Uncertainty about Carbon Impact and the Willingness to Avoid CO2 Emissions

Davide Pace (LMU Munich)
Taisuke Imai (Osaka University)
Peter Schwardmann (Carnegie Mellon University)
Joel van der Weele (Tinbergen Institute)

Discussion Paper No. 470

December 02, 2023

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Uncertainty about Carbon Impact and the Willingness to Avoid CO_2 Emissions *

Davide D. Pace[1,5] Taisuke Imai[2,5]
Peter Schwardmann[3,5] Joël J. van der Weele[4]

[1]Ludwig-Maximilians-Universität München
[2]Institute of Social and Economic Research, Osaka University
[3]Carnegie Mellon University
[4]Amsterdam School of Economics, Tinbergen Institute
[5]CESifo

December 1, 2023

Abstract

With a large representative survey ($N = 1,128$), we document that consumers are very uncertain about the emissions associated with various actions, which may affect their willingness to reduce their carbon footprint. We experimentally test two channels for the behavioural impact of such uncertainty, namely risk aversion about the impact of mitigating actions and the formation of motivated beliefs about this impact. In two large online experiments ($N = 2,219$), participants make incentivized trade-offs between personal gain and (uncertain) carbon impact. We find no evidence that uncertainty affects individual climate change mitigation efforts through risk aversion or motivated belief channels. The results suggest that reducing consumer uncertainty through information campaigns is not a policy panacea and that communicating scientific uncertainty around climate impact need not backfire.

*This paper supersedes a previous version circulated with the title "Curbing Carbon: An Experiment on Uncertainty and Information about CO_2 Emissions" by a subset of the authors. Financial support from Deutsche Forschungsgemeinschaft through CRC TRR 190 (Imai, Pace), the Bavarian Academy of Sciences and Humanities (Schwardmann), the NWO in the context of VIDI grant 452-17-004 (van der Weele), the Amsterdam Center for Behavioral Change of the University of Amsterdam, and the initiative, "A Sustainable Future," at the University of Amsterdam (van der Weele), are gratefully acknowledged. The data collection was approved by the Ethics Committee Economics and Business (EBEC) of the University of Amsterdam, with request numbers EC 20191018121005, EC 20200810100845, and EC 20210713120725.

Climate action takes place against a backdrop of considerable uncertainty. Scientists face fundamental uncertainty about the exact contribution of different human activities to aggregate CO_2 emissions and about how emissions contribute to climate change and resulting extreme weather events[1,2]. On top of that, non-experts may be uncertain about the best estimates of how a given behaviour maps onto CO_2 emissions. Understanding how uncertainty about CO_2 emissions affects climate action is therefore central to debates around the costs and benefits of communicating scientific uncertainty about climate change[3–5] and around the effectiveness of carbon labels or other information campaigns[6,7].

This paper consists of three studies. Our first study investigates the degree of consumer uncertainty about emissions impact. We survey a large representative sample of the US population on their beliefs about the CO_2 emissions associated with several common consumer products. We elicit participants' entire belief distributions for each product using incentivized elicitation methods. The survey shows that people have a high degree of subjective uncertainty about the impact of all of these products, with only 6 per cent expressing high confidence in their guesses.

We then investigate two plausible mechanisms by which such uncertainty may causally affect emission mitigation behaviour. The first of these mechanisms is risk aversion. Uncertainty about the impact of emissions may reduce climate action if people do not place much value on avoiding high amounts of (potential) emissions. Theoretically, this requires that the marginal willingness to forgo personal benefits to avoid emissions declines with each additional unit of emissions. To test this mechanism, our second study uses large-scale online experiments in which participants' choices to buy an online convenience product affect real CO_2 emissions. We find that participants indeed have an increasing but concave willingness to mitigate (WTM) CO_2 emissions. However, contrary to the predictions of standard decision theory, we do not find an effect of uncertainty, as consumption of the polluting product is similar in treatments with and without uncertainty about emissions.

The second mechanism is the use of "moral wiggle room" that allows consumers to self-deceive into believing that emissions are small and consumption is harmless, as happens in other types of ethical decision-making[8–11]. Such self-serving belief formation may also interact with carbon pricing: as CO_2-intensive products become more expensive, the temptation to form self-serving beliefs decreases[9]. In our third study, participants may again buy a virtual polluting product. They see a vague signal about the emissions associated with the product and have to update their beliefs about the emission size. The signal we use provides an opportunity for self-serving belief distortions, but only in a condition where the incentives to hold self-serving beliefs are known before the signal is seen. We find no evidence that uncertainty is exploited by participants to develop self-serving beliefs that the emissions are low, nor do we find an interaction between prices and belief formation.

We contribute to the literature by documenting the extent of individuals' uncertainty about the carbon impact of their actions and by designing novel controlled experiments to study money-emission trade-offs. Our findings on subjective uncertainty complement previous evidence that people's best guesses about the size of CO_2 emissions are very heterogeneous and often too optimistic[6,7]. Moreover, we are the first to study the explicit and isolated role of uncertainty in affecting climate action. In this way, we complement earlier studies on information provision, which focus on correcting the bias in consumer beliefs and have found conflicting evidence on carbon information provision[6,7,12,13]. We show that the variance of beliefs has little impact on climate-friendly consumption. This demonstrates the limits of information campaigns to spur climate action, but it also suggests that communicating scientific uncertainty does not negatively affect consumption patterns.

In addition, our findings deliver new insights into the willingness to make sacrifices to reduce CO_2 emissions. Most studies in this literature have relied on hypothetical choices, which have been shown to lead to large overstatements in willingness to pay for offsets[14,15]. A few studies, like ours, use incentivized willingness to pay to retire offsets or permits but have looked at a single emission quantity[16–18]. We empirically show that the valuation of such offsets is concave in emission size, in line with concurrent evidence[15]. This suggests that future welfare measurements can take such nonlinearity into account when scientists estimate the welfare benefits of climate policies, study the demand for carbon offsets in particular industries like aviation[19] or the optimal pricing of offsets more generally[20]. Finally, we contribute to the literature on motivated cognition with our null results on self-serving environmental beliefs and on the interaction between beliefs and prices[8,9,21,22].

Results

Study 1: Subjective uncertainty about CO_2 emissions. We investigate subjective uncertainty with a representative sample of the US population ($N = 1,128$). Our survey elicited participants' beliefs regarding the carbon emissions associated with 12 common consumer products and activities. The elicitation process comprises two steps. First, we asked participants to indicate their best guess about the size of the emissions associated with each product. Then, we elicited the participants' entire subjective probability distribution regarding the emissions size of each product. To do so, we partitioned the interval of possible emissions into five bins centred around the participants' initial guesses. The three central bins have a width equal to 10% of the initial guess;

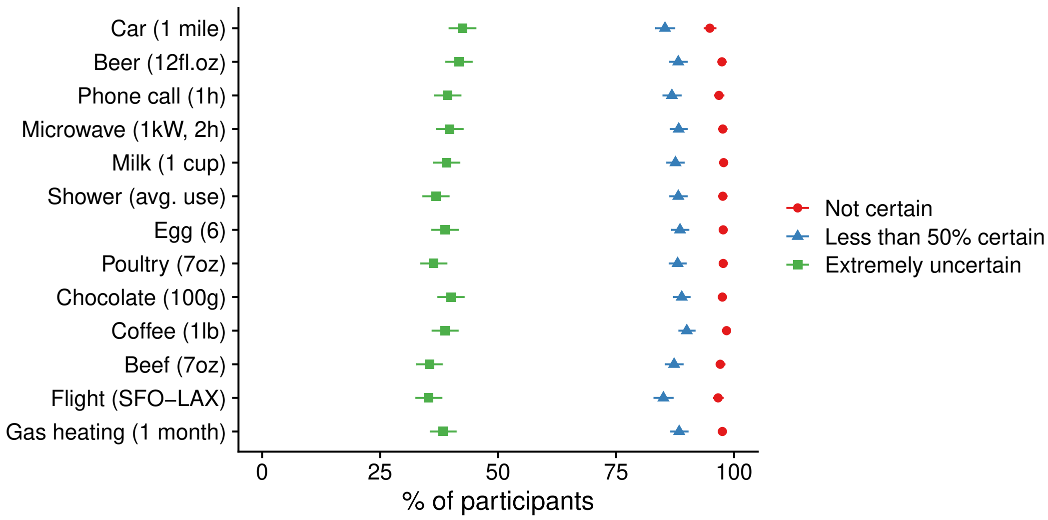


Figure 1: The percentage of participants who satisfy three subjective uncertainty criteria. A participant is *Not certain* if she puts at least one ball in a bin that does not contain her initial guess. She is *Less than 50% certain* if she puts less than 10 balls in the bin which contains her initial guess. She is *Extremely uncertain* if she puts four or fewer balls each in the bin that contains her initial guess and in the two adjacent bins. The three criteria are cumulative. The *Extremely uncertain* participants are also *Less than 50% certain*. The participants who are *Less than 50% certain* are also *Not certain*.

the two outer bins contain all the other possible emissions values. We asked the participants to allocate 20 balls across these bins to reflect the probability that the size of the emissions belongs to each of them. To incentivize truthful responses, we employed state-of-the-art techniques from experimental economics[23,24] (see Methods).

We study participants' uncertainty with three (nonexclusive) nonparametric criteria based on the ball allocation. A participant is *Not certain* if she puts at least one ball in a bin that does not contain her initial guess. She is *Less than 50% certain* if she puts less than 10 balls in the bin which contains her initial guess. She is *Extremely uncertain* if she puts four or fewer balls in each of the three central bins, indicating less than 20% confidence in being (close to) the correct answer.

Figure 1 depicts the levels of uncertainty for each product. In each case, at least 94% of the participants are *Not certain*, at least 85% are *Less than 50% sure*, and at least 35% are *Extremely uncertain*. The figure also displays limited variability across products, indicating that people's uncertainty is not domain-specific. These results show that people are aware of having very limited knowledge about the carbon footprint of common products and activities.

Study 2: Does risk aversion towards carbon impact increase polluting activity? This experiment tests whether people are risk-averse towards reducing carbon emissions. In the experiment, participants had to work on an effortful and tedious encoding task. Before engaging in the task, subjects were offered the option to obtain a computer code that would do the encoding for them, reducing their burden. However, the code was associated with the release of CO_2 emissions into the atmosphere. Across conditions, we varied the uncertainty about the size of these emissions. In addition, we measured willingness to mitigate (WTM) CO_2 emissions via an incentive-compatible multiple price list. We define the WTM as the monetary amount a participant requires in order to accept the emission of a certain quantity of CO_2. For every subject, we elicit the WTM to avoid 0, 4, 8, 12, 16, or 20 kg of CO_2 emissions. The decisions were incentivized: participants could really accrue private benefits in exchange for increasing the CO_2 emissions in the atmosphere (see Methods).

According to Expected Utility Theory, the classical theory of decisions in economics, risk aversion is driven by a concave WTM, that is, a declining marginal willingness to make sacrifices for CO_2 reductions. Intuitively, an increase in the variance of the emissions will increase the probability of both very low and very high emissions. A consumer with a concave WTM will heavily discount the very high emissions, decreasing the subjective value of offsetting these emissions, even if the expected emissions remain constant. Consequently, consumers with a concave WTM should be more likely to engage in polluting actions as the (subjective) variance of the emissions grows.

To understand whether WTM is concave, Figure 2a shows the WTM curve aggregated over all subjects. The figure shows that subjects, on average, indeed display a diminishing WTM to reduce carbon emissions. They are willing to sacrifice about £2.7 to avoid 4 kilograms of CO_2 emissions, whereas to avoid 20 kilograms of emissions, they are only willing to forgo only £4. That is, the WTM increases by less than 50% when the amount of CO_2 increases by 500%. The graph shows that this effect is robust if we exclude subjects who have decreasing valuations (a possible sign of confusion) and who are top-censored (i.e. they select the maximum WTM of £7 at least once, which could produce concavity as an artefact). Supplementary Information A.2.2 contains further details about the variation in concavity across subjects. Supplementary Information A.2.4, instead, discusses how neither cognitive uncertainty[25] nor concave moral evaluations seem to explain the concavity of the WTM. It also discusses how, given the framing of the WTM elicitation, the concavity we found is more likely due to a marginally decreasing disutility from CO_2 emissions, which are seen as a loss, rather than decreasing marginal utility from implementing offsets, which are seen as a gain.

