strongest concentration of offset evaluations in the forestry sector, followed by renewable energy and chemical processes (Figure 1a/b). Offset evaluations are split between different geographies (apart from Africa with 0 studies). Similarly, for field interventions, most studies focus on forestry and are mainly focused on Latin America as most forestry projects are being implemented in tropical forests. Overall, both offset and field interventions mainly rely on rigorous observational studies (e.g., difference-in-difference and propensity score matching methodologies). In contrast, only 8 of 61 studies use randomised controlled trials (mainly evaluating the impact of fuel-efficient cookstoves, with one exception in forestry[15]).

The offset achievement ratio

The central question of this review is: what is known in the scientific literature about the differences between ex-ante estimates and ex-post outcomes of individual carbon offsetting projects? To operationalise this question, we introduce a new, simple metric, which we call the offset achievement ratio (see Methods Section for detailed description). The offset achievement ratio compares ex-post estimates from empirical studies with ex-ante estimates made by offset project developers. Hence, if a project reduced only half of what was originally claimed, the offset achievement ratio would be 50%.

For assessing offset achievement ratios, we only include empirical impact evaluations that contain a credible control group. A credible control group has similar characteristics as the treatment group. For instance, if a project seeks to avoid deforestation, then the deforestation trends within the conservation project would be compared to a forest with similar biophysical (e.g., type of forest, distance to forest edge) and socio-economic (e.g., distance to roads) characteristics that was not protected by the offset project.

This counterfactual approach stands in stark contrast to offset verifiers, which rely on simplistic comparisons of the offset project against a historical baseline to determine whether the project achieved its intended goals. For instance, in projects Reducing Emissions from Deforestation and Forest Degradation (REDD+), historical deforestation trends are commonly used, but these are often unsuitable to gauge the impact of the project for two main reasons [16]. First, changes in underlying political and economic conditions may lead to reductions in deforestation that are wrongfully attributed to the offset project, as likely happened with Brazil's policy effort to thwart deforestation post-2004 [6,16]. Second, project developers have an incentive to inflate deforestation baselines to benefit from the sale of a larger number of offset credits, which results in questions regarding the actual emissions reductions[6].

Overall, we find that offset projects achieved considerably lower emissions reductions than claimed ex-ante. We find the lowest values for the offset achievement ratio in the renewable energy (0%) and household (0.4%) sector, followed by forestry (25.0%) and chemical processes (27.5%) (Figure 3a). In contrast to offset projects, estimates from field interventions show higher results for cookstoves (17.1%) and forestry (39.2%) but not for renewable energy (no data on chemical processes) (Figure 3b). For our estimates in Figure 3, we use the central estimates from the studies. For studies that only report an upper bound, we do not include them in our main estimates (but show them graphically in Figure 3a) as the authors make clear that the results could be as low as zero[17] (see Methods section for discussion). We discuss issues of permanence in Section *leakage, durability, and co-benefits*.

The offset achievement gap

When generalizing the estimates from offset project studies in Figure 3a, we estimate that only 12% of the total volume of existing credits constitute real emissions reductions. Hence, 88% of the current voluntary carbon market across the main four sectors may not achieve the claimed offset goals. These non-achieved emissions reductions claimed by offsets are sizeable: the volume corresponds to almost twice the annual CO_2 emissions of the entire German economy.

Forestry and renewable energy credits account for around 90% of the current market (Figure 4a). Most renewable energy credits are likely not achieving the claimed goals, whereas a share of forestry credits likely represents actual emissions reductions. (Figure 4b). While industrial credits have a higher offset achievement ratio, their overall share in the voluntary carbon market is relatively low.

Field interventions show a higher degree of ‘synthetic’ offset achievement ratio, but even applying these more optimistic estimates from field interventions, almost 80% of the current market would not constitute actual emissions avoidance or reductions. We delve into the external validity of our findings, the potential reasons for the observed low achievement and the divergence between offset projects and field interventions in the discussion section.

Leakage, durability, and co-benefits

Although the offset achievement ratio across sectors is the central focus of this study, we also evaluated whether offset project and field intervention studies address other important considerations related to carbon offsets. Our results show that studies investigating the emissions reduction potential of carbon offset projects or field interventions only partly consider leakage, durability, and co-benefits.

Some carbon offset projects may only displace carbon emissions instead of avoiding them. Only in the forestry sector, do some studies consider leakage (for which it is arguably the biggest risk) (Figure 5a). Within this sector, around ¼ of studies analyse leakage. For those that analyse leakage, 73% of these studies find no evidence of leakage and the rest a mixed picture. Leakage effects can be positive as one forestry study found additional conservation effects in nearby areas to field interventions.[18]

Another key consideration for carbon offset projects is durability, which denotes the time that the carbon offset projects avoid, reduce, or remove emissions. Avoided emissions are not per se permanent, as the avoidance may only be temporary if, for instance, a protected forest is later cut down. On average, studies in our sample analyse on average 6.5 years of intervention, with the shortest average timeframes found in cookstove studies (2 years), whereas chemical processes, renewables and forestry investigate longer intervention periods (7-11 years) (Figure 5b). Many cookstove studies rely on randomised controlled trials. Since these are costly to implement, they tend to be more short-term in nature. In addition to the relatively short intervention periods studied, almost none of the sectors considers post-intervention effects (e.g., once the payments run out). The only exceptions are a few studies in the forestry sector (13% of all forestry-related studies), which tend to show that once payments run out, conservation effects are likely to be reversed.

Lastly, co-benefits/harms are also important considerations for offset projects to assess whether a project’s impacts go beyond carbon reductions. For example, these include positive effects of cookstove projects on health (co-benefit) or an increase in poverty levels (co-harm) due to a forestry conservation project. For chemical and renewable projects, no study investigates these effects (Figure 5c). In contrast, 22% of forestry projects and 64% of cookstove projects investigate co-benefits/harms. Cookstoves projects find neutral to positive effects (especially on time saving in collecting fuelwood[19] and cooking[20], and reductions in indoor air pollution[7]). In forestry studies, co-benefits also tend to be neutral to positive (especially on socio-economic factors such as participants' subjective wellbeing[21] and poverty alleviation[2223] as well as ecological factors, such as improved agricultural productivity[24] and hydrological services[2526,27]). Only one forestry study found negative effects on the subjective well-being of project participants, mainly related to frustrations around project implementation.[28]

Discussion

Overall, our review indicates that actual emissions reductions of offset projects are substantially lower than claimed. Furthermore, there is a dearth of empirical evidence around leakage, durability, and co-benefits arising from these interventions. We next turn to potential reasons that are behind the offset achievement gap across the main four offset sectors investigated. We then turn to external validity and potential bias in our results.

Renewable Energy

Across the four sectors, studies document the lowest offset achievement ratio for renewable energy (0%). Utility-scale renewable energy projects require high up-front investments and a secure cash flow to secure funding from banks and investors. As revenue streams from offsets are often low and may fluctuate substantially, as in the CDM, revenues generated by offsets are unlikely to substantially affect the financial viability of renewable energy projects. For instance, the most prominent policy schemes for renewable energy promotion have been feed-in-tariffs, offering stable power prices for commonly 20 years[29]. These projects have been deliberately shielded from the fluctuations of power markets, which are hard to predict far in advance. While the studies in our sample analyse wind projects, the findings likely extend to other renewable energy projects which feature similar capital structures, such as utility-scale solar, hydro and biomass[30]. Ultimately, the findings of several scholars, such as Haya[31], question whether accurate, verifiable ex-ante projections can even be constructed for renewable energy projects, such as wind[32] or hydropower[33]. It is important to note, however, that existing offset studies exclusively focus on utility-scale renewable projects and may not extend to small-scale projects.

Cookstoves

Cookstove offset projects feature similarly low offset achievement ratios (0.4%), though the literature is very limited. While cookstoves are often claimed to offer win-win solutions for health and the environment, the low additionally may be explained by behavioural and cultural reasons that interfere with the correct usage and full substitution of low emissions cookstoves. These factors render the project developers' emissions reduction assumptions of cookstove offset projects commonly taken from laboratory tests highly unrealistic. These laboratory tests assess the thermal efficiency in a highly artificial environment, which often does not represent how the stove is used outside of the lab. For instance, more fuel-efficient cookstoves are typically used next to the existing stoves, therefore serving as a complement rather than a substitute [34].

While only one study assesses an official cookstove offset project (finding no emissions reductions[7]), 10 studies that have analysed field interventions show substantial variation in the achieved emissions reductions. For instance, Hanna et al.[34] conducted a large-scale RCT in India and found no environmental benefits from stove adoption. In contrast, Berkouwer and Dean[35] conducted a study with a similar set-up in Kenya finding substantial emissions reductions from the BURN stove. Various reasons could explain this divergence in findings, including price, stove design, user behaviour and maintenance.

Hence, these findings indicate that cookstove projects are not ineffectual in general, but that the effectiveness is context-dependent, and more work is needed to understand the specific drivers of effectiveness.

Forestry

Studies on forestry offsets document higher offset achievement ratios than in renewable energy and cookstoves, yet overall remain below expectations (25.0%). The studies underscore common problems in conservation projects since they may be situated in areas with low overall deforestation risk, which reduces the likelihood that these projects avoid deforestation that would have happened otherwise.

