TI 2020-059/I
Tinbergen Institute Discussion Paper

Curbing Carbon: An Experiment on Uncertainty and Information about CO2 emissions

Davide Pace[1]
Joël van der Weele[1]

[1] University of Amsterdam, Tinbergen Institute

Curbing Carbon: An Experiment on Uncertainty and Information about CO_2 emissions *

Davide D. Pace[1] and Joël van der Weele[1]

[1]University of Amsterdam, Center for Experimental Economics and political Decision making (CREED), and Tinbergen Institute

September 10, 2020

Abstract

We investigate how consumers respond to uncertainty about CO_2 emission size. In an incentivized online experiment, participants can acquire a valuable good that emits an unknown amount of CO_2. We find that beliefs about emission size are strongly predictive of purchases, even exceeding the effect of substantial changes in the price of the good. Moreover, information that makes beliefs more precise causes a 26% reduction in overall emissions, even though average beliefs are unchanged. The reduction occurs as the marginal willingness to pay for emission reduction declines with emission size, so people who are too optimistic about emissions are more responsive to information. We also test for the formation of self-serving beliefs. Contrary to theories of motivated reasoning, increasing the surplus from buying the product does not change patterns of attention or belief formation about emissions. Overall, the results suggest that information about CO_2 impact can be an important policy lever, and that willingness-to-pay for emission reductions should take into account the size of emissions.

Keywords: CO_2 emissions, sustainable consumption, economic experiments.

JEL Codes: C91, D81, Q54

*We thank the NWO for funding in the context of VIDI grant 452-17-004. We are indebted to Peter Schwardmann for his early contributions to the project and his extensive comments that significantly improved the quality of this paper. We are grateful to Marta Serra-Garcia, Nicolas Treich, Linda Thunström, and numerous seminar participants for helpful comments.

1 Introduction

How to combat climate change and move away from carbon-intensive production is a key challenge for our societies. Carbon pricing is arguably the most important tool to reduce carbon emissions (Nordhaus, 2019; Stiglitz et al., 2017), but the financial sacrifices it requires from voters makes it hard to implement politically. For instance, in France, an increase in carbon taxes on fossil fuels sprouted a protest movement known as the yellow vests; in Canada, some states are fighting a legal battle against a federally imposed carbon tax. An alternative, and cheaper approach is to increase voters' voluntary commitment through public information campaigns or (mandatory) product labeling. These measures may be effective if citizens care about protecting the environment but have wrong or imprecise beliefs about the actual emission impact of their actions. However, little is known about how beliefs and information about emissions affects purchasing decisions.

In this paper, we test a number of novel hypotheses related to the importance of consumer beliefs and the role of information about CO_2 emission size. To this end, we conduct an incentivized experiment with 1000 participants on Prolific, an online platform. We offer participants a trade-off between increasing their earnings and emitting CO_2, a decision that is framed as the acquisition of a valuable but polluting virtual product. Across treatments, we vary the information available on the magnitude of these emissions. In one condition, we communicate the exact size of CO_2 emissions – 140 kg of CO_2, which is equivalent to burning 60 liters of gasoline and costs £1.07 to offset via Carbonfund.org. In other conditions, participants could form beliefs by engaging in an attentional task, mimicking efforts like reading product labels or engaging in online search. To compare and provide a benchmark for the importance of beliefs and information with those of monetary incentives, we conduct cross-randomized treatments in which we vary the prices of the polluting product.

We find that beliefs about the CO_2 emissions associated with the virtual product matter, as they are highly predictive of purchases. For instance, subjects who report the highest emission beliefs (at most twice the true amount) only buy the product 25 percent of the time, while those with the lowest beliefs buy it 67 percent of the time. By comparison, the difference in purchases from a seven-fold price increase (from £0.25 to £1.75) is 23 percentage points. The intrinsic motivation to avoid emissions is stronger among women than among men.

Moreover, we find that providing unambiguous information about the size of emissions leads to a 26% reduction in overall emissions. This effect does not result from a shift in average beliefs, because participants in the uncertainty conditions also have correct beliefs on average. Instead, the effect of information results from a non-linear relation between beliefs and buying behavior: correcting underestimates of beliefs has a larger effect on demand than correcting overestimates. In an additional experiment, we show that the origin of this non-linearity is a decreasing marginal willingness to pay (WTP) for emission reductions. Thus, information works because people derive more disutility from emitting 140 kg of CO_2 for sure, than from doing so in expectation.

These results indicate that consumer's beliefs about emission size matter, at least when emissions are made salient, as in our experiment. This suggests that reporting product emissions and correcting misperceptions can be an effective strategy to reduce carbon emissions when consumers underestimate emission size, or when there is substantial heterogeneity in beliefs. The finding that WTP for CO_2 reduction is declining in emission size also means that WTP estimates in the literature, which focus on a single amount, are only partially informative.

We also look at the interaction of prices and beliefs. In particular, we test theories of "motivated beliefs" predicting that changes in consumer surplus may affect consumers' perceptions about the size of the emissions. If "dirty" products are cheap, consumers might self-servingly reduce attention to emission size or be more inclined to "see" convenient information that justifies buying the product. This may create an additional channel through which prices impact demand, and increases price elasticity compared to a world of full information (Hestermann et al., 2018). We test this idea by comparing the belief formation process of participants who know the surplus they can obtain from acquiring the polluting product with participants who do not know. We do not find evidence for the formation of self-serving beliefs in our setting, nor do we see an influence of the consumer surplus either on beliefs or on the effort on the attentional task. We also find no evidence that prices and information interact through other channels.

As we detail more extensively in the next section, these results contribute to the literature in three main ways. First, the evidence that reducing uncertainty about externalities decreases emissions adds to a literature on prosocial decisions under uncertainty. We document a novel mechanism for this effect, stemming from the concavity of individual WTP to avoid CO_2 emissions. Second, our findings add to a literature on ethical consumption and moral behavior in markets, as we demonstrate a substantial intrinsic motivation for avoiding emissions and show how WTP to do so depends on the size of CO_2 emissions. Finally, we contribute to a literature on the use of "moral wiggle room" and motivated reasoning in economic decisions: Our null result in an applied online setting raises questions about the scope of motivated cognition in ethical consumption.

2 Related literature

Our work contributes to several lines of literature. The first investigates the role of uncertainty in ethical decisions and social dilemmas, and has yielded contradictory results. Butera and List (2017) finds that uncertainty about the returns of contributing to a public good boosts contributions. By contrast, Exley (2015) finds that uncertainty reduces donations to the Red Cross. In line with this, Kappes et al. (2018) shows that uncertainty about the probability of harming others increases selfish behavior. However, they also find that uncertainty about the magnitude of the damage produced on a passive party induces more prosocial choices. This last result contrasts with ours, in that we find that individuals are more likely to act selfishly if there is uncertainty about the size of the emissions that they might produce.

In the field, several studies have looked at the role of feedback or information on sustainable choices. Feedback takes various forms, ranging from the behavior of peers to the environmental or personal benefits of sustainable behavior. For instance, Rodemeier and Löschel (2020) look at the impact of information on the personal benefits of energy conservation, and find that it increases energy use, as the information corrects initial overestimates in the sample. To our knowledge, few (field) experiments have considered CO_2 emissions. As an exception, Fang et al. (2020) show that home energy reports containing information about CO_2 emissions of shower use affect beliefs about impact, but do not change behavior unless combined with immediate, real-time feedback. While our paper studies more abstract decisions, we view it as complementary to such field interventions: our approach allows to carefully disentangle the role of beliefs and preferences in considering the role of information, and show how decreasing marginal WTP affects information interventions.

