The Review of Financial Studies

Willingness to Pay for Carbon Mitigation: Field Evidence from the Market for Carbon Offsets

Matthias Rodemeier
Bocconi University, IGIER, BAFFI, and IEP, Italy

This paper estimates willingness to pay (WTP) for carbon mitigation from demand for carbon offsets in a field experiment with an online supermarket. The experiment randomizes whether the firm subsidizes the price of the offset or matches the offset's impact on carbon mitigation. Consumers are price-elastic but *fully impact-inelastic*, implying that they buy the offset but their WTP for the carbon it mitigates is zero. If the firm informs consumers that it contributes to the offset costs, WTP increases to 16 EUR/tCO_2. A complementary survey shows that consumers' *stated* WTP is 238 EUR/tCO_2, far above their revealed preferences. (*JEL* D61, D82, H21, Q50, Q58)

Received: April 6, 2024; Editorial decision: February 25, 2025
Editor: Tarun Ramadorai
Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

The market for voluntary carbon mitigation has doubled in size from 2017 to 2020 and is projected to reach US $50 billion by 2030.[1] Much of this reduction comes from investments into "carbon offset" projects that engage in reforestation, which removes carbon dioxide (CO_2) from the air. Firms increasingly offer consumers the possibility to directly compensate the carbon emissions of their consumption, such as for flights or product shipping. This trend toward voluntary carbon offsetting raises an important question: What does demand for these offsets reveal about households' valuations of environmental protection?

I thank Nicola Gennaioli, Florian Heeb, Nicola Limodio, Florian Nagler, Davide Pace, Christian Rodemeier, Nicolas Serrano-Velarde, Brigitte Elise Schiavon, Gregory Sun, and participants at the AFA, Chicago, HEC Lausanne, LMU, the NBER SI, and the Paris School of Economics for insightful discussions about this paper. I also thank the editor, Tarun Ramadorai, and knowledgable referees for many helpful comments. Silvia Amalia Meneghesso and Andrea Calgaro provided outstanding research assistance. The paper was selected for the 2024 IZA Best Paper Award for Innovative Research on the Economics of Climate Change. Supplementary material can be found on *The Review of Financial Studies* website. Send correspondence to Matthias Rodemeier, matthias.rodemeier@unibocconi.it.

[1] See Blaufelder et al. (2021).

The Review of Financial Studies 00 (2025) 1–29

https://doi.org/10.1093/rfs/hhaf054 Advance Access publication September 25, 2025

Addressing this question informs both (a) cost-benefit analyses of environmental policies and (b) the value of corporate social responsibility for firm fundamentals. First, in a standard economic framework, the larger a household's willingness to pay (WTP) for carbon mitigation, the larger her benefits from policies that reduce emissions.[2] For decades, environmental economists had to resort to *hypothetical* WTP measures from contingent valuation methods in surveys. More recently, economists have started to elicit stated preferences to understand demand for sustainable assets (e.g., Stroebel and Wurgler, 2021; Giglio et al., 2025). The growing market for voluntary climate protection provides an opportunity to obtain first data points of households' *revealed* preferences (see also Pace et al., 2025).

Second, observational data has been used to understand how sustainable firm behavior and ESG rankings affect firm valuations (e.g, Hartzmark and Sussman, 2019; Berg et al., 2023). Green and Roth (2025) and Oehmke and Opp (2024) offer theories of socially responsible investing. Broccardo, Hart, and Zingales (2022) model how firm stakeholders, such as consumers, can reduce externalities by either engaging with polluting firms or boycotting them. However, there is little evidence as to how much sustainable corporate activities actually affect consumer behavior. This paper bridges the current empirical disconnect between sustainable firm investments and consumer markets.

To answer the main question of this paper, I partner with one of the largest online supermarkets in Germany and implement an experiment in their online shop, observing over 250,000 consumers and 400,000 visits to the website. In the experiment, consumers can offset the carbon emissions of grocery deliveries by buying carbon offsets. To estimate WTP, I exogenously vary both the price of the offset and the quantity of carbon compensated by the offset. Specifically, the baseline offset compensates the average emissions of a delivery: 2.4 kg of CO_2 for a price of 24 cents (i.e., mitigating 1 kg of carbon costs 10 cents). To vary price and quantity, either the price of the offset is *subsidized* by $x \in \{50\%, 75\%\}$ or the amount of carbon that the offset compensates is *matched* by $z \in \{100\%, 300\%\}$. For example, a consumer that receives a 75% price reduction only needs to pay 6 cents, and the firm covers the remaining 18 cents of the costs. A consumer that receives a 300% quantity match can offset 9.6 kg for 24 cents (instead of just 2.4 kg), and the firm covers the remaining 72 cents of the costs.

While subsidies and matches offer exogenous variation in price and impact, they also imply that the firm pays the difference between the cost of the offset and the price charged to the consumer. The increase in offset demand may then not just be driven by "intrinsic" preferences for carbon mitigation, but also by a preference to split the compensation costs with the firm. For instance,

[2] A special case under which this statement is not true is if those people who voluntarily contribute to the climate dislike environmental regulation. In Section 3, I provide evidence against this hypothesis: people with a larger stated WTP are more likely to vote for a carbon tax.

consumers might consider it fair if the firm contributes to the cost of the offset since it also benefits from the polluting transaction. Additionally, consumers may be more inclined to contribute if they learn the firm engages in sustainable investments, or they may trust the offset's effectiveness more if the firm invests its own resources in the project.[3] Therefore, I cross-randomize whether the firm informs the consumer that it shares the cost of the offset with the consumer. In the standard treatments, henceforth STANDARD, consumers who receive a subsidy or a match simply see a lower offset price or a higher offset quantity. In the information treatments, henceforth INFORMATION, the firm provides salient information to the consumer that the firm has financed the subsidy or match. This allows me to isolate the role of preferences to share the costs with the firm from the "intrinsic" valuation of carbon mitigation.

The experiment produces a number of insights. First, in STANDARD, consumers increase demand for the offset when the price falls but are *completely* inelastic to increases in the compensated quantity. Even when the offset compensates 300% more carbon than the baseline offset, demand does not increase. These results suggest that consumers buy the carbon offset, but not because of its impact on environmental protection. The conclusion is consistent with canonical theories of "warm glow" (Andreoni, 1990) and "scope-insensitivity" (Kahneman and Knetsch, 1992).

Second, in INFORMATION, where the firms' contribution is advertised, demand becomes both more price and quantity elastic. The effect of a price reduction on offsetting demand increases by up to 250% due to information provision. Doubling the impact of the offset increases demand by 11%, and quadrupling it increases demand by 22%. This finding suggests that firms can play a role in encouraging consumers to lower their carbon footprint by saliently framing corporate responsibility.[4]

The difference between STANDARD and INFORMATION delivers largely different conclusions about consumers' valuation of carbon mitigation. Using between-subject differences, I estimate that average WTP is zero in STANDARD but 16 EUR per ton of CO_2 ($p < .01$) in INFORMATION. I then estimate a structural model, using both between- and within-subject variation, to isolate the intrinsic preferences for carbon mitigation from other behavioral factors, such as warm glow, inattention to scope, and equity preferences. This model yields my final estimate of intrinsic WTP for carbon mitigation of 13 EUR/tCO_2. This is far below the current estimate by climate scientists of the Social Cost of Carbon (SCC) of 185 USD/tCO_2 (Rennert et al. 2022), as well as the SCC of the former Biden administration of 51 USD (Interagency Working

3 For a theoretical argument of this mechanism in charitable giving, see Vesterlund (2003).

4 Related studies have investigated the relative effectiveness of subsidies and matching mechanisms in increasing public good provision (e.g., Eckel and Grossman, 2003; Kesternich, Löschel, and Römer, 2016; Karlan and List, 2007; Feldman, 2010). Since my study varies the perceived contribution by the firm through an information treatment, it is the first to show that positive match elasticities are largely driven by a preference to share the costs with the firm rather than by intrinsic preferences for the public good.

Group 2021) that was used for cost-benefit analysis. This finding suggests that voluntary climate protection initiatives may internalize only a small fraction of climate externalities, which has implications for environmental policy. An important caveat to this conclusion is that the costs of carbon offsets in my experiment are small, and consumers might value carbon differently when the stakes increase.

Comparing subsidies and matches, I analyze which intervention is the most cost effective in reducing carbon emissions. I find that quantity matches are always more cost effective than subsidies even when matches do not affect demand. Subsidies are relatively expensive because they reduce the price for all consumers, but the only incremental increase in mitigation comes from *marginal* consumers, that is, from those that buy the offset due to the lower price. By contrast, matches also cause *inframarginal* consumers to mitigate more carbon, that is, those consumers that would have bought the offset even without the match. For subsidies to break even with matches, price elasticities would have to be substantially larger than they turn out to be empirically. The second result is that matches combined with INFORMATION have a “multiplier effect”: Since they increase consumers’ willingness to offset, every EUR spent by the firm on a match produces a larger reduction in total carbon emissions than if the firm used the same EUR to buy an offset directly.

Much of the existing evidence on consumers and investor preferences comes from surveys in which respondents give hypothetical answers. Through a complementary survey with customers from the same online shop, I study how much hypothetical WTP deviates from revealed preferences in the field. The mean stated WTP in the survey is 238 EUR/tCO_2. This is 1,388% larger than even my largest estimate of 16 EUR/tCO_2 in the information treatment. The result cautions against taking widely used survey responses at face value.