We found that studies diverge substantially in their offset achievement ratio assessments, even if the same forestry offset project is analysed. 20 forestry projects certified by Verra have been analysed by at least two studies. Figure 6 shows estimates for these projects. Whereas Guizar-Coutiño et al.[36] find medium to high achievement (44% offset achievement ratio across projects), West et al.[6,16] find comparatively low achievement (5%). It is noteworthy that the study estimates across these forest projects show only a low correlation between these studies (r = 0.17).

Several reasons could explain this divergence. First, studies differ on methodological grounds. West et al. [6,16] rely on synthetic control (SC) methods, which compare projects to a weighted combination of potential control units to estimate the additional emissions reductions achieved by the project. In contrast, Guizar-Coutiño et al.[36] rely on a difference-in-difference approach, which matches pixels drawn from projects to similar pixels from forests not covered by the projects. The robustness of each approach hinges on the ability to construct a credible control group to evaluate the impact of the offset project. Difference-in-difference approaches make the simplifying assumptions that project and control sites would have followed the same trend in the absence of the project (“parallel trends assumption”). In contrast, the SC method relies on a weighted combination of control units allowing to reduce bias in cases where the parallel trend assumption is violated. Yet, the SC method typically features smaller sample sizes due to a more limited set of potential control units. Difference-in-difference approaches typically draw on larger samples but cannot control for time-invariant heterogeneity. In addition, Guizar-

Coutiño et al.[36] et al rely on more fine-grained satellite data (30m) compared to West et al. [6,16]. Lastly, the somewhat different time coverage could explain some of the results, as Guizar-Coutiño et al.[36] analyse the first 5 years of projects compared to longer time frames analysed in West et al. [6,16].

The observed divergence underscores the challenge of estimating the offset achievement ratio of forestry avoidance projects. Estimates are very sensitive to the creation of the control group, a non-trivial task due to the unobservable nature of these groups and the necessity of their construction via statistical methods. Overall, while the findings diverge, both West et al. [6,16] Guizar-Coutiño et al.[36] indicate that forest protection was much less effective than assumed in the Verra projects ex-ante estimates.

The overall intervention length covered by the studies was only 7 years. This presents an additional challenge since it is expected that the offset achievement ratio would become even lower than 25.0% after more than 7 years. Offset projects contain buffer pools – a share of credits that are not sold used to cover non-permanence risks – but studies suggest that tend to be insufficient given increased risks to forests through fires. For instance, Badgley et al.[37] document that the forestry projects' buffer pools in California's cap and trade programme are almost empty after their first 10 years despite needing to protect against forest fire risk over the next 100 years. In addition, although afforestation projects have become a popular offsetting mechanism, there are no offset studies investigating the offset achievement ratios of afforestation projects.[36]

Chemical processes

Projects in chemical processes (HFC-23 and SF_6 destruction) yielded the highest offset achievement ratio in our sample. We could only find two empirically rigorous studies that evaluated the impact of HFC-23 and SF_6 in Russia. In theory, the abatement of the above-mentioned substances should offer high offset achievement ratios without financial or regulatory incentives as there is commonly no business case for these interventions. Yet, the high abatement potential of these greenhouse gases can lead to perverse incentives that increase their production in the first place. This has been shown for projects under the Joint Implementation Mechanism in Russia. While the CDM addressed some of the issues of perverse incentives, qualitative research indicates that it still represents an issue.

Divergence in results between offset projects and field interventions

We find that forestry and cookstove projects designed and implemented as field interventions with similar settings as offset projects achieve higher emissions reductions than offset projects. For forestry projects, the offset achievement ratio from study intervention projects compared to offset projects is 39% (~ 1.6 times higher). For cookstove projects, field interventions – in contrast to offset projects – achieve a significant, though relatively low, average offset achievement ratio of 17.1%. For wind projects, the average offset achievement ratio is also non-significant. For chemical processes, we did not find field interventions that fulfilled the eligibility criteria (see Codebook for details).

We can only speculate about the reasons but hypothesise that one fundamental difference might drive the difference in observed outcomes. Field interventions (in contrast to offset projects) are often designed by researchers or non-governmental organisations – especially for RCTs – that want to test the effectiveness of a particular intervention instead of maximising financial gains as with private firms developing offset projects. For instance, Delacotte et al.[38] show that NGOs tend to locate forest protection projects in higher-risk areas than private firms that only aim to sell carbon credits. These differences in motivation could in turn affect a range of factors that lead to higher observed offset achievement ratios across projects such as improved targeting, implementation, and monitoring. We further explore potential reasons for the divergence between offset studies and field interventions in Supplementary Note 2.

Bias and external validity

Our analysis should be seen as preliminary and subject to several limitations concerning the external validity of the individual studies we analyse (even though it is the largest and only cross-sectoral effort to date) and the calculation of synthetic offset achievement ratio for field intervention studies.

Our estimates about sectoral offset achievement ratios rely on the generalisation of individual project observations to overarching sectors, which neglects potentially important factors such as country, year, or implementing organisation. For some sectors, such as cookstoves and chemical processes, our study sample is relatively small which could cause a biased generalisation to the overall offset achievement ratio. In addition, carbon offset projects may also provide additional climate benefits that are not captured by existing methodologies, such as soil effects in forest carbon protocols[39,40]. Lastly, the funnel plot and Egger's test (see Supplementary Note 1 and Supplementary Fig. 5) suggest a small-study bias in our analysis, whereby studies with smaller samples find higher additionalities, suggesting the presence of publication bias.[41]

In addition, to calculate the synthetic offset achievement ratio of field interventions, we matched field interventions with similar offset projects to compare ex-post observations from field interventions to ex-ante projections by offset issuers. The matching was based on intervention type, country, and year assuming that matched projects provide suitable proxies for ex-ante projections. To increase robustness, at least two matching offset projects were selected. Further research on offset projects and field interventions is needed to increase the robustness and external validity of our offset achievement ratio estimates.

While our analysis is preliminary, the offset studies in our sample analysed offset projects with considerable ex-ante estimated credit volumes, such as 216 megatons of CO2 for forestry (equivalent to 33% of the current voluntary carbon market (VCS) forestry volume), 167 megatons for renewables (equivalent to 32% of the current VCS renewables volume), and 0.43 megatons of cookstoves (or around 1% of the current VCS cookstove volume) and 104 megatons for industry (160% of current market volume in the VCM, as these credits primarily stem from the regulated markets (e.g., CDM), which surpass current VCS volumes). Hence, for cookstove offsets, in particular, more work is needed.

Conclusion

We synthesize existing rigorous empirical studies from more than 2,000 offset projects that estimate the extent to which offset projects have achieved avoided or reduced carbon emissions.

Overall, we find low offset achievement ratios across sectors, with 0% for renewable energy, 0.4% for cookstoves, 25.0% for forestry and 27.5% for chemical processes. Based on the offset achievement ratios, we calculate that up to 88% (or ~1.1 GT of 1.3 GT CO_2) of offsets across these four sectors may not constitute real emissions reductions. The estimated share of credits without real emissions reductions corresponds to roughly twice the current annual emissions of the entire German economy. For field interventions without official credit issuance, we document higher effectiveness for cookstoves and forestry. This divergence indicates that offset projects using these interventions can likely be improved, though their overall offset achievement ratio of the field interventions still lies considerably below the emissions reduction potential that project developers commonly claim. We recognise that these results should be seen as a synthesis of the best available evidence to date but still exploratory given the low number of rigorous empirical studies that are available.

Voluntary carbon markets are expected to grow significantly over the next decades[42] and the Article 6 mechanism envisaged by the Paris Agreement will further increase demand for carbon offsets[43]. Yet, our results substantiate doubts about the environmental integrity of carbon offsets projects from the four sectors we study. Our analysis suggests that there is no one-size-fits-all solutions and specific targeting, local context adaptation, and continuous, dynamic monitoring are the cornerstones of increasing offset achievement ratios.

Yet, implementing these changes will not only increase the costs of these carbon offsets, but it will also render the underlying project-based funding model less effective. Carbon offset revenues are inherently difficult to predict as the timing, price, and quantity may change over a project's lifetime. Improvements in offset protocols, such as dynamic baselining[44], may decrease the likelihood of low offset achievement ratios, but also increase uncertainty regarding the revenues that can be generated from offset sales by a project.

Furthermore, our results underscore the recommendations from the Oxford Principles for Net Zero Aligned Carbon Offsetting[45] to move away from avoidance-based offsets towards more durable solutions. We study sectors that generate offsets based on avoided emissions and only provide short-lived storage (e.g., in forests). The inherent difficulties of ensuring effectiveness and scaling these projects while safeguarding environmental integrity, strongly support the move towards other carbon credits based on carbon removal (not avoidance) with long-lived storage. Transitioning to carbon removal with long-lived storage is particularly important if offsets are continued to be used to offset fossil fuel emissions, which remain in the atmosphere for hundreds to thousands of years. Using offsets with questionable impact and short-lived storage is therefore inadequate to properly offset these emissions.

Our analysis, therefore, underscores that current voluntary carbon markets need to be substantially improved if they are to become an important enabler of the net-zero transition.