The second literature considers consumers' WTP for ethical products when the emission size is known. In a prominent publication, Bartling et al. (2014) report results from an experimental market where some products have negative externalities on other subjects. They find that both buyers and sellers are willing to pay not to produce emissions. A separate line of research closer to our application investigates the WTP to reduce CO_2 emissions, using mostly unincentivized surveys. A review by Nemet and Johnson (2010) finds a stated WTP between \$22 and \$437 per household annually. This range is wide and likely influenced by the hypothetical nature of the questions. Two more recent studies, however, do incentivize participants to reveal their true WTP. Löschel et al. (2013) find a median WTP to buy emissions offsets for one ton of CO_2 of 0 € and an average of 12 €. Diederich and Goeschl (2014), instead, finds a median WTP of 0.30 € and a mean of 6.30 €. Both these studies elicit the WTP to offset a fixed CO_2 amount.

We contribute to this literature by showing, in an incentivized experiment, that consumers' WTP to avoid CO_2 emissions is increasing but concave in the size of the emissions. This finding mirrors WTP in other prosocial decisions: Schumacher et al. (2017) show that willingness to take prosocial acts is highly concave in the number of beneficiaries. Thus, taking into account emission size is important when evaluating WTP in research on climate change as well as sustainability more generally.

The third relevant literature focuses on the formation of *motivated beliefs*, i.e., beliefs that people hold because they directly derive utility from them, and not because they are accurate or based on evidence. In particular, people may form distorted beliefs to keep a positive view of themselves even if they behave selfishly or immorally, see surveys in Bénabou and Tirole (2016) and Gino et al. (2016). Moreover, individuals sometimes purposely avoid information that might compel them to sacrifice part of their pay-off to act generously (Dana et al., 2007; Ehrich and Irwin, 2005; Grossman and Van der Weele, 2017), and some are even willing to pay to do so (Serra-Garcia and Szech, 2019). Exley (2015) shows that people use the uncertainty about the outcome of a charitable donation as an excuse not to give. Thunström et al. (2014) demonstrate strategic ignorance about CO_2 emissions in the context of a hypothetical long-distance flight.

Of particular relevance for our setting are studies on the interactions between information avoidance, belief formation, and economic incentives. Recent theoretical contributions include Schwardmann (2019), who shows that health risk prevention and motivated beliefs are strategic complements. In a moral context, Hestermann et al. (2018) theorize that individuals are more likely to engage in denial of an externality if they earn a higher pay-off by producing it, thus increasing the elasticity of demand. On the empirical side, Festinger and Carlsmith (1959) show that participants exhibit stronger motivated beliefs about the pleasantness of a task if they have incentives to lie about it, although the effect declines for higher incentives. Recent empirical papers investigate the relationship between the price of a prosocial action and the incidence of strategic ignorance (Feiler, 2014; van der Weele, 2013; Momsen and Ohndorf, 2019b,a). Perhaps closest to this paper, Momsen and Ohndorf (2019b) find that participants are less likely to reveal information about CO_2 emissions when they can gain less from an action that could produce those emissions. Finally, Ambuehl (2016) uses a strategy similar to ours to investigate the impact of incentives on information acquisition and belief formation in the context of aversive consumption decisions, and concludes that incentives rationally skew information acquisition.

We go beyond these studies by looking systematically at the effect of price changes on beliefs about externalities, rather than information avoidance. In particular, we are the first to investigate the hypothesis that the formation of motivated beliefs amplifies the effect of price incentives. Our null result raises questions on the applicability of laboratory results on motivated cognition to climate change contexts that future research should address (see also Lind et al., 2019).[1]

[1]The focus on motivated beliefs sets us apart from a small literature in development economics that studies the interaction of monetary incentives and information in shaping consumer demand for health products. Ashraf et al. (2013) finds that information reduces the elasticity of demand for water purification products. Instead, Dupas (2009) and Meredith et al. (2013) find that the elasticity of demand for, respectively, bed nets and preventive health care for children is not affected by information provision. Our design eliminates several mechanisms that can influence the effect of information on the elasticity of demand in these studies, such as liquidity constraint, present bias, and the signaling value of price.

3 Design

In the experiment, we offer participants to buy a single unit of a virtual product.[2] If participants decide to purchase the product, they increase their pay-off by the product value of £2, minus the price at which we, the experimenters, offer it. Purchasing the product results in the emission of CO_2 into the atmosphere. We implement these CO_2 emissions via the preparation of a donation to Carbonfund.org, an NGO that finances projects to reduce CO_2 emissions and to capture CO_2 from the atmosphere. If participants buy the product, we will cancel the donation, thus increasing the carbon in the atmosphere.

We framed the experiment as a market interaction, using words like "virtual product" and "price" to make it closer to the real-life purchasing situation. After explaining the use of donations to Carbonfund.org, we refer to the consequences of buying the product as an "increase in CO_2 emissions". This wording is both truthful and makes the environmental consequences more salient and closer to everyday-life applications. Because CO_2 emissions expressed in weight units - like grams or tons - are very abstract to participants, we inform subjects that buying the good results in emitting as much CO_2 as burning 60 liters of gasoline. This emission size had offset costs of £1.07, commensurate to the other payments in the experiments.[3]

As we describe next, we implemented three price treatments and three information treatments as well as their interactions, resulting in nine treatments in total. Each participant took part in one treatment only. The link to the full instructions and comprehension questions is in Appendix D. We preregistered the experimental design and the hypotheses on AsPredicted.org; the document containing the preregistration is in Appendix E.

[2]This product is virtual in the sense that it exists only inside the experiment, it is not a physical product nor a service. However, the product is valuable to the participant because if he or she "buys" it, her pay-off from the experiment increases.

[3]On the basis of a report from the US Environmental Protection Agency (EPA (2005)), we calculate that burning 60 liters of gasoline produces 140kg of CO_2 emissions. Carbonfund.org offsets 1 metric ton of CO_2 per every \$10 (or £7.9) it receives in donations, so offsetting the products' emissions cost £1.07.

3.1 Information treatments

We are interested in the formation and impact of beliefs about the CO_2 emissions associated with the product, as well as the role of information. To study this, three treatments varied the nature of uncertainty about the size of the emissions.

In the *Info* treatment, participants receive complete information, i.e., that emissions were equivalent to burning 60 liters of gasoline. This treatment allows us to study consumer demand under full information about the magnitude of emissions. In the other two treatments, called the *Motivated* and *Unmotivated* treatments, there was uncertainty about the size of the emissions. Participants could reduce uncertainty by engaging in an attentional task, designed to mimic real-life consumption situations such as reading product labels or doing an online search for information. The task consisted of a matrix of numbers; the number that appeared most frequently indicated the size of the emissions, measured in terms of the CO_2 emissions generated by burning a liter of gasoline. We adapted this task from Ambuehl (2016), who shows that the information gathering strategy in this task is influenced by incentives for subsequent decisions.[4] Subjects had up to one minute to engage in the task, after which we elicited participants' beliefs by asking them which number appeared most often in the table. A correct answer was rewarded with a bonus of £0.10, which incentivizes subjects to report the mode of their belief distribution (Schlag et al., 2015).

The two uncertainty treatments differ in the order in which we presented the attentional task and the information about the emission size. 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 their 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. Thus, participants knew that the correct answer to the task indicated the magnitude of the CO_2 emissions, as well as the surplus they could obtain from the product. This treatment allows

[4]The task can be found in the online instructions. In total, there are 143 numbers in the table. The numbers belong to the set $N = \{0, 20, 40, 60, 80, 100, 120\}$, with 60, the most common number, appearing 35 times. 0 and 120 are the numbers that appear the most often after 60, both of them appearing 26 times. All the other numbers appear 14 times.

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 purchase of the product.

3.2 Price treatments

We also investigate how purchases depend on prices, for two main reasons. First, the effect of prices provides a natural quantitative benchmark for the effects of beliefs and information, as price incentives are a primary tool of policymakers and often discussed in the context of reducing carbon emissions.

