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

Matthias Rodemeier
Bocconi University

March 2024

Abstract

This paper leverages a large-scale field experiment with an online supermarket (N=255,000) where consumers are offered carbon offsets that compensate for emissions. Consumers are price-elastic but fully inelastic to the environmental impact of the offsets—consistent with "warm glow" utility. When the firm offers to share the offsetting costs, consumers become both more price- and impact-responsive and exhibit an implicit willingness to pay for carbon mitigation of 16 EUR/tCO2. This result highlights an important role of sustainable firm investments in encouraging consumers to lower their carbon footprint. If the firm offers carbon offsets but charges consumers the full price, cash flows fall by 2%. This negative effect can be fully compensated if the firm shares the costs with the consumer. This finding challenges the common view that corporate sustainable investment sacrifices cashflows.

JEL Codes: D61, D82, H21, Q50, Q58

Keywords: climate change, corporate sustainability, willingness to pay, carbon offsets, fairness preferences

---

Acknowledgments: I thank Nicola Gennaioli, Florian Nager, and participants at the NBER SI for insightful discussions about this paper. Silvia Amalia Meneghesso provided outstanding research assistance.
Affiliations: Bocconi University, IZA, IGIER, and BAFFI. matthias.rodemeier@unibocconi.it.
First version: December 2022. The paper was selected for the 2024 IZA Best Paper Award for Innovative Research on the Economics of Climate Change (IRECC).

1 Introduction

The market for voluntary carbon mitigation has doubled in size from 2017 to 2020 and is projected to reach USD 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 remarkable trend towards carbon offsetting raises an important question: Do firms sacrifice returns by investing into sustainable activities? What does demand for these offsets reveal about household's valuations of environmental protection?

Much of the existing literature has studied investor preferences for sustainable assets (e.g., Stroebel and Wurgler 2021, Giglio et al. 2023). Leading theories of sustainable investing assume that firms and investors need to sacrifice returns when reducing their carbon footprint (Oehmke and Opp 2023). However, there is little evidence as to how much sustainable corporate activities actually affect consumer behavior and firm cash flows. This paper bridges the current disconnect between sustainable firm investments and consumer markets.

I partner with one of the largest online supermarkets in Germany and implement an experiment in their online shop, observing over 250,000 consumers. In the experiment, consumers can offset the carbon emissions of grocery deliveries by buying carbon offsets. 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: 2.4kg of CO_2 for a price of 24 cents (i.e., mitigating 1kg of carbon costs 10 cents). In order 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 9.6kg for 24 cents (instead of just 2.4kg), 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"

[1] See, https://www.mckinsey.com/capabilities/sustainability/our-insights/a-blueprint-for-scaling-voluntary-carbon-markets-to-meet-the-climate-challenge.

preferences for carbon mitigation, but also by a preference to split the compensation costs with the firm. For instance, consumers might consider it fair if the firm contributes to the cost of the offset since it also benefits from the polluting transaction. Therefore, I cross-randomize whether the firm informs the consumer that it shared 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 fairness preferences from the "intrinsic" valuation of carbon mitigation.

Household Preferences for Carbon Offsetting. The experiment produces a number of important results. The first one is that 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 theories of warm glow (Andreoni 1990) and "scope-insensitivity" (Kahneman and Knetsch 1992).

When consumers are actively informed that the firm has subsidized or matched the offset, demand becomes quantity-elastic. Doubling the compensation amount of the offset increases its demand by 11%, and quadrupling the amount increases demand by 22%. In other words, a minimally invasive information treatment makes consumers sensitive to scope. INFORMATION also increases price elasticities. The effect of a price reduction on offsetting demand increases by up to 250% due to information provision. Thus, this finding highlights that firms can play an important role in encouraging consumers to lower their carbon footprint.

The difference between STANDARD and INFORMATION delivers largely different conclusions about consumers' valuation of carbon mitigation. Using the usual random utility model, I find that WTP is zero in STANDARD, but it is 16 EUR per ton of CO_2 ($p < 0.01$) in INFORMATION.

Effect of Offsetting on Firm Performance. What does offering carbon offsets imply for the firm? While offering carbon offsets does not affect the probability to purchase,

it does affect cashflows and profits. Specifically, if the firm offers carbon offsets to consumers but lets consumers bear the full costs, cashflow and profits both fall by approximately 2%. This negative effect can be mitigated if the firm participates in the offsetting costs, especially through matches. This finding challenges the common view that sustainable firm activities need to sacrifice cashflows.

Cost-Effectiveness of Sustainable Firm Practices: Subsidies vs. Matches. Comparing subsidies and matches, I analyze which intervention is the most cost-effective in reducing carbon emissions and find an arguably surprising result: *Quantity matches are always more cost-effective than subsidies even if matches have no effect on demand*. This is because subsidies reduce the price for all consumers, but the only incremental increase in mitigation comes from *marginal* consumers. By contrast, matches also cause *infra-marginal* consumers to mitigate more carbon. For subsidies to break even with matches, price elasticities would have to be substantially larger than they turn out to be empirically. The second important result is that matches have a "multiplier effect" when they are made salient through information: Every EUR spent by the firm on a quantity match produces a larger reduction in carbon than if that same EUR were directly invested into a carbon offset. These results may provide a motivation for governments to provide financial incentives to firms that offer carbon offsets.

Stated Versus Revealed Preferences. Much of the existing evidence on consumers and investor preferences comes from surveys in which respondents give hypothetical answers. Would we have gotten the same results from the field experiment had we just asked people for their stated preferences? 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/ton of CO_2. This is 1,388% larger than even my largest estimate of 16 EUR/tCO_2 in the information treatment. This result highlights the importance of developing tools to mitigate hypothetical bias in surveys (e.g., List and Gallet 2001).

Contributions to the Literature The paper connects to a growing literature in sustainable finance and economics that studies "green" firm practices and its implications on household behavior and asset values (e.g., Stroebel and Wurgler 2021, Giglio, Kelly, and Stroebel 2021, Bauer, Ruof, and Smeets 2021, Pedersen, Fitzgibbons, and Pomorski

2021, Hartzmark and Shue 2023, Bonnefon et al. 2022, Giglio et al. 2023, Gormsen, Huber, and Oh 2023, Heeb et al. 2023). The estimates from this paper are informative about the impact of corporate social responsibility on firm fundamentals. The common view is that sustainable firm investments come at the expense of lower cashflows and shareholder values. My results suggest that this is not the case: carbon offsets have no negative effect on firm outcomes as long as the firm shares the offsetting costs with the consumer. In addition, my findings suggest that consumers are willing to pay for carbon offsets but are fully indifferent between the impact of the offset on the environment. This finding is consistent with prior evidence in household finance that retail investors are insensitive to impact variations, as well, when deciding whether to invest in a sustainable firm (e.g., Bonnefon et al. 2022, Heeb et al. 2023).

My paper further improves on the existing empirical evidence by collecting the first data from a "natural field experiment" (Harrison and List 2004): 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. A related literature has used lab and survey experiments to measure people's environmental preferences. Perhaps most closely related, a contemporaneous survey by Pace et al. (2023) studies how subjects' uncertainty about carbon emissions affects the willingness to reduce their carbon footprint. They find that aggregate WTP is concave in emissions, which, by standard decision theory, implies that a reduction in uncertainty should affect WTP. Yet, they estimate a precise null effect of uncertainty on WTP. My findings can resolve this puzzle: The concavity of *aggregate* WTP may not come from decreasing marginal utility of mitigation but rather reflect warm glow utility of many *individual* subjects. Warm glow utility can produce an aggregate WTP curve that is concave in emissions and, at the same time, would not predict an effect of reducing uncertainty.[2]

Another strand of the literature studies the role of corporate social responsibility for public good provision. Specifically, prior studies have investigated the relative effectiveness of subsidies and matching mechanisms in increasing chartiable donations (e.g., Eckel and Grossman 2003, Kesternich, Löschel, and Römer 2016, Karlan and List 2007,

[2]Earlier important 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.

Feldman 2010). Since my study varies fairness perceptions, it is the first to show that positive match elasticities may be entirely driven by a preference to share the costs with the firm rather than by intrinsic preferences for the public good.[3] My study also builds on this literature by estimating impacts of matches and subsidies on important firm outcomes.

The rest of this paper is structured as follows. Section 2 presents the experimental design. Effects on households are discussed in Section 3 and on firm perfomance in Section 4. In Section 5, I present insights from a complementary survey. Section 6 concludes.

