CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Moral hazards and solar radiation management: Evidence from a large-scale online experiment
Authors	Schoenegger, Philipp;Mintz-Woo, Kian
Publication date	2024-04-08
Original Citation	Schoenegger, P. and Mintz-Woo, K. (2024) 'Moral hazards and solar radiation management: Evidence from a large-scale online experiment', Journal of Environmental Psychology, 102288 (11pp). https://doi.org/10.1016/j.jenvp.2024.102288
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1016/j.jenvp.2024.102288
Rights	© 2024, the Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/). - https://creativecommons.org/licenses/by/4.0/
Download date	2026-10-27 09:42:08
Item downloaded from	https://hdl.handle.net/10468/15806

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Journal of Environmental Psychology 95 (2024) 102288

Contents lists available at ScienceDirect

Journal of Environmental Psychology

journal homepage: www.elsevier.com/locate/jep

Moral hazards and solar radiation management: Evidence from a large-scale online experiment☆

Philipp Schoenegger[a], Kian Mintz-Woo[b,c,*]

[a] London School of Economics and Political Science, London, United Kingdom
[b] Philosophy and Environmental Research Institute, University College Cork, Cork, Ireland
[c] Equity and Justice Group, International Institute for Applied Systems Analysis, Laxenburg, Austria

ARTICLE INFO

Handling editor: Wokje Abrahamse

ABSTRACT

Solar radiation management (SRM) may help to reduce the negative outcomes of climate change by minimising or reversing global warming. However, many express the worry that SRM may pose a moral hazard, i.e., that information about SRM may lead to a reduction in climate change mitigation efforts. In this paper, we report a large-scale preregistered, money-incentivised, online experiment with a representative US sample (N = 2284). We compare actual behaviour (donations to climate change charities and clicks on climate change petition links) as well as stated preferences (support for a carbon tax and self-reported intentions to reduce emissions) between participants who receive information about SRM with two control groups (a salience control that includes information about climate change generally and a content control that includes information about a different topic). Behavioural choices are made with an earned real-money endowment, and stated preference responses are incentivised via the Bayesian Truth Serum. We fail to find a significant impact of receiving information about SRM and, based on equivalence tests, we provide evidence in favour of the absence of a meaningfully large effect. Our results thus provide evidence for the claim that there is no detectable moral hazard with respect to SRM.

1. Introduction

Climate change continues to pose serious challenges to societies across the globe as the international community fails to adequately address its root causes. Aside from mitigation strategies aimed at reducing greenhouse gas emissions, geoengineering approaches are increasingly being considered. These are intentional interventions in the climate system with the aim of minimising, reducing, or reversing the damaging effects of climate change. A prominent example of geoengineering is solar radiation management (SRM), which attempts to reflect back or otherwise neutralise a fraction of sunlight. This can be achieved via marine cloud brightening (Latham et al., 2012), stratospheric interventions (Hulme, 2012), or other methods. What all SRM methods have in common is that they reduce ground-level solar radiation in a way that some believe could relatively cheaply and easily reduce short-term global warming. However, such options come with technical downside risks and do not directly address the root cause or other chemical effects of greenhouse gas emissions (Mahajan, Tingley, & Wagner, 2019).

Even if SRM would work as hoped and the technical risks were contained, residual risks remain. One such risk is that of a moral hazard (Gardiner, 2017; Hale, 2012, ch 7; Svoboda, 2017), which has been called a "prominent challenge" to geoengineering (Pamplany, Gordijn, & Brereton, 2020). Moral hazard, as Baker (1996) discusses, refers to the "tendency for insurance against loss to reduce incentives to prevent or minimize the cost of loss" (239), i.e., moral hazard refers to the effect by which individuals' incentives regarding some behaviour are altered if the majority of the downside risk is borne by others, e.g., insurers. For instance, if a property is insured against fire, the property owners may be less likely to take the necessary steps to further reduce fire risks as the majority of the risk is borne by the insurance company and not them. This effect has been studied before in contexts such as health insurance (Zweifel & Manning, 2000, chap. 8), worker's compensation coverage (Butler & Worrall, 1991), natural disasters (Hudson, Botzen, Czajkowski, & Kreibich 2017), crop insurance (Quiggin, Karagiannis, & Stanton, 1993), and bank deposits (Martin, 2006).