Second, we can test how varying the consumer surplus affects the formation of beliefs and information gathering. Next to a standard substitution effect, prices may change the motivation to form the self-serving beliefs that the emissions are low. Theories of motivated cognition and cognitive dissonance provide predictions on the direction of the interaction between prices and beliefs. Hestermann et al. (2018) argue that as price decreases the temptation to buy goes up, and with it the demand for moral justification, which is achieved by developing self-serving beliefs. This self-deception resulting from lower prices increases buying beyond the direct substitution effect, and therefore raises the price elasticity of the good. We show in Appendix A how this effect can be identified empirically in the context of our study. Alternatively, Festinger and Carlsmith (1959) provides evidence that a very high surplus may in itself provide enough justification to buy, as being swayed by high incentives is not a strong signal of deviance. According to this logic, one would expect the highest demand for justification to occur for high or intermediate prices.

We implemented three price treatments which varied the price of the product between a low price (£0.25), a medium price (£1), and a high price (£1.75). We inform participants that the price has been randomly assigned to them and is uninformative about the size of emissions. We make sure participants understand this fact asking them a comprehension question on the topic.

3.3 Sample and data collection

We recruited 1154 participants using Prolific.ac, an online platform, between 9^{th} and 11^{th} May 2019. Of those, 1000 participants completed the experiment. Between the participants that concluded the experiment, 50% are 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 (33.6%), Poland (14.5%), and Portugal (13.1%). Subjects earned a fixed reward of £1.6 plus a bonus payment depending on their decisions. On average, they earned £2.04, and they took less than 13 minutes to complete the tasks. We received information on participants' gender, age, and nationality directly from Prolific.ac.[5] Following participants' decisions, we donated $911.40 to Carbonfund.org to offset CO_2 emissions, resulting in a reduction of over 9 metric tons of CO_2 in the atmosphere.

3.4 Implementation details

To ensure that the experimental subjects understood all the essential elements of the instructions, we used slides that displayed the instructions step by step. On almost every slide, an explanatory image accompanied the written text to make the instructions easier to digest. Besides, we divided the instructions into 3 or 4 sets, depending on the treatment. After each set, we asked participants to answer several comprehension questions. We did not allow subjects to continue with the experiment until they answered all the questions of each set correctly. In total, participants had to answer between 12 and 15 comprehension questions depending on the treatment. At the end of the experiment, we collected demographic information using a survey.

We took several steps to convince participants that the CO_2 emissions associated with the purchase of the product would actually be implemented. We stressed the role of the no-deception policy in obtaining ethical approval for the experiment. Moreover, we promised participants to send them the invoice of the donation to Carbonfund.org.[6]

Finally, we stressed the anonymity of participants' decisions. Experimenters have access

[5]Researchers have recently voiced concern that bots, instead of humans, might complete experiments on Amazon MTurk, another online platform routinely used to run online experiments. To the best of our knowledge, no concern about bots has been raised regarding Prolific.ac to date. However, we took two measures to further minimize the risk that undetected bots completed our experiment. First, we implemented two honey-pots, those 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. The interested reader can find more details in Appendix B.

[6]The invoice of the donation is in Appendix F and it was sent to participants on 14^{th} May 2019.

only to the Prolific identification number, a unique number that identifies participants on the platform. The experimenters do not receive, nor are allowed to ask any information that could identify participants outside the platform. To ensure that participants are aware of this policy, we reminded participants that they are anonymous and asked a comprehension question on the topic.

3.5 Follow-up experiment

We conducted a follow-up experiment to delve into the motives behind the behavior in the main experiment. As we explain in more detail below, this experiment was designed to understand how the WTP to avoid emissions depends on the emission size. In the follow-up experiment, we recruited an additional 134 participants on Prolific on the 18th October and 19th November 2019, and elicited WTP to avoid CO_2 emissions. WTP was elicited with a BDM mechanism (Becker et al., 1964). For each participant, we elicit the WTP to avoid CO_2 emissions equivalent to burning the equivalent to 5, 30, 60, 120, 180 liters of gasoline, using a multiple price list.[7] The emission size was presented in a random order to the participants and the multiple price list ranged from £0.00 to £7.00 in £0.50 increments.

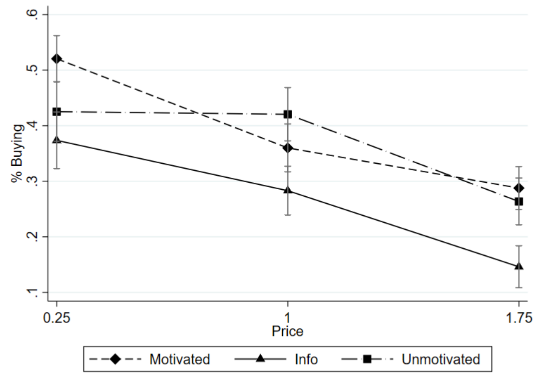
4 Results

We start with an overview of some general patterns of buying behavior across all treatments. Below, we statistically analyze the main treatment effects in more detail.

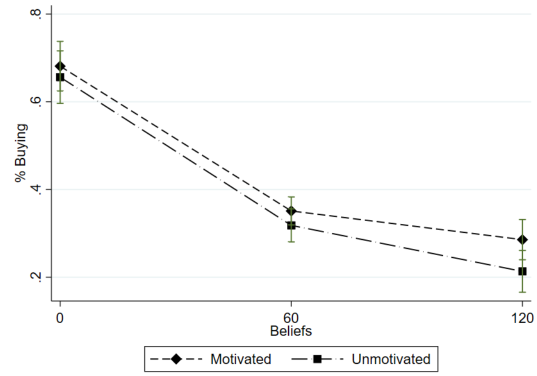
As Figure 1a shows, higher prices decrease buying in all treatments, as one would expect. Furthermore, participants display a substantial motivation to avoid CO_2 emissions. Pooling across all treatments, a minority of 34.9% of the participants bought the experimental product. It is striking that buying behavior never exceeds 52% and it is as low as 15% in some treatments, even though consumer surplus is substantial (£1.75) relative to a) the show-up fee (£1.6), b) the minimum wage enforced by Prolific (less than £1.10 for an experiment like ours), and c) the cost of compensating the CO_2 emissions (£1.07).

[7]The two sessions were identical except for one design feature: in the second session, we forced a single switching point in the multiple price list, in the first we didn't.

Figure 1: Overview of buying behavior by beliefs and prices.



(a) Average number of units bought in each treatment by price level. x-axis: price; y-axis: percentage of participant buying the product; confidence bars ± 1 standard deviation.



(b) Average number of units bought for the most common beliefs levels. x-axis: beliefs; y-axis: percentage of participant buying the product; confidence bars ± 1 standard deviation.

Figure 1b breaks down buying behavior by beliefs in the two uncertainty treatments. We focus on the three most prominent belief levels, i.e., the beliefs that emissions are most likely to be either 0, 60, or 120. Due to the nature of our attentional task, 95% of participants indicated one of these three beliefs.

Several patterns jump out. First, buying decreases when subjects believe the emissions are larger. This is a sizeable effect: Table 1 shows that between the lowest and highest beliefs, buying drops by 43 percentage points. By comparison, the buying difference between the extreme price levels prices is 23 percentage points. In our setting, beliefs about the size of the emissions thus appear to be an important motivation of participants. Second, the effect of beliefs is non-linear: buying decreases by more than 50% when beliefs about the emissions increase from 0 to 60 liters worth of gasoline, but when they increase further to 120 liters, the decrease in buying is only 33%. This observation will play an important role in our analysis below.