Contemporaneous surveys and lab experiments in economics and finance have investigated why households demand sustainable goods and invest in “green” assets. Perhaps most closely related to my study, Pace et al. (2025) finds in a survey that WTP is very concave, but not fully flat, in emissions. Another survey experiment by Heeb et al. (2023) shows that investors prefer sustainable assets but are often (though not always) insensitive to the impact of the investment on the environment.[5] My study differs from these important papers by offering the first evidence from a large-scale natural field experiment: real market participants make choices in their natural environment, not knowing they are being observed by a researcher. This may yield more accurate measures of agents’ preferences in real-world markets (List, 2007; Rodemeier, 2025).

[5] Earlier lab experiments measure people’s preferences for retiring pollution permits but do not vary the impact of the permits (Löschel, Sturm, and Vogt, 2013; Diederich and Goeschl, 2011; Diederich, 2013). This difference turns out to be pivotal in my setting: I show that ignoring the impact variation overstates true WTP for carbon mitigation *by a factor of 19* or more because it does not account for warm glow utility.

A related strand of the literature studies the role of nonstandard preferences and cognitive constraints for sustainable behavior (e.g., Andre et al., 2022; Imai et al., 2022; List et al., 2022; Rodemeier and Löschel, 2025; Löschel, Rodemeier, and Werthschulte, 2023; Schulze-Tilling, 2025; Semken, 2024). My findings illustrate that voluntary environmental contributions are partially driven by warm glow and a preference to share the contribution costs with the firm. This emphasizes a role of corporate social responsibility in promoting households' support for carbon mitigation.[6]

Finally, the paper ties into a large literature that elicits stated preferences for carbon mitigation and social discount rates in surveys (see Nemet and Johnson, 2010 for an overview of the literature). While some studies report modest WTP values of 40 USD/tCO_2 (measured in 2020-USD), many studies imply large values between 100 and 350 USD/tCO_2. My revealed preference estimates are at least an order of magnitude smaller, but I obtain similarly inflated values for stated preferences in the complementary survey. This calls for future research to develop tools to mitigate hypothetical bias in surveys.

1. Experimental Design

The experiment takes place in the webshop of one of the largest delivery services for groceries and beverages in Germany.[7] When a subject visits the website, she gets randomized into one of 10 experimental groups with equal probability. A subject is identified based on her HTTP-cookie. The experimental design involves both between- and within-subject variation in treatment. On follow-up visits, subjects are randomized again into 1 of the 10 groups.

Figure 1 visualizes the experimental design. In the treatment groups, subjects can compensate carbon emissions by buying a carbon offset. The *baseline offset* compensates 2.4 kg of CO_2 for a price of 24 cents. In the other treatments, either the price of the offset is subsidized by $x \in \{50, 75\}\%$ or the amount of carbon that the offset compensates is matched by $z \in \{100, 300\}\%$.[8]

6 On a broader level, the paper connects to the literature in climate finance (Giglio, Kelly, and Stroebel 2021) that studies whether investors value sustainability (e.g., Hartzmark and Sussman, 2019; Bauer, Ruof, and Smeets, 2021; Giglio et al., 2023; Gormsen, Huber, and Oh, 2023) and the role of beliefs about climate change and regulation (Stroebel and Wurgler, 2021; Giglio et al., 2025; Ramadorai and Zeni, 2024). In Internet Appendix C, I provide evidence that offering carbon offsets at checkout had no noticeable effects on measures of firm performance. However, since the offsets were only offered at checkout, the present results do not speak to alternative marketing techniques where offsets are saliently advertised at the outset of the website visit.

7 The experiment was implemented during February 2020.

8 Since these amounts are small, income effects and liquidity constraints are unlikely to affect demand for offsets. However, I should note that once we scale WTP for a *kilogram* of carbon to WTP for a *ton* of carbon, these factors may become important and are not considered in this paper. See Berkouwer and Dean (2022) for an example in which limited access to credit reduced the demand for energy-efficient cook stoves. Another limitation of my setting is that behavior under small stakes can yield substantially different preference parameters than under large stakes (Rabin 2000).

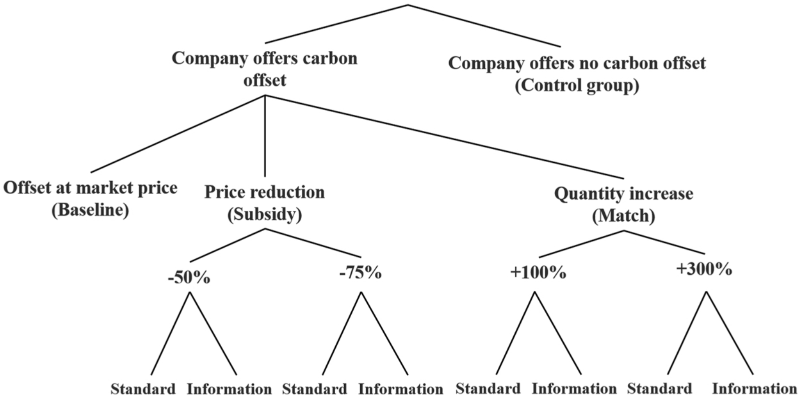

Figure 1
Experimental design
Note: This figure illustrates the experimental design. Subjects are randomized into 1 of 10 groups with equal probability upon visiting the website. Subjects are either offered a carbon offset or assigned to the control group where they see the usual shop interface. If subjects are offered an offset, it is sold at the market price (BASELINE), at a reduced price (SUBSIDY), or with an increased impact (MATCH). Subsidies reduce the price by either 50% or 75%. Matches increase the amount of carbon that is mitigated by either 100% or 300%. Finally, the experiment randomizes whether subjects receive an information treatment indicating that the firm shares the offsetting costs with the consumer (INFORMATION).

The experimental design intentionally features a symmetry between matches and subsidies. Both a 50% subsidy and a 100% match imply that the firm splits the total offset costs with the consumer 50:50. Analogously, the 75% subsidy and the 300% match imply a 25:75 split in costs between consumer and firm. This symmetry is useful because it holds the cost of mitigation per kilogram of carbon constant between a subsidy and its respective match. For instance, the carbon price for the consumer is 50 EUR/tCO_2 both when the firm subsidizes the offset by 50% and when it matches the quantity by 100%.

I also vary whether the firm advertises its own contribution to the carbon offset through an information treatment. I elaborate on this treatment further below.

Finally, after a subject makes a purchase, she is forwarded to a page that confirms the order and, in addition, asks her two questions about carbon offsetting.

1.1 Treatments

Figure 2 provides a screenshot of the baseline offset, henceforth "BASELINE." The offset is always displayed in the shopping basket of the shop, next to the list of products the subject has selected. Subjects get to that page because they want to either verify which goods they put into their shopping basket or finalize their purchase.

The offset can be added to the shopping basket by ticking the respective box next to the text *"Yes, I would like to support environmental protection and offset*

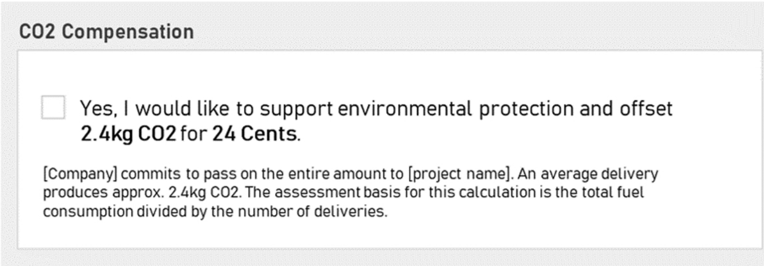

Figure 2
Carbon offset
This figure shows the baseline offset.

2.4 kg of CO_2 *for 24 cents.*" The text below informs subjects to which carbon-offsetting project the amount is donated.[9] In addition, subjects are informed that 2.4 kg of CO_2 correspond to the average emissions of one delivery.[10] This gives a reference point to consumers and helps them relate deliveries to carbon emissions. While the provided information may still be relatively abstract to consumers, we closely followed other shops when designing this treatment to replicate the typical carbon offset product in the market.

The donation goes to a reforestation project that plants trees to compensate for carbon emissions. At the time of the experiment, it cost 0.10 EUR to compensate 1 kg of CO_2 (i.e., 100 EUR/tCO_2). Thus, one average delivery that emits 2.4 kg can be compensated by 0.24 EUR.

Examples of the price and quantity variations are shown in Figure 3. Panel A shows the simple price reduction of the offset by 50%. Subjects in this group pay 12 cents for 2.4 kg of carbon instead of 24 cents. The rest of the text is identical to the baseline offset.

Panel B shows the INFORMATION treatment where the firm explicitly informs the consumer that the firm has subsidized the price by 12 cents. The additional information provides two potentially important differences relative to STANDARD. The first difference is that the consumer learns that the firm is contributing its own resources to the offsetting project and shares the burden of compensation with the consumer. This might be considered fairer by consumers and, thereby, increase demand elasticities.