2 Experimental Design

The experiment takes place in the webshop of one of the largest delivery services for groceries and beverages in Germany. 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 one 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 2.4kg 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 $z \in \{100\%, 300\%\}$.

The experimental design intentionally features an important 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. As will become more clear throughout the paper, this symmetry can be used to isolate subjects' fairness preferences from their intrinsic utility of mitigating carbon.[4]

I also vary whether the firm advertises its own contribution to the carbon offset

[3] The implication is that— different from price elasticities in markets with private goods—donation elasticities are unlikely to offer reliable measures of consumer surplus from the public good.

[4] In addition, the symmetry is useful because it holds the cost of mitigation per kg 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%.

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.

2.1 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 either because they want to verify which goods they put into the shopping basket, or to finalize the 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 2.4kg of CO_2 for 24 Cents."* The text below informs subjects to which carbon-offsetting project the amount is donated.[5] In addition, subjects are informed that 2.4kg of CO_2 correspond to the average emissions of one delivery.[6] 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 one kg of CO_2 (i.e., 100 EUR/tCO_2). Thus, one average delivery that emits 2.4kg 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 2.4kg 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

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

[6] Average emissions were calculated from historical trip data.

to consumers that the offset project is of low quality and not effective at compensating carbon.[7] A low offset price might also signal that the environmental damage of a delivery is negligible since it costs little to compensate it. By contrast, INFORMATION should avoid this negative 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, i.e., 24 Cents. However, the quantity is doubled from 2.4kg to 4.8kg of CO_2. Therefore, this treatment provides exogenous variation in the *impact* of the offset. Consumers are still informed that an average delivery produces 2.4kg, 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 in order 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 4.8kg of CO_2 is 48 Cents. The reason they are paying half of the amount is that the company pays the remaining 24 Cents.

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. Such behavior could be a problem for the identification of WTP if we made the mistake to interpreted 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

[7] See Calel et al. (2021) for empirical evidence about adverse selection in the carbon offset market.

quantity and price elasticity. 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 context is consequently no threat to identification in our experiment. A more formal version of this argument is presented in Section 3.3, where I estimate WTP and explicitly allow for any arbitrary outside option.

2.2 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 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?"

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

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

Due to technical reasons, 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.

2.3 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 Section 4, 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.

3 Consumer Behavior

3.1 Effects on Offsetting Probability

I present differences in offsetting probabilities across treatments in Figure 4a. The grey bars indicate the offsetting probabilities for standard price 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 4.3.

At the baseline price of 24 cents for 2.4kg, 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. 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 7.2kg, 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 Kahneman and Knetsch (1992)'s 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 simply 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 $p < 0.01$. 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 2.4kg and 7.2kg raises the offsetting probabilities by 1.5 and 3 percentage points (both at $p < 0.01$). These responses are large 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 significant responsiveness to the impact of the offset.

An open question is whether the impact elasticity is due to consumers' recognition of the greater impact of the offset, or whether it stems from their appreciation of the firm's contribution to the offset. The colored arrows in the graph may already foreshadow an answer to this question. As measured by the blue arrows, the effect of information is *identical* between the 50% subsidy and the 100% match. This is precisely where the firm splits the total offset costs 50:50 with the consumer. If the firm raises its contribution to 25:75, demand increases incrementally by the size of the orange arrows. While this incremental distance is slightly larger for the match than for the subsidy, the magnitudes are again remarkably similar. Further, the difference between the orange arrows is not statistically significant. This may suggest that fairness preferences for the relative split in costs between the firm and the consumer are one of the main drivers underlying the effect of information. I investigate this mechanism further below by testing different models of fairness preferences and by studying effects on beliefs.

3.2 Underlying Mechanism of Information Effect

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. However, a general tendency seems to be that the perceived damage decreases as the costs for the consumer fall. Surprisingly 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 slightly increases.[8] However, the effect is very

[8]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.

small: The 100% match increases perceived effectiveness in INFORMATION by only 6% (0.2 points on the Likert scale). Further, we do not observe any effect of information for the much larger 300% match. Even though consumers should clearly understand that the offset is more effective when the firm saliently matches the quantity, the effect on beliefs is non-monotonic in impact. These results make it unlikely that the previously observed insensitivity to scope in STANDARD and the increase in elasticities due to INFORMATION are explained by consumers' inattention towards the compensation amount.

Overall, both Panel a) and b) point towards a limited role of changes in beliefs as the underlying forces of the information effect. Thus, the INFORMATION effect does not seem to be driven by changes in beliefs about the offset. Instead, consumers mostly seem to value that the firm is contributing to the offset costs.

In Appendix C I estimate different models of fairness preferences with the data. In all of the models, the effect fairness preferences to share the costs with the firm explain the entire effect of INFORMATION on offset demand. Together with the limited effect of INFORMATION on beliefs (discussed in Section 3.2), these results provide strong evidence that consumers only increase demand in INFORMATION because they appreciate the firm's participation in the offsetting, not because they respond to the larger impact.

3.3 Willingness to Pay

In this section, I quantify WTP for carbon mitigation in STANDARD and INFORMATION. Given the prior results, the difference between WTP in the two groups is likely to measure consumer welfare from the firm's contribution instead of intrinsic preferences for mitigation.

Estimation To estimate WTP, I rely on two estimation approaches. First, I use the standard random utility model with a logit error term (McFadden et al. 1973), which has also been extensively used in contingent valuation studies for public goods (Hanemann 1984). Second, I use a local approximation of WTP that does not rely on any distributional assumptions. The latter approach is commonly used in the sufficient statistics literature (Chetty 2009).

Consumer $i \in \{1, 2, ...I\}$ 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 at a total price of p_i. I make the usual assumption that p_i and γ_i enter linearly into utility.[9] Utility is given by:

$$u_i = \alpha + \beta\gamma_i + \eta p_i + \epsilon_i. \tag{1}$$

The parameter β is the marginal utility of mitigating one ton of carbon, and η is the marginal disutility of price. α is an intercept. Idiosyncratic preferences are given by ϵ_i. If γ is measured in tons of carbon and p in euros, willingness to pay to mitigate one ton of carbon is given by $WTP = -\frac{\partial u}{\partial \gamma}/\frac{\partial u}{\partial p} = -\frac{\beta}{\eta}$ EUR. The consumer decides to buy the offset iff $u_i \geq 0$, meaning aggregate demand for the offset is given by $D(p, \gamma) = 1 - G(-\beta\gamma - \eta p)$.

Logistic Distribution Under the usual assumption that ϵ_i follows a logistic distribution, the probability that consumer i chooses to buy the offset, denoted π_i, can be written in closed form as

$$\pi_i = \frac{1}{1 + \exp(-\beta\gamma_i - \eta p_i)}. \tag{2}$$

The model parameters β and η can be estimated by maximum likelihood.

Linear Approximation An alternative approach that does not require a distributional assumption about ϵ_i is to linearly *approximate* WTP by reduced-form elasticities. Specifically, note that the derivatives of aggregate demand with respect to price and carbon quantity are $\frac{\partial D}{\partial p} = \eta g(-\beta\gamma + \eta p)$ and $\frac{\partial D}{\partial \gamma} = \beta g(-\beta\gamma + \eta p)$, such that WTP is given by $WTP = -\frac{\beta}{\eta} = -\frac{\partial D}{\partial \gamma}\frac{\partial D}{\partial p}$.

Denote the demand responses to price and carbon quantity variations by $\Delta_p D$ and $\Delta_\gamma D$, respectively. The demand derivatives can be approximated by $\Delta_p D/\Delta p \approx \eta g(-\beta\gamma + \eta p)$ and $\Delta_\gamma D/\Delta\gamma \approx \beta g(-\beta\gamma + \eta p)$, where the approximation requires that Δp and $\Delta\gamma$ are small, or alternatively, that demand is locally linear, in which case $g(\epsilon)$ is locally constant. WTP can therefore be approximated by

$$WTP \approx -(\Delta_\gamma D/\Delta\gamma)/(\Delta_p D/\Delta p). \tag{3}$$

[9] In an unreported regression, I allow for nonlinearities in the utility function and cannot reject that they are statistically zero.

WTP Estimates Estimation results for both the logistic regression and OLS are shown in Table 4. I estimate the model for STANDARD and INFORMATION separately. Subjects with the baseline offset are included in both estimations. Regression coefficients in columns 1 and 2 come from a logistic regression, while coefficients in columns 3 and 4 are produced by OLS. Implied WTP is the ratio of the quantity and price coefficient, multiplied by (minus) 1.