Table 1: Number of participants that buy the product per each price and beliefs level

			Price			
			£0.25	**£1**	**£1.75**	Total
	0	# Buying	37	34	18	89
		Observations	44	46	43	133
		% Buying	**84.1%**	**74.0%**	**41.9%**	**66.9%**
Beliefs	**60**	# Buying	88	68	49	205
		Observations	212	228	225	665
		% Buying	**41.5%**	**30.0%**	**21.8%**	**30.8%**
	120	# Buying	20	15	9	44
		Observations	58	57	58	173
		% Buying	**34.5%**	**26.3%**	**15.5%**	**25.4%**
	Total	# Buying	145	117	76	338
		Observations	314	331	326	971
		% Buying	**46.1%**	**35.3%**	**23.3%**	**34.8%**

Number of participants that buy the product per each price and beliefs level. Note that the drop in buying due to the change in price is smaller than the drop in buying due the change in beliefs.

4.1 The effect of information

We now turn to analyze the effect of providing information about emissions in more detail. To this end, we compare behavior in the *Info* and the *Unmotivated* treatment. We focus on the *Unmotivated* treatment as beliefs cannot be conditioned on prices. All qualitative results go through when we instead compare the *Info* with the *Motivated* treatment (See Appendix C.1).

Figure 1a shows that, for any given price level, buying is lower in the *Info* treatment than in both other treatments. Pooling the data across price levels, a Fisher's exact test corroborates this finding: we reject the null hypothesis of equal buying behavior between the *Info* and *Unmotivated* treatment (two-sided $p = 0.013$). To confirm that this result is

Table 2: Comparison between Info and Unmotivated treatment

	(1)	(2)
	Units	Units
Info Treatment	-0.095*	-0.032
	[-0.17,-0.018]	[-0.13,0.064]
Beliefs=60		-0.36***
		[-0.51,-0.22]
Beliefs=120		-0.44***
		[-0.59,-0.28]
Dummies for beliefs		✓
Controls	✓	✓
Observations	568	568

95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All the models are linear probability models with robust standard errors. Data from the Unmotivated and the Info treatment only. Baseline: *Unmotivated treatment*. Dependant variable: purchasing decision (equal to 1 if the participant purchased the product). Controls list: sex, age, student status, education (6 categories), nationality (27 categories). Dummies for Beliefs: 6 dummies corresponding to beliefs of 20, 40, 60, 80, 100, 120. Dummy corresponding to beliefs of 0 excluded because of collinearity.

not driven by differences in subject characteristics, Table 2 reports the results of an OLS regression where we control for sex, age, student status, education, and nationality. The results in column (1) indicate that being in the *Info* treatment (relative to the baseline of the *Unmotivated* treatment) decreases the likelihood that a participant buys the good by 9.5 percentage points (26 percent). Note that the difference in buying behavior is unlikely to be explained by participants in the *Unmotivated* treatment having wrong beliefs on average. In fact the average belief in the *Unmotivated* treatment is 62.7, not significantly different from 60 ($p = 0.25$).

To study the role of beliefs in the effect of information, we include dummies for the different belief levels in Column (2) of Table 2.[8] The dummy coefficients demonstrate the non-linear effects of beliefs on buying noted above. Moving from believing that emissions are most likely to be 0 to believing that they are most likely equal to 60, reduces the probability of buying the product by 36 percentage points ($p < 0.001$). Moving from belief 60 to 120 decreases the probability of buying by only 8 additional percentage points. This latter difference is not statistically significant ($p = 0.26$). Column (2) also shows that after including the dummies for beliefs, the treatment effect goes down by about two thirds, and becomes insignificantly different from 0.

We conclude that the effect of information is driven by the asymmetric effect of beliefs: correcting low beliefs about the emission size leads to a high decrease in buying, while correcting high beliefs leads only to a modest increase. Thus, providing information works because people who are too optimistic about the emission size are more responsive to information.[9]

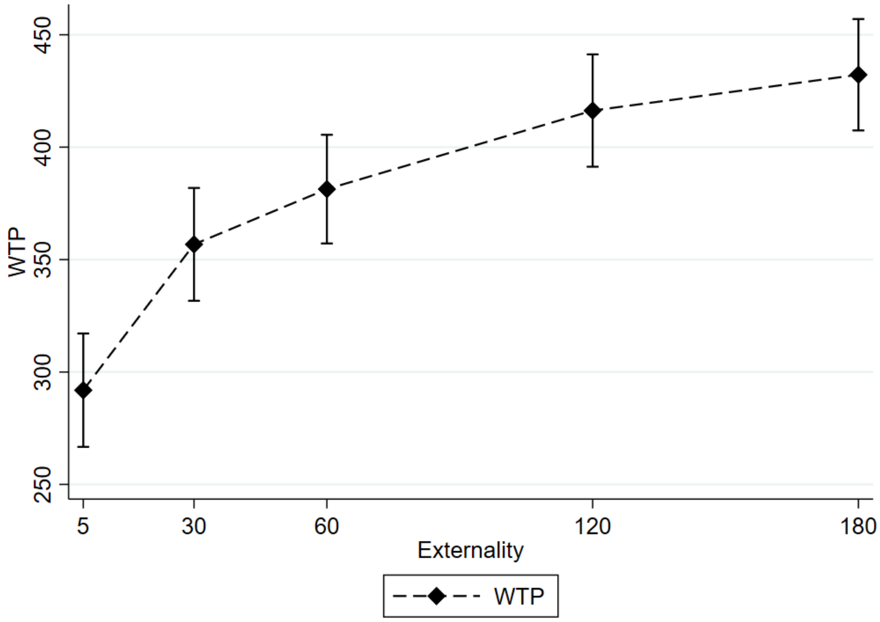


Figure 2: Average WTP to avoid different levels of CO2 emissions. X-axis: CO2 emissions expressed in liters of gasoline-burning equivalent. Y-axis: average WTP in pennies. The set of possible answers is $\{0, 50, ..., 700\}$ for each of the 5 levels of emissions.

4.1.1 Concave WTP to avoid CO_2 emissions

While these results are intriguing, they leave open several interpretations. For instance, the asymmetric effect of beliefs can be explained by a decline in the marginal WTP for additional units of emissions. Alternatively, there may be omitted variables that are corre-

[8] Here, we assume that all participants in the *Info* treatment have a belief of 60, as this is the information that they have been explicitly provided with. We also include dummies for the few participants who held other beliefs, but omit the estimates from the table, as they are imprecisely estimated.

[9] Our estimates are likely to provide an underestimate of the effect of beliefs. Even subjects who report a belief of 60 in the *Unmotivated* treatment are unlikely to be entirely sure about their guess. Their belief distribution may thus have additional weight on 0 and 120 as emission levels. If these weights have asymmetric effects, as our data suggest, sharpening these beliefs by providing more information may, in fact, cause further reductions in buying. In our specification, this effect is picked up by the treatment dummy.

lated with both different belief levels and buying behavior, although this seems unlikely.[10] To distinguish between these explanations, we collected additional data on Prolific for 134 individuals on individual willingness to pay (WTP) to avoid CO_2 emissions of different magnitude. We excluded 9 participants from the analysis because they acted like they had inconsistent preferences switching more than once along at least one multiple price list. Hence the analysis that follows is based on 125 participants. The design of this follow-up experiment is detailed in Section 3.5.

Figure 2 shows the average WTP data. WTP is increasing and concave in the size of the emissions. Moving from emissions equivalent to burning 5 liters of gasoline to emissions equivalent to 30 liters, a six-fold growth, increases the WTP by 65 pence, while moving from 5 to 180 liters, a jump seven times as large as the previous one, causes the WTP to grow by only 141 pence.

More formally, we estimate a Tobit model to take into account that the WTP is censored at £7.00. In the model, we include the emissions size and the square root of the emission size as independent variables to test whether the WTP increases less than linearly. Columns (1) and (2) of Table 3 reports the estimates of model. The square root term is highly significant both excluding and including the demographic controls, showing that indeed WTP increase concavely. Columns (3) and (4) exclude subjects who have zero WTP for all emission sizes, and are likely to be insensitive to information about CO_2 emissions. Excluding these participants does not affect the significance of the square root term and hence shows that the concavity of the WTP is not driven by these individuals. Thus, we find strong evidence that the aggregate WTP is increasing and concave in the size of the emissions.