Second, the information may change attention to the offset and beliefs about the offset's effectiveness. The lower price in STANDARD relative to BASELINE may signal to consumers that the offset project is of low quality and not effective at compensating carbon. A low offset price might also signal that the environmental damage of a delivery is negligible since it costs little to compensate for it. By contrast, INFORMATION should avoid this negative

9 The project name is not mentioned in this paper to protect the company's anonymity.

10 Average emissions were calculated from historical trip data.

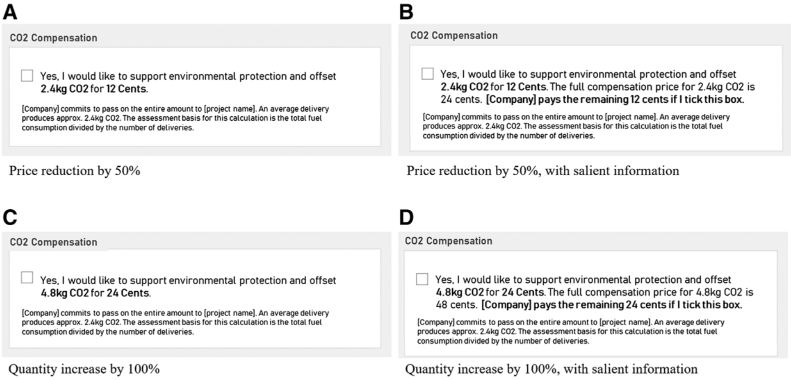

Figure 3
Examples of treatment variation
This figure shows examples of price and quantity variations. Panels A and C illustrate the variations in the standard treatments, whereas panels B and D illustrate the variations in the information treatments.

signal of low prices because subjects should be aware that the actual price of the offset is higher than the costs they have to cover. In addition, consumers might trust the offset project more if they learn that the company donates its own resources to the project.

Panel C shows an example of a quantity match. The price is equal to the one of the baseline offset, that is, 24 cents. However, the quantity is doubled from 2.4 kg to 4.8 kg of CO_2. Therefore, this treatment provides exogenous variation in the *impact* of the offset. Consumers are still informed that an average delivery produces 2.4 kg, such that they have the same reference point as in BASELINE. This should help them realize that they compensate 2 instead of 1 delivery in expectation. In general, note that any exogenous change in quantities implies, by definition, that the compensation amount deviates from the emissions of the average subject. However, this is precisely the required variation needed to identify WTP for the compensated amount of carbon.

Panel D shows the corresponding quantity match in INFORMATION. Subjects receiving the salient quantity match are informed that the full compensation price for 4.8 kg of CO_2 is 48 cents. The reason they are paying half of the amount is that the company pays the remaining 24 cents.

1.1.1 The role of the outside option for identification. Even if consumers do not choose to offset carbon emissions in the experiment, they may still reduce their carbon footprint through alternative measures outside of the webshop. This could include buying offsets on other platforms or avoiding other emission-intensive activities.[11] Such behavior could be a problem for the

[11] It should also be emphasized that the experimental design is specific to carbon offsetting as an *add-on service* offered by firms. While this practice is widespread in real-world consumer markets, the experimental design and

identification of WTP if we made the mistake of interpreting a consumer's probability to offset as the reduced-form analog to her willingness to pay. For instance, we could falsely assume that consumers with a low offset probability have a lower willingness to pay than those with a high offset probability, even though the former group might choose to offset much more carbon outside of the web shop.

My experimental design is robust to these misinterpretations and identifies WTP for carbon mitigation despite the fact that consumers have individual-specific outside options. As I explain further below, I identify WTP by the (absolute) ratio of the aggregate quantity and price *elasticities*. These elasticities are unambiguously identified in my setting because the treatment assignment is, by randomization, orthogonal to both subjects' preferences and their individual outside options. The fact that consumers may choose to reduce their carbon footprint in other contexts is consequently no threat to identification in our experiment. A more formal version of this argument is presented in Section 2.4, where I estimate WTP and explicitly allow for any arbitrary outside option.[12]

1.2 Post-purchase survey

If a subject has placed a delivery, she gets forwarded to the order confirmation page, where she is asked two questions (see Figure A1 in the Internet Appendix). The first question elicits subjects' belief about the environmental damage of a delivery if the emissions are not compensated:

> *"How large do you think are the negative consequences of your delivery for the environment if the carbon emissions of the delivery are not compensated?"*

Possible answers are presented on a 7-point Likert scale from 1 ("very low") to 7 ("very high"). The idea behind the question is that consumers might interpret a low offset price as a signal that the environmental damage of a delivery is low because it costs little to compensate a delivery.

The second question elicits beliefs about the effectiveness of the offset:

> *"How effective do you think our carbon offset program is in reducing these negative consequences?"*

findings may not directly be portable to settings where carbon emissions are directly embedded in products or financial assets, such as in Heeb et al. (2023).

12 Relative to the previous literature, the method employed in this paper differs in the way it identifies WTP. Heeb et al. (2023) and Pace et al. (2025) use the Becker–DeGroot–Marschak mechanism where subjects make multiple choices between options with varying degrees of carbon emissions, and one choice is randomly implemented at the end. This method, while widely accepted, is not applicable in natural field experiments or observational data. To overcome this challenge, I randomize prices and carbon emissions and then use this variation to estimate a random utility model in the spirit of McFadden et al. (1973) and Hanemann (1984). Understanding how different elicitation methods across studies affect WTP would be an important avenue for future research.

Possible answers are presented on a scale from 1 (“not helpful at all”) to 7 (“very helpful”).

This question is intended to test (a) whether subjects interpret a low price as a signal of low effectiveness of the offset, and (b) whether effectiveness beliefs increase as the compensated quantity increases.

Because of technical constraints, subjects using a mobile device are not forwarded to these questions after placing an order. In addition, subjects in the control group who are not offered carbon offsets cannot answer the two survey questions because they have not been offered the offset previously.

1.3 Sample

I observe 406,984 website visits by 255,376 subjects. These subjects place a total of 108,478 orders during the experimental period. Table 1 reports summary statistics for the 10 experimental groups. Here, a subject’s treatment group is defined as the one she has been assigned to during her first visit during the experimental period. Each of the 10 experimental groups consists of approximately 25,000 subjects. The balance in the number of subjects across treatments provides support for successful randomization.

The expected travel time of a delivery van is around 14 minutes across groups. The expected service time refers to the time the driver is expected to need in order to unload the delivery van. This number is larger for orders with a larger number of goods or more bulky products. Expected service time is approximately 7 minutes and balanced across experimental groups.

In the control group, the purchase probability is around 27%, and the average subject visits the website 1.6 times during the experimental period. Both of these numbers are roughly the same for the treatment groups. Note that these differences do not need to balance because they are potentially endogenous to the treatment variation.

The reported offsetting probabilities are conditional on placing an order. For the control group, the offsetting probability is zero by construction. In the other groups, the offsetting probability is positive and varies substantially across treatments.

In Internet Appendix C, I show that the treatments have no effect on the probability of buying at the store. Therefore, differences in offsetting demand across treatments have a causal interpretation since the treatments do not induce selection from the sample of website visitors into the subsample of buyers.

2. Results

2.1 Between-subject effects on offsetting behavior

Figure 4 presents between-subject effects on the offsetting probability among the subsample of buyers. I condition on subjects’ first visit to the website during the experimental period since this is their first exposure to an offset treatment. The gray bars indicate the offsetting probabilities for standard price

Table 1
Summary table

Variable	Control	Baseline: 0.24€ at 2.4kg	–0.12€	–0.18€	–0.12€, information
Number of website visits	1.593	1.596	1.595	1.588	1.602
	(1.365)	(1.828)	(1.390)	(1.350)	(1.415)
Order (1= yes)	0.329	0.330	0.333	0.327	0.333
	(0.470)	(0.470)	(0.471)	(0.469)	(0.471)
Offset (1= yes)	0.000	0.135	0.143	0.157	0.163
	(0.000)	(0.342)	(0.350)	(0.364)	(0.369)
Expected travel time (in min)	14.508	14.366	14.498	14.509	14.561
	(9.397)	(9.433)	(10.110)	(9.582)	(9.825)
Expected service time (in min)	7.201	7.260	7.304	7.282	7.296
	(3.817)	(3.650)	(4.071)	(3.815)	(3.773)
N	25,564	25,427	25,654	25,556	25,643

Variable	-0.18€, information	+2.4kg	+7.2kg	+2.4kg, information	+7.2kg, information
Number of website visits	1.584	1.591	1.598	1.592	1.598
	(1.617)	(1.526)	(1.492)	(1.462)	(1.449)
Order (1= yes)	0.332	0.333	0.334	0.330	0.331
	(0.471)	(0.471)	(0.472)	(0.470)	(0.471)
Offset (1= yes)	0.185	0.128	0.133	0.150	0.164
	(0.388)	(0.334)	(0.339)	(0.357)	(0.371)
Expected travel time (in min)	14.525	14.442	14.428	14.470	14.685
	(9.546)	(9.319)	(9.464)	(9.562)	(9.832)
Expected service time (in min)	7.371	7.334	7.305	7.285	7.302
	(3.855)	(3.781)	(4.048)	(3.921)	(4.159)
N	25,375	25,564	25,762	25,642	25,189

Note: This table presents the mean of observable variables in different treatment conditions. Standard deviations are reported in parentheses.