☆ We thank Miguel Costa-Gomes, Stephen Gardiner, Basil Halperin, David Reinstein, Patrick Smith, and Theron Pummer, as well as Alex Wong and Samuel Kaufmann from SilverLining for helpful comments on this paper. We also greatly appreciate the reviews of anonymous referees at this and another journal.

* Corresponding author.
E-mail addresses: contact.schoenegger@gmail.com (P. Schoenegger), mintzwoo@iiasa.ac.at (K. Mintz-Woo).

https://doi.org/10.1016/j.jenvp.2024.102288
Received 11 July 2023; Received in revised form 1 March 2024; Accepted 28 March 2024
Available online 3 April 2024
0272-4944/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Lin (2013) suggests that a similar kind of moral hazard applies to geoengineering. In Lin's account, geoengineering research is like an insurance policy, the government researching geoengineering is like an insurer, and the public is like the insured. With these similarities, analogous moral hazard is potentially present. On Lin's account, if the public is made aware that the government has backup policies to reducing emissions (i.e. geoengineering), then the public may end up with less motivation or behaviour to reduce emissions. This is of significant importance for public policy; if research into, and information about, geoengineering reduce motivation, that is a strong reason to refrain from researching or adopting geoengineering. Conversely, if the posited moral hazard is absent or undetectable, then there could be reason to develop research and disseminate information about geoengineering without worrying about this "prominent challenge."

While the moral hazard objection has received significant attention (e.g., Pamplany et al.'s (2020) recent review found 33 papers on the topic) and is plausibly a central concern in the social science literature on this topic, empirical work on moral hazard remains relatively sparse. Further evidence could help indicate whether or to what extent moral hazard is generated by introducing geoengineering in general (and SRM in particular) to the public, which is largely unaware that there are alternatives to conventional mitigation (Mahajan et al., 2019). If individuals were made aware of SRM, this could reduce motivation to act, e.g. because such information lessens the perceived threat of climate change (Campbell-Arvai, Hart, Raimi, & Wolske, 2017), or because it weakens resolve (Austin & Converse, 2021).

Overall, the results presented in the literature so far are mixed.[1] In line with the theoretical predictions outlined by Lin (2013), Raimi, Maki, Dana, & Vandenbergh (2019) find that reading about geoengineering leads to a reduction in mitigation support in a US sample (irrespective of the framing of the problem), though the magnitude and significance of the effects varied depending on the description of SRM. Contrary to this finding, however, Cherry, Kallbekken, Kroll, and McEvoy (2021) find that information about SRM leads to an increase in support for a national carbon tax. This is corroborated by a similar result in a German sample by Merk, Pönitzsch, & Rehdanz (2015), who find that reading about SRM increases willingness to invest in mitigation. Further, Fairbrother (2016) finds no effect of receiving an introduction to SRM on the willingness to pay taxes. Lastly, in a climate disaster game, Andrews, Delton, and Kline (2022) find no moral hazard amongst "citizens", but that "policymakers" somewhat anticipate that "citizens" will be subject to moral hazard.

We believe that the present mixed results are inconclusive and that this is, at least in part, because of the inconsistent methods employed in the literature. Specifically, we identify three methodological features that have not been consistently employed in past work. These are research purpose covertness, salience control, and behavioural measures. First, keeping the research purpose covert is important to minimize experimenter demand effects, especially when it comes to a potentially politically divisive topic like climate change. For example, if participants are aware that the study is about climate change, they might self-select into it only if they have strong opinions on climate change or they might behave differently throughout the study. Second, controlling for salience ensures credible identification of a geoengineering-specific effect as opposed to a climate-related effect. A salience control is like a treatment control in containing information on the same topic; here, the salience control includes information about climate change. This is so results can reveal if simply being presented with information about climate change might by itself have an effect that could be mistaken for a treatment effect. Finally, including behavioural measures is central to upholding higher external validity of study results since the actual outcome of interest in the real-world is behaviour, e.g., as compared to more widely used hypothetical choices and preferences elicited via surveys. Ideally, research in this area (and related fields) would include all three of these methodological features.

While most work in this area does indeed have one or more of these methodological features, we find that none of the papers have all three. Below we outline the five most relevant papers, indicating whether their methods ensure covert study purpose,[2] control for salience, or include behavioural measures, see Fig. 1. These results show that while most individual studies include some of these features, none of them include all three.