Not only the aggregate WTP function is concave, but a majority of individual functions are concave as well. To show this, we construct a test of concavity for each individual, based on the fact that any convex combination of a concave WTP function points will lay below the function. The details of this analysis are given in Appendix C.2. Depending on how strict we construct our test to be, we find that between 49% and 67% percent of

[10] Apart from the fact that we control for several subject characteristics, participants in the *Unmotivated* treatment see the attentional task before knowing anything about the existence of the CO_2 emissions. We can not think of plausible individual characteristics that correlate with beliefs and influence participants' buying behavior.

Table 3: Testing concavity on aggregate

	(1)	(2)	(3)	(4)
	WTP	WTP	WTP	WTP
Emission size	-1.07**	-1.27**	-1.28**	-1.54**
	[-1.82,-0.32]	[-2.05,-0.50]	[-2.20,-0.36]	[-2.50,-0.59]
(Emission size)$^{\frac{1}{2}}$	33.5***	37.4***	40.7***	46.2***
	[19.0,47.9]	[22.4,52.3]	[23.4,57.9]	[28.2,64.2]
Controls		✓		✓
Observations	625	600	525	500

p-values in brackets.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Tobit models. Dependant variable WTP censored at 700 penny. Independent variables emissions size and square root of emissions size. Column (3) and (4) only include subjects with positive WTP for at least one emission size. Controls: age, gender, education (5categories), political identification (5 categories), frequency of car usage (5 categories). For 5 participants the demographic controls were not recorded participants have concave WTP, while less than 2% have a convex WTP. The concavity of the WTP at the individual level can explain why uncertainty increases buying in our experiment. People derive more disutility from an action that pollutes X emissions for sure than from an action that produces X in expectation.

4.2 Do subjects form motivated beliefs?

To investigate the formation of motivated beliefs, as discussed in the Section 3, we compare the *Unmotivated* treatment and *Motivated* treatment. The formation of self-serving beliefs, or "self-deception", is ruled out by design in the former treatment, as participants engaged with the attentional task before knowing the design of the experiment. By contrast, in the *Motivated* treatment, subjects knew the details of the decision and their economic interests.

Before turning to the results, note that participants have room for self-deception. First, the task is not trivial: only 51% of participants answer the belief question correctly in the *Unmotivated* treatment, even though they spent, on average, 50 seconds on the task screen. Thus, there was an ambiguity that would allow subjects to see what they want to see. Furthermore, our analysis so far shows that beliefs are an important determinant of buying behavior, implying that self-deception about emission size would indeed provide subjects with a powerful reason to buy the good.

Figure 4 shows the distribution of beliefs in the *Unmotivated* and *Motivated* treatment. It provides no evidence of the formation of motivated beliefs. A Mann-Whitney test fails to reject the null hypothesis of equal distribution of beliefs in the *Motivated* and the *Unmotivated* treatment ($p = 0.66$). Table 4 provides a regression analysis of beliefs in both treatments on the *Motivated* treatment dummy, including individual characteristics as controls. 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 the theories of motivated beliefs would predict. However, the coefficient is insignificantly different from 0.

We also test for differences in purchasing behavior between the two treatments with a Fisher's exact test. The test fails to reject that subjects are equally likely to buy the product in the *Motivated* and in the *Unmotivated* treatment ($p = 0.483$, two-sided test).

Figure 3: Belief distribution in the *Motivated* and in the *Unmotivated* treatment.

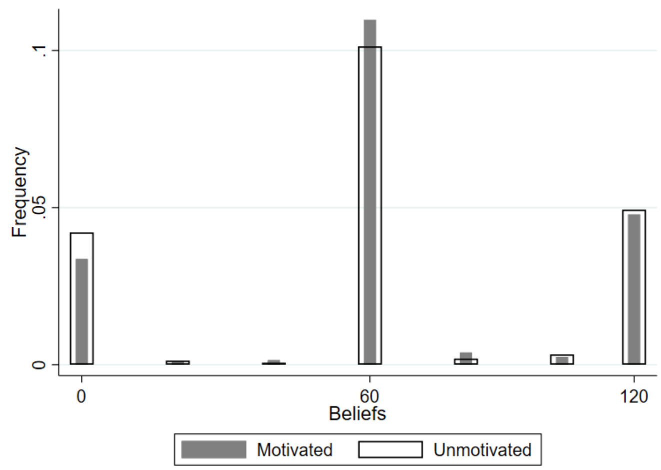


Figure 4: x-axis: beliefs; y-axis: frequency of beliefs. Separated histograms for the *Motivated* and the *Unmotivated* treatment

In column (2) of Table 4, we regress buying decisions on a treatment dummy and controls for individual characteristics. We find no evidence for a difference between treatments.

4.3 The effect of prices on buying and beliefs

We expected prices to have a standard substitution effect on buying as well as a more subtle indirect effect via beliefs. With regard to the former, Figure 1a demonstrates that subjects decrease their buying when the price of the good rises. This finding is confirmed statistically in the regression Table 5. Pooling across all treatments, we regress the purchasing decisions on the price of the good, we find a coefficient of -0.15 ($p < 0.001$). If we repeat this test for each of the different information treatments, we find similar results. Thus, we confirm that the emission of CO_2 follows standard economic laws, both under certainty and uncertainty.

We now turn to the indirect effect of prices on purchases via beliefs. As we hypothesized in Section 3, we expected an increase in the surplus (a decrease in prices) to increase the motivation to form self-serving beliefs in the attentional task. We tested this hypothesis by adding an interaction between *Motivated* and price in column (3) of Table 4. In line

Table 4: Comparison between the *Motivated* and the *Unmotivated treatment*

	(1)	(2)	(3)	(4)
	Beliefs	Units	Beliefs	Units
Motivated treatment	1.92	0.040	-4.41	0.051
	[-4.36,8.20]	[-0.033,0.11]	[-16.6,7.82]	[-0.093,0.20]
Price			-4.24	-0.13^{**}
			[-12.3,3.86]	[-0.22,-0.040]
Price* Motivated treatment			6.17	-0.022
			[-4.07,16.4]	[-0.14,0.094]
Observations	695	695	695	695

95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Models include observations from the *Motivated* and the *Unmotivated treatment*. The *Unmotivated treatment* is the baseline. The first model and third models are linear regressions, the second and the fourth are a linear probability models. All models are estimated with robust standard errors. Dependant variable: beliefs in columns (1) and (3); Purchasing decisions (1 if the participant purchased the product) in columns (2) and (4). Controls list: sex, age, student status, education (6 categories), nationality (27 categories).

with the lack of evidence for motivated beliefs, as documented in the previous subsection, we find no evidence that beliefs in the *Motivated* treatment depend on the price of the product. For completeness, column (4) tests for the interaction of price and *Motivated* treatment on buying behavior, where again we find no significant coefficients.

Overall, we find no evidence that the participants engage in motivated reasoning, or that there is causal effect of the size of the surplus. Although the estimates go in the hypothesized direction, the p-values are high, so our experiment yields a relatively clear null result.

Finally, we look at the interaction between prices and information. To do so we compare the *Motivated* and the *Information* treatment. While we found no evidence for motivated beliefs, our original hypothesis behind an interaction effect, there may be other reasons

Table 5: Downward sloping demand curves

	All observations	Motivated tr.	Info tr.	Unmotivated tr.
	Units	Units	Units	Units
Price	-0.15***	-0.13**	-0.19***	-0.12*
	[-0.20,-0.10]	[-0.21,-0.049]	[-0.27,-0.096]	[-0.22,-0.029
Controls	✓	✓	✓	✓
Observations	970	402	275	293

95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Linear probability models with robust standard errors. Dependent variable: purchasing decision (equal to 1 if the participant purchased the product). Controls list: sex, age, student status, education (6 categories), nationality (27 categories). Missing observations due to participants skipping one or more questions in the final questionnaire.

why the two policies interact. For instance, Rodemeier and Löschel (2020) find that information causes people to pay less attention to price effects. We test for this interaction in column (4) Table 6. We do not find statistically significant evidence for an interaction effect.