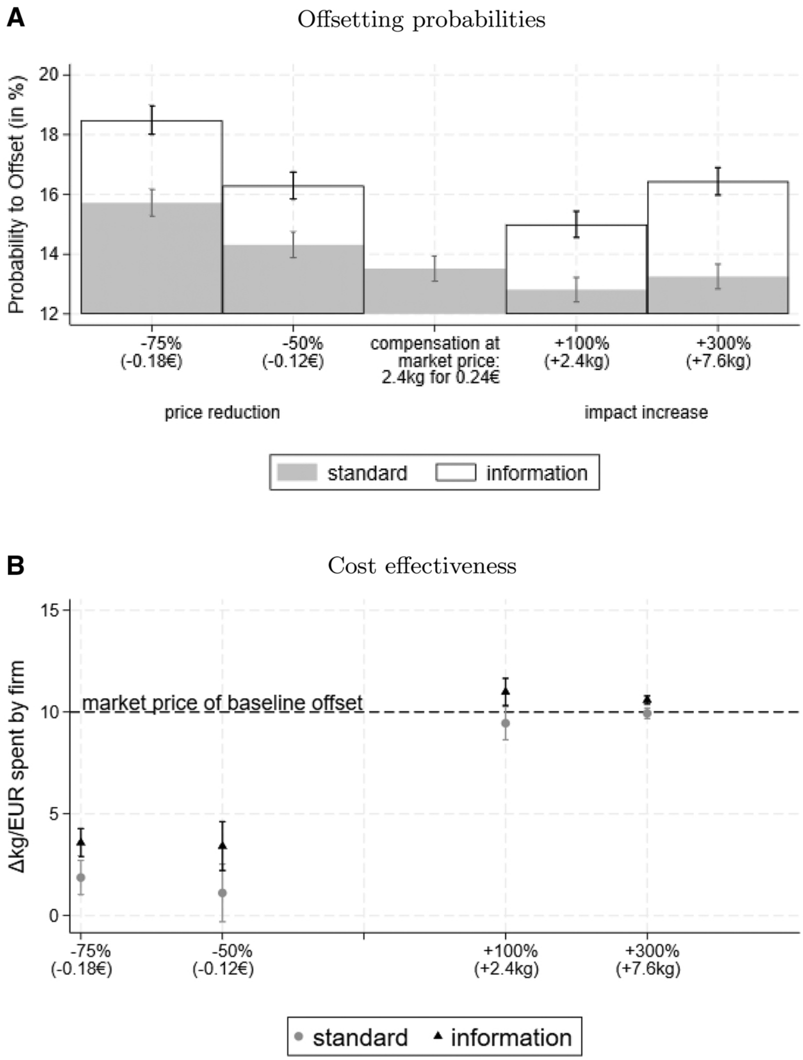

Figure 4
Main results
Panel A represents the offsetting probabilities across treatments. Gray bars represent standard treatment groups, transparent bars represent information treatment groups. Panel B plots the increase in compensated kilograms per EUR spent by the firm, relative to the baseline offset. The dotted line represents the market price of the baseline offset (10 kg/EUR).

and quantity variations, as well as for the baseline offset. The transparent bars show offsetting probabilities for the salient price and quantity variations. Figure 4b shows the cost effectiveness of each intervention, which I discuss later in Section 2.5.

At the baseline price of 24 cents for 2.4 kg, 13.5% of customers choose to buy the offset. If the offset price falls by 12 and 18 cents, the offsetting probability increases by 0.8 and 2.2 percentage points, respectively. This implies a convex demand curve with price elasticities of –0.12 and –0.31. In other words, the incremental demand response to a given price reduction is larger at lower price levels. The larger price reduction is statistically significant at conventional levels.

We observe an even more pronounced pattern for quantity variations. Increasing the amount of carbon compensated by the offset does not increase demand in STANDARD. The offsetting probabilities are even slightly lower than BASELINE when quantities increase but these differences are not statistically different from zero. Using these results to identify elasticities would imply that consumers are completely inelastic to compensated quantities. Even when the compensated quantity is increased by 7.2 kg, which is a large relative increase of 300% relative to BASELINE, the offsetting probability does not change. Taking these point estimates at face value yields the conclusion that consumers buy carbon offsets, but not because of how much carbon they offset. WTP for voluntary carbon mitigation is zero. This conclusion is in line with models of "warm glow" (Andreoni 1990) in which people receive binary utility from the act of giving but do not care about the impact of their donation. Similarly, the results are in line with the Kahneman and Knetsch (1992) finding that people's *hypothetical* willingness to pay for a public good is "insensitive to scope" in hypothetical choices (e.g., rescuing a bird vs. rescuing an entire species). However, results may also suggest that consumers have an imperfect understanding of the product they are buying and do not understand kilograms of carbon as a measure of impact. In addition, consumers may be inattentive to the impact of the offset.

Offsetting behavior changes substantially when consumers are explicitly informed about the firm contribution. Demand becomes more price-elastic and consumers suddenly become sensitive to scope. The price reductions now increase demand by 2.8 and 5 percentage points, respectively. Both effects are highly statistically significant with $p < .01$. Put differently, making the price variations salient increases its effects by 250% and 127% for the 12 and 18 cent reductions. The price elasticities are now -0.53 and -0.84. In INFORMATION, increasing the compensated carbon by 2.4 kg and 7.2 kg raises the offsetting probabilities by 1.5 and 3 percentage points (both at $p < .01$). These responses are relative treatment effects of 11% and 22% compared to baseline. The point estimates imply quantity elasticities of 0.22 and 0.19. Consequently, in the presence of salient matches, consumers exhibit a responsiveness to the impact of the offset. Note, however, that these responses are still relatively small, considering that the firm doubled and quadrupled the amount of compensated carbon.

An open question is whether the increase in the impact elasticity in INFORMATION is due to consumers' recognition of the greater impact of the

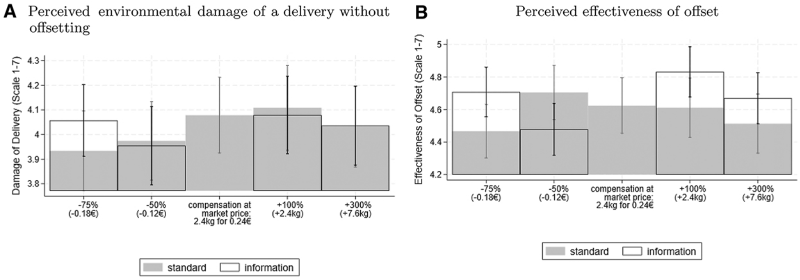

Figure 5
Post-experimental survey
Panel A illustrates subjects' beliefs about the size of the environmental damage of one delivery that is not compensated by an offset. Panel B shows beliefs about the effectiveness of the offset in mitigating environmental damages.

offset, or whether it stems from their appreciation of the firm's contribution to the offset. I explore these aspects next.

2.2 Effect of information on beliefs

Figure 5 illustrates differences across treatments in consumer perceptions elicited in the post-purchase survey. Looking at panel A, we do not find statistically significant differences between STANDARD and INFORMATION in the perceived environmental damage of an uncompensated delivery. A general tendency seems to be that the perceived damage decreases as the costs for the consumer fall. However, this is also true when consumers know that the true offset price is higher than what they pay. Overall effects are relatively noisy and small in absolute size.

Panel B plots the perceived effectiveness of an offset across treatments. While it is hard to draw stark conclusions from the figure, a couple of tendencies emerge. Perhaps most importantly, in STANDARD, the perceived effectiveness of the offset barely increases as the quantity of mitigated carbon increases. This implies that consumers may not understand quantity increases. Information seems to reduce this misperception for the 100% match as perceived effectiveness increases by 6% (0.2 points on the Likert scale).[13] The effect of information for the much larger 300% match is qualitatively similar, but not statistically significant. These results are not fully conclusive but point to the possibility that the previously observed insensitivity to scope in STANDARD may partially be explained by consumers' inattention toward the compensation amount.

[13] These results are relevant to an early model in philanthropy by Vesterlund (2003) arguing that information about a fundraiser's own contribution to the charity increases donors' perceived quality of the charity.

2.3 Within-subject effects

One potential mechanism by which subjects are impact-inelastic is that they have a difficult time understanding kilograms of carbon as a measure of impact. We can explore this mechanism by exploiting within-subject variation in impact on repeated visits to the website. If subjects learn about kilograms of carbon as a measure of impact as they see varying compensation amounts, we would expect them to become impact-elastic over time. To test this, I run the following panel regression:

$$\text{Offset}_{it} = \eta_1 Subsidy_{it} + \eta_2 \times Subsidy_{it} \times I_{it} + \beta_1 Match_{it} + \beta_2 \times Match_{it} \times I_{it} + \lambda_i + \mu_t + \epsilon_{it}, \quad (1)$$

where $Subsidy_{it} \in \{0,1\}$ and $Match_{it} \in \{0,1\}$ are indicators that pool price and impact treatments, respectively. The interactions with $I_{it} \in \{0,1\}$ measure how price and quantity responses change with INFORMATION. I include subject fixed effects, λ_i, as well as time-fixed effects μ_t.[14] Therefore, the corresponding treatment effects constitute behavioral changes *within-subject* rather than between-subject.

Table 2 reports results. For comparison, I first report the main between-subject treatment effects from Section 2.1 in column 1. Column 2 conditions on the subsample of 9,102 buyers for which repeated purchases can be observed and estimates Equation (1). The first notable result is that INFORMATION has a smaller effect within- than between-subject. The second result is that the within-subject effect of matches is now *positive* and amounts to an average increase in offsetting of 2 pp. This effect is not far away from the subsidy coefficient in the same column (2.6 pp). Since the sample is much smaller, coefficients are generally less precisely estimated. Yet, the match coefficient has a p-value of $p = 0.054$, placing it just at the edge of statistical significance at conventional levels. The effect is still not economically large considering that matches doubled or quadrupled the mitigated impact. However, it suggests that repeated interactions with offsets may have helped subjects learn about differences in impact and start appreciating it even in STANDARD. This result is consistent with models on attention and memory, hypothesizing that experiences with abstract attributes in the past generate an anchor for valuations in the future (e.g., Bordalo, Gennaioli, and Shleifer, 2020).

2.4 Structural model

In this section, I organize the previous findings through a unified framework. I use this framework to quantify the relative importance for carbon offsetting of (a) intrinsic WTP for carbon mitigation, (b) warm glow from donating,

[14] Time fixed effects control for the day of the visit as well as for the number of visits. The latter effect is absorbed by constructing a variable that, for each observation, indicates how many website visits have been done by the subject.