Methods

Analysis framework: We developed a framework which we use to systematically assess whether offsets achieve the intended goal of reducing or avoiding carbon emissions. Based on this framework, we searched the academic literature. The framework has the following four components: sectoral classification of carbon offsets, the development of criteria for the evaluation of offsets, the systematic review process, and the analysis of offset achievement ratios.

i. Sectoral classification of carbon offset

To set the scope of our analysis, we rely on voluntary carbon market data provided by the Berkeley Carbon Trading project[46]. We assess offset achievement ratios of carbon offsets that represent all major offset sectors. To our knowledge, this is the most comprehensive open-access database maintained on voluntary carbon markets (see Figure 1).

We assess seven major sectors and 14 sub-sectors collectively comprising sectors that account for more than 90% of issued carbon offsets on the voluntary markets (Figure 1b). For the classification of offset projects into different categories we rely on the Berkeley Carbon Trading offset typology [14]. As there are many small sub-sectors, we concentrate our literature search on all major sectors and sub-sectors, which collectively cover ~90% of issued credits (Table 1 and Figure 1).

We base our keyword search on this list of carbon offset sectors and the relevant sub-sectors. The full list of keywords can be found in Supplementary Tables 1 & 2.

ii. Criteria and outcomes for the evaluation of the offset

This study relies on a systematic review methodology to assess the carbon reduction and avoidance impacts of various offset projects. We only include studies that are either experiments (where researchers assign treatment) or rigorous observational studies (in which researchers leverage plausible exogeneous sources of variation to estimate project impacts) (see Supplementary Table 3 for inclusion and exclusion criteria). The fundamental difference between typical offset projects and these rigorous studies is that they include a credible control group that can plausibly answer the question: What would have happened if the project had not been implemented? Typically, offset projects use historical baselines as the control group, which is an imperfect approximation of project impact as contemporaneous socio-economic changes may drive the apparent project impact, not the project itself.[6,16]

iii. Systematic review process

In line with a large body of systematic reviews[9], we employ the Context-Intervention-Mechanism(s)-Outcome(s) (CIMO) framework to define keywords and select the studies for the systematic review. The

CIMO framework includes the definition of the central research question, inclusion, and exclusion criteria to select studies from the large pool of potentially relevant studies, as well as a description of the final sample. We proceed as follows. After having defined the keywords and inclusion and exclusion criteria (see two sub-sections before), we use the AI-supported systematic review tool AS Review[9] to filter for relevant studies (e.g., using experimental or rigorous observational research methodologies) from 64,993 potentially relevant studies identified in the first step (Supplementary Figure 1 and Supplementary Table 1 & 2 for search terms). We then download the full text of the studies identified using AS Review and manually check for relevance. Then, two researchers independently extract the reported additionalities from individual projects and other relevant aspects of the study detailed in our Codebook. For field interventions that did not officially issue offsets, we compute a 'synthetic' additionality (i.e., the additionality if these projects had used assumptions of similar, real-world offset projects to issue offsets; see next section for details). In total, our final sample comprises more than 2,000 offset projects, and 130 effect sizes from 61 studies (see Extended Data 1). The detailed ROSES flow diagram for systematic reviews can be found in Supplementary Figure 1 and all included studies in Supplementary Table 6.

iv. Analysis of the offset achievement ratio

Carbon offsets are typically issued by comparing the actual carbon reductions of a project to a hypothetical baseline scenario if the project had not been implemented. This counterfactual baseline scenario is typically based on extrapolating historical emission trends. Yet, historical baselines are commonly an imperfect guide to future emissions. It is, therefore, critical to contrast the ex-ante estimated emissions reductions to the ex-post achieved emissions reductions by offset projects. We call this the 'offset achievement ratio', which is the share of achieved emissions reductions based on credible academic studies relative to the claims made by project developers ex-ante.

For field interventions without official offset issuance, we approximate the offset achievement ratio by developing our approach to compute a 'synthetic' offset achievement ratio if these field interventions had issued offsets and had employed standard assumptions from similar, real-world offset projects. We first discuss assessing the offset achievement ratio of official offset projects, followed by our approach to assessing field interventions. Lastly, we discuss how we synthesise the offset achievement ratio is related to the commonly used concept of additionality and how we integrate different project estimates (lower, medium, and upper bound estimates) (see Figure 3, which contains upper bound estimates),

Official offset projects

Offset projects commonly report two distinct metrics in their project documentation. To illustrate our approach, let's assume we analyse a project that seeks to reduce deforestation and forest degradation (REDD+). A REDD+ project commonly reports ex-ante projection of baseline emissions C_{BL} (e.g., emissions through continued deforestation in the area, commonly a continuation of historical trends) and expected emissions reductions C_E due to the project (e.g., increased protection of forest leading to lower deforestation rates) (see Figure 7). For simplicity, we assume that there are two points in time: t= 0 (before the project) and t = 1 (at the point of evaluation, after the project has been implemented). We assume that the offset achievement ratio stays constant over time.

Yet, the true carbon emissions reductions from an offset project can only be assessed ex-post (i.e., after project implementation, see Figure 8). Two metrics are important in this regard: Counterfactual emissions C_C – which describes the true baseline emissions that would have occurred without the project. Rigorous empirical studies use a variety of statistical methods – such as propensity score matching and difference-in-difference econometrics – to create a credible control group. For a REDD+ project that could mean that the project area is compared to a similar plot of land, which faces a similar level of deforestation pressure but has not been enrolled in the programme.

In theory, the ex-ante projected baseline emissions could be the same as the ex-post estimated counterfactual emissions ($C_{BL} = C_C$), but as many exogenous factors change during project implementation, the projected and true baseline emissions likely diverge. For instance, an unexpected fall in international beef prices might decrease deforestation pressures suddenly, which in turn would decrease the "true" emissions baseline (Figure 7, see ex-ante baseline projection B_L and ex-post counterfactual C_C).

The second important metric is C_P, which is the true carbon reductions that the project led to. Again, in theory, $C_E = C_P$ could be true, but the project might either be more or less effective in decreasing carbon emissions than projected ex-ante.

We then compute the offset achievement ratio (OAR) of project *i* as:

$$OAR = (C_P - C_C) / C_E \qquad (1)$$

where:

$$C_E = C_{BL} * EI \qquad (2)$$

EI is an effectiveness index and describes the effectiveness of the project in reducing carbon emissions relative to the baseline. The EI differs from project to project, but it commonly ranges between 0.5 and 1. If the project has an EI of 1, it is assumed that the offset project completely eliminates carbon emissions relative to the baseline. The average assumed EI in project design documents is 100% for renewable energy (relative to the grid factor) and chemical processes (relative to the baseline emissions), 75% reduction relative to the baseline for forestry and 65% for cookstoves projects.

A project that has an offset achievement ratio of 0%, did not lead to any emissions reductions, whereas a project with an OAR of 100% fully yielded the expected emissions reductions. A project with an offset achievement ratio of -100% led to emissions increases proportionally to the size of the initially claimed reductions. In cases where the emissions savings are not reported, we use a corresponding measure from the study that linearly correlates with emissions savings (e.g., reductions in deforestation rates between a project and counterfactual scenario or reduction in fuelwood use by households for cookstoves projects).

For instance, let's assume (C_P - C_C) was an emissions reduction of -20 t CO_2 (studies commonly report the aggregated effect size between the project and a counterfactual (i.e., true baseline) scenario instead of separate effect sizes for the baseline and counterfactual scenario)). The claimed emissions reduction was -40t CO_2.

Hence:

$$OAR_i = (C_P - C_C) / C_E = -20t / -40t = 50\% \quad (3)$$

If (C_P - C_C) = 0, then the OAR is by definition 0. Hence, for studies that show no difference between the baseline and counterfactual scenario, and do not report E, we collect no data on E as E would not change the offset achievement ratio in these cases.

If a project claims to have offset 1 megaton of emissions, but had an OAR of 50%, then only 0.5 megatons were reduced. We call the absolute difference between what was claimed and achieved, the offset achievement gap.

Field interventions

To compute the potential OAR of field interventions that did not officially issue offsets, we approximate the potential OAR. To calculate the OAR by offset sector, we compared actual emissions reductions (from ex-post evaluations of the research studies) with ex-ante estimates of emissions reductions from relevant offset project reports. The matching of research studies with offset projects was conducted in four steps:

1. We developed an algorithm that matched each field intervention with an official offset project from the Berkeley Voluntary Registry Offsets Database[46] in the same sub-sector (e.g., *REDD+*, *cookstoves*), country, and intervention years of the study and randomly shuffled the filtered sample of offset projects using package pandas (Version 1.2.5) in Python (Version 3.8).
2. We manually went through the sample from step 1 in chronological order. For each project, we evaluated if the project resembled the study setting of the research study (e.g., by checking if a similar cookstove was used for the intervention). If the project was not found suitable, we moved to the next project in the sample set. If the project was found suitable, we extracted ex-ante estimates of

baseline emissions and project emissions (during crediting period) from the project documents to calculate the estimated percentage of emissions reduction of the project.

3. Next, the project documents were retrieved from the websites of the credit issuers (e.g., Verra, Gold Standard). For each study, we extracted ex-ante estimates of emissions reductions from two different projects. If the estimates deviated from each other by more than 20 per cent, we also included a third project.
4. Finally, we used the mean from the expected projects' emissions reduction estimates as a comparison to the ex-post emissions reductions calculated in the research studies. For instance, if an official cookstove project implemented in the same country at the same time assumed that emissions would be reduced relative to the baseline by 60% but the field interventions only found a 20% reduction, then the offset project would have an offset achievement ratio of 20% / 60 % = 33%.