4.4 Women are less likely to buy the product

Finally, we look at gender effects. These analyses were not preregistered, nor do we have clear directional hypotheses. The findings could nevertheless be of interest to policymakers seeking to target information. Table 7 shows regressions of buying behavior on a gender dummy controlling for beliefs about the size of the emission size, price of the product, and the provision of information. In column (1), we estimate that females are 23 percentage points less likely to buy the product than males ($p < 0.001$). Overall, females are a striking 37% less likely to buy the product than males.

In the remaining columns of Table 7, we attempt to pinpoint the source of the difference in response to information, beliefs, or attention. Column (2) decomposes the behavioral response to information and prices across males and females by introducing interaction terms. The coefficients for the interaction terms are not statistically significant, although

Table 6: Comparison between Info and Motivated treatment

	(1)	(2)
	Units	Units
Price	-0.14***	-0.14***
	[-0.20,-0.082]	[-0.20,-0.085]
Info treatment	-0.10	-0.082
	[-0.23,0.025]	[-0.21,0.049]
Price* Info Treatment	-0.013	-0.038
	[-0.11,0.087]	[-0.14,0.066]
Controls		✓
Observations	1000	970

95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All the models are linear probability models with robust standard errors. Data from the Motivated and the Info treatment only. Baseline: *Motivated treatment*. Dependant variable: purchasing decision (equal to 1 if the participant purchased the product). Controls list: sex, age, student status, education (6 categories), nationality (27 categories).

Table 7: Women are less likely to buy but they do not have different beliefs.

	(1) Units	(2) Units	(3) Beliefs	(4) timeSec
Female	-0.23***	-0.21***	0.16	-4.92
	[-0.35,-0.11]	[-0.34,-0.092]	[-8.46,8.78]	[-11.0,1.14]
Info treatment	-0.072*	-0.044		
	[-0.14,-0.0012]	[-0.14,0.055]		
Female* Info treatment		-0.058		
		[-0.18,0.064]		
Price	-0.19***	-0.19***	2.58	-0.16
	[-0.26,-0.12]	[-0.26,-0.12]	[-3.84,8.99]	[-4.67,4.35]
PriceGender	0.074	0.073		
	[-0.020,0.17]	[-0.021,0.17]		
Dummies for beliefs	✓	✓		
Controls	✓	✓	✓	✓
Observations	970	970	401	694

95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The first and the second models are linear probability models, the third and the fourth are linear regression models. All the models are estimated with robust standard errors. The first and the second models include observations from all treatments; the third from the Motivated treatment only; the fourth from the Motivated and the Unmotivated treatment; Dependant variables: purchasing decision (equal to 1 if the participant purchased the product) for column (1) and (2); Beliefs for column (3); Time for column (4). Control list: age, student status, education (6 categories), nationality (27 categories). The variable Female takes value 0 for men and 1 for women.

the effect is imprecisely estimated. Column (3) investigates whether women have different beliefs on average, by regressing beliefs on a female dummy. Column (4) does the same for the time spent on the attentional task. We find no clear results from these last two analyses. We conclude that the most plausible explanation for the gender effect is that women have a stronger preference against generating CO_2 emissions.

5 Conclusions

We investigated the role of beliefs and information under uncertainty about CO_2 emission impact, using a series of experimental money-emission trade-offs in an online experiment. In general, we find strong intrinsic motivation to avoid emissions, especially among women. Beliefs about CO_2 emissions are strongly predictive of buying behavior, even more so than the price of the product. Even though beliefs are correct on average, information reduces emissions through a hitherto unexplored channel: because WTP is decreasing in the size of emissions, it is more effective on people who are too optimistic about the carbon impact of their behavior. We find no evidence that the surplus associated with the product influences attentional processes or the formation of self-serving beliefs about emissions.

Our results indicate that information can be an important policy tool in the fight against climate change, both in correcting beliefs and making them more precise. We also demonstrate that existing measures of WTP miss an important dimension, namely the size of the emissions. While the climate literature typically focuses on measuring WTP for a given unit of emissions, more effort should be directed to the changes in WTP for broader set of metrics. Our null finding on motivated cognition suggests that policymakers can implement carbon pricing and information policies without considering the interaction effects on beliefs. However, in light of some contradictory findings in related laboratory experiments (Momsen and Ohndorf, 2019a,b), and the possibility of other channels of complementarity (Rodemeier and Löschel, 2020), further research in this direction is warranted.

We believe our results offer ample opportunities for follow-up research. While our controlled setting allows us to disentangle the different mechanisms behind the role of information and beliefs, the experiments take place in a stylized online environment using a virtual good. To be a stronger guide to policy-making, the results are in need of corroboration in the field, using actual goods, a broader set of informational interventions, and more diverse samples. For instance, the effect of information for any given product will depend on the public's original beliefs about that product. Rodemeier and Löschel (2020) show that information may *reduce* sustainable behavior if most consumers initially overestimate the impact of efficiency measures. Finally, it is important to study the effect of information policies not just on behavior, but also effect on political support for climate change action.

References

Ambuehl, Sandro, "An Offer You Can't Refuse? Incentives Change What We Believe," SSRN Scholarly Paper ID 2830171, Social Science Research Network, Rochester, NY October 2016.

Ashraf, Nava, B Kelsey Jack, and Emir Kamenica, "Information and subsidies: Complements or substitutes?," *Journal of Economic Behavior & Organization*, 2013, *88*, 133–139.

Bartling, Björn, Roberto A Weber, and Lan Yao, "Do markets erode social responsibility?," *The Quarterly Journal of Economics*, 2014, *130* (1), 219–266.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral science*, 1964, *9* (3), 226–232.

Bénabou, Roland and Jean Tirole, "Mindful economics: The production, consumption, and value of beliefs," *Journal of Economic Perspectives*, 2016, *30* (3), 141–64.

Butera, Luigi and John A List, "An economic approach to alleviate the crises of confidence in science: With an application to the public goods game," Technical Report, National Bureau of Economic Research 2017.

Dana, Jason, Roberto A Weber, and Jason Xi Kuang, "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness," *Economic Theory*, 2007, *33* (1), 67–80.

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

Dupas, Pascaline, "What matters (and what does not) in households' decision to invest in malaria prevention?," *American Economic Review*, 2009, *99* (2), 224–30.

Ehrich, Kristine R and Julie R Irwin, "Willful ignorance in the request for product attribute information," *Journal of Marketing Research*, 2005, *XLii* (August), 266–277.

EPA, "Average Carbon Dioxide Emissions Resulting from Gasoline and Diesel Fuel," *US EPA, February*, 2005.

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

Fang, Ximeng, Lorenz Goette, Bettina Rockenbach, Matthias Sutter, Verena Tiefenbeck, Samuel Schoeb, Thorsten Staake et al., "Complementarities in Behavioral Interventions: Evidence from a Field Experiment on Energy Conservation," Technical Report, CRC TR 224 Discussion Paper 149, University of Bonn and University of Mannheim, Germany 2020.

Feiler, Lauren, "Testing models of information avoidance with binary choice dictator games," *Journal of Economic Psychology*, 2014, *45*, 253–267.

Festinger, Leon and James M Carlsmith, "Cognitive consequences of forced compliance.," *The journal of abnormal and social psychology*, 1959, *58* (2), 203.