Table 2
Between- and within-subject treatment effects

	Offsetting Probability ×100	
	(1)	(2)
subsidy	1.560***	2.579**
	(0.460)	(1.084)
× information	2.394***	1.231
	(0.396)	(0.876)
match	–0.326	1.959*
	(0.450)	(1.015)
× information	2.563***	0.373
	(0.378)	(0.899)
Baseline Probability ×100	13.52	12.90
Time FE	Yes	Yes
Subject FE	No	Yes
R^2	0.01981	0.74159
N	76,229	9,102

*Note:*This table reports treatment effects on the offsetting probability among subjects that purchased at the shop. Column 1 reports between-subject effects during the first visit. Column 2 shows within-subject effects. Specifically, in column 2 we use the subsample of buyers that purchase more than once and include subject fixed effects. Standard errors are clustered on the subject level and presented in parentheses. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

(c) inattention to and learning about impact, and (d) preferences for a fair division of costs between consumer and firm.

As a baseline model, I use the usual random utility model (McFadden et al., 1973), which has also been extensively used in contingent valuation studies for public goods (Hanemann, 1984). I then add behavioral extensions to this model such as inattention to scope and contrast effects as in Bordalo, Gennaioli, and Shleifer (2020), as well as fairness preferences about the division of compensation costs, following the influential work of Fehr and Schmidt (1999).[15]

2.4.1 The baseline model. Consumer $i \in \{1,2,\dots I\}$ can choose between buying a carbon offset and an outside option, where utility from the outside option is normalized to zero. The carbon offset compensates γ_i units of carbon. The consumer pays price $p_i = \rho - s_i$ where ρ is the market price of the offset and s_i the price reduction sponsored by the firm. I make the usual assumption that p_i and γ_i enter linearly into utility.[16] I let ω measure the warm glow utility from donating which is independent of the donation's effectiveness (Andreoni, 1990). Idiosyncratic preferences are given by ϵ_i and follow a smooth distribution function G. Utility in the benchmark model is then given by:

$$u_i = \omega + \beta\gamma_i + \eta p_i + \epsilon_i. \tag{2}$$

[15] Bushong, Rabin, and Schwartzstein (2021) provide another important theory of contrast effects. Their model is not directly applicable to my setting as they model static rather than sequential choices.

[16] In an unreported regression, I allow for nonlinearities in the utility function and cannot reject that they are statistically zero. See also Internet Appendix B.

The consumer buys the offset iff $u_i \geq 0$. Aggregate demand for the offset is $D = 1 - G(\xi)$ with $\xi = -\beta\gamma - \eta p - \omega$. I do not impose a distributional assumption on ϵ_i and rather linearly *approximate* WTP by reduced-form elasticities (see Chetty, 2009).[17] Specifically, the derivatives of aggregate demand with respect to price and carbon quantity are $\frac{\partial D}{\partial p} = \eta g(\xi)$ and $\frac{\partial D}{\partial \gamma} = \beta g(\xi)$, such that WTP is given by $WTP = -\frac{\beta}{\eta} = -\frac{\partial D}{\partial \gamma} / \frac{\partial D}{\partial p}$. Note that this approach makes the assumption that the consumer believes that the offset in fact compensates the promised amount. In Internet Appendix B, I relax this assumption by introducing subjective uncertainty over the offset's effectiveness (e.g., because of distrust in offsets) and by allowing for risk aversion. I find that the main estimates derived in this section are robust to a reasonable range of subjective probabilities that are in line with survey responses.[18]

Finally, the money-metric utility gain from warm glow is given by $-\frac{\omega}{\eta}$. We can estimate the normalized utility parameters through a linear model of the form:

$$D_i = \alpha + \tau \gamma_i + \zeta p_i + \xi_i \tag{3}$$

$$\approx -\frac{\omega}{\eta}\zeta - \frac{\beta}{\eta}\zeta \gamma_i + \zeta p_i + \xi_i. \tag{4}$$

This baseline model is estimated by ordinary least squares.

2.4.2 Incorporating inattention and fairness preferences. Motivated by our previous findings, I extend the model by considering the possibility that subjects pay attention to impact only with probability $\theta \in [0,1]$. Any $\theta < 1$ mutes the unbiased offsetting elasticity with respect to γ_i.

Further, I allow consumers to have equity concerns $F(p_i, s_i, I_i) = \phi \max\{s_i I_i - p_i, 0\} + \rho \max\{p_i - s_i I_i, 0\}$ about how costs should be divided between the firm and the consumer.[19] The parameterization of $F(p_i, s_i, I_i)$ follows Fehr and Schmidt (1999) but, in addition, accommodates our empirical setting where the firm's contribution is either salient $I_i = 1$ or shrouded $I_i = 0$.[20] If the firm's contribution is salient and exceeds the consumer price, $s > p_i$, the consumer gets marginal utility ϕ for every EUR by which the firm's subsidy exceeds the consumer price. Conversely, for every EUR by the firm for which a

17 The traditional approach is to assume a logit distribution for ϵ_i and estimate the model by maximum likelihood. If demand is low, as is the case in the experiment, a logit distribution imposes too much probability mass on negative values, implying that a large share of consumers would get negative utility from buying the offset.

18 See Calel et al. (2025) and Kim, Li, and Wu (2024) for evidence on adverse selection problems in the offsetting market.

19 The axiomatic foundation of the fairness model aligns with that of a standard neoclassical model, as the firm's contribution can simply be interpreted as a consumption good that provides utility.

20 Another difference is that I hypothesize that consumers have fairness preferences over the division of donation costs, while Fehr and Schmidt (1999) study a subject's preference over the division of money for own consumption.

salient subsidy is below the consumer price, $s_i < p_i$, the consumer gets marginal utility ρ. Both ϕ and ρ may be positive or negative, representing either a gain or a loss from an unequal division of costs. For equal splits, $p_i = s$, the consumer gets no fairness utility, $F(p_i, p_i, 1) = 0$. If the firm's contribution is shrouded, we have $F(p_i, s_i, 0) = \rho p_i$.[21]

The final specification that incorporates these behavioral extensions is

$$u_i = \omega + \theta\beta\gamma_i + \eta p_i + F(p_i, s_i, I_i) + \epsilon_i. \quad (5)$$

The utility parameters can again be approximated by a linear model:

$$D_i = \alpha + \tau\gamma_i + \zeta p_i + \psi \max\{s_i I_i - p_i, 0\} + \varphi \max\{p_i - s_i I_i, 0\} + \xi_i \quad (6)$$

$$= -\frac{\omega}{\eta}\zeta - \theta\frac{\beta}{\eta}\zeta\gamma_i + \zeta p_i - \frac{\phi}{\eta}\zeta \max\{s_i I_i - p_i, 0\} - \frac{\rho}{\eta}\zeta \max\{p_i - s_i I_i, 0\} + \xi_i \quad (7)$$

The term $-\frac{\rho}{\eta}$ is of particular policy interest: It measures the consumer's willingness to donate (beyond their baseline donation) for every EUR donated by the firm. Put differently, $-\frac{\rho}{\eta}$ is a multiplier effect of sustainable firm engagement: Every EUR invested by the firm causes an additional donation of $-\frac{\rho}{\eta}$ EUR by consumers (conditional on the firm paying weakly more than the consumer).

2.4.3 Identification and estimation. For identification, I leverage both between- and within-subject variation. I assume that subjects are inattentive to impact when it is varied between-subject, $\theta = 0$, but attentive in within-subject variation, $\theta = 1$, because the impact of the previously seen offset serves as an anchor that helps subjects to value the attribute Bordalo, Gennaioli, and Shleifer (2020).[22] Put differently, β is identified only through within-subject variation.

Warm glow utility, ω, is the constant in the utility function measuring utility from offsetting that is unrelated to the impact γ. The fairness term $\max\{s_i I_i - p_i, 0\}$ is identified from treatments in which the firm's contribution

21 This model interprets the INFORMATION effect as primarily resulting from consumer preferences to share the compensation amount with the firm. This formulation abstracts from some alternative mechanism. One such mechanism is that subjects receiving the INFORMATION treatment may consider the firm as overall more environmentally sustainable. This mechanism is generally plausible but appears less likely in our setting for three reasons. First, under this mechanism we would expect that the INFORMATION treatment affects firm outcomes, such as the buying probability or revenues, by increasing consumers' willingness to engage with the firm. Yet, we find no effects along these dimensions. Second, this alternative mechanism would imply that consumers think the firm is overall more environmentally friendly. For instance, the firm may use electric vehicles rather than conventional fuel-powered vehicles or adopt renewable energy sources for its operations. However, in the post-purchase survey from Section 2.2, we find no effect of INFORMATION on beliefs about the environmental harm of a transaction. Third, survey evidence in Section 3 provides strong evidence that consumers have explicit equity preferences to split the compensation costs with the firm. Despite this evidence in favor of equity concerns, it is important to remember that alternative mechanisms could be at play.

22 If this assumption is violated because consumers are still not fully attentive to within-subject variation, my estimates represent an upper bound of θ.