We now turn to our main outcome, the purchase of the polluting computer code, which simplifies the real effort task. Before purchase, participants were randomized into two treatments. In the *Information* treatment, the participants knew that the emissions from buying the product were equal to 4kg of CO_2. Instead, in the *Uncertainty* treatment, they knew that there was a 40% chance that the emissions were equal to

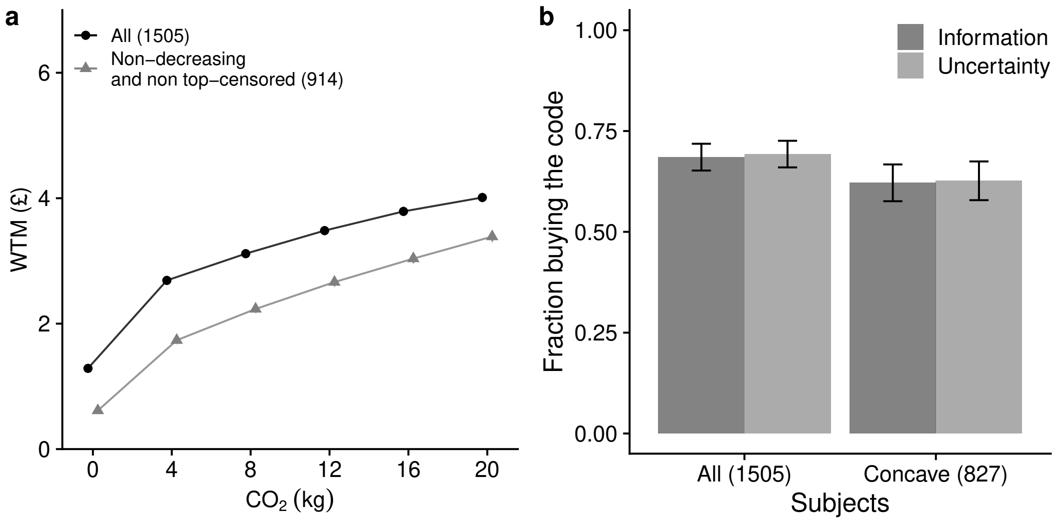


Figure 2: (a) Aggregate WTM. (b) Purchasing decision and uncertainty. *Notes*: In panel (b), the light-grey bars correspond to the group of subjects whose WTMs exhibit concavity. Bars indicate 95% CI.

0kg, a 20% chance that the emissions were 4kg, and a 40% chance that they were 8kg. Thus, the expected emissions are the same in both cases, but the variance is larger in the *Uncertainty* treatment.

In line with the reasoning above and our finding that WTM is concave, we hypothesize that uncertainty increases the fraction of subjects who purchase the polluting product. We do not find support for this hypothesis. The left bars in Figure 2b show that purchasing decisions are similar in the two treatments (68.5% vs. 69.3%, $z = 0.3168$, two-sided $p = 0.7514$). It is possible, however, that there are offsetting effects for subjects with concave and convex WTM. The right bars in Figure 2b show the treatment difference only for participants with a concave WTM. Again, we find little evidence for the hypothesized effect (62.2% vs. 62.7%, $z = 0.1494$, $p = 0.8813$).

To provide further statistical backup, we run several regression models, in which we regress the purchasing decision on the treatment, several concavity scores, and the interaction of these two (Supplementary Table 4). We also include the average WTM and several demographic controls. We find little evidence for our hypothesis. Only a single specification of concavity score produces a significant interaction with the information treatment dummy, yet this effect is not robust to other concavity measures. However, we do find a statistically significant negative effect of the average WTM on purchasing decisions, showing that the WTM data is predictive of subjects' decisions to get the computer code. We also find some demographic effects, as young people are less likely to buy than old, women less than men, and left-wing less than right-wing (Supplementary Figure 2).

Study 3: Do participants form self-serving beliefs? We now turn to a second potential channel through which uncertainty may matter: the formation of motivated beliefs. As in study 2, participants made trade-offs between monetary gains and CO_2 emissions. These trade-offs were framed as consumer decisions: participants could buy a virtual product, represented as a nondescript wrapped package on the screen. The purchase increased their pay-off by the product value of £2, minus the price at which we, the experimenters, offered it. However, purchasing the product resulted in the emission of CO_2 into the atmosphere.

We conducted an online experiment ($N = 714$) that consisted of a *Motivated* and an *Unmotivated* treatment. Both treatments featured uncertainty about the size of the emissions associated with the product. Unlike in the risk aversion experiment discussed above, we did not state explicit probabilities, as there is evidence that more ambiguous settings are conducive to the formation of motivated beliefs[26,27]. Participants could reduce uncertainty by engaging in an attentional task, somewhat akin to reading a product label, after which we elicited their beliefs about the emission size.

To study motivated beliefs, we vary whether participants complete the attentional task before or after knowing about the emission-money trade-off. Manipulating the timing of knowledge of the incentive scheme is a standard design feature in experiments studying motivated cognition[28–30]. Thus, in the *Unmotivated* treatment, subjects were presented with the task before they knew any other details of the experimental design. In this way, participants have no self-serving motive to distort their attention or beliefs in the direction of their economic interest. By contrast, in the *Motivated* treatment, participants engaged in the task after reading the full experimental instructions. Therefore, they knew that higher numbers corresponded to higher CO_2 emissions and indicated a more "inconvenient" trade-off. We hypothesize that the latter treatment will lead to motivated beliefs, i.e. a lower estimate of the impact of the emissions.

Moreover, we hypothesized that motivated beliefs are more pronounced when the payoffs from the product are higher[9,22]. To test this last hypothesis, we implemented three price treatments that varied the price of the product: a low price (£0.25), a medium price (£1), and a high price (£1.75). The information treatments are orthogonal to the price treatments, creating a 2×3 design (see Methods for details). While the trade-offs in our experiment are not about concrete, branded products, this simple design allows us to fully control the price of the products and the uncertainty about emissions while ensuring that beliefs remain, on average, correct. The average belief in the *Unmotivated treatment* is 62.7, not significantly different from 60— the real value of the emissions ($t(303) = 1.31$, 95% CI $[58.1, 67.3]$, $p = 0.25$, two-sided).

We find no evidence of the formation of self-serving beliefs. The left panel of Figure 3 shows the distribution of beliefs in both treatments, where the spikes are driven by the

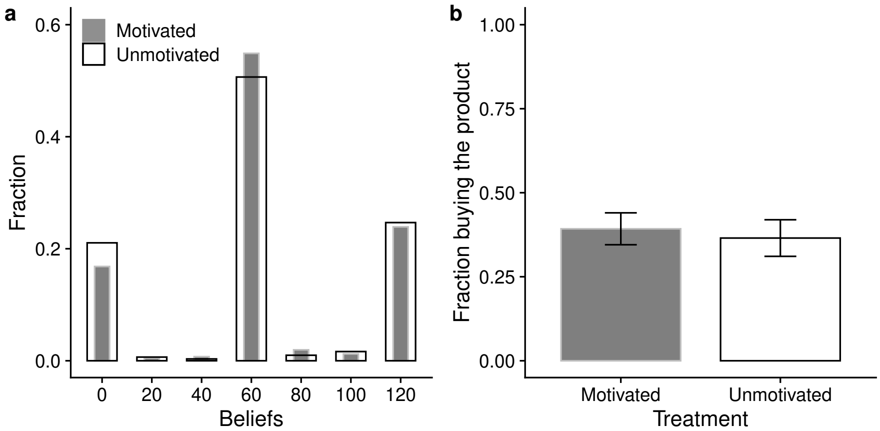


Figure 3: (a) Distributions of beliefs in the *Motivated* and *Unmotivated* treatment. (b) Fraction of participants buying the product in the *Motivated* and *Unmotivated* treatment. *Notes*: In panel (b), bars indicate 95% CI.

nature of the perceptual task (see Methods). We cannot reject the hypothesis that beliefs about emissions are the same in both treatments (Fisher's Exact test, $p = 0.66$, two-sided). Moreover, Table 1 provides the results of regressing beliefs from both treatments on a dummy for the *Motivated* treatment controlling for individual characteristics. The coefficient in column (1) is positive, indicating that, if anything, participants in the *Motivated* treatment believe that emissions are larger. Hence, the direction of the effect is the opposite of what theories of self-serving beliefs and cognitive dissonance would predict. However, the coefficient is not significantly different from zero ($t(659) = 0.60$, 95% CI $[-4.36, 8.20]$, $p = 0.548$, two-sided). Furthermore, we find no evidence that the *Motivated* treatment causes people to spend less time looking at the information as we discuss in Supplementary Information A.3.2.

This null result obtains despite substantial ambiguity: only 51% of participants answered the belief question correctly in the *Unmotivated* treatment, even though they spent, on average, 50 seconds on the task screen. Moreover, the standard deviation of beliefs is 0.65 times the average, indicating that the degree of uncertainty in the experiment is sizable. Thus, there was room for motivated subjects to perceive emissions to be lower than they actually were.

We also test for differences in purchasing behaviour between the two treatments. The right panel of Figure 3 shows that behaviour is similar in the two treatments. Both a Fisher's exact test and a t-test based on column (2) of Table 1 fail to reject that subjects are equally likely to buy the product in the *Motivated* and in the *Unmotivated* treatment (Fisher's Exact test: $p = 0.483$; t-test: $t(695) = 1.07$, 95% CI $[-3.33, 11.34]$, $p = 0.285$; both tests are two-sided). By contrast, a higher product price has a strong independent impact on purchasing behaviour: Column (4) of Table 1 shows that an increase in the price of one pound leads to a 13 percentage point decline in purchases.

Table 1: Comparison between the *Motivated* and the *Unmotivated* treatments.

	Beliefs		Units	
	(1)	(2)	(3)	(4)
Motivated	1.921	−4.406	0.040	0.051
	(3.199)	(6.226)	(0.037)	(0.074)
Price		−4.242		-0.128^{**}
		(4.127)		(0.045)
Price × Motivated		6.172		−0.022
		(5.215)		(0.060)
Controls	Yes	Yes	Yes	Yes
Observations	695	695	695	695
R^2	0.050	0.052	0.092	0.122

Notes: Models include observations from the *Motivated* and the *Unmotivated* treatments, where the latter is the baseline. The regressions exclude 19 participants for which we failed to record the demographic characteristics. Dependent variables are beliefs in columns (1) and (2) and purchasing decisions (1 if the participant purchased the product) in columns (3) and (4). Control variables: age, gender, student status, education (6 categories), frequency of car usage (5 categories), and nationality (27 categories). Robust standard errors are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Discussion

We document that people are uncertain about the CO_2 emissions associated with common consumer products. Yet we find no evidence that uncertainty affects climate action through either risk aversion or the formation of motivated beliefs about the magnitude of emissions.

Our findings add to a growing literature on the effect of information provision on climate-related behaviours. While information campaigns are popular among politicians, previous studies have found mixed evidence on the effect of correcting beliefs about emissions: some labelling studies find small reductions in emissions from informing consumers about the carbon impact of e.g. meat products[6,13,31,32], whereas other studies do not find such an effect[7]. Our findings suggest that while consumer information may reduce uncertainty, this alone may be insufficient to spur voluntary reductions in emissions. At the same time, our results also suggest that scientists can be upfront about the uncertainty of their estimates without fear of providing an excuse for polluting behaviour.

Another set of implications stems from our finding that consumers' willingness to mitigate the first units of CO_2 is much higher than for subsequent units. The first is that consumers' behaviour will be sensitive to reference points and to the framing of emission impact. For instance, framing multiple emission events separately may lead to a higher willingness to avoid or offset them than framing them as a single event. Future research could explicitly test this prediction. A second implication concerns the use of people's WTM to offset emissions to calculate the benefits of climate policies. The literature has so far relied on linear extrapolations of the WTM for a single emission amount in order to compute the benefits of reducing one ton of CO_2 in the atmosphere[16–18]. Our results strongly suggest that future studies should measure individual WTM to avoid several emission amounts and should use nonlinear models to produce accurate estimates of these benefits. It is reasonable to expect that the two implications above hold, albeit we did not find that the concavity of the WTM curve generates risk aversion toward emissions. The participants' WTM is significantly associated with them buying the polluting product. Hence, the WTM is a meaningful predictor of consumption behaviour and environmental preferences.

One limitation of our research is that, while we do not find evidence for the formation of motivated beliefs about emission size, it is possible that consumers find other dimensions for motivated reasoning about climate change[21]. Further research could investigate these additional dimensions.

Methods

General remarks

The implementation of the CO_2 emissions in the experiments. In both of our experiments, participants were asked to make decisions that could result in the emission of CO_2. To ensure that these emissions are consequential, we prepared a monetary transfer to Carbonfund.org, an organization that offsets CO_2 emissions. Every time a participant made a decision resulting in CO_2 emissions, we decreased the amount of our transfer. We explicitly communicated this procedure to the participants. We always framed these decisions as choices between private benefits and emitting CO_2 to enhance the external validity and maximize the salience of the emissions. We took several measures to assure participants of the tangible nature of these CO_2 emissions associated with the purchase of the product. We emphasized the role of the no-deception policy in obtaining ethical approval for the experiment. Additionally, we promised participants to send them the invoice for the donation to Carbonfund.org (see Supplementary Information B.5) and actually did so. To ensure that the participants believed that their actions had environmental consequences, we asked them whether they believed that we would buy the CO_2 offsets as described in the instructions. We find that between 80% and 84% of the participants expressed trust in our intentions.