Gino, Francesca, Michael I Norton, and Roberto A Weber, "Motivated Bayesians: Feeling moral while acting egoistically," *Journal of Economic Perspectives*, 2016, *30* (3), 189–212.

Grossman, Zachary and Joël J Van der Weele, "Self-image and willful ignorance in social decisions," *Journal of the European Economic Association*, 2017, *15* (1), 173–217.

Hestermann, Nina, Yves Le Yaouanq, and Nicolas Treich, "An economic model of the meat paradox," 2018.

Kappes, Andreas, Anne-Marie Nussberger, Nadira S Faber, Guy Kahane, Julian Savulescu, and Molly J Crockett, "Uncertainty about the impact of social decisions increases prosocial behaviour," *Nature Human Behaviour*, August 2018, *2* (8), 573–580.

Lind, Jo Thori, Karine Nyborg, and Anna Pauls, "Save the planet or close your eyes? Testing strategic ignorance in a charity context," *Ecological Economics*, July 2019, *161*, 9–19.

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

Meredith, Jennifer, Jonathan Robinson, Sarah Walker, and Bruce Wydick, "Keeping the doctor away: experimental evidence on investment in preventative health products," *Journal of Development Economics*, 2013, *105*, 196–210.

Momsen, Katharina and Markus Ohndorf, "Information Avoidance, Selective Exposure, and Fake (?) News-A Green Market Experiment," Technical Report 2019.

_ **and** _ , "When do people exploit moral wiggle room? An experimental analysis in a market setup," 2019.

Nemet, Gregory F and Evan Johnson, "Willingness to Pay for Climate Policy: A Review of Estimates," SSRN Scholarly Paper ID 1626931, Social Science Research Network, Rochester, NY June 2010.

Nordhaus, William, "Climate change: The ultimate challenge for Economics," *American Economic Review*, 2019, *109* (6), 1991–2014.

Rodemeier, Matthias and Andreas Löschel, "The welfare effects of persuasion and taxation: Theory and evidence from the field," *ZEW-Centre for European Economic Research Discussion Paper*, 2020, (20-019).

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

Schumacher, Heiner, Iris Kesternich, Michael Kosfeld, and Joachim Winter, "One, two, many—Insensitivity to group size in games with concentrated benefits and dispersed costs," *The Review of Economic Studies*, 2017, *84* (3), 1346–1377.

Schwardmann, Peter, "Motivated health risk denial and preventative health care investments," *Journal of Health Economics*, 2019.

Serra-Garcia, Marta and Nora Szech, "The (in) elasticity of moral ignorance," 2019.

Stiglitz, Joseph E, Nicholas Stern, Maosheng Duan, Ottmar Edenhofer, Gaël Giraud, Geoffrey M Heal, Emilio Lèbre la Rovere, Adele Morris, Elisabeth Moyer, Mari Pangestu et al., "Report of the high-level commission on carbon prices," 2017.

Thunström, Linda, Klaas van't Veld, Jason F Shogren, and Jonas Nordström, "On strategic ignorance of environmental harm and social norms," *Revue d'économie politique*, 2014, *124* (2), 195–214.

van der Weele, Joël J, "Inconvenient Truths : Determinants of Strategic Ignorance in Moral Dilemmas," *SSRN Electronic Journal*, 2013.

Appendices

A Taxes and information are substitutes, a theoretical derivation

In this section, we show that two testable behavioral assumptions are sufficient to prove that taxes and information are substitutes. Those assumptions are:

1. The marginal utility of consumption is decreasing in the believed magnitude of the negative emissions produced by consuming a product.

2. Consumers believe that the size of the emissions is higher when their surplus shrinks.

Formally assume:

$$U = u(c, y, \beta)$$

where y is the numeraire commodity, c is the good producing the emissions, and β is the believed size of the emissions. β is a statistic of the beliefs distribution over the possible value of the emissions. To stay close to the experimental design, we can think of β as the mode of the beliefs distribution. A higher β indicates that the consumer believes the emissions to be bigger. We assume U to be increasing and concave in both consumption goods and let p be the price of c.

We further assume that $U_{c\beta} < 0$, such that $c_\beta(p,\beta) < 0$. This entails that the marginal utility of consumption is decreasing in the believed size of the emissions. This is the first of the two assumptions we need: the consumer cares about the emissions he thinks he is producing.

Finally, we let $\beta = \beta(p)$ and $\beta'(p) \geq 0$. $\beta'(p) > 0$, indicates that consumers self-deceive: they adjust their beliefs depending on their monetary incentives. This formalizes the second assumption. Consumers' beliefs are increasing when consumer surplus decreases. We can, then, model an information policy that prevents self-deception, as an intervention that imposes $\beta'(p) = 0$.

We will now show that these two assumptions suffice to derive that taxes and information are substitutes. To see this, assume that a tax causes p to rises from p_1 to p_2, with $p_2 > p_1$, and that the information policy is not in place. In this case, $\beta'(p) > 0$, and β increases from β_1 to β_2 with $\beta_2 > \beta_1$. The change in consumption is then given by: [11]

$$c(p_2,\beta_1) - c(p_1,\beta_2) = \underbrace{c(p_2,\beta_1) - c(p_1,\beta_1)}_{\text{direct effect}} + \underbrace{c(p_2,\beta_2) - c(p_2,\beta_1)}_{\text{indirect effect}}$$

The expression on the RHS decomposes the total effect of the price change in a direct effect and an indirect effect due to the change in beliefs. The direct effect is negative because c has become relatively more expensive than y. The indirect effect is negative as well because $c_\beta(p,\beta) < 0$. As a consequence, the change in beliefs reinforces the behavioral response to the change in incentives. This implies that the marginal effect of introducing a tax is higher when self-deception is possible. A policy that imposes $\beta'(p) = 0$, would eliminate the indirect effect described above.

Obviously, the opposite is also true: the effect of information policy is higher when no tax is in place. In fact, the motivated bias in beliefs is higher when the consumer surplus is larger.

[11] To simplify the exposition of the proof, we assume that the utility function is quasi-linear. This assumption is innocuous: the wealth effect does not depend on beliefs in the model.

B Why we are sure that no bot completed the study

Two aspects of our design minimize the risk that a bot could complete our experiment. First, our instructions are not machine-readable. Hence, a computer script would need to answer all the comprehension questions randomly. This would result in an exceptionally high number of answer attempts. We keep track of the number of attempts, and none of the participants that reached the end of the experiment needed more than 65 attempts to answer to the all the comprehension questions correctly.[12] Moreover, we included two honey-pots in the program. Those 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. We considered answering to any of the two honey-pot as sufficient evidence that the participant is a bot. We do not find any participant that answers any of these two questions. Putting together the evidence from the number of attempts and the honey pots, we conclude that no bot completed our experiment.

C Additional Analysis

C.1 Robustness test on the effect of information

As a robustness check for the effect of information on buying behavior, we compare behavior in the Info treatment with that in the Motivated treatment.

Figure 1a shows that for any given price level, buying is lower in the INFO treatment than in both other treatments. Pool the data across price levels, a Fisher's exact test corroborates this finding: we reject the null hypothesis of equal buying behavior between the INFO and Motivated treatment (two-sided $p = 0.001$). To obtain further insights, Table 8 shows the results of OLS regression of buying on the Motivated treatment, where the Info treatment is the baseline. The results in column (1) indicate that being in the *Motivated* treatment increases buying by 12 percentage points.

We next explore the role of beliefs in the effect of uncertainty. Figure 1b shows how buying behavior depends on beliefs in the treatments with uncertainty. We focus on the

[12] 99% of our subjects needed less than 24 attempts to answer all the questions correctly. The minimal number of attempted needed is 3 or 4 depending on the treatment.