The flow diagram and the relevant steps can be found in Supplementary Figure 2.

The offset achievement ratio and additionality

Studies employ different approaches to assess whether a project reaches its intended goal of reducing or avoiding carbon emissions. The literature typically employs the concept of 'additionality', which asks what would have happened in the absence of the project. Additionality, therefore, is used to conceptualise the real carbon emissions savings (see Figure 7), which we then divide by ex-ante estimates from project design documents (if the study does not already do so) to calculate the offset achievement ratio.

As the concept of additionality underpins the offset achievement ratio, we therefore briefly discuss these different types, their stringency, and the focus of our study. As the ex-ante estimates are determined by the project developers and are therefore standardised, the additional carbon savings estimated by studies differ substantially (see below).

In Supplementary Figure 3, we differentiate between four types of additionality:

1. **Financial additionality** (Voluntary carbon market leads to financing that the projects would otherwise not have raised via other private or public sources of finance and only this funding makes the project viable. Even with the funding project might still not be implemented
2. **Project additionality** (The sole reason for the existence of the project is the funding/revenue stream that the project acquired via voluntary carbon markets)
3. **Emissions additionality** (Project leads to emissions avoidance, abatement or removal that would not have happened without the project)
4. **Marginal additionality** (Each sale of a carbon offset leads to a decrease in CO_2 emissions in the project)

To illustrate these different types of additionality, assume that we want to evaluate the additionality of a biomass power plant financed by carbon credits. The first, and least stringent criterion, for assessing the additionality of the project, is to ask whether the project could have acquired sufficient financing even without the carbon credits. If the revenue generated through the (prospective) sale of carbon credits was sufficient to make the project financially viable, then the next question becomes whether the project was eventually built (project additionality). If the project was financed and built due to carbon credits, then the question becomes to what extent, the biomass power plant is reducing emissions in the power grid into which it is delivering its electricity. For instance, if the grid is already zero emissions due to large shares of hydropower and conventional renewables such as wind and solar, then the emissions additionality of the project would be zero, as no additional emissions are displaced. Yet, if the biomass plant feeds into a grid dominated by coal-fired electricity, the emissions additionality is clear. Lastly, the most stringent form of additionality, is whether each additional sale of credits leads to an additional decrease of carbon emissions. For instance, if the biomass-fired power plant cannot maintain its operation (e.g., maintenance, buying additional biomass) were it not for the sale of carbon credits, then even the marginal additionality would be fulfilled.

Yet, the studies in our review typically fall into two types of camps. First, studies investigating the additionality of renewable energy typically assess financial[3] and project additionality[17]. While clearing the hurdle of financial and project additionality are necessary conditions for emissions additionality, they are not sufficient. To establish emissions additionality, a detailed power system model would be needed, to assess the exact emissions displaced in the grid, which depends on the exact production volume and time of the wind power plant, its exact location in the supply curve, the grid operator, and many other factors. Hence, the additionality assessments of renewable energy projects should be considered less stringent than studies that assess emissions additionality. In contrast, chemical processes, cookstoves, and forestry projects assess emissions additionality. As there is typically no business case to implement these projects otherwise, financial and project additionality can be assumed to be true (at least, in most cases). Hence, these projects assess emissions additionality, by considering the tailpipe emissions from industrial plants, emissions associated with changes in deforestation levels or fuelwood use. Emissions additionality assessments can therefore be considered more stringent than financial/project additionality. No study in our sample considers marginal additionality, likely due to the complexities of measuring that type of additionality.

Central and upper bound estimates

Studies typically report central estimates. We consider central estimates those empirical estimates that consider two sources of low offset achievement ratios:

1. **Wrong baseline:** The study assesses what the real, counterfactual baseline would have been if the project had not been implemented. Typically, the ex-ante baseline is compared to a credible, ex-post baseline

2. **Wrong project impact:** The study assesses what the real project impact was after the project had been implemented. Typically, the ex-ante, expected emissions reductions associated with the project are compared to a credible, ex-post project impact assessment.

Please note that these two sources of low offset achievement ratio correspond to comparing the real carbon savings to the ex-ante expected carbon savings in Figure 7. For studies that only consider whether the baseline has been inappropriately set, but do not analyse whether the project itself was additional, we consider these estimates to be an upper bound (see, for instance, ref[47]). These are upper-bound estimates, as the project impact could be as low as zero. Similarly, if studies explicitly state that their estimates could be as low as zero, we also record those as upper bound (see ref [17].)

Declarations

Contributions

All authors developed the research idea. B.P. conducted the empirical analysis with support from M.T., B.P. analysed and visualized the data and wrote the manuscript with support from all authors, while A.K., L.D.A. and V.H. edited the final draft.

Corresponding author

Correspondence to Benedict Probst.

Data availability

The data is available upon reasonable request from the corresponding author.

Code availability

The code is available upon reasonable request from the corresponding author.

References

1. Net Zero Tracker. Net Zero Tracker. https://zerotracker.net/ (2022). IPCC. IPCC AR6 WGIIII Annex I: Glossary. (2022).
2. Chan, G. *Essays on Energy Technology Innovation Policy*. http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467190%0AThis (2015).
3. Michaelowa, A. Interpreting the Additionality of CDM Projects: Changes in Additionality Definitions and Regulatory Practices over Time. in *Legal Aspects of Carbon Trading* 248–271 (Oxford University PressOxford, 2009). doi:10.1093/acprof:oso/9780199565931.003.0012.
4. Schneider, L. Assessing the additionality of CDM projects: practical experiences and lessons learned. *Climate Policy* **9**, 242–254 (2009).

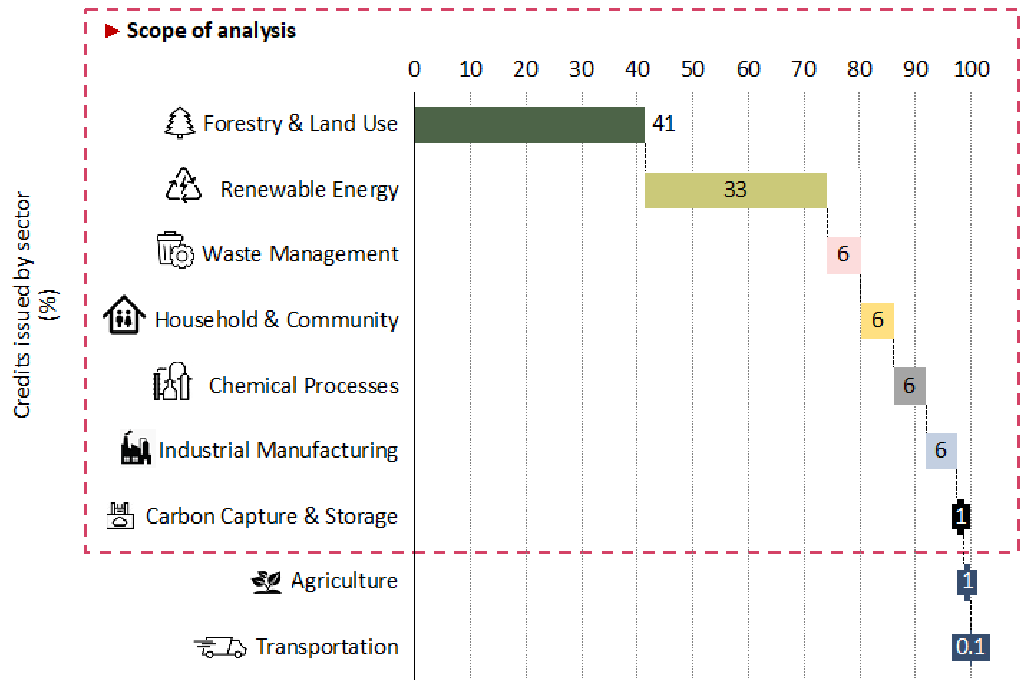
5. West, T. A. P., Börner, J., Sills, E. O. & Kontoleon, A. Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. *Proc Natl Acad Sci U S A* **117**, 24188–24194 (2020).
6. Aung, T. W. *et al.* Health and Climate-Relevant Pollutant Concentrations from a Carbon-Finance Approved Cookstove Intervention in Rural India. *Environ Sci Technol* **50**, 7228–7238 (2016).
7. Peñasco, C., Anadón, L. D. & Verdolini, E. Systematic review of the outcomes and trade-offs of ten types of decarbonization policy instruments. *Nat Clim Chang* (2021) doi:10.1038/s41558-020-00971-x.
8. van de Schoot, R. *et al.* An open source machine learning framework for efficient and transparent systematic reviews. *Nat Mach Intell* **3**, 125–133 (2021). Cames, M. *et al. How additional is the Clean Development Mechanism?* https://ec.europa.eu/clima/sites/clima/files/ets/docs/clean_dev_mechanism_en.pdf (2016) doi:CLIMA.B.3/SERI2013/0026r.
9. Roopsind, A., Sohngen, B. & Brandt, J. Evidence that a national REDD+ program reduces tree cover loss and carbon emissions in a high forest cover, low deforestation country. *Proc Natl Acad Sci U S A* **116**, 24492–24499 (2019).
10. Duchelle, A. E., Simonet, G., Sunderlin, W. D. & Wunder, S. What is REDD+ achieving on the ground? *Curr Opin Environ Sustain* **32**, 134–140 (2018).
11. Wunder, S., Börner, J., Ezzine-de-Blas, D., Feder, S. & Pagiola, S. Payments for Environmental Services: Past Performance and Pending Potentials. *Annu Rev Resour Economics* **12**, 209–234 (2020).
12. Berkeley Carbon Trading Project's Voluntary Registry Offsets Database. Scopes & Types. 1–16 (2021).
13. Jayachandran, S. *et al.* Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science (1979)* **357**, 267–273 (2017).
14. West et al. *Carbon credits from tropical forest conservation projects unlikely to mitigate climate change*. (2022).
15. Calel, R., Colmer, J., Dechezleprêtre, A. & Glachant, M. Do Carbon Offsets Offset Carbon? *SSRN Electronic Journal* (2021) doi:10.2139/ssrn.3950103.
16. Giudice, R., Börner, J., Wunder, S. & Cisneros, E. Selection biases and spillovers from collective conservation incentives in the Peruvian Amazon. *Environmental Research Letters* **14**, 045004 (2019).
17. Jeuland, M. A., Pattanayak, S. K., Samaddar, S., Shah, R. & Vora, M. Adoption and impacts of improved biomass cookstoves in rural Rajasthan. *Energy for Sustainable Development* **57**, 149–159 (2020).
18. Brooks, N. *et al.* How much do alternative cookstoves reduce biomass fuel use? Evidence from North India. *Resour Energy Econ* **43**, 153–171 (2016).
19. Carrilho, C. D., Demarchi, G., Duchelle, A. E., Wunder, S. & Morsello, C. Permanence of avoided deforestation in a Transamazon REDD+ project (Pará, Brazil). *Ecological Economics* **201**, 107568 (2022).