Data quality. We took several measures to maximize data quality as described in Supplementary Information B.4.

Study 1: survey about subjective uncertainty

Our survey measures consumers' prevailing beliefs about the CO_2 emissions generated in the production of common consumer goods. The survey then continues with additional questions that we describe and analyze elsewhere[7].

Our elicitation methods employed incentive-compatible payment schemes developed in the experimental economics literature while maintaining simplicity and participant-friendliness in both instructions and the interface to allow for a representative sample to participate. Below, we elaborate on each of the elicitation procedures in more detail. Supplementary Information B.4 presents additional information about the steps we took to maximize the data quality.

Belief elicitation

In the survey, we elicited participants' beliefs about the CO_2 emissions generated by driving one mile by car. We then elicited beliefs about 12 common consumer products and activities listed in Table 2, including food items, household appliance use, and transportation. To enhance understanding, we provided participants with information about product specifications and the type of emissions considered by the estimate.[a] We retrieved these estimates from top-tier academic journals or from the estimates the UK government uses for its environmental regulations. We disclosed these scientific sources only at the end of the experiment. To make the values more meaningful to subjects, we refrained from eliciting emissions in grams and instead asked about the number of miles one must drive by car to emit the same amount of CO_2 as the product in question, an approach in line with previous studies[6].

We divided the belief elicitation into two parts. We first elicited a point estimate for the modal value of the emissions. Participants indicated how much CO_2 each of the 12 products in Table 2 emitted relative to driving one mile by car. Participants answered all 12 questions on one page, and the order of the products was randomized across participants. A screenshot of the interface is in Supplementary Information B.1. In the rest of the experiment, the same order was used every time participants answered additional questions about these products. To help participants keep track of their guesses and the rankings of the products, we presented an interactive box summarizing their (current) answers at the bottom of the page, including the ranking of the products by estimated impact. We incentivized a correct point estimate with a £4 bonus. We considered an estimate correct if it was within a 5% interval from the scientific estimate. This incentive scheme truthfully elicits the mode of the subjective probability distribution about

[a] Participants could learn a detailed breakdown of the factors considered in the scientific calculations of CO_2 emissions size. See Supplementary Table 12.

Table 2: List of consumer products and activities.

	Quantity	Emission size Estimate	Unit	Source
Beer	12 fl oz	1.46	mile	Poore and Nemecek[1]
Phone call	1 hour	1.55	mile	Smith et al.[33]
Microwave	1000W, 2 hour	1.76	mile	UK BEIS[34]
Milk	1 cup	2.60	mile	Poore and Nemecek[1]
Egg	6 eggs	4.81	mile	Poore and Nemecek[1]
Poultry meat	7 oz	6.78	mile	Poore and Nemecek[1]
Shower	Average usage	3.90	mile	Hackett and Gray[35]
Dark chocolate	100g	16.03	mile	Poore and Nemecek[1]
Coffee beans	1 lb	44.41	mile	Poore and Nemecek[1]
Beef	7 oz	68.39	mile	Poore and Nemecek[1]
Flight	SFO to LAX	304.60	mile	UK BEIS[34]
Gas heating	One month	606.68	mile	Padgett et al.[36]

the scientific estimate[23].[b]

To understand the participants' confidence in their answers, we then elicited the subjective probability distribution of the size of CO_2 emissions. For each product, we presented five "bins" around the point estimate the participant reported in the first part and asked the participant to allocate 20 balls into these five bins. We told participants that each bin represents an interval that might contain the scientific estimate and that they should allocate the balls to represent their level of confidence that the estimate is, in fact, in that bin. A screenshot of the interface is in Supplementary Information B.1. We incentivized the elicitation by randomly selecting one of the bins and scoring the answer according to a randomized quadratic scoring rule. This mechanism encourages participants to truthfully reveal their belief that the scientific estimate falls in a particular bin[37]. To keep things simple and avoid information overload, we did not provide participants with the exact details of the scoring rule but made them available with a mouse click. We told subjects that they would maximize their expected earnings by answering truthfully, an approach suggested by Danz et al.[24]

Implementation We recruited 1,430 participants on Prolific (`https://www.prolific.co/`) between the 3rd and 6th December 2020, and 1,128 completed all the belief elicitation questions described in this paper. We restricted participation to US residents, and we aimed to collect a sample representative for age, gender, and ethnicity.[c] Our sample

[b]We did not incentivize the questions about the CO_2 emissions and the social cost of driving one mile by a car as we realized that answers to these questions can be straightforwardly obtained on Google.

[c]We compared the demographic characteristics of study participants and information from US Census Bureau[38], and confirmed that our sample is representative for gender and ethnicity, but not

is, on average, 42.6 years old (SD = 15.4), and 48.5% of the participants self-identified as male. Supplementary Table 1 shows the demographic characteristics of the sample. In total, participants were asked to answer 21 comprehension questions.

At the end of the survey, we randomly selected one question from the entire study per participant. There was a 30% probability that one of the belief questions analyzed above was selected for payment. The average duration of the survey was 51 minutes, but all the questions described here were in the first half of the study.

Study 2: Risk Aversion experiment

The experiment consists of two parts. In the first part, we measure the participants' valuations of CO_2 emissions. In the second part, we ask participants whether they want to get a valuable but polluting product. The instructions for this experiment are in Supplementary Information C.2. We preregistered the experimental design and the hypotheses on AsPredicted.org (see Supplementary Information B.6).

Valuation of CO_2 emissions. We measured the participants' valuation of CO_2 emissions by presenting them with trade-offs between money and emissions. Specifically, participants were offered a choice between Option A and Option B. Opting for Option A meant forgoing any monetary gain but preventing the generation of CO_2 emissions. In contrast, choosing Option B allowed them to earn money but resulted in CO_2 emissions. For a given CO_2 emission level, the participants had to make 15 choices where the amount of money they could earn increased from £0 to £7 in 50 pence increments. These decisions were embedded in a Multiple Price List (see Supplementary Information C.2), which enforced a single switching point between Option A and Option B. This switching point gives us the participants' valuation for a given amount of CO_2 emissions. To gauge the participants' certainty regarding their valuations, we employed the "cognitive uncertainty" elicitation method developed by Enke and Graeber[25]. This elicitation was skipped if the participants never switched from Option A to Option B. In total, the participants saw six Multiple Price Lists, each corresponding to one of the following emissions levels: 0 kg, 4 kg, 8 kg, 12 kg, 16 kg, and 20 kg. Half of the participants saw these lists in ascending order of emission size, while the other half saw them in descending order.

Consumption decision. In the second part of the experiment, the participants had the opportunity to receive a valuable yet polluting product, framed as a "computer code", which significantly speeded up the completion of a time-consuming and laborious real-effort task. The task involved typing 15 strings, each consisting of 15 characters, in

for age (Supplementary Table 2).

reverse order (Supplementary Figure 6). The participants were required to transcribe these strings flawlessly to complete the task: any mistakes incurred resulted in an error message indicating the specific strings that required correction before they could proceed with the experiment. Those who obtained the computer code were presented with a task in which the code generated all the correct answers, reducing their task to simply clicking one button to submit them. To ensure task completion without interruption and to make the task sufficiently "annoying," we implemented an attention check. The participants saw a warning sign every 30 seconds, and upon its appearance, they had a 5-second window to click on a button on the screen to confirm their active engagement with the task. The participants were excluded from the experiment if they failed to click on the button within the specified time window in more than four instances.

At first, all the participants were required to complete the real-effort task once without the help of the computer code. In the second round, they had the chance to obtain the code to speed up the completion time. On average, the participants spent 8 minutes and 40 seconds to complete the first task.

Obtaining the code came at the cost of emitting CO_2. There are two treatments that vary the information participants had about the size of the CO_2 emissions. In the *Information* treatment, participants were informed that purchasing the product would result in emissions equivalent to 4kg of CO_2. In contrast, in the *Uncertainty* treatment, participants were informed of a probability distribution: a 40% probability of emissions being 0kg, a 20% probability of emissions being 4kg, and a 40% probability of emissions being 8kg.

Survey. Before the participants complete (or are allowed to skip) the second real effort task, they complete a questionnaire. A battery of questions elicits people's moral evaluations of emissions-money trade-offs. Participants indicate the morality acceptability of emitting 4, 12, and 20kg of CO_2 for either £1 or £5 using a Likert scale, where 1 indicates a decision that is "morally very inappropriate" and 7 a decision that is "morally very appropriate". Further questions in the survey ask about attitudes toward climate change and demographic characteristics.

Sample and data collection. We recruited a total of 1,935 participants through the online platform Prolific.co on January 5^{th}, 2023. Following the preregistration, 1,505 participants who successfully completed the final survey were included in the analysis. Among them, 753 participants were assigned to the *Information* treatment, while the remaining 752 were assigned to the *Uncertainty* treatment. Fifty per cent of the participants identified as females, and the average age is 39 years old (min = 18, max = 79, SD = 12.51). We restricted participation to individuals based in the UK. Subjects earned a fixed reward of £3, with the potential for a bonus payment based on their decisions. On average, they earned £3.66, and they took about 34 minutes to complete the tasks. Following the participants' decisions, we donated \$160 to Carbonfund.org to offset CO_2 emissions, resulting in a reduction of almost 13 metric tons of CO_2 in the atmosphere. The experiment maintained participant anonymity, as we identified participants only via alphanumerical IDs provided by the platform. Participants knew that they were taking part in an experiment.

Study 3: Motivated Belief experiment

In this experiment, we offered participants the opportunity to buy a single unit of a virtual product.[d] If participants decided to purchase the product, they increased their pay-off by the product value of £2, minus the price at which we, the experimenters, offered it. Importantly, purchasing the product entailed the emission of CO_2 into the atmosphere. We framed the experiment as a market interaction, employing terminology such as "virtual product" and "price," to make it closer to a real-life purchasing situation.

We informed participants that purchasing the product would result in CO_2 emissions equivalent to burning 60 litres of gasoline. This emission size had offset costs of £1.07, commensurate to the other payments in the experiments.[e]

As we describe next, we orthogonally implemented three price treatments and two information treatments, resulting in a total of six treatments. Each participant was assigned to a single treatment. The instructions for this experiment are in Supplementary Information C.3. We preregistered the experimental design and the hypotheses on AsPredicted.org (see Supplementary Information B.6).

Uncertainty treatments. Our primary focus is on the formation and impact of beliefs about the CO_2 emissions associated with the product, as well as the role of emission information. To study this, we employed two treatments, called the *Motivated* and *Unmotivated* treatments, that varied the nature of uncertainty about the size of the emissions.

In both treatments, there was uncertainty about the size of the emissions. The participants had the opportunity to reduce this uncertainty by engaging in an attentional task designed to mimic real-life consumption scenarios, such as reading product labels or conducting an online search for information. The task involved examining a matrix

[d] This product is virtual in the sense that it exists only inside the experiment; it is not a physical product nor a service. Nevertheless, the product is valuable to the participants since their pay-off from the experiment increases if they "buy" it.

[e] Using a report from the US Environmental Protection Agency[39], we calculate that burning 60 litres of gasoline produces 140kg of CO_2 emissions. At the time of the experiment, Carbonfund.org offset 1 metric ton of CO_2 per every \$10 (or £7.9) it receives in donations, so offsetting the products' emissions cost £1.07.

of numbers, with the most frequently appearing number representing the emission size, measured in terms of the CO_2 emissions generated by burning a litre of gasoline. We adapted this task from Sandro Ambuehl's work[40], which shows that the information-gathering strategy in this task is influenced by incentives for subsequent decisions.[f] Participants had up to one minute to engage with the task, after which we elicited their beliefs by asking them which number they believed was the most frequently found in the table. Providing the correct answer was rewarded with a bonus of £0.10, which incentivized participants to report the mode of their belief distribution[23].[g]

These two treatments differ in the order in which we presented the attentional task and the information about the emission size. In the *Unmotivated* treatment, participants were presented with the task prior to receiving any instruction about the possibility of emitting CO_2. In this way, we eliminated any self-serving motives that might lead participants to distort their attention or beliefs in the direction of their economic interests. By contrast, in the *Motivated* treatment, participants engaged in the task after they had read the full experimental instructions. Consequently, they were aware that the correct answer to the task indicated the magnitude of the CO_2 emissions as well as of the surplus they could obtain from the product. This treatment enables us to test if the surplus from buying the product has a causal effect on participants' belief formation in the attentional task and, in turn, on the product's purchase.