To address this methodological literature gap, this present study adopts all three methodological features in an attempt to reduce the heterogeneity in methods that may explain some of the mixed results in the literature. Our design thus has the following methodological strengths: First, from recruitment all through the end of the study, participants were told (truthfully) that the study is concerned with a multitude of topics. They could not know that we were only interested in their attitudes and behaviours relating to climate change. As such, the actual purpose of the study was kept fully opaque, allowing us to mitigate experimenter demand worries to a significant extent. Second, by the introduction of our Salience Control, we disentangled a potential effect driven simply by making climate change salient—again a feature that was not always properly controlled for in all previous work. Third, we employed a variety of outcome measures, in terms of both stated preferences and behaviours. This allowed us to capture a wide spectrum of participant responses. This also made it less likely that our design

Fig. 1. Heatmap of experimental design features. *Notes:* Authors' summary of previous literature's methodological features, showcasing experiment design heterogeneity.

[1] While we lack the space to be comprehensive, the experimental literature on carbon dioxide removal shows similarly mixed results (Campbell-Arvai et al., 2017; Austin & Converse, 2021). As we do, Hart, Campbell-Arvai, Wolske, and Raimi (2022) control for salience in the case of carbon dioxide removal and, as we do with SRM, find no moral hazard effect.

[2] Study purpose (c)overtness can be a feature of both the recruitment of participants and the study itself, e.g., when the study only asks about attitudes to climate-related questions. Because many papers do not detail how they recruited participants (including their advertised study title and description), we employ a dual criterion. For a study to count as having an overt research purpose, the paper has to report either an overt recruitment or an overt study design. For a study to count as having a covert research paper, the paper has to have covert recruitment and study design (if both are reported) or only study purpose be covert.

omitted plausible outcomes while also leading to higher external validity as the central response of interest is actual behaviour, i.e., we did not only rely on hypothetical measures, giving our study a higher level of external validity.

2. Methods

2.1. Hypotheses

Our study includes three conditions: a Treatment condition and two control conditions. The Treatment condition exposes participants to information about SRM, one control condition ("Salience Control") exposes participants to information about climate change mitigation generally, and the other control ("Content Control") condition is about a different topic entirely. We collect measures for both policy support as well as behaviours related to climate mitigation. The two main null hypotheses that we preregistered are:

Null Hypothesis I. Information about geoengineering does not reduce (or increase) policy support for mitigation measures.

Null Hypothesis II. Information about geoengineering does not reduce (or increase) behaviours related to mitigation measures.

As preregistered, we understand the following patterns of data as providing the attendant evidence in favour of (or against) the existence of a moral hazard in the context of SRM: Strong evidence in favour of the existence of a moral hazard effect would involve the rejection of both null hypotheses. Weak evidence in favour of the existence of a moral hazard effect would involve the rejection of only one of the two null hypotheses. Failing to reject both null hypotheses, observation of an increase, or finding evidence in favour of a null would be taken as evidence against the existence of a moral hazard effect.

2.2. Participants

We preregistered this study on the Open Science Framework,[3] and it has received ethics approval. For this study, we recruited 2500 participants via Prolific, adopting Prolific's representative quota sampling option. After excluding participants based on attention and comprehension checks, the final sample consisted of a total of 2284 participants. The age of participants ranged from 18 to 92 years, with a mean age of $M = 44.12$ ($SD = 15.86$). In terms of gender distribution, 1168 participants identified as Female (51.14%). The majority of participants identified as White (75.8%) while 57.4% identified as liberal. In terms of education, 30.5% completed only High School, with 48% of participants having completed an undergraduate degree and the remainder holding a postgraduate or professional degree. We collected these data between March 5, 2022 and March 12, 2022 on Prolific. Participants received a participation reward of £1.25.[4] During the experiment, they could earn an endowment of up to £0.45 depending on their choices.

To calculate this sample size, we conducted an a priori power analysis via G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). Our main analysis is a multiple regression model with treatment dummies and control variables. In order to have the high level of 95% power to detect the smallest effect size of interest at the global effect of $f^2 = 0.01$ (a conventionally very small effect) with an alpha level at 0.01 (adjusting for multiple comparisons) in a multiple regression model with over 10 predictors (which include treatment dummies and control variables), the required total sample size is 1785. To adjust for exclusions, we recruit 750 participants per control condition and 1000 in the treatment (totalling 2500), which is within the maximum deliverable representative sample size of 2500 and allows this study to have a high level of power even if many participants would need to be excluded. To avoid self-selection, we advertised this study as a study regarding current topics and did not mention the focus on climate change.