As shown in column 1, using the variation in STANDARD to estimate utility parameters in the logistic regression, we find that only the disutility of price, η, is significant. Utility from the compensation amount, β, is indistinguishable from zero, suggesting consumers do not value the carbon-mitigating attribute of the offset. As a result, WTP for mitigating a ton of carbon is statistically zero. Column 2 shows that consumers receive larger price disutility and larger, statistically significant utility from the carbon-offsetting attribute of the offset. The utility parameters translate into a WTP estimate of 16.44 EUR/tCO_2. This estimate is highly statistically significant with $p < 0.01$.

The linear approximation produces almost identical results. WTP is zero in STANDARD and 15.99 EUR/tCO_2 ($p < 0.01$) in INFORMATION. Results are similar to the logistic regression because, empirically, aggregate demand turns out to be fairly linear (i.e., $g(\epsilon)$ is locally constant).

4 Firm Outcomes

4.1 Demand for Deliveries

To analyze whether the treatments affected demand for deliveries, I estimate a linear probability model, regressing whether a subject placed an order on the treatment vectors. Table 2 reports the regression results. In column 1, I include both between- and within-subject variation. In column 2, I only consider between-subject variation, i.e., a subject's first visit to the website during the experimental period.

The probability of ordering at the shop is 27% for the whole sample and 33% during the first visit. All treatment coefficients in both columns are economically small and tightly estimated null effects. This suggests that offering website visitors a carbon offsetting program does not affect demand for deliveries. Below, I also show that offsets do not affect product demand and revenues.

A reassuring implication from these results is that differences in offsetting behav-

ior conditional on placing an order have a causal interpretation because treatments do not cause systematic selection from the sample of website visitors to the subsample of customers.

Another upside is that the results do not suggest that offering carbon offsets has a negative effect on delivery demand by steering consumers' attention to the polluting attribute of a delivery. This was one of the main concerns of the firm prior to the experiment.

4.2 Cash Flows and Profits

Table 3 reports treatment effects on cashflows *net of the firm's offsetting costs* and profits conditional on buying at the shop. I.e., I first subtract any costs that the firm had for subsidizing or matching the offset. To ease comparability, I use as the control group the Baseline offset that costs 24 Cents and compensates 2.4kg. Row 1 suggest that offering carbon offsets without sharing the costs with consumers reduces cashflows by 1.9% ($p < 0.05$). Importantly, this negative effect seems to be mitigated when the firm offers subsidies or matches. While the coefficients are sometimes imprecisely estimated due to the high variance in the outcome variable, a general pattern is that they are positive and large. Matches almost fully offset the negative effect on cashflows, witht the exception of one point estimate. Subsidies seem to only partially undo the negative effect, increasing revenue by around 1% on average but not statistically significant.

A potential explanation is that some consumers dislike paying for the emissions of the delivery alone. While this is not enough to discourage them from buying at the shop, it does slightly reduce their spending.

We observe similar patterns for profits in column 2. Offering an offset to consumers without contributing to it reduces profits by 2%($p < 0.05$). If the firm matches the offset, this negative effect mostly vanishes. Subsidies have smaller positive point estimates and, as a result, are less precisely estimated.

Overall, these results indicate that sustainable firms investments into carbon offsetting has no negative impact on cashflows and profits. The only negative effect arises if the firm is not investing its own funds into the offset project and lets the consumer pay everything. These results cast doubt on the ubiquitous assumption that corporate sustainability comes at the cost of lower cash flows and shareholder value.

4.3 Cost-Effectiveness of Sustainable Firm Practices: Subsidies vs. 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 the firm but also for public policies that support corporate sustainability. If carbon offsets offered by firms to consumers are very cost-effective, then it may be efficient for governments to sponsor these programs.[10]

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 4a visualizes the results. 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 10kg/EUR). Perhaps surprisingly, quantity matches are always more cost-effective than subsidies, *even when matches have no impact on demand*. The reason for this stark result 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 1kg/EUR and 2kg/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 3.40kg/EUR and 3.60kg/EUR, respectively. The ratio is always below the market price of 10kg/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 10kg/EUR and are thereby more than 2.5 times more cost-effective than subsidies.

[10] This is particularly true if privately-offered carbon offsets are cost-effective but do not increase sales, such that the incentive to offer offsets may be missing in equilibrium.

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 11kg of carbon, i.e., 10% more than if the same EUR were invested directly into the baseline carbon offset.

This result suggests that policymakers may leverage corporate social responsibility to more efficiently invest into offset projects. A potential takeaway is that public funds should not be offered to firms to subsidize offsets they offer to consumers but rather for quantity matches.

5 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.[11] A translated version of the survey can be found in Appendix F.

In order 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 saw 2.4kg in the first question, see 4.8kg in the second. Analogously, subjects who first saw 4.8kg, next see 9.6kg. This creates both within- and between-subject variation in the compensation amount and allows me i) to estimate the distribution of stated WTP, and ii) to test if 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 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

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

before answering the WTP questions. Treatment subjects are informed i) that an average delivery emits 2.4kg of CO_2 (as in the field experiment), ii) that one would have to drive 11km in an average car to emit the same amount of carbon as the delivery, iii) 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 Appendix E, I describe the sample in more detail and discuss observable characteristics. While I cannot exclude that subjects select on unobservables into the survey, observable statistics are fairly representative of the firm's customer population.

Results Table 5 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.[12] 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, Nemet and Johnson 2010 Carlsson et al. 2012, Achtnicht 2012): numbers range from 40 to 350 USD/tCO_2 (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%.[13]

Stated preferences do not significantly change when consumers receive the information that the firm contributes to the offset. One coefficient is even marginally significantly negative, although this is not a robust finding as other coefficients are positive. 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 6a illustrates that the modal consumer thinks the company should pay

[12] 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.

[13] 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 into 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.

half the compensation costs, indicating that consumers do value the firm's contribution positively.

The education treatment generally has positive coefficients, although none of them is statistically significant. This suggests a limited role of information provision for WTP in line with prior studies (Imai et al. 2022, Pace et al. 2023).[14]

There is no statistically significant effect of raising the compensation amount by 2.4kg 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 4.8kg and to 9.6kg (both $p < 0.01$). Thus, consumers become highly 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 4.8kg as for 9.6kg. While this could point to extreme concavity in the WTP function, it is likely another symptom of the same behavioral phenomena. Specifically, the results suggest that consumers might not be able to compare magnitudes unless they are presented right after each other. This would be supported by theories of "relative thinking" (Bushong, Rabin, and Schwartzstein 2021) and salience (Bordalo, Gennaioli, and Shleifer 2013, Kőszegi and Szeidl 2013, Bordalo, Gennaioli, and Shleifer 2013) and may provide a new explanation for scope-insensitivity. The interpretation obviously matters for welfare: If consumers are insensitive to scope due to relative thinking, then their choices do not reflect true preferences. This contrasts with more traditional models of warm glow in which consumers do not respond to scope because they do not receive any utility from it.

A limitation of this interpretation is that within-subject differences can also be explained by experimenter demand effects in which subjects try to comply with the objective of the researcher. While recent evidence suggests that these effects are likely small (De Quidt, Haushofer, and Roth 2018), they cannot be ruled out within the scope of this survey.

[14]To complement this result, 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.

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. 33% of subjects oppose a carbon tax, while 67% endorse it. Subjects' political preference for carbon taxation is a strong predictor of hypothetical WTP. Figure 6b 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.

6 Conclusion

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

I find that consumers are elastic to price but fully inelastic to simple variations in impact. This result indicates that consumers buy the offset but do not value the carbon it mitigates. A simple but powerful 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 finding highlights that firms can play an important role in encouraging consumers to lower their carbon footprint.

Since I find no positive effect of carbon offsetting on on firm performance, there may be a lack of incentives for firms to invest sustainably. This provides a rationale for policymakers to support corporate sustainability, which in turn has a multiplier effect on consumer demand for lower carbon emissions.

Stated preferences from a complementary survey heavily diverge from revealed preferences in the experiment. Additional tests of scope-insensitivity point to models of relative thinking and salience as new and unexplored mechanisms. The development of

techniques aimed at obtaining agents' environmental valuations in the presence of behavioral models is an important avenue for future research.

References

Achtnicht, Martin. 2012. "German car buyers' willingness to pay to reduce CO2 emissions." *Climatic change* 113 (3): 679–697.

Andreoni, James. 1990. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The economic journal* 100 (401): 464–477.

Bauer, Rob, Tobias Ruof, and Paul Smeets. 2021. "Get real! Individuals prefer more sustainable investments." *The Review of Financial Studies* 34 (8): 3976–4043.