At the end of the experiment, we collected demographic information using a survey.

Price treatments. We also investigate the relationship between purchasing decisions and prices. The effect of prices provides a natural quantitative benchmark for assessing the effects of beliefs and information, as price incentives are a primary tool used by policymakers and are often discussed in the context of reducing carbon emissions.

We implemented three price treatments that varied the price of the product: a low price of £0.25, a medium price of £1, and a high price of £1.75. Participants were informed that the price was randomly assigned and held no informational content regarding emission size. We made sure of the participants' understanding of this aspect by asking them a comprehension question on the topic.

[f]The task can be found in Supplementary Information B.3. The matrix contained a total of 143 numbers drawn from the set $\{0, 20, 40, 60, 80, 100, 120\}$. The number 60, the most frequently occurring, appeared 35 times, with 0 and 120 being the next most frequent, each appearing 26 times. All other numbers appeared 14 times each.

[g]Note that the experiment had a third treatment which gave people full information about the carbon impact of the product. This treatment is analyzed and described elsewhere[41]. We do not analyze it here, as the comparison between the info and the uncertainty treatments is complicated by the fact that we do not have full information about the belief distributions. Hence, we cannot be sure that the expected value of emissions is constant in the different treatments. Experiment 1 in this study, therefore, provides higher quality evidence on the effect of precise information on behaviour.

Sample and data collection. We recruited 714 participants using Prolific.co, an online platform, between 9^{th} and 11^{th} May 2019.[h] Of those, 304 participants were assigned to the *Unmotivated* treatment (87 faced a £0.25 price, 107 a price of £1.00, and 110 a price of £1.75), and 410 participants were assigned to the *Motivated* treatment (146 faced a £0.25 price, 125 a price of £1.00, and 139 a price of £1.75). Demographic information for 19 participants was not successfully recorded (11 and 8 from the *Unmotivated* and *Motivated* treatment, respectively). These 30 subjects are included in the analysis when we run non-parametric tests, but they are excluded in the regression analysis, which includes the demographic controls.

Fifty per cent of the participants identified as females, 42% are students, and the average age is 29 years old. We accepted only EU nationals as participants. The most represented countries in our sample are the UK (34.7%), Poland (14.54%), and Portugal (12.1%). Participants earned a fixed reward of £1.60, with a potential bonus payment contingent on their decisions. On average, they earned £2.04, and they took less than 13 minutes to complete the tasks. We obtained participants' demographic information, including gender, age, student status, and nationality, directly from Prolific.co. Following the participants' decisions, we donated \$911.40 to Carbonfund.org to offset CO_2 emissions, resulting in a reduction of over 90 metric tons of CO_2 in the atmosphere.

The experiment maintained participant anonymity, as we identified participants only via alphanumerical IDs provided by the platform. Participants knew that they were taking part in an experiment.

References

[1] Joseph Poore and Thomas Nemecek. Reducing food's environmental impacts through producers and consumers. *Science*, 360(6392):987–992, 2018.

[2] Robert S Pindyck. What we know and don't know about climate change, and implications for policy. *Environmental and Energy Policy and the Economy*, 2(1): 4–43, 2021.

[3] David V. Budescu, Han-Hui Por, Stephen B. Broomell, and Michael Smithson. The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change*, 4(6):508–512, 2014.

[h] Researchers have recently voiced concern that bots, instead of humans, might complete online experiments. To the best of our knowledge, no such concerns have been raised regarding Prolific.co. Nevertheless, we took two precautionary measures to further minimize the risk of undetected bots completing our experiment. First, we incorporated two "honey-pots," which are questions that a human participant would not be able to see, but that a bot that reads the source code of the experimental program should identify as questions to be answered. Second, we kept track of the number of attempts participants needed to answer the comprehension questions correctly. We did not find any evidence that bots completed our study. Additional details are in Supplementary Information B.4.

[4] Baruch Fischhoff and Alex L. Davis. Communicating scientific uncertainty. *Proceedings of the National Academy of Sciences*, 111(supplement_4):13664–13671, 2014.

[5] Stephen B. Broomell and Patrick Bodilly Kane. Public perception and communication of scientific uncertainty. *Journal of Experimental Psychology: General*, 146 (2):286, 2017.

[6] Adrian R. Camilleri, Richard P. Larrick, Shajuti Hossain, and Dalia Patino-Echeverri. Consumers underestimate the emissions associated with food but are aided by labels. *Nature Climate Change*, 9(1):53–58, 2019.

[7] Taisuke Imai, Davide D. Pace, Peter Schwardmann, and Joël J. van der Weele. Correcting consumer misperceptions about CO_2 emissions. CESifo Working Paper No. 10138, 2022.

[8] Roland Bénabou and Jean Tirole. Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30(3):141–164, 2016.

[9] Nina Hestermann, Yves Le Yaouanq, and Nicolas Treich. An economic model of the meat paradox. *European Economic Review*, 129:103569, 2020.

[10] Rafael Di Tella, Ricardo Perez-Truglia, Andres Babino, and Mariano Sigman. Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *American Economic Review*, 105(11):3416–3442, 2015.

[11] Christine L. Exley. Excusing selfishness in charitable giving: The role of risk. *Review of Economic Studies*, 83(2):587–628, 2015.

[12] Matthias Rodemeier and Andreas Löschel. The welfare effects of persuasion and taxation: Theory and evidence from the field. Available at SSRN: `http://dx.doi.org/10.2139/ssrn.3587339`, 2020.

[13] Paul Lohmann, Elisabeth Gsottbauer, Anya Doherty, and Andreas Kontoleon. Do carbon footprint labels promote climatarian diets? Evidence from a large-scale field experiment. *Journal of Environmental Economics and Management*, 114:102693, 2022.

[14] John A. List and Craig A. Gallet. What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental and Resource Economics*, 20:241–254, 2001.

[15] Matthias Rodemeier. Willingness to pay for carbon mitigation: Field evidence from the market for carbon offsets. IZA Discussion Paper 15939, 2023.

[16] Gregory F. Nemet and Evan Johnson. Willingness to pay for climate policy: A review of estimates. Available at SSRN: `https://dx.doi.org/10.2139/ssrn.1626931`, 2010.

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

[18] Johannes Diederich and Timo Goeschl. Willingness to pay for voluntary climate action and its determinants: Field-experimental evidence. *Environmental and Resource Economics*, 57(3):405–429, 2014.

[19] Hannes Cordes, Stefan Baumeister, and Minna Käyrä. Factors influencing the willingness to pay for aviation voluntary carbon offsets: A literature review. *European Journal of Tourism Research*, 36:3602, 2024.

[20] Marc N. Conte and Matthew J. Kotchen. Explaining the price of voluntary carbon offsets. *Climate Change Economics*, 1(02):93–111, 2010.

[21] Shahzeen Z. Attari, David H. Krantz, and Elke U. Weber. Energy conservation goals: What people adopt, what they recommend, and why. *Judgment and Decision Making*, 11(4):342–351, 2016.

[22] Peter Schwardmann. Motivated health risk denial and preventative health care investments. *Journal of Health Economics*, 65:78–92, 2019.

[23] Karl H. Schlag, James Tremewan, and Joël J. van der Weele. A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, 18 (3):457–490, 2015.

[24] David Danz, Lise Vesterlund, and Alistair J. Wilson. Belief elicitation and behavioral incentive compatibility. *American Economic Review*, 112:2851–2883, 2022.

[25] Benjamin Enke and Thomas Graeber. Cognitive uncertainty. *Quarterly Journal of Economics*, 138(4):2021–2067, 2023.

[26] Emily C. Haisley and Roberto A. Weber. Self-serving interpretations of ambiguity in other-regarding behavior. *Games and Economic Behavior*, 68(2):614–625, 2010.

[27] Jan Engelmann, Maël Lebreton, Peter Schwardmann, Joël van der Weele, and Li-Ang Chang. Anticipatory anxiety and wishful thinking. Tinbergen Institute, 2019.

[28] Linda Babcock, George Loewenstein, Samuel Issacharoff, and Colin Camerer. Biased judgments of fairness in bargaining. *American Economic Review*, 85(5):1337–1343, 1995.

[29] Peter Schwardmann and Joël van der Weele. Deception and self-deception. *Nature Human Behaviour*, 3(10):1055–1061, 2019.

[30] Uri Gneezy, Silvia Saccardo, Marta Serra-Garcia, and Roel van Veldhuizen. Bribing the self. *Games and Economic Behavior*, 120:311–324, 2020.

[31] Anna Schulze Tilling. Changing consumption behavior with carbon labels: Causal evidence on behavioral channels and effectiveness. Mimeo, University of Bonn, 2023.

[32] David Bilén. Do carbon labels cause consumers to reduce their emissions? Evidence from a large-scale natural experiment. Mimeo, Gothenburg University, 2022.

[33] Anna Jo Bodurtha Smith, Imogen Tennison, Ian Roberts, John Cairns, and Caroline Free. The carbon footprint of behavioural support services for smoking cessation. *Tobacco Control*, 22(5):302–307, 2013.

[34] UK BEIS. Greenhouse gas reporting: Conversion factors 2019, 2020. UK Department for Business, Energy & Industrial Strategy. `https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2019`.

[35] M. J. Hackett and Nicholas. F. Gray. Carbon dioxide emission savings potential of household water use reduction in the UK. *Journal of Sustainable Development*, 2 (1):36–43, 2009.

[36] J. Paul Padgett, Anne C. Steinemann, James H. Clarke, and Michael P. Vandenbergh. A comparison of carbon calculators. *Environmental Impact Assessment Review*, 28(2-3):106–115, 2008.

[37] Karl H. Schlag and Joël van der Weele. Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality. *Theoretical Economics Letters*, 3:38–42, 2013.

[38] US Census Bureau. National population by characteristics: 2020-2022, 2022. United States Census Bureau. `https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-detail.html`.

[39] United States Environmental Protection Agency. Emission facts: Average carbon dioxide emissions resulting from gasoline and diesel fuel. EPA420-F-05-001, February 2005. `https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1001YTF.txt`, 2005.

[40] Sandro Ambuehl. An offer you can't refuse? Incentives change what we believe. Available at SSRN: https://dx.doi.org/10.2139/ssrn.2830171, 2016.

[41] Davide D. Pace and Joël van der Weele. Curbing carbon: An experiment on uncertainty and information about CO_2 emissions. Tinbergen Institute Discussion Paper No. 2020-059, 2020.

Supplementary Information

Uncertainty about Carbon Impact and the Willingness to Avoid CO_2 Emissions

Davide D. Pace Taisuke Imai
Peter Schwardmann Joël J. van der Weele

Contents

A Additional Results

A.1 Study 1: Survey about Subjective Uncertainty

Supplementary Table 1: Demographic characteristics.

Age			*Education*		
18-27	223	0.198	Less than high school	8	0.008
28-37	267	0.237	High school degree	107	0.104
38-47	195	0.173	Some University but no degree	288	0.281
48-57	186	0.165	Bachelor Degree	373	0.364
58+	257	0.228	Postgradute degree	249	0.243
Gender			*Household income*		
Female	525	0.504	- \$5,000	26	0.025
Male	505	0.485	\$5,000 - \$15,000	68	0.066
Other	12	0.012	\$15,000 - \$30,000	129	0.126
Ethnicity			\$30,000 - \$45,000	130	0.127
Asian	70	0.067	\$45,000 - \$60,000	136	0.133
Black	141	0.135	\$60,000 - \$75,000	114	0.111
Mixed	30	0.029	\$75,000 - \$90,000	88	0.086
White	776	0.745	\$90,000 - \$105,000	83	0.081
Other	25	0.024	\$105,000 - \$120,000	90	0.088
Party affiliation			\$120,000 - \$135,000	30	0.029
Republican	152	0.148	\$135,000 - \$150,000	37	0.036
Republican-leaning independent	68	0.066	\$150,000 -	94	0.092
Independent	202	0.197			
Democratic-leaning independent	147	0.143			
Democratic	456	0.445			
Political orientation					
Conservative	100	0.098			
Somewhat conservative	225	0.220			
Somewhat liberal	324	0.316			
Liberal	376	0.367			

Notes: 1,128 participants completed the belief elicitation task.

Supplementary Table 2: Representativeness of the sample.

	Sample	Population	
Age			
18-27	0.198	0.172	
28-37	0.237	0.176	
38-47	0.173	0.160	
48-57	0.165	0.162	$\chi^2(4) = 64.658$
58+	0.228	0.330	$p < 0.001$
Gender			
Female	0.510	0.504	$\chi^2(1) = 0.1453$
Male	0.490	0.496	$p = 0.7031$
Ethnicity			
Asian	0.071	0.064	
Black	0.143	0.142	$\chi^2(2) = 0.7821$
White	0.786	0.794	$p = 0.6763$

Notes: Population-level data is retrieved from US Census Bureau (2022).