20. Sims, K. R. E. & Alix-Garcia, J. M. Parks versus PES: Evaluating direct and incentive-based land conservation in Mexico. *J Environ Econ Manage* **86**, 8–28 (2017).
21. American, S. *et al.* Only One Tree from Each Seed? Environmental Effectiveness and Poverty Alleviation in Mexico ' s Payments for Ecosystem Services Program Author (s): Jennifer M . Alix-Garcia , Katharine R . E . Sims and Patricia Yañez-Pagans Published by : American Econ. **7**, 1–40 (2015).
22. Correa, J. *et al.* Evaluating REDD+ at subnational level: Amazon fund impacts in Alta Floresta, Brazil. *For Policy Econ* **116**, 102178 (2020). von Thaden, J., Manson, R. H., Congalton, R. G., López-Barrera, F. & Salcone, J. A regional evaluation of the effectiveness of Mexico's payments for hydrological services. *Reg Environ Change* **19**, 1751–1764 (2019).
23. Clements, T. & Milner-Gulland, E. J. Impact of payments for environmental services and protected areas on local livelihoods and forest conservation in northern Cambodia. *Conservation Biology* **29**, 78–87 (2015).
24. Jones, K. W. *et al.* Measuring the net benefits of payments for hydrological services programs in Mexico. *Ecological Economics* **175**, 106666 (2020).
25. Montoya-Zumaeta, J. G., Wunder, S., Rojas, E. & Duchelle, A. E. Does REDD+ Complement Law Enforcement? Evaluating Impacts of an Incipient Initiative in Madre de Dios, Peru. *Frontiers in Forests and Global Change* **5**, (2022). UNEP. *Feed-in Tarrifs as Policy Instrument for Promoting Renewable Energies and Green Economies in Developing Countries*. (2012). IRENA. *Renewable power generation costs in 2021*. (2022).
26. Barbara Kresch Haya. Carbon Offsetting: An Efficient Way to Reduce Emissions or to Avoid Reducing Emissions? An Investigation and Analysis of Offsetting Design and Practice in India and China. (2010).
27. He, G. & Morse, R. Addressing carbon Offsetters' Paradox: Lessons from Chinese wind CDM. *Energy Policy* **63**, 1051–1055 (2013).
28. Haya, B. & Parekh, P. Hydropower in the CDM: Examining additionality and criteria for sustainability. *Berkeley Energy and Resources Group Working Paper* (2011).
29. Hanna, R., Duflo, E. & Greenstone, M. Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves. *Am Econ J Econ Policy* **8**, 80–114 (2016).
30. Berkouwer, S. B. & Dean, J. T. *Credit, Attention, and Externalities in the Adoption of Energy Efficient Technologies by Low-Income Households*. *American Economic Review* vol. 112 (2022).
31. Guizar-Coutiño, A., Jones, J. P. G., Balmford, A., Carmenta, R. & Coomes, D. A. A global evaluation of the effectiveness of voluntary REDD+ projects at reducing deforestation and degradation in the moist tropics. *Conservation Biology* **36**, 1–13 (2022).
32. Badgley, G. *et al.* California's forest carbon offsets buffer pool is severely undercapitalized. *Frontiers in Forests and Global Change* **5**, (2022).
33. Delacote, P., Le, G. & Simonet, G. Revisiting the location bias and additionality of REDD + projects : the role of project proponents status and certification. **67**, (2022).

34. Haya, B. *et al.* Managing uncertainty in carbon offsets: insights from California's standardized approach. *Climate Policy* **20**, 1112–1126 (2020).
35. Bento, A., Kanbur, R. & Leard, B. On the importance of baseline setting in carbon offsets markets. *Clim Change* **137**, 625–637 (2016).
36. Khanna, T. M. *et al.* A multi-country meta-analysis on the role of behavioural change in reducing energy consumption and CO2 emissions in residential buildings. *Nat Energy* (2021) doi:10.1038/s41560-021-00866-x.
37. Blaufelder, C., Levy, C., Mannion, P. & Pinner, D. *A blueprint for scaling voluntary carbon markets to meet the climate challenge*. (2021).
38. IETA. *The Economic Potential of Article 6 of the Paris Agreement and Implementation Challenges*. https://www.ieta.org/resources/International_WG/Article6/CLPC_A6%20report_no%20crops.pdf (2019).
39. Verra. S&P Global Commodity Insights: "Verra approves dynamic baseline methodology for forest carbon projects. https://verra.org/press/sp-global-commodity-insights-verra-approves-dynamic-baseline-methodology-for-forest-carbon-projects/ (2022).
40. Allen, M. *et al.* The Oxford Principles for Net Zero Aligned Carbon Offsetting. *University of Oxford* 15 (2020).
41. Haya, B., Elias, M. & So, I. *Voluntary Registry Offsets Database, Berkeley Carbon Trading Project, v6. November*. https://gspp.berkeley.edu/faculty-and-impact/centers/cepp/projects/berkeley-carbon-trading-project/offsets-database (2022).
42. Badgley, G. *et al.* Systematic over-crediting in California's forest carbon offsets program. *Glob Chang Biol* **28**, 1433–1445 (2022).
43. Mobarak, A. M., Dwivedi, P., Bailis, R., Hildemann, L. & Miller, G. Low demand for nontraditional cookstove technologies. *Proceedings of the National Academy of Sciences* **109**, 10815–10820 (2012).
44. Agurto Adrianzén, M. Improved cooking stoves and firewood consumption: Quasi-experimental evidence from the Northern Peruvian Andes. *Ecological Economics* **89**, 135–143 (2013).

Figures

a)



b)

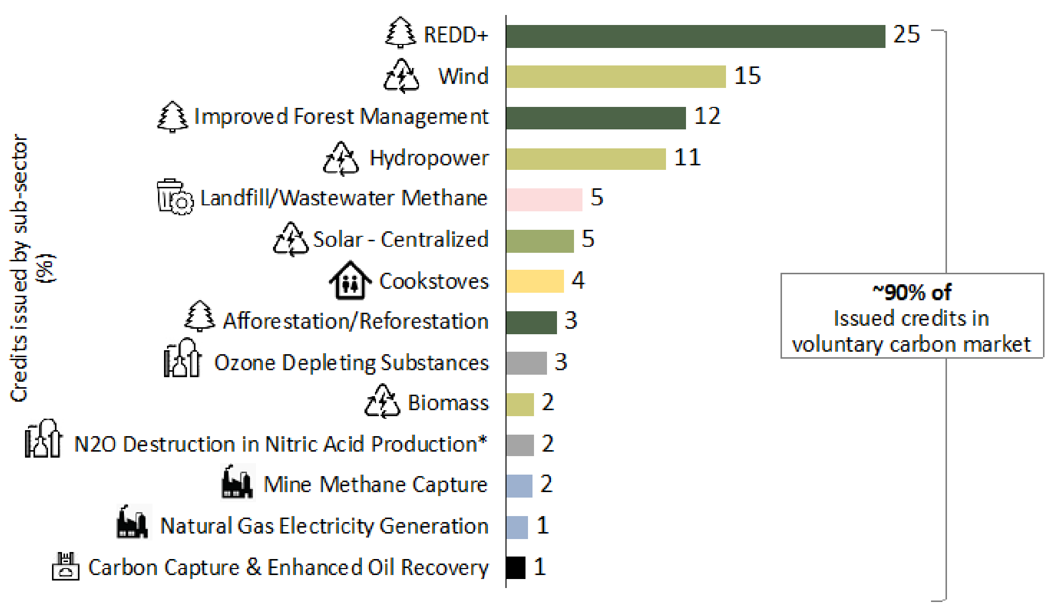


Figure 1

Credit issuance in the voluntary carbon market. a) Issued credits in voluntary carbon markets by sector and b) by sub-sector from 1996 – November 2022. The numbers next to the bars are in %. The scope of our analysis includes all major sectors. Each sector comprises a range of sub-sectors. Collectively, we cover sub-sectors accounting for ~90% of the credits issued. Please note that adding the sectors in Figure 1a accounts for more than 90% as many sectors contain a range of small sub-sectors, which were outside the scope of our analysis. Based on the Berkeley Carbon Trading dataset (v6, November 2022). Clean Development Mechanism (CDM) credits are included only if they were transferred to a voluntary registry. [10] Numbers mentioned in the text may differ slightly from those in Figure due to rounding. REDD+ refers to projects related to Reducing Emissions from Deforestation and Forest Degradation.