Bonnefon, Jean-François, Augustin Landier, Parinitha R Sastry, and David Thesmar. 2022. "The moral preferences of investors: Experimental evidence." Technical Report, National Bureau of Economic Research.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. "Salience and Consumer Choice." *Journal of Political Economy* 121 (5): 803–843.

Brouwer, Roy, Luke Brander, and Pieter Van Beukering. 2008. "A convenient truth: air travel passengers' willingness to pay to offset their CO2 emissions." *Climatic change* 90 (3): 299–313.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein. 2021. "A model of relative thinking." *The Review of Economic Studies* 88 (1): 162–191.

Calel, Raphael, Jonathan Colmer, Antoine Dechezleprêtre, and Matthieu Glachant. 2021. "Do carbon offsets offset carbon?" Working Paper.

Carlsson, Fredrik, Mitesh Kataria, Alan Krupnick, Elina Lampi, Åsa Löfgren, Ping Qin, Susie Chung, and Thomas Sterner. 2012. "Paying for mitigation: A multiple country study." *Land Economics* 88 (2): 326–340.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1 (1): 451–488.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and bounding experimenter demand." *American Economic Review* 108 (11): 3266–3302.

Diederich, Johannes, and Timo 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, Johannes Heiko. 2013. "Determinants of the Voluntary Provision of Public Goods: Experimental Investigations." Ph.D. diss., Ruprecht Karl University of Heidelberg.

Eckel, Catherine C, and Philip J Grossman. 2003. "Rebate versus matching: does how we subsidize charitable contributions matter?" *Journal of Public Economics* 87 (3-4): 681–701.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde. 2022. "The preference survey module: A validated instrument for measuring risk, time, and social preferences." *Management Science*.

Feldman, Naomi E. 2010. "Time is money: Choosing between charitable activities." *American Economic Journal: Economic Policy* 2 (1): 103–130.

Giglio, Stefano, Bryan Kelly, and Johannes Stroebel. 2021. "Climate finance." *Annual Review of Financial Economics* 13:15–36.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, Zhenhao Tan, Stephen Utkus, and Xiao Xu. 2023. "Four facts about ESG beliefs and investor portfolios." Technical Report, National Bureau of Economic Research.

Gormsen, Niels Joachim, Kilian Huber, and Sangmin Oh. 2023. "Climate capitalists." *Available at SSRN*.

Hanemann, W Michael. 1984. "Welfare evaluations in contingent valuation experiments with discrete responses." *American journal of agricultural economics* 66 (3): 332–341.

Harrison, Glenn W., and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42 (4): 1009–1055.

Hartzmark, Samuel M, and Kelly Shue. 2023. "Counterproductive sustainable investing: The impact elasticity of brown and green firms." Technical Report, Working Paper, Boston College.

Heeb, Florian, Julian F Kölbel, Falko Paetzold, and Stefan Zeisberger. 2023. "Do investors care about impact?" *The Review of Financial Studies* 36 (5): 1737–1787.

Hersch, Joni, and W Kip Viscusi. 2006. "The generational divide in support for environmental policies: European evidence." *Climatic Change* 77 (1): 121–136.

Imai, Taisuke, Davide D Pace, Peter Schwardmann, and Joël J van der Weele. 2022. "Correcting Consumer Misperceptions About CO 2 Emissions."

Kahneman, Daniel, and Jack L Knetsch. 1992. "Valuing public goods: the purchase of moral satisfaction." *Journal of environmental economics and management* 22 (1): 57–70.

Karlan, Dean, and John A List. 2007. "Does price matter in charitable giving? Evidence from a large-scale natural field experiment." *American Economic Review* 97 (5): 1774–1793.

Kesternich, Martin, Andreas Löschel, and Daniel 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* 137:70–78.

Kőszegi, Botond, and Adam Szeidl. 2013. "A model of focusing in economic choice." *The Quarterly journal of economics* 128 (1): 53–104.

List, John A, and Craig A Gallet. 2001. "What experimental protocol influence disparities between actual and hypothetical stated values?" *Environmental and resource economics* 20 (3): 241–254.

Löschel, Andreas, Bodo Sturm, and Carsten Vogt. 2013. "The demand for climate protection—Empirical evidence from Germany." *Economics Letters* 118 (3): 415–418.

McFadden, Daniel, et al. 1973. "Conditional logit analysis of qualitative choice behavior."

Nemet, Gregory F, and Evan Johnson. 2010. "Willingness to pay for climate policy: a review of estimates." no. 2010-011.

Oehmke, Martin, and Marcus M Opp. 2023. "A theory of socially responsible investment." *Swedish House of Finance Research Paper*, no. 20-2.

Pace, Davide, Taisuke Imai, Peter Schwardmann, and Joel van der Weele. 2023. "Uncertainty about Carbon Impact and the Willingness to Avoid CO2 Emissions." Working paper.

Pedersen, Lasse Heje, Shaun Fitzgibbons, and Lukasz Pomorski. 2021. "Responsible investing: The ESG-efficient frontier." *Journal of financial economics* 142 (2): 572–597.

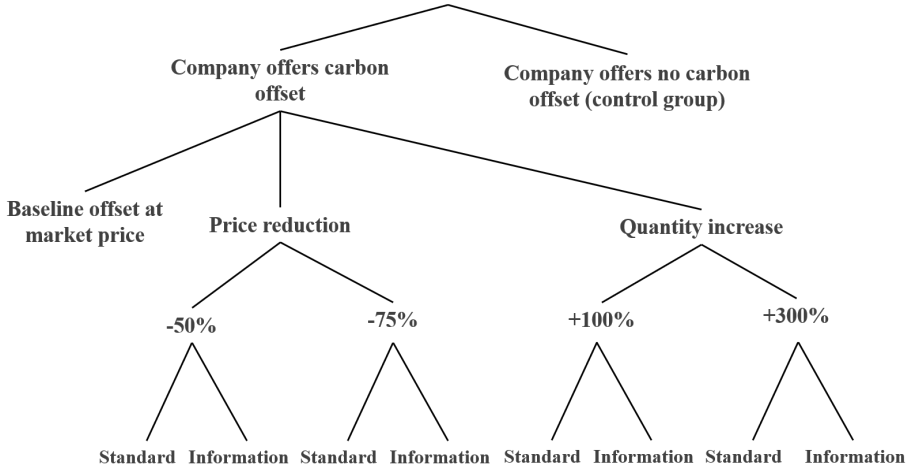
Stroebel, Johannes, and Jeffrey Wurgler. 2021. What do you think about climate finance?

Vesterlund, Lise. 2003. "The informational value of sequential fundraising." *Journal of public Economics* 87 (3-4): 627–657.

Viscusi, W Kip, and Richard J Zeckhauser. 2006. "The perception and valuation of the risks of climate change: a rational and behavioral blend." *Climatic change* 77 (1): 151–177.

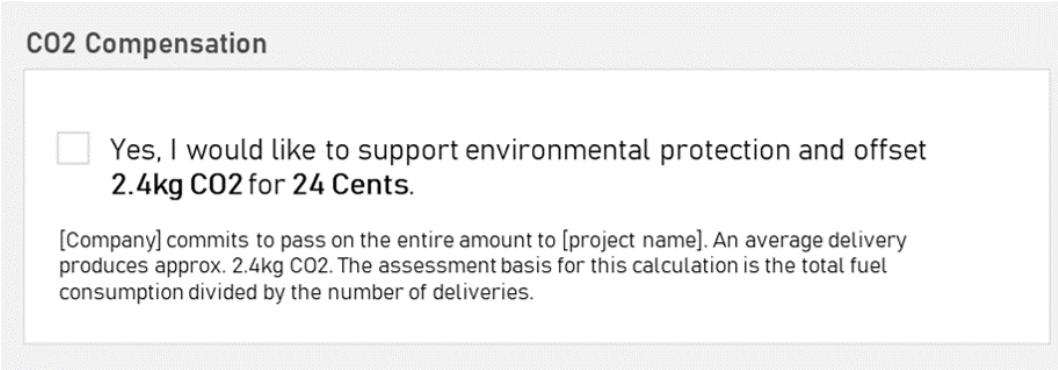
Figures

Figure 1: Experimental Design



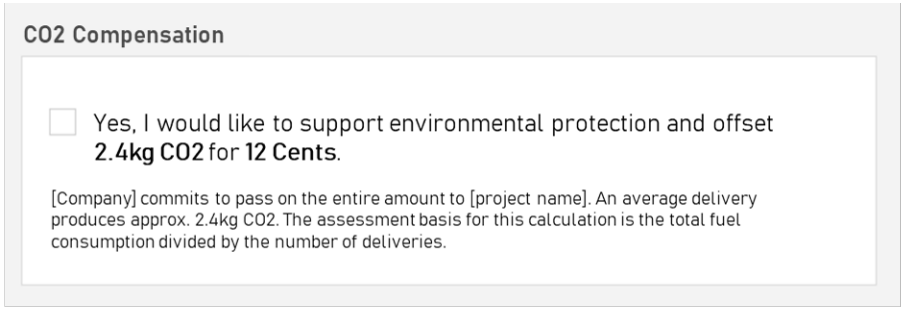
Note: This figure illustrates the experimental design. Subjects are randomized into one of ten groups with equal probability upon visiting the website.