A.2 Study 2: Risk Aversion Experiment

A.2.1 Demographic characteristics

Supplementary Table 3: Demographic characteristics.

			Treatment		
		All	*Information*	*Uncertainty*	
Age bracket					
18-27	277	0.184	0.178	0.190	$\chi^2(4) = 1.06$
28-37	482	0.320	0.331	0.310	$p = 0.90$
38-47	354	0.235	0.230	0.241	
48-57	235	0.156	0.155	0.157	
58+	157	0.104	0.106	0.102	
Gender					
Male	744	0.494	0.490	0.499	$\chi^2(2) = 1.47$
Female	754	0.501	0.507	0.495	$p = 0.48$
Other	7	0.005	0.003	0.007	
Political view					
Left	295	0.196	0.190	0.202	$\chi^2(4) = 1.06$
Center-left	445	0.296	0.303	0.289	$p = 0.90$
Center	493	0.328	0.321	0.334	
Center-right	214	0.142	0.145	0.140	
Right	58	0.039	0.041	0.036	
Education					
Less than high school	27	0.018	0.013	0.023	$\chi^2(4) = 7.18$
High school	330	0.219	0.211	0.227	$p = 0.127$
Some University	176	0.117	0.120	0.114	
Bachelor	658	0.437	0.465	0.410	
Postgraduate	314	0.209	0.191	0.226	
Income					
- £5,000	27	0.018	0.020	0.016	$\chi^2(8) = 15.68$
£5,000 - £15,000	134	0.089	0.092	0.086	$p = 0.047$
£15,000 - £30,000	329	0.219	0.220	0.217	
£30,000 - £45,000	321	0.213	0.218	0.209	
£45,000 - £60,000	246	0.163	0.151	0.176	
£60,000 - £75,000	183	0.122	0.139	0.104	
£75,000 - £90,000	124	0.082	0.084	0.081	
£90,000 - £105,000	60	0.040	0.024	0.056	
£105,000 -	81	0.054	0.052	0.056	

A.2.2 Shape of the individual-level WTM curve

We elicited willingness to mitigate (WTM) across six emission levels: 0, 4, 8, 12, 16, and 20 kilograms of CO_2 emissions. In this analysis, we focus on characterizing the shape of individual-level WTM curves. Let (e_i, w_i) denote the pair of emission size e_i and the reported WTM $w_i \in [0, 7]$, for each $i = 1, \ldots, 6$. It is important to note that in this analysis, we consider the entire range of 0-20 kg, as opposed to the more constrained range of 0-8kg (used in Supplementary Table 4 below) when classifying the shape of these curves.

Step 1. For each participant, we construct a piecewise linear WTM curve using linear interpolation, consisting of five line segments. The WTM curve has five line segments. The slope of the ith line segment, denoted as s_i, is given by:

$$s_i = \frac{w_{i+1} - w_i}{e_{i+1} - e_i}.$$

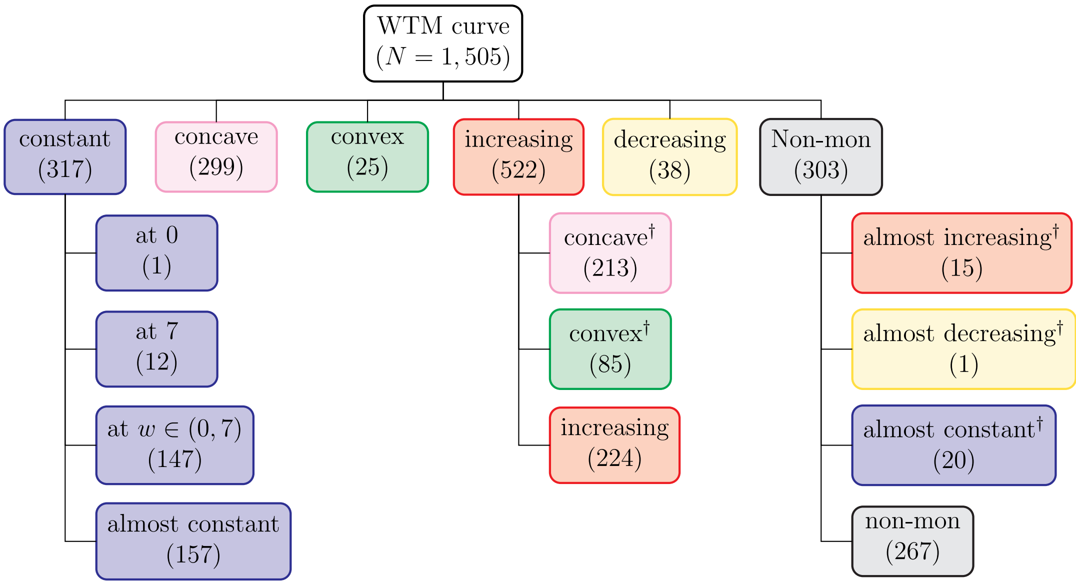
We apply the following rule sequentially to classify the shape of the WTM curve.[1] We say that a WTM curve is

- *constant* if $s_i = 0$ for all i;
- *almost constant* if $\max w_i - \min w_i \leq 0.5$, i.e. the step size of the MPL;
- *decreasing* if $s_i \leq 0$ for all i with at least one strict inequality;
- *concave* if $s_{i+1} \leq s_i$ for all i with at least one strict inequality;
- *convex* if $s_{i+1} \geq s_i$ for all i with at least one strict inequality;
- *increasing* if $s_i \geq 0$ for all i with at least one strict inequality;
- *non-monotonic* if it does not fall into any of the above categories.

In our dataset, we identified 317 WTM curves as (almost) constant, 38 as decreasing, 299 as concave, 25 as convex, and 522 as increasing. The remaining 303 WTM curves exhibit a non-monotonic behavior.

Step 2. Let us direct our attention to the subset of 522 participants whose WTM curves exhibit an increasing trend while not falling into the categories of concave or convex shapes. Among these participants, 66 individuals have their WTM values censored at a maximum of £7. Let $\bar{w}$ represent the largest observed WTM value. If $\bar{w} = 7$, we define $\bar{e}$ as the smallest emission level e_i for which $w_i = 7$. If $\bar{w} < 7$, on the other hand, we set

[1] This means that *concave* and *convex* WTM curves in this classification are non-decreasing, and *increasing* WTM curves are neither concave nor convex. We classify *linear* WTM curves as *concave*.



Supplementary Figure 1: Classification of individual-level WTM curves.

$\bar{e} = e_6$. Next, we draw a chord connecting two points: (e_1, w_1) and $(\bar{e}, \bar{w})$. We say that a WTM curve is *concave*† (*convex*†) if the points (e_i, w_i) for which $e_i \leq \bar{e}$ lie above (below) the chord. In our dataset, we identified 213 WTM curves as concave† and 85 as convex†.

Step 3. Finally, we turn to the remaining 303 participants whose WTM curves exhibit non-monotonic behavior.

First, we say that a WTM curve is *almost constant*† if the difference between the largest and smallest WTM values does not exceed £1, which corresponds to two steps in the MPL. This relaxation captures the shape of an additional 20 WTM curves.

Second, we say that a WTM curve is *almost increasing*† (*almost decreasing*†) if the piecewise linear WTM curve has only one line segment with a negative (positive) slope, and the relative change of WTM on that segment is "not too large".[2] This relaxation captures the shape of an additional 16 WTM curves.

Classification summary. Allowing some margin of error, we have established a comprehensive and mutually exclusive classification of individual-level WTM curves as follows: 337 are constant, 512 are concave, 110 are convex, 239 are increasing, 39 are decreasing, and 267 are non-monotonic.

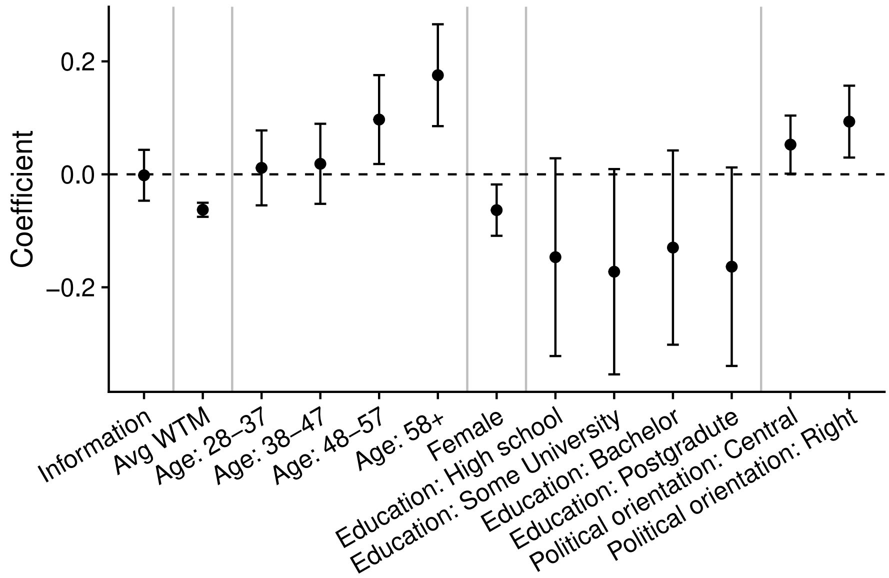
[2]Suppose the sign of the slopes changes on the segment connecting (e_j, w_j) and (e_{j+1}, w_{j+1}). We require the absolute relative change to be less than 10%, i.e., $|(w_{j+1} - w_j)/w_j| \leq 0.1$.

A.2.3 Purchasing decisions

Supplementary Table 4: Effect of uncertainty on purchasing decisions.

	(1)	(2)	(3)	(4)	(5)	(6)
Information	-0.008	-0.008	0.011	0.014	0.00002	0.002
	(0.024)	(0.034)	(0.032)	(0.027)	(0.034)	(0.032)
Concavity				-0.032		-0.067
				(0.026)		(0.061)
Info × Concavity				-0.069^{*}		-0.043
				(0.035)		(0.064)
Average WTM				-0.037^{***}		-0.034^{*}
				(0.009)		(0.015)
Concavity alt					0.022	
					(0.033)	
Info × Concavity alt					-0.074	
					(0.046)	
Constant	0.571^{***}	0.437^{*}	0.644^{***}	0.709^{***}	0.516^{**}	0.697^{***}
	(0.127)	(0.194)	(0.147)	(0.118)	(0.170)	(0.118)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,505	827	678	1,162	912	1,118
R^2	0.054	0.058	0.075	0.100	0.052	0.102

Notes: Models (1) to (5) are linear regressions. Model (6) is an IV regression. The dependent variable is `product_bought`, a dummy equal to 1 if the participant bought the convenience product. `Information` is a treatment dummy equal to 1 in the *Information* treatment. `Concavity` is given by $\text{WTM}_4 - (\text{WTM}_0 + \text{WTM}_8)/2$, `Average WTM` is given by $(\text{WTM}_0 + \text{WTM}_4 + \text{WTM}_8)/3$, and `Concavity_alt` is given by dividing `Concavity` by $(\text{WTM}_8 - \text{WTM}_0)/2$. Column 1 uses all the observations. Column 2 includes only the participants with a strictly concave WTM in the interval 0-8kg. Column 3 includes only the participants with a strictly convex or linear WTM in the 0-8kg interval. Column 4 excludes the participants whose WTM is censored or decreasing in the 0-8kg interval. Column 5 further excludes the participants for whom $\text{WTM}_0 = \text{WTM}_8$, since `Concavity alt` is not defined for them. Column 6 instruments `Concavity`, `Info*Concavity`, and `Average WTM` with their equivalent variables coming from the unincentivized WTM elicitation. This column excludes the participants with decreasing or censored WTM in any of the elicitations. List of control variables common to all regressions: age, gender (male, female, other), political affiliation (5 categories), education (6 categories), income (7 categories), and time needed to complete the first real effort task. Robust standard errors are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.



Supplementary Figure 2: Estimated coefficients from a linear regression. *Notes*: The dependent variable is `product_bought`, which is a dummy equal to 1 if the participant bought the computer code. The intercept is not shown for a better visual rendering. The baseline categories are: "18-27" for `Age`, "Male" for `Gender`, "Less than high school" for `Education`, and "Left wing" for `Political orientation`. Bars indicate 95% CI.

A.2.4 Psychological mechanisms behind concavity

We empirically explore two potential psychological mechanisms that may give rise to a concave WTM curve. The first mechanism relates to individuals' inability to appreciate increasingly large (and unfamiliar) amounts of emissions. The second mechanism considers the possibility that the concavity in WTM arises from concave moral judgments about the acceptability of causing different levels of emissions. Our data does not support either of these two mechanisms. However, based on the framing of the elicitation questions, we propose that the concavity we found is more likely due to a marginally decreasing disutility from CO_2 emissions — which are seen as a loss — rather than decreasing marginal utility from implementing offsets — which are seen as a gain.