(through either subsidies or matches) is salient, that is, from INFORMATION groups. The second term, $\max\{p_i - s_i I_i, 0\}$, is identified from price variation in STANDARD where the consumer perceives to pay more than the firm.[23]

I estimate utility parameters jointly from demand moments between-subject, D_i^b, and within-subject variation, D_{it}^w, using the following two equations:

$$D_i^b = \zeta p_i - \underbrace{\frac{\omega}{\eta}\zeta}_{\text{warm glow}} \underbrace{-\frac{\phi}{\eta}\zeta \max\{s_i I_i - p_i, 0\} - \frac{\rho}{\eta}\zeta \max\{p_i - s_i I_i, 0\}}_{\text{equity concerns}} + \xi_i^b, \quad (8)$$

$$D_{it}^w = \left(\zeta p_i - \underbrace{\frac{\beta}{\eta}\zeta}_{\text{WTP}} \gamma_i \underbrace{-\frac{\phi}{\eta}\zeta \max\{s_{it} I_i - p_{it}, 0\} - \frac{\rho}{\eta}\zeta \max\{p_{it} - s_{it} I_{it}, 0\}}_{\text{equity concerns}} \right) \underbrace{\Omega}_{\text{attenuation effect}} + \underbrace{\lambda_i + \mu_t}_{\text{fixed effects}} + \xi_{it}^w. \quad (9)$$

The important differences between the two equations are that (a) only D_{it}^w is used to estimate WTP, $\frac{\beta}{\eta}$, (b) D_{it}^w includes subject and time fixed effects, λ_i and μ_i, which means that $\frac{\beta}{\eta}$ represents a within-subject treatment effect, (c) D_{it}^w does not include the term $\frac{\omega}{\eta}$ because constant terms are absorbed by the fixed effects, and (d) D_{it}^w controls for a general attenuation in treatment effects over time by multiplying all coefficients by Ω.

Point D deserves additional explanation. Treatment coefficients often tend to become smaller within-subject but this decrease over time may not be due to a change in subjects' underlying deep primitives. Instead, attenuation could be caused by other idiosyncratic factors that are unrelated to economic fundamentals. For example, the INFORMATION treatment may be more salient to consumers on the first visit than on follow-up visits to the shop (see, e.g., Allcott and Rogers, 2014). The parameter Ω absorbs the average attenuation effect across treatments.[24]

I jointly estimate Equations (8) and (9) by the Generalized Method of Moments (GMM) and find the optimal weight matrix using the two-step GMM estimator.

[23] Note that the term is always zero in INFORMATION because, in the experiment, whenever the firm provides information, the subsidy is at least as large as the price paid by the consumer ($s_i \geq p_i$ if $I_i = 1$). Since the experimental design only varies information in which the firm saliently pays more than the consumer, one might argue that consumers in STANDARD never think about the fact that they pay more than the firm. In this case, a better model would be one in which $F = \max\{s_i I_i - p_i, 0\}$, so the term $\max\{p_i - s_i I_i, 0\}$ is dropped from the estimation. In an unreported estimation, I find that this alternative specification yields almost identical results.

[24] This adjustment is not crucial to the estimation results. An alternative estimation without allowing for an dynamic attenuation effect (i.e., forcing $\Omega = 1$) yields a WTP of 12.92EUR/tCO_2, which is virtually identical to the estimate in Table 3.

Table 3
Willingness to pay for carbon mitigation, warm glow, and fairness

	(1)	(2)	(3)
	Baseline Model		Behavioral Model
	Standard	Information	Entire Sample
WTP in €/tCO_2 ($-\frac{\gamma}{\eta}$)	–0.18	15.99***	12.84**
	(5.59)	(2.51)	(5.13)
Warm glow utility in € ($-\frac{\omega}{\eta}$)	1.27***	0.74***	0.86***
	(0.24)	(0.07)	(0.14)
Fairness disutility per € above firm contribution ($-\frac{\phi}{\eta}$)			0.08
			(0.16)
Fairness utility per € below firm contribution ($-\frac{\rho}{\eta}$)			0.21***
			(0.06)
Price coefficient (ζ)	0.13***	–0.25***	–0.21***
	(0.03)	(0.03)	(0.04)
Dynamic Attenuation Effect (Ω)			0.84***
			(0.17)
Sample	Buyers' first visit	Buyers' first visit	Buyers in panel
Time & Zip-Code FE	Yes	Yes	Yes
N for between-subject estimation	42440	42186	76229
N for within-subject estimation			9102

Note: This table reports coefficients from the empirical specifications in Equation (4) and (7). Coefficients in columns 1 and 2 are obtained by estimating the baseline model (Equation (4)) from the sample in STANDARD and INFORMATION, separately. Column 3 reports coefficients from the extended model that allows for inattention and fairness preferences (Equation (7)). The first four coefficients are measured in EUR. The final coefficient, ζ, measures the change in demand in percentage points for every one EUR change in price. Standard errors in parentheses. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

2.4.4 Results. Estimation results are shown in Table 3. In the baseline specification, I estimate the model for STANDARD and INFORMATION separately to illustrate the difference in WTP estimates. Subjects with the baseline offset are included in both estimations.

As shown in column 1, using the variation in STANDARD to estimate utility parameters, we find that WTP for carbon mitigation is indistinguishable from zero, suggesting consumers do not value the carbon-mitigating attribute of the offset. Instead, warm glow utility is highly statistically significant and amounts to, on average, 1.27 EUR ($p<.01$). Column 2 uses variation in INFORMATION and yields a WTP estimate of 16 EUR/tCO_2 ($p<.01$). Warm glow utility amounts to 0.74 EUR ($p<.01$) per offset. The empirical moments in STANDARD and INFORMATION, therefore, deliver vastly different implications for why consumers buy carbon offsets: In one case consumers seem to be entirely driven by warm glow, while in the other case they exhibit strong intrinsic preferences for the effectiveness of the offset.[25]

Column 3 suggests that inattention to the impact of the offset might have muted consumers' true WTP in STANDARD. Identifying WTP from

[25] The reason warm glow utility is lower in column 2 is that INFORMATION increases the price elasticity (as reported in the second-to-last column), which in turn implies that the utility value of the regression constant α becomes smaller.

our within-subject variation, we obtain a WTP of 12.84 EUR/tCO_2 ($p <$ .05). Warm glow now amounts to 0.86 EUR per offset ($p < .01$), which is somewhere between the estimate in STANDARD and INFORMATION. In terms of fairness preferences, the coefficient of $-\frac{\phi}{\eta}$ is close to zero and statistically insignificant, suggesting that subjects do not get direct disutility from paying more than the firm. Instead, strong evidence indicates that consumers appreciate when the firm pays weakly more. The coefficient $-\frac{\rho}{\eta}$ implies that every EUR donated by the firm above the consumer's contribution increases willingness to pay for the offset by 0.21 EUR ($p < .01$). This is a sizeable multiplier effect of sustainable firm investment.

Finally, we observe a significant dynamic attenuation of treatment effects in general. The parameter $\Omega = 0.84$ implies that within-subject elasticities are 84% of between-subject elasticities.

In conclusion, the results indicate that (a) scope-insensitivity in STANDARD is at least partially driven by inattention to impact, and (b) the increased responsiveness to scope in INFORMATION is partially driven by consumers' equity concerns. Point A is broadly in line with the results discussed in Section 2.2, where perceived effectiveness did not change between-subject as we varied the impact. Point B suggests an important role of corporate sustainable engagement in promoting voluntary contributions to climate protection among households.

2.5 Cost effectiveness of sustainable firm practices: Subsidies versus matches

What is the cost-effectiveness profile of subsidies and quantity matches, and how does information change this profile? This question is not just important for policy makers but also for stakeholders who seek to maximize the impact of sustainable firm investments.

To quantify cost effectiveness, I calculate the difference in compensated carbon between an intervention (subsidy or match) and the baseline offset. I then divide this number by the total monetary contributions made by the firm on that intervention. We can interpret this number as the incremental increase in compensated carbon of the intervention per EUR spent by the firm.

Panel B in Figure 4 visualizes the results, again conditioning on the first interaction that each buyer had with a treatment during the experiment. The dotted gray line marks the market price if the firm directly buys the offset instead of offering it to consumers (i.e., the baseline price of 10 kg/EUR). Quantity matches are always more cost effective than subsidies, *even when matches have no impact on demand*. The reason for this is that with subsidies, the only incremental increase in compensated carbon comes from marginal consumers. By contrast, with matches, the increase in compensated carbon also comes from inframarginal consumers since every offset now compensates a larger amount. Price elasticities would have to be much larger for subsidies to be more cost effective than quantity matches.

In terms of magnitudes, we see that subsidies in STANDARD increase the compensated quantity by approximately 1 kg/EUR and 2 kg/EUR per invested EUR for the 12 and 18 cent subsidies, respectively. Only the latter is statistically significant from zero. INFORMATION, instead, increases the benefit-cost ratio of both subsidies substantially. The cost-effectiveness ratio becomes 3.4 kg/EUR and 3.6 kg/EUR, respectively. The ratio is always below the market price of 10 kg/EUR. This means the firm could offset more carbon if they used the money spent on subsidies and purchased carbon offsets directly instead.

By contrast, quantity matches in STANDARD just break-even with the market price of 10 kg/EUR and are thereby more than 2.5 times more cost effective than subsidies. The quantity matches in INFORMATION are able to offset more carbon per EUR spent, implying that matches can have a multiplier effect. In particular, every EUR spent by the firm compensates around 11 kg of carbon, that is, 10% more than if the same EUR were invested directly into the baseline carbon offset.

This result suggests that firms may leverage consumer preferences for corporate social responsibility to more efficiently invest in offset projects.

3. Revealed versus Stated Household Preferences

To further understand consumers' preferences, I implement a second survey several months after the field experiment. Customers receive an email from the company inviting them to take an opinion survey. The survey investigates how stated preferences for carbon mitigation respond to changes in the impact of carbon offsets, to an education treatment about carbon offsetting, as well as to the firm's contribution to the offset. It also sheds light on people's preferences for a carbon tax as an alternative protective policy.[26] A translated version of the survey can be found in Internet Appendix I.