Figure 2: Carbon Offset

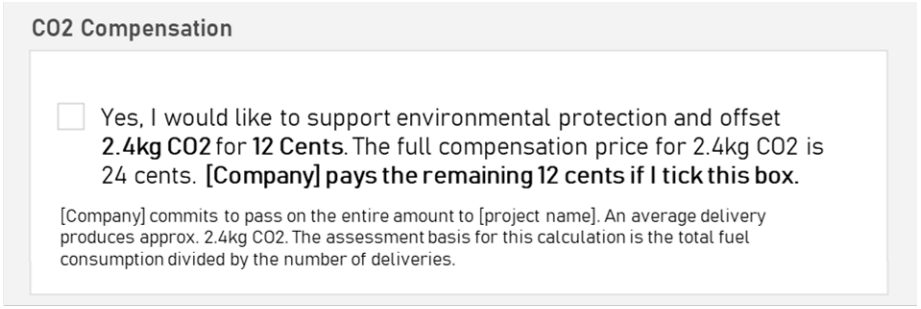


Note: This figure shows the baseline offset.

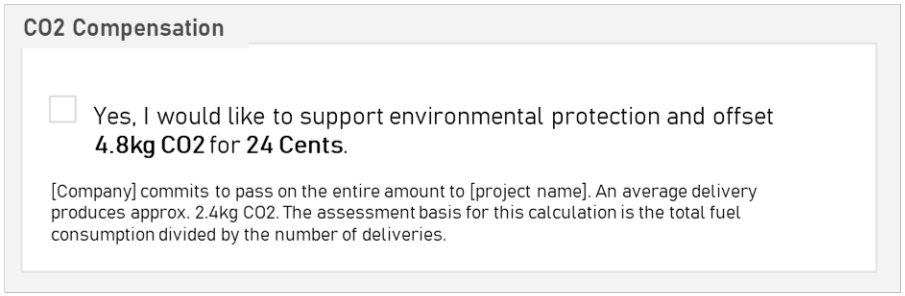
Figure 3: Examples of Treatment Variation



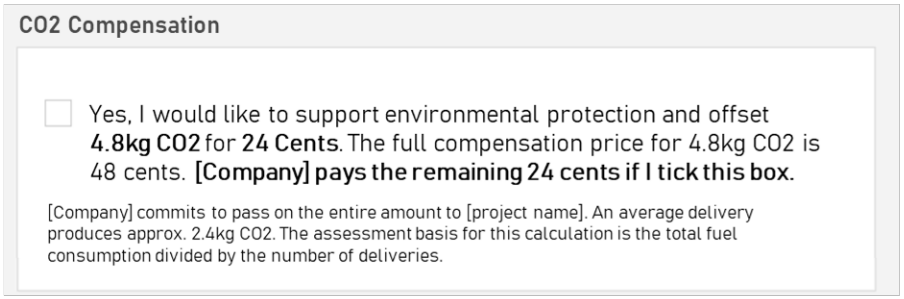
a) Price reduction by 50%



b) Price reduction by 50%, with salient information



c) Quantity increase by 100%

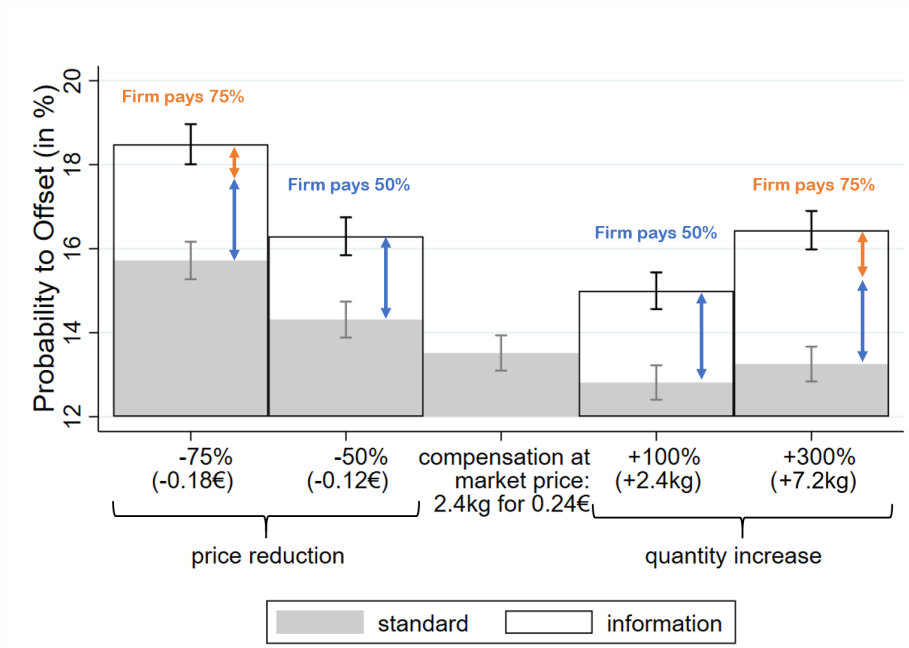


d) Quantity increase by 100%, with salient information

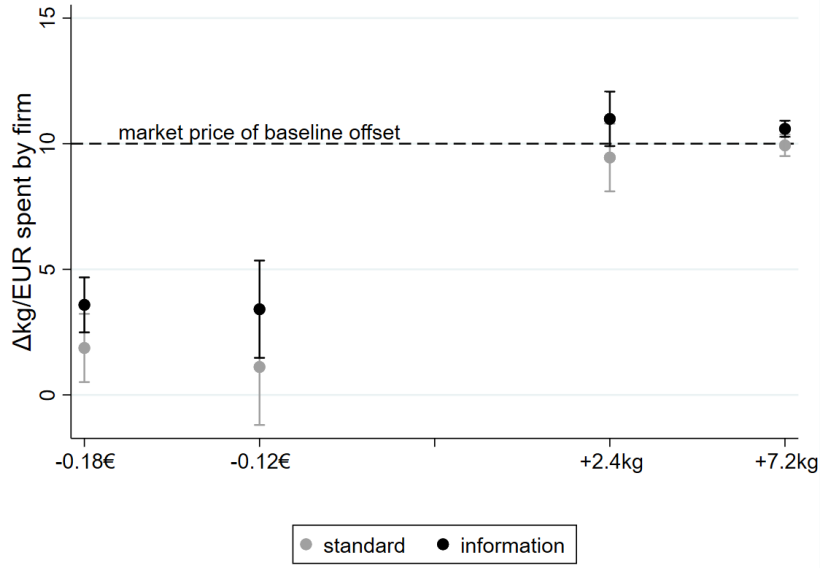
Note: This figure shows examples of price and quantity variations. Panel a) and c) illustrate the variations in the standard treatments, whereas panel b) and d) illustrate the variations in the information treatments.