Increasing cognitive uncertainty. People may perceive the questions involving larger emission quantities as more challenging due to the inherent complexity of visualizing the precise scale of higher levels of emissions. This heightened level of complexity can lead participants to experience greater cognitive uncertainty when deciding their WTM, making them less sensitive to variations in increases in emission sizes. This relation between cognitive uncertainty and valuation can generate a concave WTM curve within the framework of an "anchoring and adjustment" model, in which the weight attributed to the anchor increases with cognitive uncertainty (Enke and Graeber, 2023). The anchor, in this case, is the default behaviour of not compensating for CO_2 emission produced in everyday life.

The anchoring and adjustment model predicts the concavity of WTM under two conditions: (a) individuals generally do not engage in emissions offsetting, making an anchor value of £0 a plausible assumption, and (b) cognitive uncertainty increases with emission size. We find empirical support for both of these underlying assumptions. Specifically, over 82.7% of our participants reported to "Never" or "Rarely" compensate for their emissions. Furthermore, in a regression that controls for demographic characteristics, we find that cognitive uncertainty increases with emission size ($t(1231) = 8.132$, 95% CI $[0.067, 0.109]$, two-sided $p < 0.001$).

However, our analysis does not reveal any substantial evidence of a relationship between cognitive uncertainty and the concavity of the WTM curve. In Supplementary Table 5, we present the results of a regression in which the WTM is regressed on a) cognitive uncertainty, b) emission levels, c) the square of the emission levels, d) the interaction between cognitive uncertainty and the emission levels, and e) the interaction between cognitive uncertainty and the square of the emission levels. We include the square of the emissions to account for potential nonlinear associations between emissions and WTM. The interaction between the square of emissions and cognitive uncertainty is included to explore whether higher levels of uncertainty are linked to more pronounced

Supplementary Table 5: Concavity of WTM and cognitive uncertainty.

	(1)	(2)
Emissions	0.16859***	0.16906***
	(0.01090)	(0.01204)
Emissions^2	−0.00327***	−0.00322***
	(0.00043)	(0.00048)
Cognitive uncertainty	0.00776	0.00940
	(0.00785)	(0.00858)
Cognitive uncertainty × Emissions	0.00164	0.00126
	(0.00198)	(0.00213)
Cognitive uncertainty × Emissions^2	−0.00009	−0.00008
	(0.00009)	(0.00009)
Constant	0.10094	0.14302
	(0.58873)	(0.65010)
Controls	Yes	Yes
Observations	7,392	6,222
Clusters	1,232	1,037
R^2	0.1494	0.1485

Notes: The dependent variable is WTM. The first column includes only the participants that have an uncensored WTM for all 6 emission amounts. The second column further excludes the participants who said that offset all their emissions or that they "often" offset their emissions. List of control variables common to all regressions: age, gender (male, female, other), political affiliation (5 categories), education (6 categories), income (7 categories), and time needed to complete the first real effort task. Standard errors clustered at the individual level are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

concavity, which a negative coefficient for the interaction term would indicate. Column (1) of Supplementary Table 5 shows that the coefficient indeed appears negative, but it is small in magnitude and fails to reach statistical significance ($t(1231) = -0.987$, 95% CI $[-2.6 \cdot 10^{-4}, 8.5 \cdot 10^{-5}]$, two-sided $p = 0.324$). Column (2) restricts the sample to subjects who "Never" or "Rarely" compensate for their emissions and confirms the null result.

Another approach to assess the relationship between an increase in cognitive uncertainty and a concave WTM involves examining whether individuals with greater uncertainty as emissions rise are more likely to exhibit a concave WTM curve. To do this, we define $CU_j(e)$ as the cognitive uncertainty of participant j at emission level e. The increase in cognitive uncertainty can then be quantified as:

$$\Delta_{CU} = CU_j(\bar{e}) - CU_j(0),$$

where $\bar{e}$ denotes the highest emission level for which the participant reported an uncensored WTM. We regress the concave-WTM dummy on Δ_{CU} and find that there is no statistically significant correlation between the two variables ($t(1082) = -0.431$, 95% CI

$[-0.004, 0.003]$, two-sided $p = 0.666$).[3]

Based on these two analyses, we conclude that there is insufficient support for the idea that cognitive uncertainty is a driver of concavity in the WTM curve.

Concave moral valuations. Another potential psychological channel that may explain a concave WTM relates to concave moral judgments. Individuals might perceive emitting 4kg of CO_2 as considerably morally worse than emitting 0kg, while the moral distinction between emitting 4kg and emitting 20kg might seem relatively minor. Such concave moral evaluations might, in turn, influence and shape the participants' WTM.

Let $\mu_j(e, k) \in \{1, 2, \ldots, 7\}$ denote the moral evaluation assigned by participant j to emitting $e$kg of CO_2 in exchange for £k, where the range spans from "morally very appropriate" (1) to "morally very inappropriate" (7). These evaluations are collected for each $e \in 4, 12, 20$ and $k \in 1, 5$. We aggregate these moral judgments by computing their average over the two values of k, yielding $m_j(e) = (\mu_j(e, 1) + \mu_j(e, 5))/2$. This composite measure is labelled as "Morality." Finally, we compute the variable ϕ_j as:

$$\phi_j = m_j(12) - \frac{m_j(4) + m_j(20)}{2}.$$

A positive value of ϕ_j indicates that the moral valuation of participant j is concave. The average ϕ is 0.107, which is positive and statistically significant ($t(1504) = 7.08$, 95% CI $[0.078, 0.137]$, two-sided $p < 0.001$), suggesting that moral judgments are indeed concave.

To investigate whether the presence of concave moral valuations is linked to a concave WTM, we regress the "concavity" dummy, which is equal to 1 if a participant exhibits a concave WTM, on the variable ϕ_j. The results presented in Supplementary Table 6 indicate that concavity in moral valuations has limited predictive power regarding the concavity of WTM.

A proposed interpretation based on the questions frame. We propose that the concavity of the WTM data reflects decreasing marginal disutility from emitting CO_2— emissions perceived as losses— rather than decreasing marginal utility from offsets, which are viewed as gains. This argument relies on the framing of our WTM questions.

In the WTM elicitation, participants are presented with a choice between "Option A", which entails no emissions and no monetary pay-off, and "Option B", which entails positive emissions and a monetary bonus. The reference point consists of no emissions. Participants were informed that emissions would be implemented as follows: 1) we set

[3] We follow the classification of individual WTM curve discussed in Section A.2.2. Note that the concave-WTM is a dummy variable taking a value of 1 when the WTM curve is characterized as either "concave" or "concave†" in the classification. In this analysis, we excluded participants whose WTM curves were classified as decreasing or non-monotonic. Additionally, participants with only an uncensored WTM value at $e = 0$ were also excluded, as Δ_{CU} is undefined for this subgroup.

Supplementary Table 6: Concavity of WTM and morality.

	(1)	(2)	(3)	(4)
Concavity of moral judgment (ϕ)	0.022	0.020	0.016	0.045
	(0.020)	(0.020)	(0.021)	(0.026)
Constant	0.212	0.202	0.233	0.153
	(0.121)	(0.125)	(0.128)	(0.143)
Controls	Yes	Yes	Yes	Yes
Observations	1,504	1,430	1,373	1,100
R^2	0.040	0.040	0.042	0.040

Notes: The dependent variable is `concave`, a dummy taking a value of 1 if the WTM curve is either "concave" or "concave†" based on the classification discussed in Section A.2.2. Samples are increasingly restrictive, from left to right. Column (1) includes all participants. Column (2) excludes participants who failed the attention check embedded in the moral judgment elicitation. Column (3) excludes participants whose $m_j(e)$ are decreasing in e. Column (4) excludes participants whose WTM curve is either decreasing or non-monotonic. List of control variables common to all regressions: age, gender (male, female, other), political affiliation (5 categories), education (6 categories), income (7 categories), and time needed to complete the first real effort task. Robust standard errors are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

aside a portion of the funds to be donated to Carbonfund.org, but 2) we would reduce the donation if they chose Option B. This procedure establishes the default as the donation to go through, with participants having the option to deviate from this default by choosing Option B. Moreover, the instructions asked the participants to indicate "the minimum bonus you require to accept the CO_2 emissions".

The questions are framed similarly to Willingness To Accept (WTA) elicitations, which ask participants to indicate the compensation required to engage in something they dislike—in our case, allowing the emission of CO_2. Given this framing, the concavity observed in our WTM data suggests that individuals have a marginally decreasing disutility from emitting CO_2: their aversion to emissions increases less at a rate less than proportional to the size of emission. This marginally decreasing disutility is surprising given that most economic models assume convex utility functions in losses.

The preceding discussion raises concerns regarding whether the concavity of WTM is driven by the way we ask the questions. Recent evidence, however, suggests this is not the case. Rodemeier (2023) successfully replicates the concavity result using a Willingness To Pay (WTP) framework, which asks individuals how much they are willing to pay out of their own pocket to offset emissions. In a WTP framework, offsets are considered as gains. The fact that the WTM exhibits concavity in both the loss and the gain domains suggests that decisions regarding emissions should be modelled with a reference-dependent model characterized by increasing insensitivities as outcomes move away from the reference point.

A.3 Study 3: Motivated Belief Experiment

A.3.1 Demographic characteristics

Supplementary Table 7: Demographic characteristics.

			Treatment		
		All	*Motivated*	*Unmotivated*	
Age bracket					
18-27	371	0.530	0.536	0.522	$\chi^2(4) = 2.64$
28-37	207	0.296	0.308	0.279	$p = 0.620$
38-47	67	0.096	0.084	0.111	
48-57	40	0.057	0.055	0.061	
58+	15	0.021	0.017	0.027	
Gender					
Male	354	0.496	0.493	0.502	$\chi^2(1) = 0.03$
Female	359	0.504	0.507	0.498	$p = 0.872$
Student					
No	417	0.590	0.587	0.593	$\chi^2(1) = 0.01$
Yes	290	0.410	0.413	0.407	$p = 0.932$
Education					
Less than high school	17	0.024	0.022	0.027	$\chi^2(4) = 0.24$
High school	190	0.273	0.274	0.270	$p = 0.993$
Some University	145	0.208	0.204	0.213	
Bachelor	231	0.331	0.334	0.328	
Postgraduate	114	0.164	0.165	0.162	
Income					
- £5,000	44	0.069	0.075	0.062	$\chi^2(8) = 3.88$
£5,000 - £15,000	146	0.230	0.232	0.228	$p = 0.868$
£15,000 - £30,000	183	0.289	0.290	0.287	
£30,000 - £45,000	103	0.162	0.149	0.180	
£45,000 - £60,000	77	0.121	0.130	0.110	
£60,000 - £75,000	48	0.076	0.066	0.088	
£75,000 - £90,000	22	0.035	0.036	0.033	
£90,000 - £105,000	7	0.011	0.014	0.007	
£105,000 -	4	0.006	0.008	0.004	

Notes: The table includes observations from the *Motivated* and the *Unmotivated* treatment ($N = 714$). Missing observations: 14 in age bracket, 1 in gender, 7 in student, 1 in education, 73 in income.

A.3.2 Dwell time in the attention task

In this analysis, we investigate whether the *Motivated* treatment has any impact on the time subjects spend completing the attention task. This variable is of particular importance because recent findings in economics and neuroscience suggest that dwell time on a piece of information causally increases the weight given to that information in subsequent decisions (Pärnamets et al., 2015; Amasino, Pace and van der Weele, 2021; Engelmann, Hirmas and van der Weele, 2021).

In the context of our study, we observed no substantial differences between the *Motivated* and *Unmotivated* treatments. If anything, the participants in the *Motivated* treatment tended to spend more time on the attention task. As shown in column (1) of Supplementary Table 8, in which we regress the time the participants spent on the task (in seconds) on a dummy for the *Motivated* treatment and on demographic controls, we observe that in the *Motivated* treatment spend 1.6 seconds more on the task. However, this difference is not statistically significant ($t(655) = 0.53$, 95% CI $[-4.29, 7.47]$, two-sided $p = 0.595$).

Column (2) confirms this finding, focusing solely on the 91% of participants who completed the task in less than 70 seconds. This subset represents individuals for whom we can be most confident that they did not take any breaks between receiving the information and providing their responses (the information was displayed for up to 60 seconds).

Supplementary Table 8: Time spent on the attention task.