2.3. Endowment and exclusions

We randomly selected participants into one of three conditions: Treatment, Salience Control, and Content Control; the Treatment and Salience Control provided participants with a text that includes climate, varying only by whether SRM or climate mitigation was offered as the solution. This allows the Salience Control condition to control for the salience of just thinking about climate generally. The Content Control condition received no information about climate at all, just an added text about another potentially contentious topic, racism. All participants were shown texts on three different topics, including texts on abortion and CRISPR in addition to each group's treatment specific texts. All three texts within each group were presented randomly to avoid order effects. After one of the texts, participants were presented with an attention check that instructed them to respond with 'Disagree' on a Likert-scale item asking them how they had enjoyed reading the texts so far. For a visual depiction of which text was shown in which condition and the overall experimental design, cf. Fig. 2.[5]

Directly under each text, participants were quizzed on the content of the text to ensure that they read the texts carefully and to verify comprehension of the Treatment text. Participants received £0.10 for each correctly answered question (for a total of up to £0.30),[6] making up the earnings for their experimental endowment. In case the question was answered incorrectly, participants were told this after answering all three quiz questions and were provided with the text and the same question again, explaining to them that they had answered this question incorrectly while giving them the opportunity to retake the quiz. If they answered the question incorrectly again, we coded this as failing the comprehension check and removed them from all analyses. We excluded all participants who failed either the comprehension check, the attention check, or both.

In total, we excluded 216 participants for failing either the attention check or one or more comprehension questions at the second attempt (or both). All analyses in this paper are reported with the remaining 2284 participants. The results in all preregistered analyses are robust to this exclusion decision; results do not differ if all participants are included.

2.4. Experimental manipulation

Depending on the condition, those in the Treatment condition received a text about SRM, while those in the Salience Control condition received a text about climate mitigation, and those in the Content Control condition received a text about racism with all conditions receiving two other texts (about abortion and CRISPR), see Fig. 1.

The Treatment condition's specific text introduced SRM. This text included both an introductory paragraph on climate change generally and a specific paragraph explaining SRM, outlining potential risks and upsides. This functions as our central intervention and is phrased

[3] Preregistration available at the Open Science Framework in anonymised form: https://osf.io/n6vt3/?view_only=3358f4414543401caf79d3331e9240d9.

[4] All participant rewards on Prolific are denominated in GBP (£). As such, even a US sample like the one we draw on in this paper is well acquainted with this currency and we do not anticipate this to have any impact on our results.

[5] We also collected importance measures for each item but did not preregister any analyses with them. As such, while they are represented at the study design for completeness, they will not be discussed in the results section.

[6] We have chosen relatively small stakes because previous research has shown that in donation contexts, participant behaviour is relatively invariant to stake sizes, with the primary exception being cases of extremely high stakes, in which hyper-altruistic behaviour (donating all of one's endowment) vanish (Brañas-Garza, Jorrat, Kovářík, & López, 2021).