Figure 4: Main Results

(a) Offsetting Probabilities



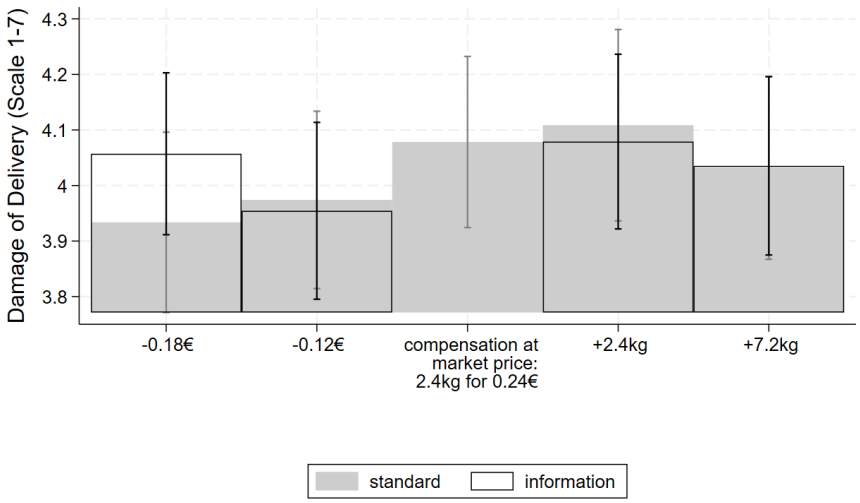
(b) Cost Effectiveness



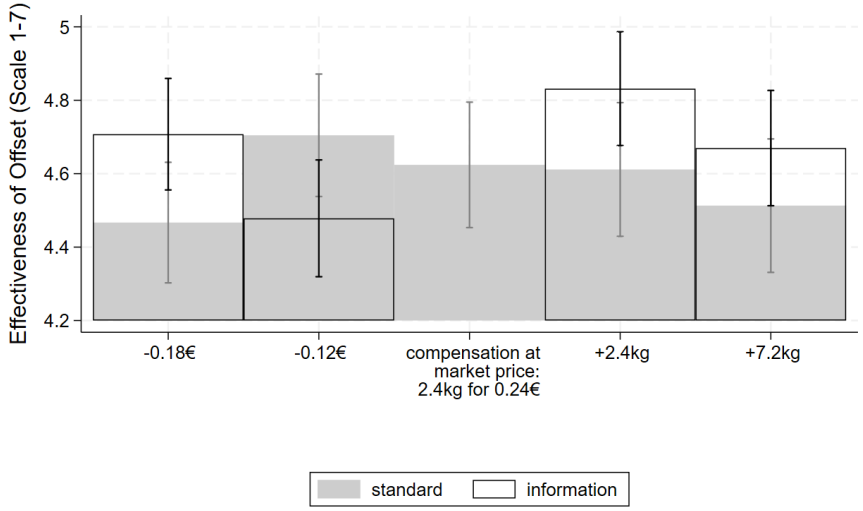
Notes: Panel a) represents the offsetting probabilities across treatments. Gray bars represent standard treatment groups, transparent bars represent information treatment groups. Panel b) plots the increase in compensated kilograms per EUR spent by the firm, relative to the baseline offset. The dotted line indicates the market price of the baseline offset (10kg/EUR).

Figure 5: Post-Experimental Survey

(a) Perceived Environmental Damage of a Delivery w/o Offsetting



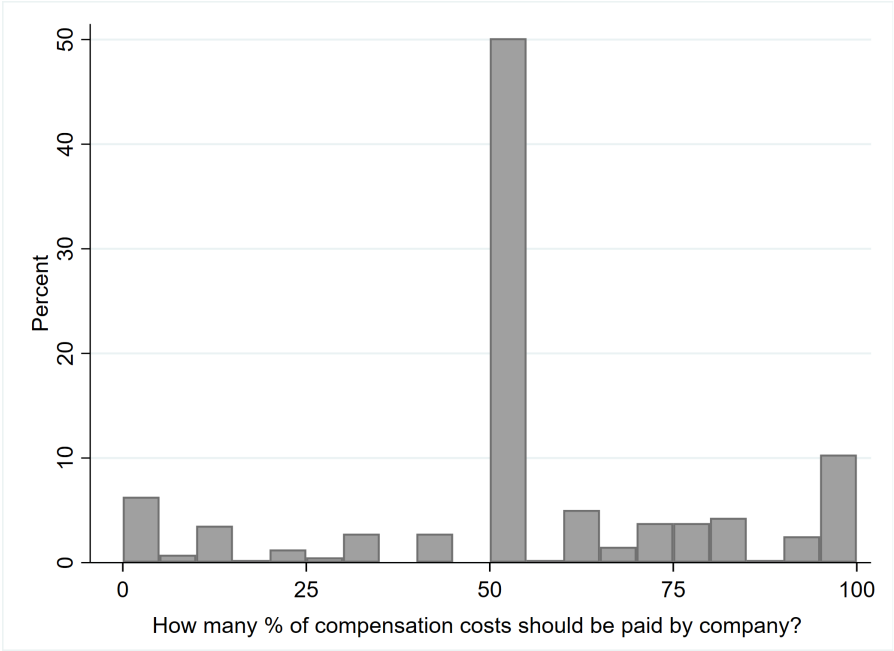
(b) Perceived Effectiveness of Offset



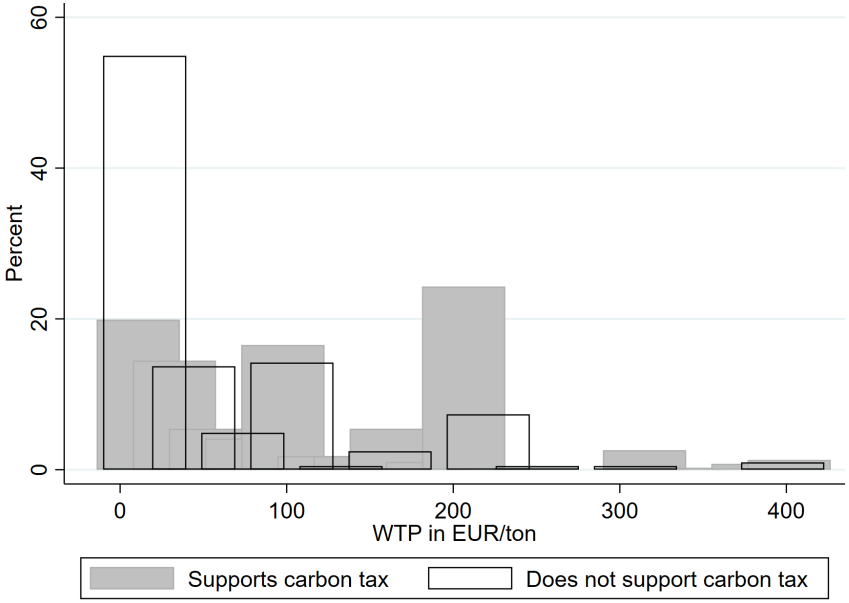
Notes: Panel a) illustrates subjects' beliefs about the size of the environmental damage of one delivery that is not compensated by an offset. Panel b) shows beliefs about the effectiveness of the offset in mitigating environmental damages.

Figure 6: Email Survey

(a) Fairness Preferences



(b) WTP and Support for Carbon Tax



Notes: 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 CO_2 among the supporters of the tax (in gray) and the opponents (transparent).

Tables

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.

Table 2: Demand for Deliveries

	(1) Order Probability	(2) Order Probability
Baseline: 24 Cents, 2.4kg	0.001	0.001
	(0.003)	(0.004)
-0.12€	0.003	0.004
	(0.003)	(0.004)
× information	0.003	0.004
	(0.003)	(0.004)
-0.18€	0.000	-0.002
	(0.003)	(0.004)
× information	0.002	0.003
	(0.003)	(0.004)
+2.4kg	0.003	0.004
	(0.003)	(0.004)
× information	0.001	0.001
	(0.003)	(0.004)
+7.2kg	0.003	0.005
	(0.003)	(0.004)
× information	0.000	0.002
	(0.003)	(0.004)
Constant: No offset offered	0.265***	0.329***
	(0.002)	(0.003)
N	406,980	255,374

Note: This table reports treatment effects on the probability to place an order among website visitors. The first column includes all website visits, whereas the second column only includes the first visit of a subject during the experimental period. Standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

Table 3: Cashflow and Profits

	(1) Cashflow (in % to Control)	(2) Profits (in % to Control)
Control: No Offset	0.0188**	0.0200**
	(0.0089)	(0.0093)
-0.12€	0.0141	0.0130
	(0.0088)	(0.0091)
× information	0.0119	0.0124
	(0.0090)	(0.0094)
-0.18€	0.0108	0.0106
	(0.0088)	(0.0090)
× information	0.0050	0.0066
	(0.0085)	(0.0093)
+2.4kg	0.0160*	0.0199**
	(0.0086)	(0.0092)
× information	0.0172**	0.0188**
	(0.0087)	(0.0090)
+7.2kg	0.0258***	0.0299***
	(0.0093)	(0.0101)
× information	0.0061	0.0075
	(0.0087)	(0.0092)
N	107,061	107,061

Note: This table reports treatment effects on cash flows net of offsetting costs and on profits. The first column includes all website visits, whereas the second column only includes the first visit of a subject during the experimental period. Standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