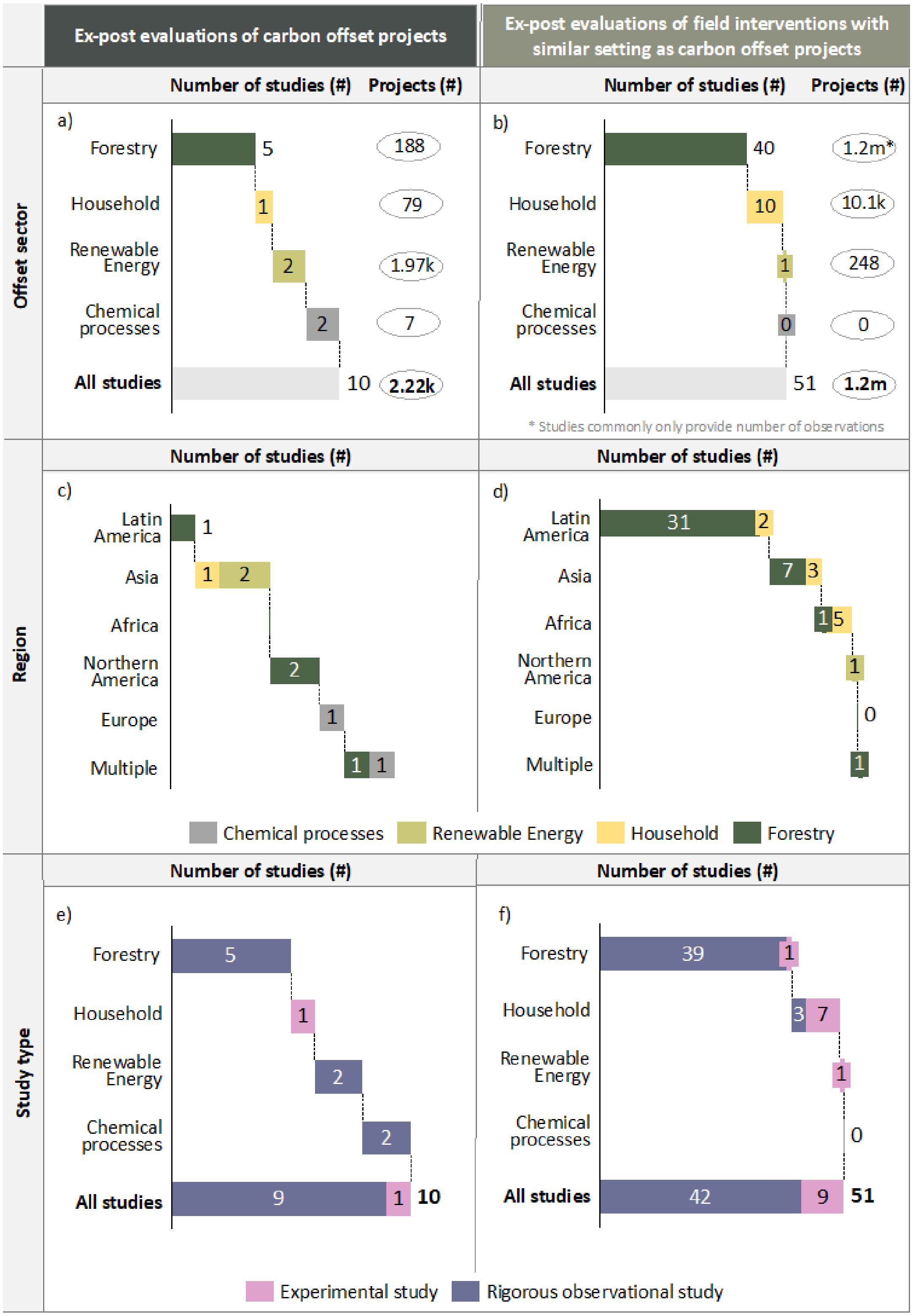


Figure 2

Overview of studies in the systematic review. a/b) Distribution of studies across offset sectors, c/d) across regions, and e/f) methodology types. Note: k refers to the number in thousands, and m refers to the number in millions. See Supplementary Table 5 for a descriptive overview of the sample.

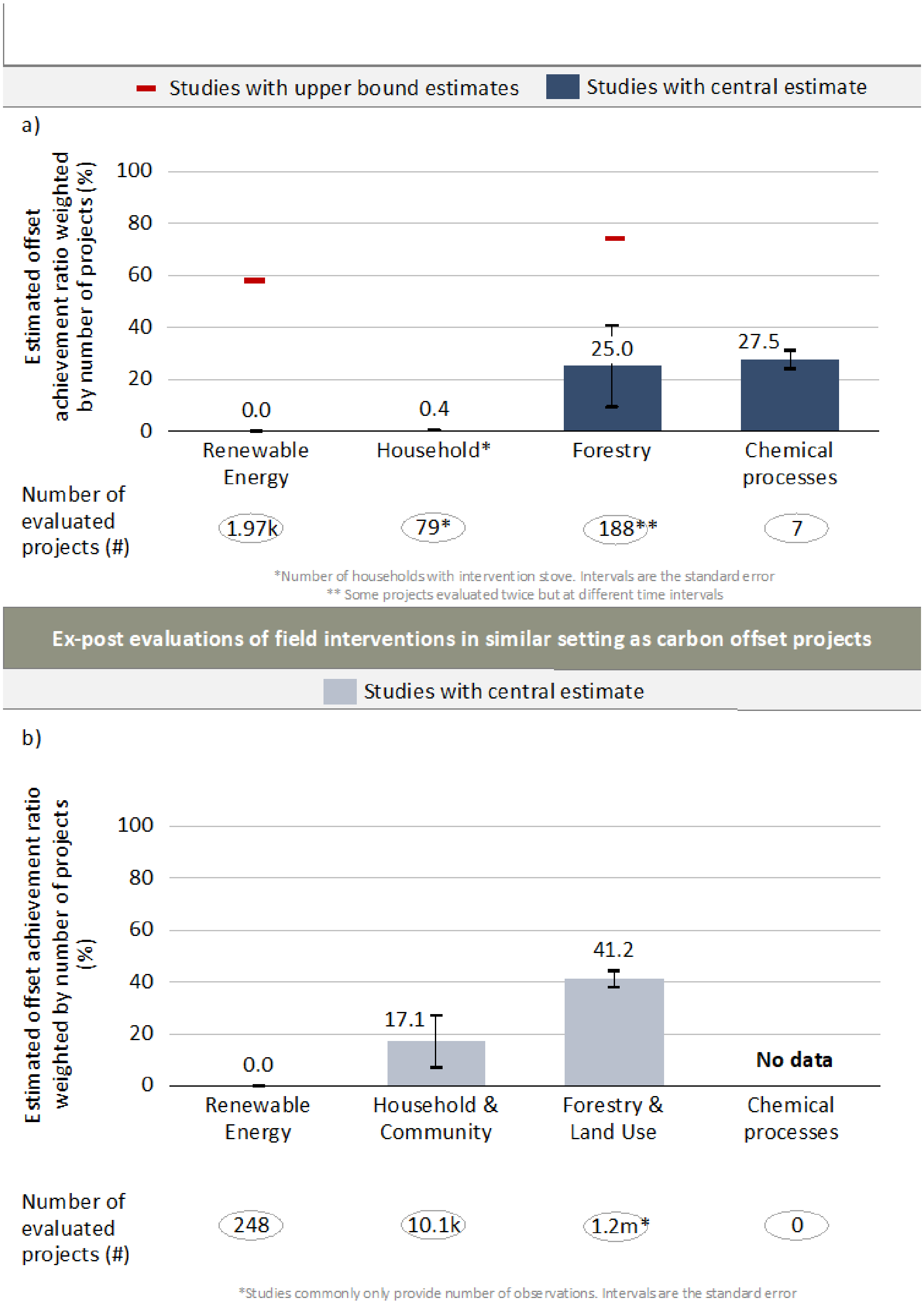


Figure 3

Estimated offset achievement ratios across sectors with a) estimates from carbon offset projects, and b) from field interventions in a setting comparable to carbon offsets but without the official issuance of offsets. The offset achievement ratio is weighted by the number of offset projects analysed in each study. Intervals are standard errors.