	(1)	(2)
Motivated treatment	1.593	0.082
	(2.994)	(1.508)
Controls	Yes	Yes
Observations	694	632
R^2	0.067	0.079

Notes: The models are linear regressions with dependent variable the seconds the participants spent on the attention task. The models include the observations from the *Motivated* and the *Unmotivated* treatment with the *Unmotivated* as the baseline. The first column includes all the participants for which we recorded the demographic data, except one for which the program did not record the time spent on the task. The second column only includes the participants who spent less than 70 seconds to complete the attention task. Control variables: sex, age, student status, education (6 categories), frequency of car usage (5 categories), nationality (27 categories). Robust standard errors are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

A.4 Robustness Checks

A.4.1 Subsample by the level of understanding

Supplementary Tables 9 (for the Risk Aversion experiment) and 10 (for the Motivated Belief experiment) demonstrate the robustness of the null result even when excluding participants who made mistakes in the comprehension questions.

Supplementary Table 9: Effect of uncertainty on getting the computer code (Risk Aversion experiment).

	(1)	(2)	(3)	(4)
Information treatment	-0.005	0.005	-0.004	-0.004
	(0.036)	(0.027)	(0.025)	(0.024)
Constant	0.538^{*}	0.576^{***}	0.587^{***}	0.550^{***}
	(0.213)	(0.148)	(0.131)	(0.130)
# Mistakes	0	≤ 3	≤ 6	≤ 12
Controls	Yes	Yes	Yes	Yes
Observations	668	1,221	1,364	1,460
R^2	0.089	0.060	0.054	0.054

Notes: All models are linear regressions with dependent variable `product_bought`: an indicator variable equal to 1 if the participant bought the computer product. Models include the observations from the *Information* and the *Uncertainty* treatments of the Risk Aversion experiment. The *Uncertainty* treatment is the baseline. Column (1): participants who made no mistakes in the comprehension questions. Column (2): participants who made 3 or fewer mistakes. Column (3): participants who made 6 or fewer mistakes. Column (4): participants who made 12 or fewer mistakes. Control variables: gender (male, female, other), age, education (5 categories), political identification (5 categories), income (9 categories), time needed to complete the real effort task. Robust standard errors are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Supplementary Table 10: Effect of the *Motivated* treatment on purchasing and beliefs (Motivated Belief experiment).

A: Units	(1)	(2)	(3)	(4)
Motivated treatment	0.073	0.020	0.031	0.034
	(0.060)	(0.041)	(0.039)	(0.038)
# Mistakes	0	≤3	≤6	≤12
Controls	Yes	Yes	Yes	Yes
Observations	324	579	635	670
R^2	0.119	0.108	0.101	0.099
B: Beliefs	**(1)**	**(2)**	**(3)**	**(4)**
Motivated treatment	2.284	3.179	3.124	2.727
	(4.715)	(3.550)	(3.378)	(3.277)
# Mistakes	0	≤3	≤6	≤12
Controls	Yes	Yes	Yes	Yes
Observations	324	579	635	670
R^2	0.110	0.049	0.048	0.049

Notes: All the models are linear regressions. Dependent variable: (A) a dummy variable equal to 1 if the participant bought the virtual product, (B) beliefs about the size of the CO_2 emissions associated with the virtual product. The models include the observations from the *Motivated* and the *Unmotivated* treatment. The *Unmotivated* treatment is the baseline. Column (1): participants who made no mistakes in the comprehension questions. Column (2): participants who made 3 or fewer mistakes. Column (3): participants who made 6 or fewer mistakes. Column (4): participants who made 12 or fewer mistakes. Control variables: sex, age, student status, education (6 categories), frequency of car usage (5 categories), nationality (27 categories). Robust standard errors are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

A.4.2 Subsample by the level of trust

Supplementary Table 11 presents the results after excluding participants who expressed scepticism about our commitment to actually paying for the CO_2 offsets. The null results remain consistent in this robustness check.

Supplementary Table 11: The effect of trust in researchers.

	Risk Aversion	Motivated Belief	
	(1) Purchase	(2) Purchase	(3) Belief
Information treatment	−0.005		
	(0.027)		
Motivated treatment		0.028	2.552
		(0.041)	(3.523)
Controls	Yes	Yes	Yes
Observations	1,217	579	579
R^2	0.056	0.112	0.053

Notes: All models are linear regressions. Dependent variables: in column (1), an indicator variable equal to 1 if the participant bought the computer product; in column (2), an indicator variable equal to 1 if the participant bought the virtual product; in column (3), beliefs about the size of the CO_2 emissions associated with the virtual product. The first column includes the observations from the *Information* and the *Uncertainty* treatments of the Risk Aversion experiment. The second and third columns include the observations from the *Motivated* and the *Unmotivated* treatment from the Motivated Beliefs experiment. The *Unmotivated* treatment is the baseline. The *Uncertainty* treatment is the baseline. In all columns we exclude the participants who indicated low level of trust towards the CO_2 offsets taking place are excluded from this analysis. In the Risk Aversion experiment, these are the participants that answered with a 1, 2, or 3 to the question: "Do you trust that the researchers will indeed buy CO_2 offsets as described in the instructions?". Where 1 means "not at all" and 5 means "completely". In the Motivated Beliefs experiment these are the subjects that answered "No" rather than "Yes" to the same question. Control variables in column (1): gender (male, female, other), age, education (5 categories), political identification (5 categories), income (9 categories), time needed to complete the real effort task. Control variables in columns (2) and (3): sex, age, student status, education (6 categories), frequency of car usage (5 categories), nationality (27 categories). Robust standard errors are reported in parentheses. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

B Experimental Materials

B.1 Belief Elicitation in Study 1

Point estimates of the emission sizes. When asking about the CO_2 emissions generated by driving, we allowed the participants to express their guesses either in ounces or grams so they could use the more familiar unit of measure (Supplementary Figure 3).

For all the other products, we elicited the point estimates on a single interface that allowed the participants to go back and modify their previous answers easily. The order of the products on the interface was randomized at the individual level.

The 12 questions were graphically displayed (Supplementary Figure 4). The product in each question was represented by clip art, below which the name of the product and its size appeared. The participants could see which emissions were taken into account by the scientific estimate by hovering the mouse cursor on an info icon ℹ shown above each question. The list of products, their amount, and the emissions to be considered were all described in the instructions as well.

The participants' answers were summarized in an interactive box displayed at the bottom of the page. The box appeared as soon as a participant filled in the first question on the screen and it stayed visible until the moment the participant confirmed her answers. The "Confirm" button appeared inside the summary box to draw the participant's attention to the box itself.

The summary box graphically showed a participant's guesses on a line. Crucially, we designed the line to avoid any anchoring effects. No number appeared on it if the participant had not entered any guesses. Moreover, the scale of the line adjusted dynamically depending on the highest guess.

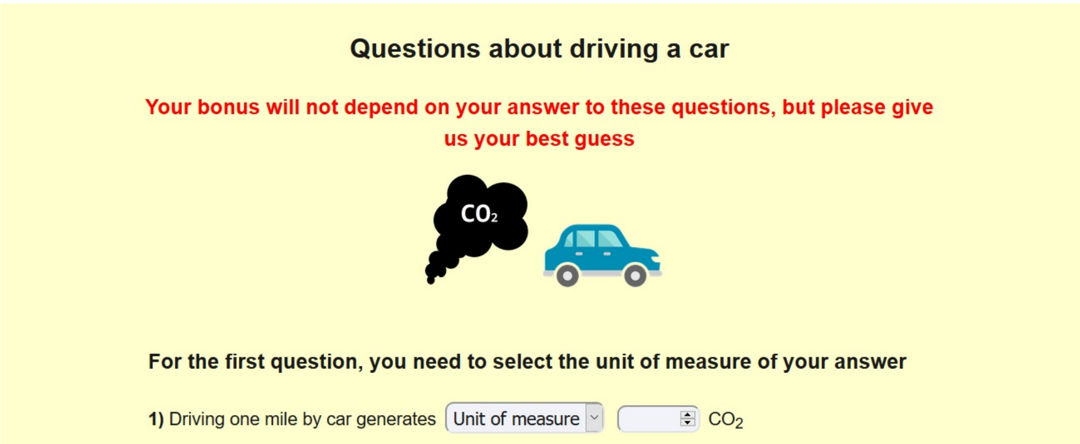
Belief distribution. The elicitation interface showed the name and the quantity of the product and reminded the participants of their point estimates. The participants could see which emissions were taken into account by the scientific estimate by hovering the mouse cursor on an info icon ℹ.

The interface displayed five bins for each question (Supplementary Figure 5). The participant's point estimate for the product, call it m, was taken as the midpoint of the central bin. The central bin covers numbers from $0.95m$ to $1.05m$. The two bins on both sides of the central bin cover numbers from $0.85m$ to $0.95m$ and from $1.05m$ to $1.15m$. Finally, the farthest two bins cover numbers below $0.85m$ and above $1.15m$, respectively.

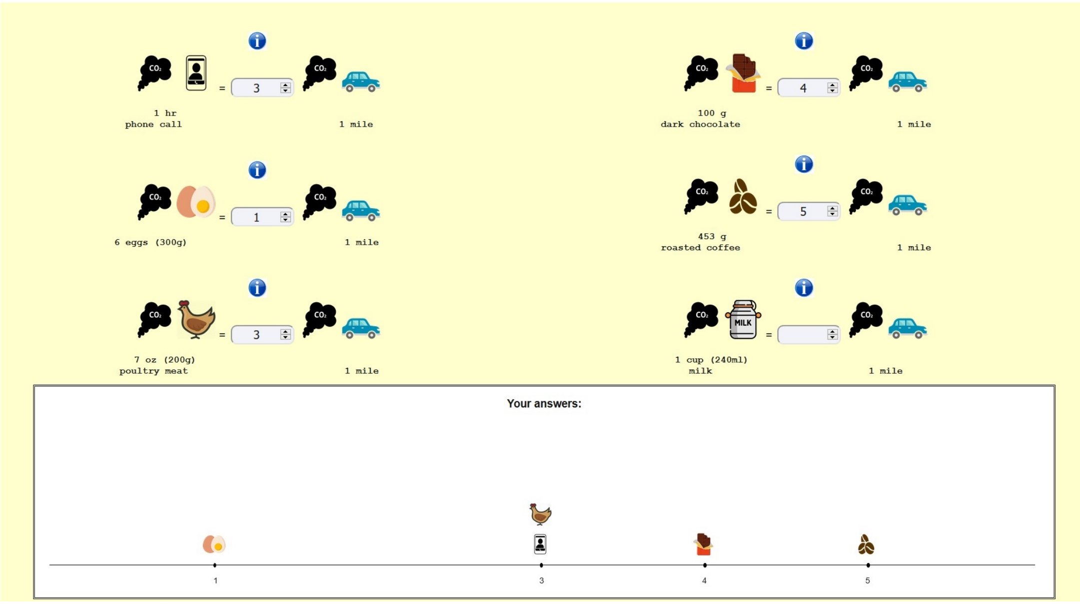
The interface showed a box containing the 20 balls the participants had to allocate among the bins. The participants could move the balls to a bin by (i) moving a slider below the bin, (ii) directly typing the number of balls they wanted to move in a text field below the bin, or (iii) clicking on the arrows next to the text field. The participants could move all the balls back to the box by pressing the button "Reset".

Supplementary Table 12: Comments on the calculation of CO_2 emissions.

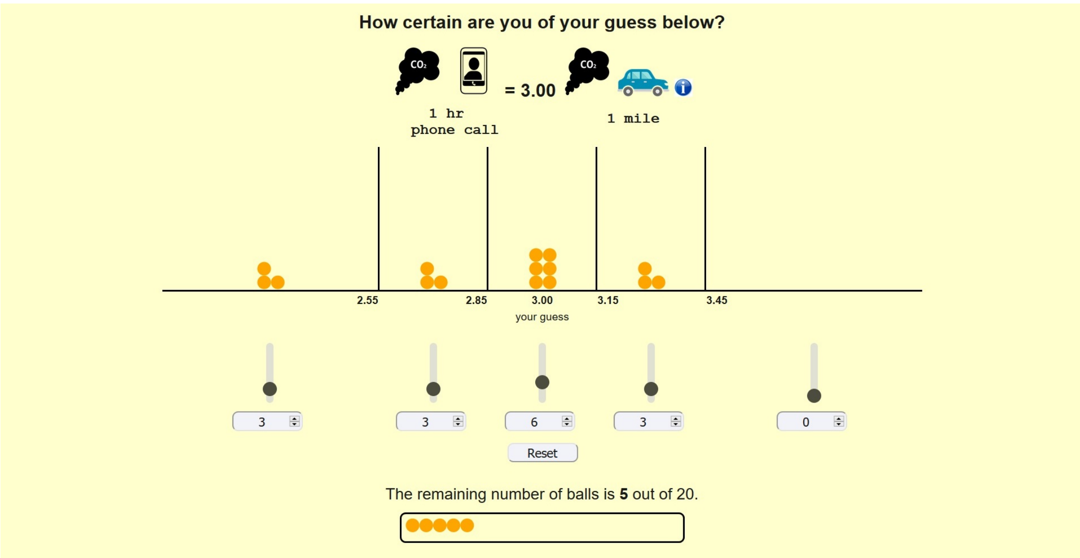
Product	Comment
Beer	It takes into account all the emissions, starting with the production and ending with the distribution of the products to the consumer.
Phone call	It takes into account the CO_2 emissions generated to operate the phone and the communication network.
Microwave	It takes into account only the emissions generated by the power plants that produce the energy used by the microwave.
Milk	It takes into account all the emissions, starting with the production and ending with the distribution of the products to the consumer.
Egg	It takes into account all the emissions, starting with the production and ending with the distribution of the products to the consumer.
Poultry meat	It takes into account all the emissions, starting with the production and ending with the distribution of the products to the consumer.
Shower	It takes into account the emissions generated by warming up the water and all the emissions connected to the water delivery and cleaning.
Chocolate	It takes into account all the emissions, starting with the production and ending with the distribution of the products to the consumer.
Coffee	It takes into account all the emissions, starting with the production and ending with the distribution of the products to the consumer.
Beef	It takes into account all the emissions, starting with the production and ending with the distribution of the products to the consumer.
Flight	It takes into account only the emissions generated by burning the plane fuel.
Gas heating	It is the average of the estimates of 10 different carbon footprint calculators.