Table 4: Willingness to Pay for Carbon Mitigation

	(1)	(2)	(3)	(4)
	Logit		OLS	
	Standard	Information	Standard	Information
Quantity	-0.29	31.42***	-0.02	4.07***
	(6.08)	(5.75)	(0.71)	(0.76)
Price	-1.05***	-1.91***	-0.13***	-0.25***
	(0.22)	(0.21)	(0.03)	(0.03)
Constant	-1.63***	-1.46***	0.16***	0.19***
	(0.04)	(0.03)	(0.04)	(0.03)
WTP in €/tCO_2	-0.27	16.44***	-0.18	15.99***
	(5.83)	(2.47)	(5.59)	(2.51)
WTP for offset itself (in €)	-1.56***	-0.76***	1.27***	0.74***
	(0.35)	(0.10)	(0.24)	(0.07)
N	42,440	42,186	42,440	42,186

Note: This table reports regression coefficients and implied WTP for carbon mitigation. Coefficients in column 1 and 2 are from a logistic regression, coefficients in column 3 and 4 from OLS. Implied WTP is the absolute ratio of the quantity and price coefficients. The second-to-last row shows WTP for the offset, computed by the absolute ratio between the regression *constant* and the price coefficient. Standard errors for WTP are obtained by the delta method and reported in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

Table 5: Hypothetical WTP

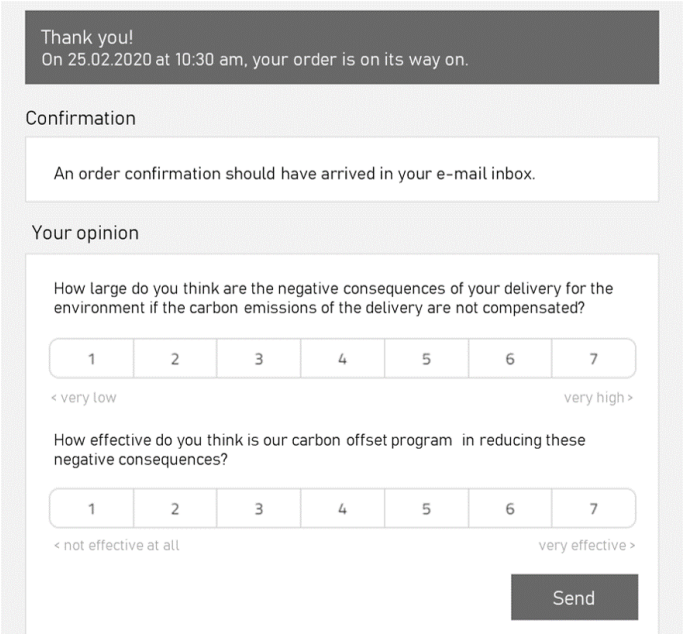
	(1) Total WTP (in Cents)
Quantity increase between-subject:	
+2.4kg	-7.857
	(9.891)
+4.8kg	-0.996
	(9.984)
Quantity increase within-subject:	
+2.4kg	31.243***
	(2.559)
+4.8kg	32.149***
	(3.463)
Between-subject variation in Education and Fairness Treatments:	
+4.8kg, Education	10.457
	(12.936)
+2.4kg, Education & Firm Contribution	-0.184
	(9.977)
+2.4kg, Firm Contribution	-23.774**
	(9.312)
+4.8kg, Firm Contribution	-8.936
	(9.479)
+4.8kg, Education & Firm Contribution	-9.073
	(9.579)
Constant (baseline offset: 2.4kg)	57.001***
	(6.751)
N	1,617
WTP in EUR/tCO_2	237.50***
	(28.13)

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 prior to the WTP elicitation. "Firm contribution" indicates whether subjects were informed that the firm contributes to the match. Robust standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

Online Appendix

A Additional Figures

Figure A1: Two-Question Survey Directly after Purchase



B Additional Tables

Table B1: Summary Statistics of Survey Sample

	Variable	N	Percent
Gender			
	male	514	67.72
	female	228	30.04
	diverse	4	0.53
	no answer	13	1.71
Age			
	18-19	6	0.79
	20-29	120	15.81
	30-39	195	25.69
	40-49	164	21.61
	50-59	140	18.45
	60-79	102	13.44
	> 70	24	3.16
	no answer	8	1.05
Occupation			
	employed	517	68.12
	unemployed	10	1.32
	apprentice	4	0.53
	homemaker	11	1.45
	retired	91	11.99
	student	42	5.53
	other	65	8.56
	no answer	19	2.50
	N	820	100.00

Note: This table reports frequencies of gender, age, and occupational status among participants in the email survey that are included in the analysis.

C Isolating the Role of Fairness Preferences

To isolate the role of fairness preferences from changes in beliefs, I leverage specific features of the experimental design. As I discuss below, the experiment allows us to test two general types of fairness models. In the first one, consumers care about the relative share that the firm pays. In the second one, the *absolute* size of the contribution (in EUR) rather than the relative share matters.

To see how we can test the first model, recall that there exists a symmetry in experimental design between matches and subsidies. Specifically, the experiment has been designed such that each subsidy level has a match for which the split between the consumer and the firm is identical (50:50 or 25:75). Therefore, there is a common factor between a subsidy and its respective match that we can control for. Put differently, the price the consumer pays is *not* perfectly collinear with the split between the firm and the consumer.[15] This allows us to isolate the effect of fairness utility on demand from other factors, such as changes in beliefs about the offset quality. The identifying assumption is that for a given split (e.g., 50:50) a subsidy does not change beliefs differently than the respective match. This assumption seems reasonable since consumers learn essentially the same information in both cases. Specifically, they learn the split between the firm and the consumer and that the price of one kilogram of carbon is 10 Cents.

Given the identifying assumption, we can isolate the underlying mechanisms of the information effect. The empirical specification for this model is[16]

$$y_i = \alpha + \eta p_i + \beta q_i + \underbrace{F_1 \times \mathbb{1}_i(50:50) \times I_i}_{\text{Fairness from 50:50 split}} + \underbrace{F_2 \times \mathbb{1}_i(25:75) \times I_i}_{\text{Fairness from 25:75 split}} + \epsilon_i. \tag{4}$$

Here, F_1 and F_2 measure the change in demand due to the 50:50 and 25:75 split, respectively. Note that these coefficients represent the *isolated* effect on demand that is solely driven by fairness preferences, not by changes in effectiveness beliefs. If the

[15] While the price *per ton of CO_2* is perfectly collinear with the split between firm and consumer, the price the consumer pays for the offset is not. The latter price is what matters for identification.

[16] Alternatively, one could also specify a utility function that incorporates fairness preferences and estimate the underlying parameters, e.g. by estimating a conditional logit with utility function $u_i = \alpha + \eta p_i + \beta q_i + F_1 \times \mathbb{1}_i(50:50) \times I_i + F_2 \times \mathbb{1}_i(25:75) \times I_i + \epsilon_i$. This section focuses on the reduced-form linear model because it is primarily interested in approximating effect sizes of fairness preferences rather than quantifying structural parameters. Section 3.3 estimates structural parameters and also shows how the linear model above can be used for a structural interpretation.

information treatment raised consumers' awareness of the increased effectiveness of a matched offset, then this change in belief should result in an increased quantity coefficient β, even after controlling for fairness preferences. If, on the other hand, information only increased quantity-elasticities due to fairness utility, then β should become zero when controlling for the split between consumers and firm.

The second fairness model follows a similar strategy. If consumers care about the absolute monetary amount spent by the firm, the experiment is rich enough to control for this, even if fairness utility is nonlinear in absolute contributions. With some abuse of notation, the following specification represents the second model:

$$y_i = \alpha + \eta p_i + \beta q_i + \underbrace{(F_1 \times C_i + F_2 \times C_i^2) \times I_i}_{\text{Fairness from absolute contribution}} + \epsilon_i.$$

Here, C is the absolute contribution by the firm measured in EUR. F_1 and F_2 now represent the linear and quadratic terms of the fairness utility function, respectively.

Regression results of both models are presented in Table B2. In the first model, the 50:50 and 25:75 split increase demand by, on average, 1.7 and 3.2 percentage points, respectively. Both effects are highly significant ($p < 0.01$). Consulting the price coefficient of -0.12, the fairness effects correspond to respective subsidies of 0.14 and 0.26 EUR. To put this into perspective, the latter represents 27% of the total offset costs of the 25:75 match (0.96 EUR).

Importantly, once we control for the split between firm and customer, the quantity coefficient becomes statistically indistinguishable from zero. This suggests that even though consumers become quantity-elastic under INFORMATION, this is entirely driven by fairness utility and not by a perceived higher impact.