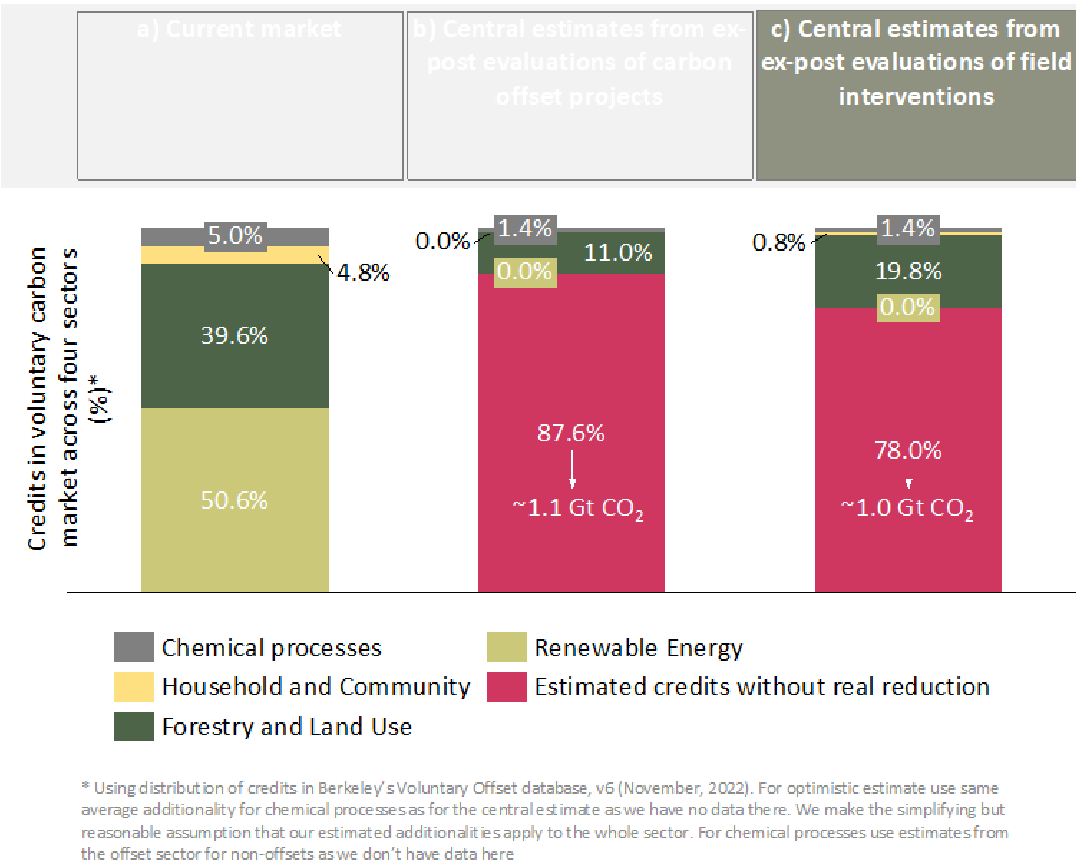


Figure 4

Current and estimated distribution of credits in the voluntary carbon market. a) Current market distribution according to Berkeley's Voluntary Offset database, b) central estimates from ex-post evaluations of carbon offset projects, and c) central estimates from ex-post evaluations of field interventions in similar settings to carbon offset projects.

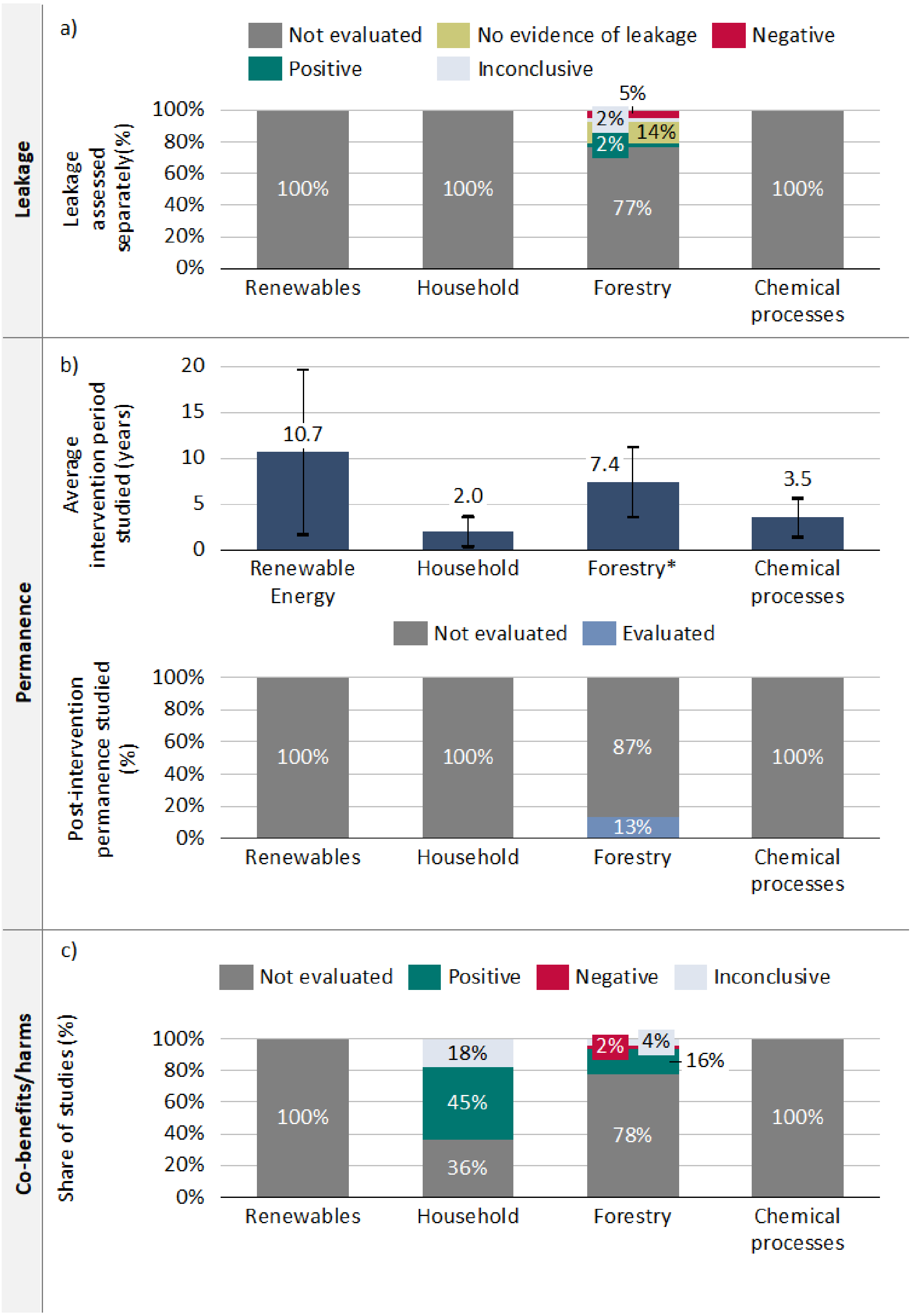


Figure 5

a) Leakage, b) intervention period (and standard deviation) and share of the post-intervention period studied, and c) co-benefits/harms reported in studies in our study sample.