3.1 Survey design

To elicit subjects' stated preferences, they receive two questions that elicit their hypothetical WTP. First, they are asked how much they are willing to pay to compensate $x \in \{2.4, 4.8\} kg$ of CO_2, where the amount they see is randomly assigned. Directly after that, they are asked how much they would be willing to pay to compensate a higher amount $y \in \{4.8, 9.6\} kg$ of CO_2. Subjects who see 2.4 kg in the first question, then see 4.8 kg in the second. Analogously, subjects who first see 4.8 kg, next see 9.6 kg. This creates both within- and between-subject variation in the compensation amount and allows me (a) to estimate the distribution of stated WTP, and (b) to test whether subjects are inattentive to scope between- and within-subject.

In addition, I randomize a treatment in which subjects receive additional information in the second question on WTP that the firm *matches the*

[26] For privacy constraints, I cannot match survey participants to the observations in the field experiment.

compensation amount on its own cost to $Y \in \{4.8, 9.6\}$kg of CO_2. This treatment allows us to investigate the effect on stated WTP of a quantity match by the company.

Finally, I investigate whether education about carbon offsetting affects WTP. I randomize a treatment in which subjects see three stylized facts about carbon emissions before answering the WTP questions. Treatment subjects are informed (a) that an average delivery emits 2.4 kg of CO_2 (as in the field experiment), (b) that one would have to drive 11 km in an average car to emit the same amount of carbon as the delivery, and (c) that one would have to plant 5 beech trees, on average, to compensate 2,000 deliveries. Subjects are then randomly asked about one of these facts in a follow-up question to test their understanding.

In Internet Appendix G, I describe the sample in more detail and discuss observable characteristics. Subjects are more likely to be male, slightly younger than the average German citizen, and less likely to be unemployed. While I cannot exclude that subjects select on unobservables into the survey, observable statistics are fairly representative of the firm's customer population. They also match the typical profile of a U.S. customer that shops groceries online, according to market research data (Capital One, 2024).

3.2 Results

Table 4 reports results from an OLS regression of WTP on the treatments. As is common in the literature that measures WTP with open-ended questions, I adjust for outliers by only considering the 90th percentile of WTP answers.[27] Column 1 is stated WTP in cents. The constant implies that subjects in the first question state a WTP of 57 cents. This translates into 238 EUR/tCO_2 as reported at the bottom of the table. The estimate falls into the range of prior estimates from contingent valuation studies (e.g., Hersch and Viscusi, 2006; Viscusi and Zeckhauser, 2006; Nemet and Johnson, 2010; Brouwer, Brander, and Van Beukering, 2008; Carlsson et al., 2012; Achtnicht, 2012): numbers range from 40 to 350 USD/tCO_2 (measured in 2020-USD). Overall, the stated preference approach used in the survey does not capture the revealed preference estimate from the experiment. If we were to take 16 EUR/tCO_2 as our preferred estimate, the survey results would overstate WTP by 1,388%.[28]

Stated preferences do not significantly change when consumers receive the information that the firm contributes to the offset. One coefficient is even

---

[27] More specifically, I use the 90th percentile of WTP *per* tCO_2. It is important to normalize in this context as subjects have been offered different compensation amounts. If we do not exclude outliers, stated WTP estimates become more inflated due to some unreasonably large extreme values.

[28] A limitation is that I do not observe which customers answered the survey because participation was fully anonymous. However, even if there is systematic selection in the survey, the results provide an important insight: a survey with stated preferences yields estimates 11 times larger than estimates from a field experiment with the entire customer base that makes actual consumption choices. Whether this is driven by hypothetical bias or selection, we can conclude that the survey yields inflated estimates for the sample of interest.

Table 4
Hypothetical WTP

	(1) Total WTP (in EUR)
Quantity increase between-subject:	
+4.8kg	6.861
	(9.707)
Quantity increase within-subject:	
+2.4kg	31.243***
	(2.559)
+4.8kg	32.149***
	(3.463)
Between-subject variation in Education and Fairness Treatments:	
2.4kg + Education	7.857
	(9.891)
4.8kg + Education	18.314
	(12.724)
2.4kg + Firm Contribution	–15.917*
	(9.056)
4.8kg + Firm Contribution	–1.079
	(9.187)
2.4kg + Education & Firm Contribution	7.672
	(9.738)
4.8kg + Education & Firm Contribution	–1.216
	(9.290)
Constant (baseline offset: 2.4kg)	57.001***
	(6.751)
WTP in EUR/tCO_2	237.50***
	(28.13)
R^2	0.0478
N	1,617

Note: This table reports treatment effects on hypothetical WTP as absolute WTP in cents. Subjects stated their WTP in an open-end question. The second-to-last row shows implied WTP in EUR/ton of CO_2. The treatment "Education" indicates whether subjects received an education treatment about carbon offsetting before the WTP elicitation. "Firm contribution" indicates whether subjects were informed that the firm contributes to the match. Robust standard errors are in parentheses. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

marginally significantly negative, though this is not a robust finding as other coefficients are insignificant. In a follow-up question, subjects were asked what share of the carbon compensation costs of the delivery should be paid by the firm. Possible answers were between 0% and 100%. Figure 6(a) illustrates that the modal consumer thinks the company should pay half the compensation costs, indicating that consumers do value the firm's contribution positively.

The education treatment generally has positive coefficients, though none of them are statistically significant. This suggests a limited role of pedagogic information provision for WTP in line with prior studies (Imai et al., 2022; Pace et al., 2025).[29]

29 To complement this result, Internet Appendix D shows subjects' answers to the belief questions and suggests that, without the education treatment, subjects overestimate the carbon emissions of the average delivery, the equivalent kilometers that one needs to drive with a conventional car, and the number of trees necessary to compensate for 2,000 deliveries. The education treatment reduces the average overestimation for the last two questions. Consequently, subjects realize that it takes less to compensate for a delivery than they thought, which may explain the positive coefficients on WTP.

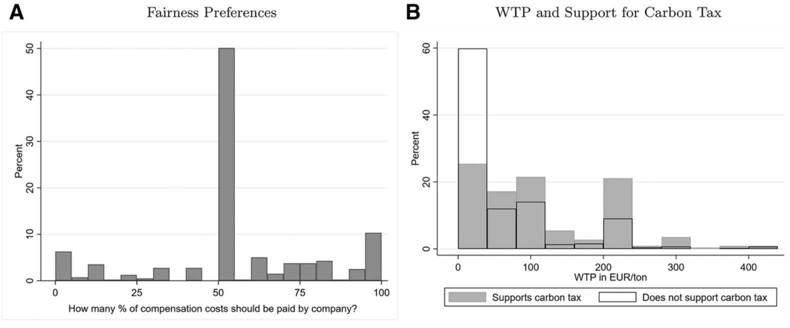

Figure 6
Email Survey
Panel A illustrates the distribution of subjects' answers to the question of what share of the carbon compensation costs should be paid by the firm. Panel B shows the distribution of WTP in EUR/ton of CO2 among the supporters of the tax (in gray) and the opponents (transparent).

There is no statistically significant effect of raising the compensation amount by 2.4 kg of CO_2 *between subject*. This again implies that consumers are fully quantity-inelastic even for hypothetical choices, in line with the seminal result by Kahneman and Knetsch (1992). However, WTP increases by 65% (+32 cents) when the compensation amount is raised *within-subject*. This is true for both the increase to 4.8 kg and to 9.6 kg (both $p < .01$). Thus, consumers again become quantity-elastic when they realize that the compensation amount is larger. Another interesting observation is that even in the within-subject design, consumers are scope-insensitive *in differences*: the effect of the quantity increase seems to be the same for 4.8 kg as for 9.6 kg. While this could point to extreme concavity in the WTP function, it is likely another symptom of the same cognitive manifestation related to contrast effects. The stated values support the findings from the field experiment that consumers do not seem to be able to compare magnitudes unless they are presented right after each other. The robustness of this result calls for further research to develop methods that allow us to estimate environmental preferences while accounting for inattention to scope.

Finally, I investigate how preferences for voluntary climate protection relate to political support for a carbon tax. At the end of the survey, subjects were asked whether they would support a carbon tax. Thirty-three percent of subjects oppose a carbon tax, while 67% endorse it. Subjects' political preference for carbon taxation is a strong predictor of hypothetical WTP. Figure 6(b) plots the empirical distribution of WTP in EUR/ton of CO_2 for supporters and opponents of the tax. I exclude values above the 90th percentile to adjust for outliers and increase the readability of the graph. Around 55% of subjects who oppose a carbon tax have a WTP below 20 EUR/tCO_2, while 32% have a WTP of zero. By contrast, only 20% of carbon tax supporters have a WTP below 20 EUR/tCO_2 and 6% a WTP of zero. The modal opponent of a carbon tax has a stated WTP of zero, while the modal supporter has a stated WTP of around 208 EUR/tCO_2. Overall, the probability distribution is shifted to the right for supporters relative to opponents of the tax. This suggests that hypothetical WTP—while overstating true WTP—still has strong predictive power regarding stated political preferences for environmental policies.

4. Conclusion

What does the market for voluntary climate protection imply about people's environmental preferences? This paper investigates this question with a natural field experiment to estimate how demand for carbon offsets responds to exogenous variations in subsidies and matches by the firm.

I find that consumers are elastic to price variations but fully inelastic to between-subject variations in impact. This result indicates that consumers buy the offset but do not value the carbon it mitigates. Using additional within-subject variation, I find that this scope-insensitivity may not purely reflect indifference but also an incomplete mental representation of impact as an attribute. Furthermore, a simple intervention that advertises the firm's participation in the offset makes subjects sensitive to impact and implies a WTP of 16 EUR/tCO_2. This suggests that firms can play a meaningful role in motivating consumers to reduce their carbon footprint.