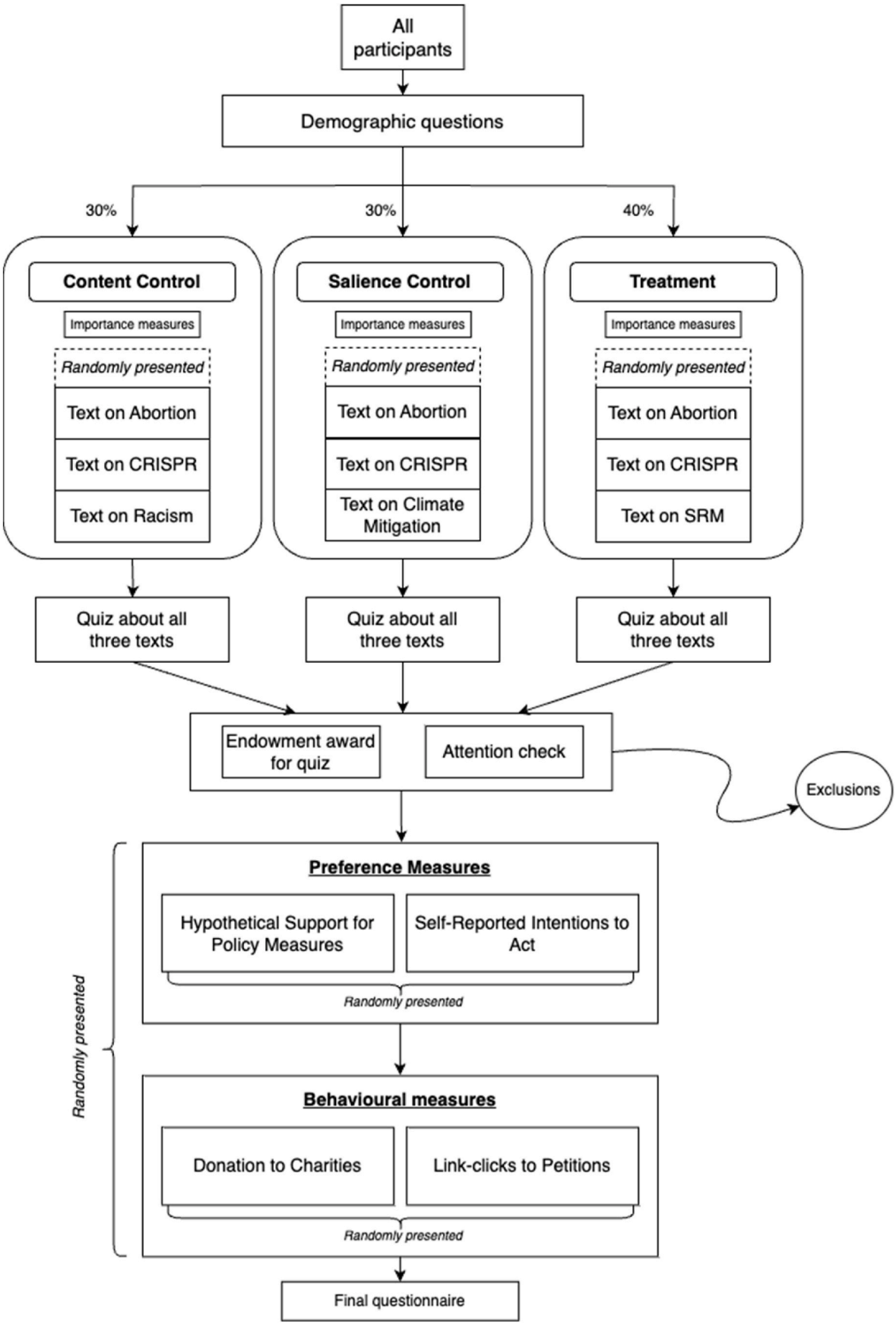

Fig. 2. Experimental procedure. *Notes:* Full experimental procedure, including sample sizes of all conditions and an overview of all measures collected.

neutrally to mimic the type of information most likely to be received in a 'real world' context. The Salience Control's specific text was the same text on the topic of standard climate change impacts as in the Treatment. This paragraph was followed by another paragraph focused exclusively on standard mitigation techniques. This text consisted of material which we expect subjects were likely to be familiar with, allowing us to control for the salience of climate change generally. For full treatment texts, see Appendix E.

2.5. Outcome measures

For each of the topics shown to participants, we collected two types of data across numerous topics: *stated preference measures* and *behavioural measures*. Our primary variables of interest relate to direct climate change mitigation preferences and behaviours, and all additional measures (e.g., one's desire to attend a social justice march or one's donation to a global poverty charity) do not enter into our analyses. Similarly, this applies to behaviours and preferences that may impact climate change in less direct ways. We randomised the order in which people were presented with questions about stated preference measures and behavioural measures to control for order effects, minimising this potential source of bias. We incentivised behavioural measures by having participants make choices with their previously earned real endowment, thus increasing ecological validity.

The first set of items (*stated preference measures*) had two components: First, we collected *stated policy preferences* on US food and drug regulation (FDA), US Senate rules (filibuster), and a carbon tax, with policy preferences on the carbon tax being our preregistered variable of interest. As before, we collected these additional measures to keep the purpose of our study opaque while truthfully stating the topic of our study to participants. Their support was measured on a 5-point Likert scale ranging from '1 – Strongly oppose' to '5 – Strongly support'.

Second, we collected *reported intentions to act* from participant responses on a number of items where we asked them to state on a 5-point Likert scale ranging from '1 – Very unlikely' to '5 – Very likely' whether they were planning to undertake any of the following actions within the next twelve months: attend a protest march to address social justice, donate to charity to reduce global poverty, reduce carbon emissions, quit one's job, or stop eating meat. Our two preregistered variables of interest in this section were the support for a carbon tax and the self-reported intention to personally reduce carbon emissions within the next twelve months. We did not combine any of these measures but rather treated them individually.