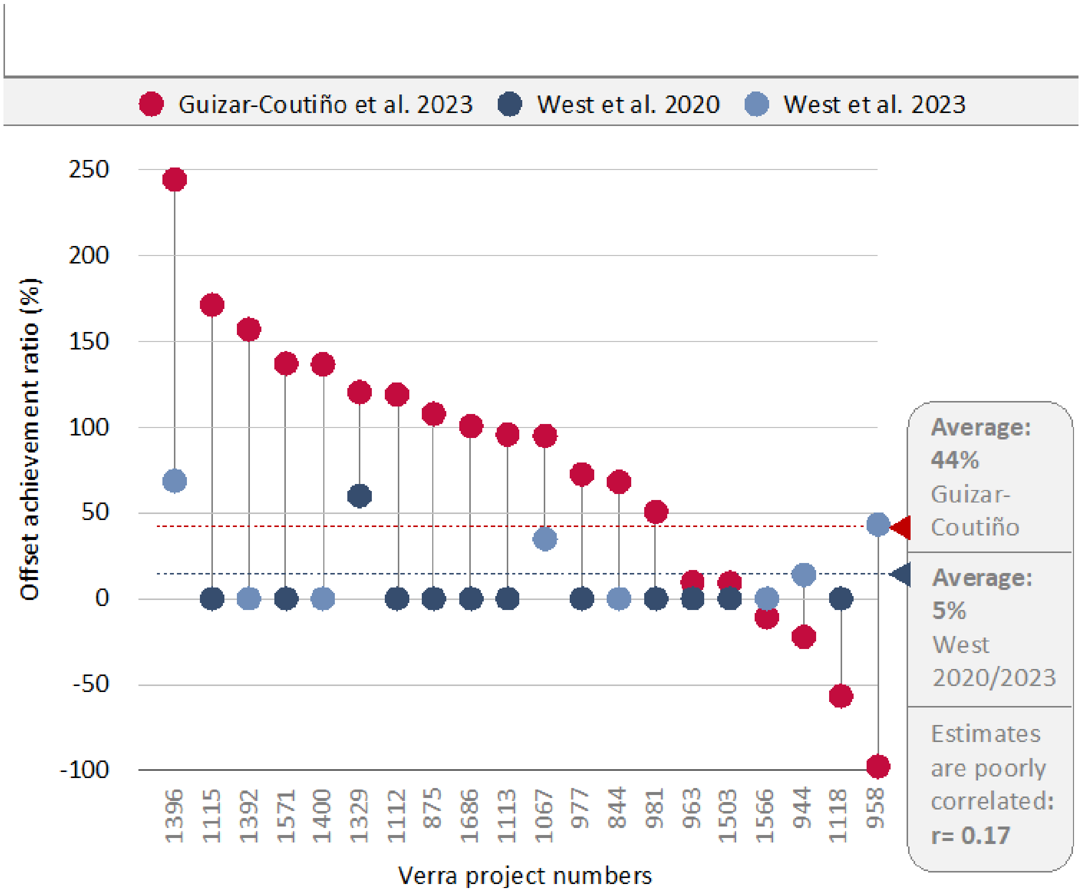


Figure 6

Comparison of offset achievement ratio for the same forestry projects across studies. Source: Authors, based on studies mentioned in Figure. For Guizar-Coutiño we divide the achieved emissions reductions reported in the paper by the ex-ante predicted emissions reductions in the project design documents by Verra. The average achievement ratio in each study is weighted by expected project emission reduction in the first 10 years. Hence, projects that are expected to avoid more CO2 are weighted more strongly.

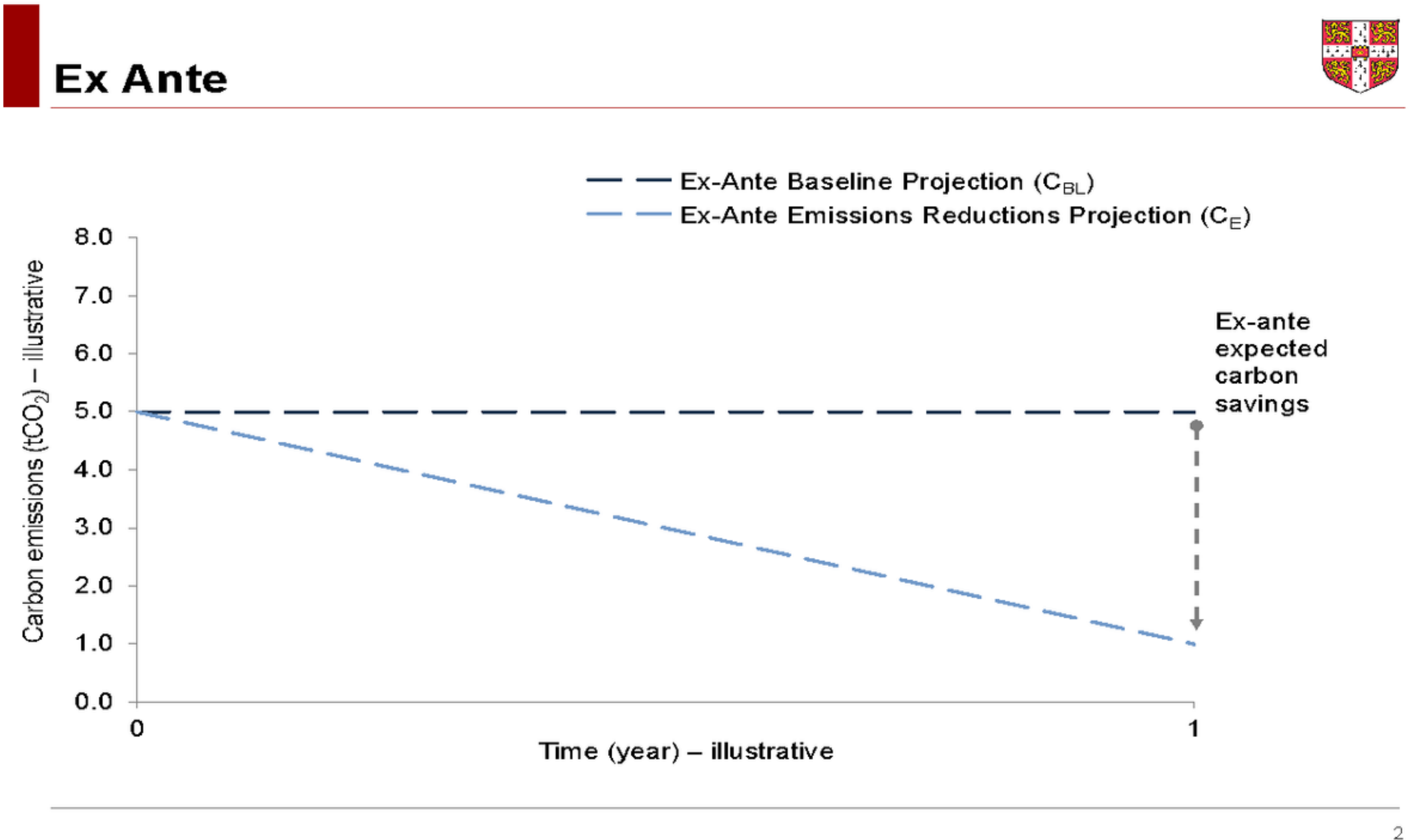


Figure 7

Illustrative ex-ante projections of carbon savings through a carbon offset project. Source: author

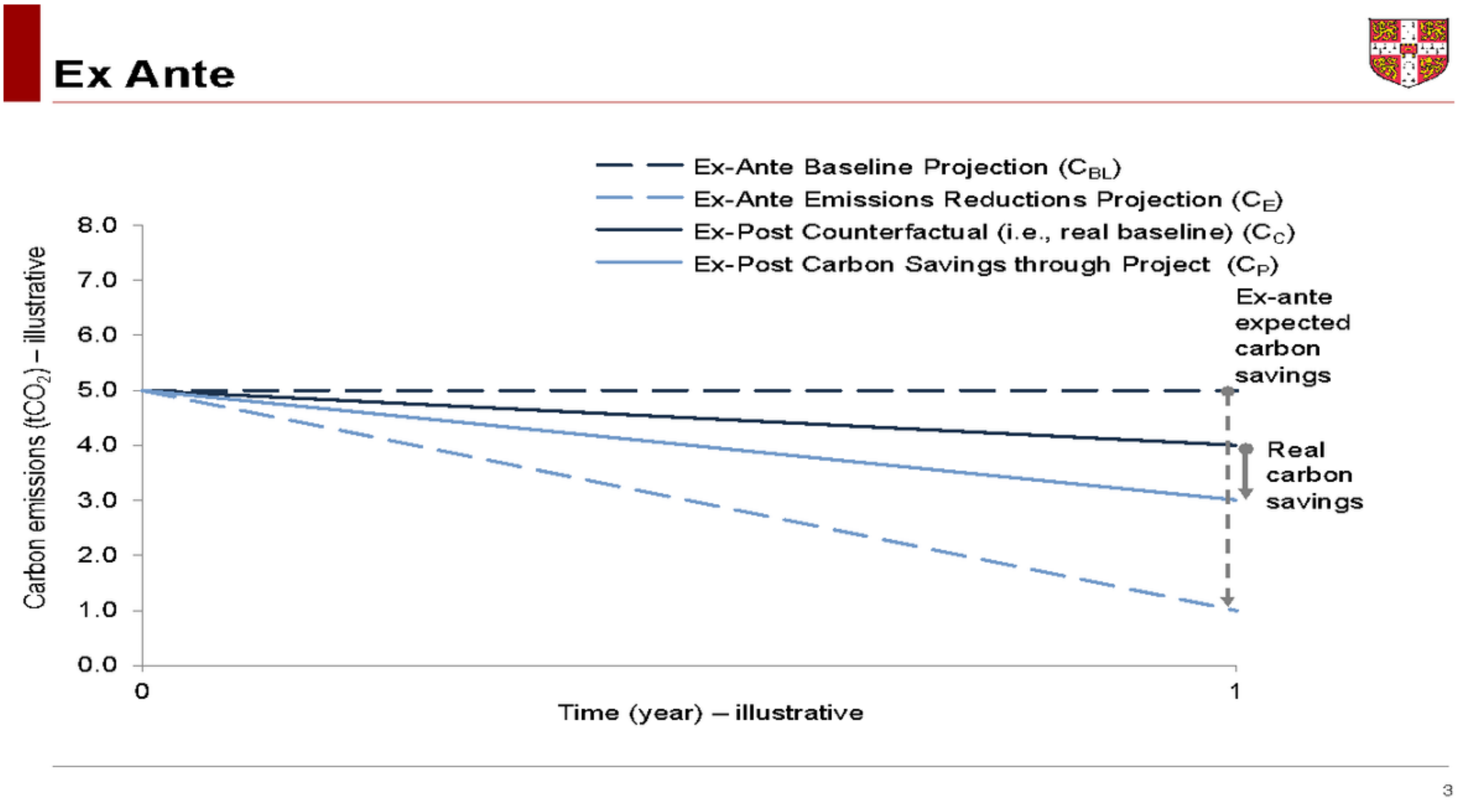


Figure 8

Illustrative ex-post assessment of carbon savings through a carbon offset project. Source: author

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 230707datavf.xlsx
- 230707SIvf.docx