The paper further shows how firms can mitigate carbon cost effectively. I show that demand is too price-inelastic for subsidies to be cost effective. Inelastic demand implies that subsidies induce transfers to all inframarginal consumers but do not induce a sufficiently large increase in mitigation. By contrast, matches induce sufficient increases in mitigation because even inframarginal consumers mitigate more. The takeaway is that, in the present context, cost-effective carbon offsetting relies on matches rather than subsidies.

Finally, stated preferences from the complementary survey overestimate revealed preferences observed in the field. The difference between stated and revealed preferences suggests two key directions for future research. First, the experiment involves relatively small stakes, both in terms of the carbon amount to be mitigated and the total compensation price. While the survey also elicits willingness to pay (WTP) for small stakes to ensure comparability, understanding consumer valuation of carbon mitigation in higher-stakes decisions remains crucial. It is well-known that extrapolation from small- to large-stakes settings can yield misleading preference parameters (Rabin, 2000). Second, additional tests of scope insensitivity in both the field and the survey point to contrast effects and memory-based anchoring as potential, yet unexplored, mechanisms. Understanding the role of cognition for environmental preferences presents a promising avenue for research.

Code Availability: The replication code is available in the Harvard Dataverse at https://doi.org/10.7910/DVN/OYA8IN.

References

Achtnicht, M. 2012. German car buyers' willingness to pay to reduce co2 emissions. *Climatic change* 113:679–97.

Allcott, H., and T. Rogers. 2014. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review* 104:3003–37.

Andre, P., T. Boneva, F. Chopra, and A. Falk. 2022. Misperceived social norms and willingness to act against climate change. *Econtribute Discuss. Pap*.

Andreoni, J. 1990. Impure Altruism and Donations to Public-Goods: A Theory of Warm-Glow Giving. *Economic Journal* 100:464–77.

Bauer, R., T. Ruof, and P. Smeets. 2021. Get real! Individuals prefer more sustainable investments. *Review of Financial Studies* 34:3976–4043.

Berg, F., J. F. Koelbel, A. Pavlova, and R. Rigobon. 2023. Esg confusion and stock returns: Tackling the problem of noise. Working Paper, National Bureau of Economic Research.

Berkouwer, S. B., and J. T. Dean. 2022. Credit, attention, and externalities in the adoption of energy efficient technologies by low-income households. *American Economic Review* 112:3291–330.

Blaufelder, C., C. Levy, P. Mannion, and D. Pinner. 2021. A blueprint for scaling voluntary carbon markets to meet the climate challenge. *Working Paper*, McKinsey & Company, New York, NY. Accessed: 2025-07-03.

Bordalo, P., N. Gennaioli, and A. Shleifer. 2020. Memory, attention, and choice. *The Quarterly Journal of Economics* 135:1399–442.

Broccardo, E., O. Hart, and L. Zingales. 2022. Exit versus voice. *Journal of Political Economy* 130:3101–45.

Brouwer, R., L. Brander, and P. Van Beukering. 2008. A convenient truth: air travel passengers' willingness to pay to offset their co2 emissions. *Climatic Change* 90:299–313.

Bushong, B., M. Rabin, and J. Schwartzstein. 2021. A model of relative thinking. *The Review of Economic Studies* 88:162–91.

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

Capital One. 2024. Online grocery shopping statistics. URL: https://capitaloneshopping.com/research/online-grocery-shopping-statistics/.

Carlsson, F., M. Kataria, A. Krupnick, E. Lampi, Å. Löfgren, P. Qin, S. Chung, and T. Sterner. 2012. Paying for mitigation: A multiple country study. *Land Economics* 88:326–40.

Chetty, R. 2009. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics* 1:451–88.

Diederich, J., and T. Goeschl. 2011. Willingness to pay for individual greenhouse gas emissions reductions: evidence from a large field experiment. *Discussion Paper Series* 517, University of Heidelberg, Department of Economics, Heidelberg.

Diederich, J. H. 2013. Determinants of the voluntary provision of public goods: Experimental investigations. PhD, Ruprecht Karl University of Heidelberg.

Eckel, C. C., and P. J. Grossman. 2003. Rebate versus matching: does how we subsidize charitable contributions matter? *Journal of Public Economics* 87:681–701.

Fehr, E., and K. M. Schmidt. 1999. A Theory Of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 114:817–68.

Feldman, N. E. 2010. Time is money: Choosing between charitable activities. *American Economic Journal: Economic Policy* 2:103–30.

Giglio, S., B. Kelly, and J. Stroebel. 2021. Climate finance. *Annual Review of Financial Economics* 13:15–36.

Giglio, S., T. Kuchler, J. Stroebel, and X. Zeng. 2023. Biodiversity risk. Working Paper, Yale.

Giglio, S., M. Maggiori, J. Stroebel, Z. Tan, S. Utkus, and X. Xu. 2025. Four facts about esg beliefs and investor portfolios. *Journal of financial economics* 164:103984–.

Gormsen, N. J., K. Huber, and S. Oh. 2023. Climate capitalists. *Available at SSRN*.

Green, D., and B. N. Roth. 2025. The allocation of socially responsible capital. *Journal of Finance* 80:755–81.

Hanemann, W. M. 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics* 66:332–41.

Hartzmark, S. M., and A. B. Sussman. 2019. Do investors value sustainability? A natural experiment examining ranking and fund flows. *Journal of Finance* 74:2789–837.

Heeb, F., J. F. Kölbel, F. Paetzold, and S. Zeisberger. 2023. Do investors care about impact? *Review of Financial Studies* 36:1737–87.

Hersch, J., and W. K. Viscusi. 2006. The generational divide in support for environmental policies: European evidence. *Climatic Change* 77:121–36.

Imai, T., D. D. Pace, P. Schwardmann, and J. J. van der Weele. 2022. Correcting Consumer Misperceptions About CO2 Emissions. CESifo Working Paper.

Interagency Working Group. 2021. Technical support document: Social cost of carbon, methane, and nitrous oxide interim estimates under executive order 13990. Working Paper, Interagency Working Group on Social Cost of Greenhouse Gases, United States Government.

Kahneman, D., and J. L. Knetsch. 1992. Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management* 22:57–70.

Karlan, D., and J. A. List. 2007. Does price matter in charitable giving? evidence from a large-scale natural field experiment. *American Economic Review* 97:1774–93.

Kesternich, M., A. Löschel, and D. Römer. 2016. The long-term impact of matching and rebate subsidies when public goods are impure: Field experimental evidence from the carbon offsetting market. *Journal of Public Economics* 137:70–8.

Kim, S., T. Li, and Y. Wu. 2024. Carbon offsets: Decarbonization or transition-washing? *Available at SSRN 4762799*.

List, J., M. Rodemeier, S. Roy, G. Sun, *et al* 2022. Judging nudging: Understanding of the welfare effects of nudges versus taxes. NBER Working Paper.

List, J. A. 2007. Field experiments: a bridge between lab and naturally occurring data. *The BE Journal of Economic Analysis & Policy* 6:0000102202153806371747–.

Löschel, A., M. Rodemeier, and M. Werthschulte. 2023. Can self-set goals encourage resource conservation? field experimental evidence from a smartphone app. *European Economic Review* 160:104612–.

Löschel, A., B. Sturm, and C. Vogt. 2013. The demand for climate protection—empirical evidence from germany. *Economics Letters* 118:415–8.

McFadden, D., *et al* 1973. Conditional logit analysis of qualitative choice behavior.

Nemet, G. F., and E. Johnson. 2010. Willingness to pay for climate policy: a review of estimates. Working Paper 2010-011, University of Wisconsin-Madison.

Oehmke, M., and M. M. Opp. 2024. A theory of socially responsible investment. *The Review of Economic Studies* 92:1193–225.

Pace, D. D., T. Imai, P. Schwardmann, and J. J. van der Weele. 2025. Uncertainty about carbon impact and the willingness to avoid co2 emissions. *Ecological Economics* 227:108401–.

Rabin, M. 2000. Risk aversion and expected-utility theory: A calibration theorem. *Econometrica* 68:1281–92.

Ramadorai, T., and F. Zeni. 2024. Climate regulation and emissions abatement: Theory and evidence from firms' disclosures. *Management Science* 70:8366–85.

Rennert, K., F. Errickson, B. C. Prest, L. Rennels, R. G. Newell, W. Pizer, C. Kingdon, J. Wingenroth, R. Cooke, B. Parthum, *et al* 2022. Comprehensive evidence implies a higher social cost of CO2. *Nature* 610:687–92.

Rodemeier, M. 2025. Buy baits and consumer sophistication: Field evidence from instant rebates. *American Economic Journal: Economic Policy* 17:30–59.

Rodemeier, M., and A. Löschel. 2025. Information nudges, subsidies, and crowding out of attention: Field evidence from energy efficiency investments. *Journal of the European Economic Association.*

Schulze-Tilling, A. 2025. The effectiveness of carbon labels. Working paper, University of Bonn.

Semken, C. 2024. The marginal impact of emission reductions: estimates, beliefs and behavior. Working paper.

Stroebel, J., and J. Wurgler. 2021. What do you think about climate finance?

Vesterlund, L. 2003. The informational value of sequential fundraising. *Journal of Public Economics* 87:627–57.

Viscusi, W. K., and R. J. Zeckhauser. 2006. The perception and valuation of the risks of climate change: a rational and behavioral blend. *Climatic Change* 77:151–77.