Supplementary Figure 3: Beliefs about CO_2 emissions from driving one mile by car.



Supplementary Figure 4: Beliefs about CO_2 emissions from consumer products and activities.



Supplementary Figure 5: Belief distribution.

B.2 Work Task in Study 2

The task involved typing 15 strings, each consisting of 15 characters, in reverse order. The participants were required to transcribe these strings flawlessly to complete the task: any mistakes incurred resulted in an error message indicating the specific strings that required correction before they could proceed with the experiment. Participants saw a warning sign ATTENTION CHECK every 30 seconds, and upon its appearance, they had a 5-second window to click the I AM HERE button to confirm their active engagement with the task (right panel in Supplementary Figure 6). The participants knew that they would be excluded from the experiment if they failed more than 4 of these attention checks.

Supplementary Figure 6: Work task.

B.3 Attention Task in Study 3

The task involved finding the most frequently appearing number in a matrix of numbers. The matrix contained a total of 143 numbers, drawn from the set $\{0, 20, 40, 60, 80, 100, 120\}$ (Supplementary Figure 7). The number 60, the most frequently occurring, appeared 35 times, with 0 and 120 being the next most frequent, each appearing 26 times. All other numbers appeared 14 times each. Participants earned £0.10 if they answered 60.

0	40	60	120	0	120	100	60	60	0	120
40	20	120	120	60	80	80	20	20	60	120
0	0	0	80	100	20	60	60	20	20	40
40	0	120	60	100	100	80	80	60	60	60
120	120	0	100	20	0	60	0	20	100	60
100	80	0	60	120	120	60	40	60	80	20
0	0	40	0	0	60	120	60	60	100	40
120	60	80	120	80	120	60	40	0	60	120
60	0	60	100	60	120	20	100	40	0	80
60	60	60	120	60	0	120	40	0	60	120
120	20	0	40	120	100	0	60	100	20	60
0	40	120	60	0	120	20	80	40	60	60
80	100	80	20	80	100	120	0	40	0	120

Supplementary Figure 7: The matrix of numbers presented in the attention task.

B.4 Measures to Assure Data Quality

Instructions and comprehension questions. To ensure a comprehensive understanding of the essential elements of the instructions by our experimental subjects, we created slides that presented the information step by step. Most of the slides were accompanied by explanatory images to facilitate a more intuitive comprehension of the instructions. Furthermore, we organized the instructions into several sets. Following the completion of each set, participants were required to answer a series of comprehension questions. Importantly, we did not allow subjects to proceed with the experiment until they had successfully answered all questions in a given set. In total, participants were required to answer 7 questions related to the beliefs elicitation in Study 1, 19 questions (5 of which concern the work task and the computer code) in Study 2 and 15 in the *Motivated* and *Unmotivated* treatments in Study 3.

How we made sure that no bot completed the study. Our design incorporates two elements that mitigate the risk of an automated script ("bot") completing our experiment. Firstly, our instructions are not machine-readable. As a result, a computer script would need to provide random answers to the comprehension questions, leading to an exceptionally high number of answer attempts. This is further exacerbated by several comprehension questions requiring participants to input precise numerical values. We kept track of the number of attempts. In the Risk Aversion experiment, none of the participants who reached the end of the experiment needed more than 50 attempts to answer the comprehension questions, and 95% of them needed less than 22 attempts (the minimum number of attempts was 10). In the Motivated Belief experiment, no participant needed more than 65 attempts and 95% less than 14 attempts (the minimum number of attempts was 3).

Secondly, we incorporated three "honey-pots" within our two experiments. Those are questions hidden from human participants but discernible to a bot that reads the source code of the experimental program and identifies them as questions to be answered. We considered answering either honey-pot as sufficient evidence that the participant is a bot. We found no participant who answered these two "hidden" questions.

Combining the evidence from the number of attempts and the honey-pots, we can confidently conclude that no bot completed our experiment.

B.5 Invoice

Invoice 10622 for order 38268

Order Date: May 14, 2019

Billing Address

Davide Pace
Roetersstraat 11
1018WB Amsterdam
Netherlands

Shipping Address

Davide Pace
Roetersstraat 11
1018WB Amsterdam
Netherlands

Shipping Method

No shipping

SKU	Product	Quantity	Price
general-donation	General Donation **Name for e-certificate(s):** Davide Pace on behalf of the University of Amsterdam	1	$911.40
		Subtotal:	$911.40
		Payment Method:	Credit Card
		Total:	$911.40

Customer Details

- **Email:** d.d.pace@uva.nl

> *This donation is the result of participants decisions in the experiments "Decision Making 6-13" of the University of Amsterdam*

B.6 Preregistrations

ASPREDICTED

Wharton CREDIBILITY LAB
UNIVERSITY of PENNSYLVANIA

CONFIDENTIAL - FOR PEER-REVIEW ONLY

Curbing carbon: Does information about climate impact reduce emissions? (#109190)

Created: 10/11/2022 07:59 AM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

We have two main research questions.
1)Does quantitative information about CO2 emission reduce the acquisition of an emission-intensive product?
2)If so, is the effect of information explained by the declining marginal willingness to pay for mitigating CO2 emissions?

3) Describe the key dependent variable(s) specifying how they will be measured.

Participants can buy a product that results in an uncertain amount of CO2 emissions. The key dependent variable is the consumption of the product (1: if the participant buys; 0: otherwise). In addition, we elicit willingness to pay to mitigate (WTM) CO2 emissions of various sizes, to explain the reaction to information (see below). For each emission size, we elicit WTM twice to rule out measurement error (see point 8).

4) How many and which conditions will participants be assigned to?

There are 2 conditions. Each participant will participate in one condition.

Info Treatment: participants know exactly the size of the CO2 emissions of the product (4 kg).
Uncertainty Treatment: participants know CO2 emissions associated with the product are 0 kg with a probability of 0.4, 4 kg with a probability of 0.2, or 8 kg with a probability of 0.4.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Hypothesis 1: Participants in the Info treatment buy fewer units of the good than participants in the Uncertainty treatment.
We test this hypothesis by means of a Fisher's exact test. We will also perform OLS regressions to control for subject characteristics.

Hypothesis 2: The effect of information increases with higher concavity of WTM (declining marginal willingness to mitigate).
For each individual, we compute a measure of concavity as the WTM 4 kg of CO2 emissions minus the average of the WTM 0 kg and 8 kg. We then regress the buying decision on the Info treatment dummy, our concavity measure, and the interaction of this concavity measure and the Info treatment dummy. In the regression, we control for subject characteristics and the average WTM across all three relevant levels. For this analysis, we exclude subjects whose WTM is top censored for at least one emission amount in the interval [0kg; 8kg] and whose WTM to mitigate is not weakly increasing, i.e., who don't satisfy the law of demand.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We exclude observations only if we have evidence that the respondent is not a human (we will run the experiment online). In addition, we will run robustness checks where we exclude participants who indicate that they don't believe we will actually implement the CO2 emissions.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

We will collect 1500 complete observations on prolific. The sample size in the analysis might be larger because we will include subjects who have completed the demographic questionnaire but have not finished the experiment.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

We will provide various robustness checks. First, we will do robustness using an alternative specification of the concavity measure, dividing our original measure by (WTM(8 kg) - WTM(0 kg))/2 to correct for the increase in the WTM. We will also look at the treatment differences for the subsets of participants whose willingness to mitigate is more or less concave than the median.
Second, to combat measurement error in our concavity measure, we will do a second, unincentivized elicitation of WTM. Following the approach in Gillen et al. (Journal of Political Economy, 2019) we will instrument one measure with the other to eliminate any variation that is not common to both measures.

Taxes, beliefs, and the demand for goods with negative externalities (#23181)

Created: 05/08/2019 05:11 AM (PT)
Shared: 07/10/2019 06:08 AM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Policy-makers have two main instruments to change consumer demand for goods that produce a negative externality. They can change the price using taxes and subsidies, or they can provide information about the externality. We have three main research questions
1) What is the effect of prices and information on consumption?
2) Do higher prices and information reduce self-serving beliefs about the externality?
3) Does information reduce the effect of price policies by eliminating self-serving beliefs?

3) Describe the key dependent variable(s) specifying how they will be measured.
Participants can buy a good that results in an uncertain externality (CO_2 emissions). Thus, the two key dependent variables are:
1) Consumption: this is a binary variable (1: if the participant buys the good; 0: otherwise),
2) Beliefs: this is an integer between 0 and 120. It represents participants' beliefs about the magnitude of the externality they may produce (measured as the equivalent of liters of gasoline).

4) How many and which conditions will participants be assigned to?
Three treatments differ in the way participants are informed about the size of the externality:
Info Treatment: participants know exactly the size of the externality.
Motivated Treatment: The answer to a puzzle gives participants the magnitude of the externality. Participants solve the puzzle after knowing the relation with the externality.
Unmotivated Treatment: The answer to a puzzle indicates to participants the magnitude of the externality. Participants solve the puzzle before knowing the relation with the externality

In three cross-cutting conditions, we vary the price of the good in the set {0.25, 1, 1.75}, measured in British pounds.

Overall, this leads to 9 conditions, all subjects participate only in one condition.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
Hypothesis 1: Participants in the Info treatment buy fewer units of the good than participants in the Motivated treatment.

We test this hypothesis by means of a Fisher's exact test, pooling all price levels. We will perform regressions to control for subject characteristics.

Hypothesis 2: Participants in the Unmotivated treatment have higher beliefs and buy fewer units of the good than participants in the Motivated treatment.

We compare this with a non-parametric rank sum test (beliefs) and Fisther exact test (consumption), pooling all price levels. We will perform regressions to control for subject characteristics.

Hypothesis 3: In the Motivated treatment, demand is decreasing in prices.

We test this in a linear regression, using a one-sided t-test.

Hypothesis 4: In the Motivated treatment, beliefs are increasing in prices.

We test hypothesis using a linear regression and a one-sided t-test.

Hypothesis 5: Conditional on hypothesis 4 being confirmed, price-sensitivity of demand in the Info treatment is lower than that in the Motivated treatment.

We test this in a linear probability model, using a one sided t-test. Note that if the relationship between beliefs and prices is different than in

References

Amasino, Dianna, Davide Pace, and Joël J. van der Weele. 2021. "Fair Shares and Selective Attention." Tinbergen Institute Discussion Paper 2021-066/I.

Engelmann, Jan, Alejandro Hirmas, and Joël J. van der Weele. 2021. "Top Down or Bottom Up? Disentangling the Channels of Attention in Risky Choice." Tinbergen Institute Discussion Paper 2021-031/I.

Enke, Benjamin, and Thomas Graeber. 2023. "Cognitive uncertainty." *Quarterly Journal of Economics*, 138(4): 2021–2067.

Pärnamets, Philip, Petter Johansson, Lars Hall, Christian Balkenius, Michael J. Spivey, and Daniel C. Richardson. 2015. "Biasing Moral Decisions by Exploiting the Dynamics of Eye Gaze." *Proceedings of the National Academy of Sciences*, 112(13): 4170–4175.

Rodemeier, Matthias. 2023. "Willingness to Pay for Carbon Mitigation: Field Evidence from the Market for Carbon Offsets." IZA Discussion Paper 15939.

US Census Bureau. 2022. "National Population by Characteristics: 2020-2022." United States Census Bureau. `https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-detail.html`.