We obtain similar results for the second fairness model. The first-order effect is an increase in demand of 94 basis points for every 10 Cents the firm spends. The slope falls by 31 basis points for every further 10 Cents increase, implying demand is convex in price. Both fairness coefficients are highly significant ($p < 0.01$). Controlling for fairness utility again yields a quantity coefficient indistinguishable from zero (and even negative in sign).

Together with the limited effect of INFORMATION on beliefs (discussed in Section 3.2), these results provide strong evidence in favor of scope-insensitivity in carbon offsetting. Consumers only increase demand in INFORMATION because they appreciate the firm's participation in the offsetting, not because they respond to the larger impact.

Table B2: Isolating Fairness Preferences

	(1) Relative Contribution	(2) Absolute Contribution
Quantity (β)	0.001	-0.001
	(0.007)	(0.007)
Price (η)	-0.122 ***	-0.182 ***
	(0.025)	(0.023)
Fairness Parameters:		
F_1	0.017 ***	0.094 ***
	(0.003)	(0.017)
F_2	0.032 ***	-0.031 ***
	(0.004)	(0.013)
Constant (α)	0.132 ***	0.130 ***
	(0.003)	(0.003)
N	76229	76229

Note: This table reports price and quantity coefficients while holding fixed fairness preferences. Column 1 assumes that consumers receive fairness preferences from the relative share that the firm contributes to the total offset costs. Column 2 assumes that fairness preferences are a (nonlinear) function of the absolute contribution. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

D Additional Results from Email Survey

Table E1 reports results from a regression of the education treatment on the belief questions. 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 amount of trees necessary to compensate for 2,000 deliveries. The last column shows how certain subjects were in their answers, where larger values indicate more certainty. The education treatment results in a substantial and statistically significant rise in subjects' certainty by close to 100% compared to the control group.

Table E1: Answers to Belief Questions in Email Survey

	(1) Delivery	(2) Car	(3) Trees	(4) Certainty
Education treatment about carbon offsetting	-0.409	-11.363***	-85.991***	2.323***
	(0.826)	(3.577)	(22.477)	(0.139)
Constant	5.463***	30.288***	144.844***	2.354***
	(0.543)	(2.471)	(15.893)	(0.095)
N	285	266	218	769

Note: This table reports answers to the belief questions in the email survey. Robust standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

Demographics and Risk Preferences

I regress hypothetical WTP on basic demographics elicited in the survey. I also include an established measure of risk preferences developed by Falk et al. (2022) that has been shown to predict actual risk preferences in incentivized questions. The question asked to subjects is: *"Please tell us, in general, how willing or unwilling you are to take risks."* Potential answers are on a Likert scale from 1 ("not at all willing to take risks") to 10 ("very willing to take risks").

Table E2 reports results. In terms of demographics, relatively few variables are a strong predictor of WTP. The constant represents WTP for an employed, male subject, between 40-49 years of age. On average, female subjects have a substantially higher WTP by around 31 Cents. In addition, retired subjects have a higher WTP of 20 Cents. This may be surprising as it is often claimed that older people have a lower incentive to protect the climate as they will be less exposed to future damages. Subjects that answered "other" to the employment question have an 18-cents lower WTP.

Interestingly, risk preferences are a strong predictor of hypothetical WTP. For every 1 point increase on the "willingness to take risk"-scale, WTP increases by 4 Cents, a relative increase of 11% relative to the constant. Note that the direction of the relationship between WTP and risk preferences partially depends on how much uncertainty subjects have about climate change versus how uncertain they are about the effectiveness of carbon offsets. On the one hand, it seems reasonable to assume that more risk-averse individuals have a stronger willingness to pay for carbon mitigation since there is large uncertainty

about future climate damages. On the other hand, the effectiveness of carbon offsets itself is uncertain, such that more risk-averse individuals may be less willing to donate to these projects. The present results may indicate that the second effect dominates.

To investigate this relationship visually, Figure E1 plots the correlation between risk preferences and average WTP. Specifically, each data point represents average WTP for a given level of risk preferences. The red line provides a linear prediction of the relationship.

While the relationship does not appear linear visually, it seems positive for most intervals. Thus, more risk-seeking consumers state a higher willingness to invest into carbon offsets. While correlations should always be interpreted cautiously, these patterns suggest that uncertainty may constitute an important barrier to voluntary climate protection.

Table E2: WTP and Demographics

	(1) Total WTP (in EUR)
Willingness to take risk	0.043***
	(0.010)
Age:	
18-19	0.016
	(0.287)
20-29	0.097
	(0.077)
30-39	-0.012
	(0.063)
50-59	0.009
	(0.069)
60-79	-0.125
	(0.090)
> 70	-0.064
	(0.156)
Gender:	
diverse	-0.169
	(0.295)
female	0.308***
	(0.049)
Employment Status:	
unemployed	0.001
	(0.186)
apprentice	-0.340
	(0.295)
housewife/husband	-0.277
	(0.181)
retired	0.202**
	(0.096)
other	-0.180**
	(0.079)
student	-0.180
	(0.111)
Constant (40-49 years, male, employed)	0.349***
	(0.080)
N	1,466

Note: This table reports correlations between WTP, risk preferences, and demographics. The constant represents WTP for an employed, male subject, between 40-49 years of age. Robust standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

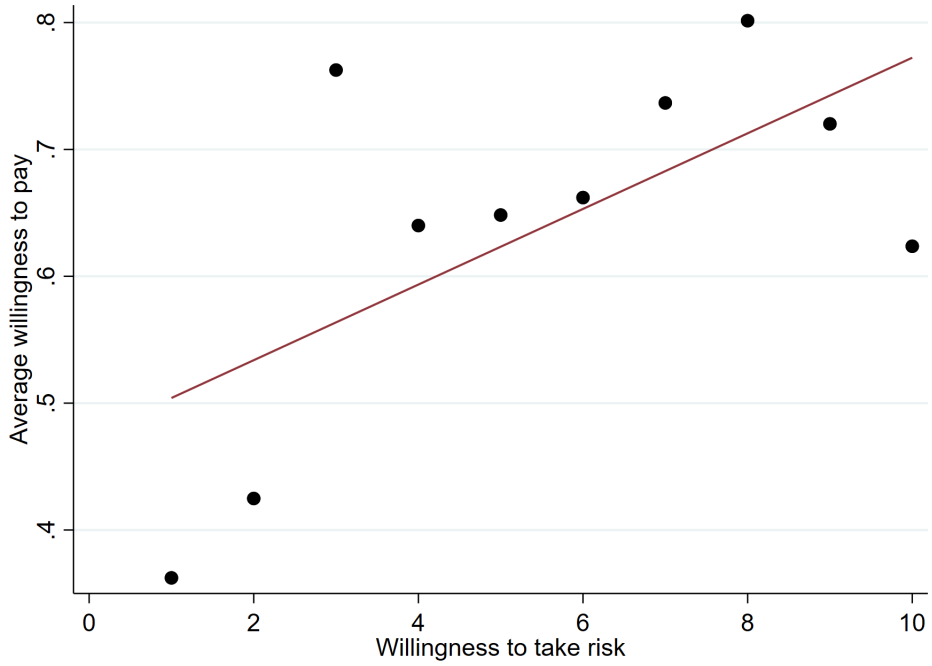


Figure E1: Correlation between Risk Preferences and WTP

E Sample Characteristics

Table B1 in the Appendix reports observable characteristics of the sample of respondents.[17] Around 68% are male, 30% female, 0.5% diverse, and 1.7% do not report a gender. 36% are between 20 and 40 years old, which is similar to the German average (31%).[18] Subjects between 40 and 60 years are slightly over-represented compared to the national average (40% vs. 33%), while subjects between 60-79 are underrepresented (13% vs. 27.5%).

Consistent with the age distribution, fewer subjects are retired than in the German population (12% vs. 32%). 5.5% are students compared to the national average of 3.5%. Around 1.3% are unemployed compared to 5% nationally.

As we would expect from online shop customers, the sample is overall slightly younger and more likely to have an occupation than the German population. According to the firm, the statistics on gender and age are very representative of their customer base. This

[17] I present statistics for the sample included in the analysis, excluding outliers, as described further below.

[18] For national statistics see https://www.destatis.de/EN/.

is important when we want to compare stated preferences from the survey with revealed preferences from the field experiment. While we cannot exclude that subjects select on unobservables into the survey, it is reassuring that observable statistics are fairly representative of the firm's customer population.

F Email Survey

The email survey, with all instructions and questions translated from German into English, can be found under this link: click here.