Table 8: Participants buy more units in the Motivated treatment

	(1)	(2)
	Units	Units
Info treatment	-0.14***	-0.10*
	[-0.21,-0.068]	[-0.19,-0.019]
Beliefs=60		-0.33***
		[-0.45,-0.20]
Beliefs=120		-0.40***
		[-0.54,-0.27]
Dummies for beliefs		✓
Controls	✓	✓
Observations	677	677

95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All the models are linear probability models with robust standard errors. Data from the Motivated and the Info treatment only. Baseline: *Motivated treatment*. Dependant variable: Units. Controls list: sex, age, student status, education (6 categories), nationality (27 categories).

three most prominent belief levels, i.e., the beliefs that the emissions are either 0, 60, or 120. Due to the nature of our visual task, 95% of participants indicated one of these three beliefs. As is readily apparent, buying decreases when subjects believe the emissions are larger. It is also clear that this effect is non-linear: buying decreases by about 50% when beliefs about the emissions increase from 0 to 60 liters worth of gasoline, but when they increase further to 120 liters, the decrease in buying is only 24%.

The average belief in the Unmotivated treatment is 62.7, insignificantly different from 60 ($p = 0.25$). This means that the average beliefs are very similar to those in the *Information* treatment, where subjects were informed that the true emissions were equivalent to 60 liters. Thus, we can exclude that differences in buying behavior between the *Information* treatment and the *Motivated* treatment is due to correcting a systematic and unmotivated bias in beliefs. In column (4), we include a dummy for the most prominent belief levels. These dummies demonstrate the non-linear effects of beliefs on buying. Moreover, the treatment effect now goes down by about 33% and becomes less precisely estimated. Thus, we confirm the finding that at least part of the effect is driven by the asymmetric effect of beliefs: correcting low beliefs about the emissions leads to a high decrease in buying, while correcting high beliefs leads only to a modest increase. Thus, providing information works because people who are too optimistic about the emissions are much more responsive to it. Note that our estimates are likely to provide an underestimate of the effect of beliefs. Even subjects who report a belief of 60 in the motivated treatment are unlikely to be entirely sure about their guess. Their belief distribution may thus put additional weight on 0 and 120 as emissions levels. If these weights have asymmetric effects, as our data suggest, sharpening these beliefs by providing more information may in fact, cause further reductions in buying. In our specification, this effect is picked up by the treatment dummy.

C.2 Test for the non-linearity of the WTP at the individual level

In this Appendix, we illustrate the test we run determine whether a participant has a concave WTP. The idea of the test is based on the Jensen's inequality: if a function is concave, any convex combination of its points will lay below the function, vice versa, if the function is convex, any convex combination of its points will lay above the function. Operationally, we built the convex combination of two points - $(0; 0)$ and $(max(WTP);\ emissions(max(WTP))$. We chose $(0; 0)$ under the very reasonable assumption that every participant has a WTP of £0.00 to avoid CO_2 emissions equivalent to 0 liters of gasoline. $max(WTP)$, instead, is the highest, but not censored WTP that the individual reports. $emissions(max(WTP)$ is the corresponding emissions level. We consider an individual WTP concave (convex) if the convex combination is above (below) all the emissions levels smaller than $emissions(max(WTP))$. Note that this empirical strategy allows us to test the concavity of the WTP only for values smaller than £7.00, the maximum value in the multiple price lists. For this reason, we can only run the test on 71% of our sample. The remaining 29% has a WTP of £7.00 or higher already for an externality level of 30 liters of gasoline.

Table 9 display the results of the tests. Column (1) shows that 49% of the subjects that we can test exhibit a concave WTP, while only 1% have a convex WTP. Check the robustness of this finding. The tests for these columns are constructed in the same way as the test of column (1) except for the set of subjects that are considered eligible for the test. Column (2) shows that 63% of the subjects have a concave WTP if we exclude participants that have a WTP of £0.00 regardless of the size of the emissions. Those are participants that are unluckily be responsive to information. Column (3) instead shows that the percentage of concave WTP increases to 63% if we only include participants that have a not censored WTP for every emissions level. Those are the participants for which we have the most information about their WTP. Finally, Column (4) highligths that the results are not driven by participants that show a WTP that is not weakly monotone. Those are participants that, one could argue, got confused or where not paying much attention. All in all, our test provides very strong evidence that a majority of people a WTP to avoid CO_2 emissions that is concave in the size of the emissions.

C.3 Treatment effect of the time spent looking at the attentional task.

On average, participants spend 51.3 seconds looking at the attentional task. In the *Motivated* treatment, participants spend 2.7 seconds more than in the Unmotivated treatment. This difference is small and not significant ($p = 0.33$). Table 10 summarizes this result.

D Instructions

The instructions of this experiment can be found following this link

Table 9: Test for concave WTP at the individual level.

	Test (1)	Test (2)	Test (3)	Test (4)
Total Participant	125	125	125	125
Eligible	**89 (71%)**	**69 (55%)**	**44 (35%)**	**36 (29%)**
Concave	44	44	28	24
% Concave over eligible	**49%**	**63%**	**63%**	**67%**
Convex	1	1	0	0
% Convex over eligible	**1%**	**1%**	**0%**	**0%**
Censored for every externality ≥ 30	X	X	X	X
Selfish	✓	X	X	X
Censored for at least one externality level	✓	✓	X	X
Not Monotone	✓	✓	✓	X

Per each participant, the tests above checks weather the linear combination of 0 and the highest level of not censored WTP lays above or below the WTP for smaller externality levels. A participant is considered to have a "Concave" WTP if the linear combination is below the corresponding WTP for at least 75% of the externality levels for which it is possible to run the test. Parallelly, a participant is considered to have a "Convex" WTP if the linear combination is below the corresponding WTP for at least 75% of the externality levels for which it is possible to run the test. The test differ for their exclusion restrictions. A ✓ indicates that the corresponding group is included in the test, an X indicates that the corresponding group is excluded by the test. Groups definition: 1) "Censored for every externality ≤ 30": participants that have a WTP larger or equal to £7.00 for every externality level above 30 liters of gasoline. 2) "Selfish": participants whose WTP is £0.00 for every level of the externality 3) "Censored for at least one externality level": participant that have a WTP above 700 larger or equal to £7.00 for at least one externality level. 4) "Not Monotone": participants whose WTP is not weakly increasing.

Table 10: No treatment effect on the time spent on the puzzle.

	(1)	(2)
	time	time
Motivated Treatment	2.79	1.59
	[-2.86,8.44]	[-4.29,7.47]
Controls		✓
Observations	713	694

95% confidence intervals in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All the models are linear regression models with robust standard errors. Dependant variable: time spent on the puzzle in seconds. Controls list: sex, age, student status, education (6 categories), nationality (27 categories).

E Preregistration

In this Appendix, we report the text of the preregistration of the experiment we uploaded on AsPredicted.org

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 hypothesis 4, we may do additional analysis (see under point 8).

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).

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 1000 observations. Every participant has 40% probability of being in the Motivated Treatment, 30% probability of being in the Unmotivated Treatment, and 30% probability of being in the Info Treatment. Participants are assigned to one of the 3 price conditions within these treatments with equal probabilities.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
1) We test if, in the Motived Uncertainty Treatment, the time spent on the puzzle depends on the price of the good.
2) We estimate the effect of beliefs on purchasing behavior. We do so using an Instrumental Variable approach and using data from Motivated Uncertainty Treatment and from the Unmotivated Uncertainty Treatment. We repeat the analysis with price dummies to test if beliefs affect consumption at all price levels.
3) We test whether questionnaire items like age, gender and self-reported concern about climate change correlate with our main variables of interest.
4) If the relation between beliefs and prices is different than that hypothesized under Hypothesis 4, we will do additional analysis to see how prices determine beliefs and how beliefs affect price sensitivity.

F Invoice Carbonfund.org

In this Appendix we report the invoice of the donation to Carbonfund.org. The invoice was shared with participants on $14^t h$ May.

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*