Both types of stated preferences were incentivised with The Bayesian Truth Serum (BTS) (Prelec, 2004). The BTS is a mechanism to encourage and incentivise honest responses from participants regarding their subjective preferences, since these can be influenced by various biases or a tendency to provide socially desirable answers. This mechanism works by collecting two types of data: (i) Participants respond to a survey item directly, e.g., indicating their agreement with a statement. (ii) They complete a prediction task where they are asked to estimate the frequency of each answer, e.g., how many percent of the participants will choose 'Disagree' on a Likert-scale item. Participants are rewarded financially for answers that are surprisingly common, i.e., where the mean of (i) exceeds the mean of (ii). In other words, questions where the "actual frequency is greater than its predicted frequency" (Weaver & Prelec, 2013, p. 289). This draws on the Bayesian insight that people are likely to underestimate the prevalence of their own views in the general population. Due to this, surprisingly common answers are more likely to be honest beliefs. The BTS also creates an environment where truth-telling is a Bayesian Nash equilibrium, where, assuming everyone else is responding honestly, telling the truth is the optimal strategy for any participant (Prelec, 2004). The Bayesian Truth Serum has been utilized in diverse domains, including forecasting energy commodity prices (Zhou, Page, Perrons, Zheng, & Washington, 2019) as well as Likert-scale self-report items (Schoenegger, 2023; Schoenegger & Verheyen, 2022) and has been validated in the context of online studies, showing improvements in honest reporting of private information (Frank, Cebrian, Pickard, & Rahwan, 2017).

In our study, we paid out a bonus of £0.15 to those participants who scored in the top third on the surprisingly common metric. The specific additional task that participants were asked to complete that allowed us to calculate the surprisingly common metric was the prediction task, where we asked participants to estimate the mean proportion of answers that other participants would give to a question. Specifically, our instructions were: "For this question, please estimate the average proportion of responses other participants will give. For example, if you think that 10% of other participants will select 'Neutral', then you should enter '10' under 'Neutral'. Importantly, all entries have to sum to '100'."

The second set of items (*behavioural measures*) also had two components: First, we presented participants with three charities that were all drawn from *The Founders Pledge*, a charitable initiative aiming to promote effective charitable giving. This was done to hold constant the factors that may influence donor behaviour such as brand recognition, trust, or previous familiarity with the organisation. We collected their donation choices with respect to the following three charities: The Global Health and Development Fund (global poverty), the Climate Change Fund (climate change), and the Patient Philanthropy Fund (long-term future of humanity), where participants were able to choose to donate to one of these charities or not donate at all. Their donation was capped at the endowment they previously earned, i.e., £0.30.

Second, we collected behaviour measures relating to participant interest in *signing petitions* to address pressing social issues. We presented participants with three real and active petitions and measured whether they clicked the links to those petitions. We chose to use actual petitions and only measured clicks (as opposed to creating new petitions and measuring actual signing) as the former had higher ecological validity while also preserving participant anonymity to a much higher degree (we did not track whether they signed the petitions). Further, we believe that interest in a petition is theoretically interesting in itself and connects up directly to the research question of a potential geoengineering moral hazard by showing direct interest in petitions that impact the respective question at hand. The three petitions included a petition on access to abortions, climate change action, and a reform of the filibuster. Our exclusive variables of interest in this section were the frequency and size of donation to the climate charity and clicks on the climate petition.

Lastly, we collected data on a number of *additional variables*. Those were used as control variables in our main preregistered analyses. In addition to the demographic variables (age, gender, ethnicity), we also collected data on level of education, political identity, belief in anthropocentric climate change, previous knowledge of geoengineering, subjective financial well-being, rurality/urbanicity, trust in government, and trust in science (on top of further variables aimed at keeping the study's purpose opaque such as views on political polarisation and abortion, as well as previous knowledge of CRISPR and the Senate filibuster).

2.6. Analysis methods

We collect our primary outcome variables as follows: frequency of donation to a climate change charity is coded as a dummy, frequency of link clicks to a climate change petition is also coded as a dummy, size of donation with a maximum of £0.30 if a donation had been made to a climate change charity, support for a carbon tax on a 5-point Likert scale, and intention to reduce emissions on a 5-point Likert scale.

For our main regression analyses, we have three condition dummies, though we only enter two in the analysis (Treatment and Salience Control). We control for a variety of factors. First, we control for standard demographic characteristics like age, gender (with '1 = Female'), ethnicity (with 'White' as the comparison group), and education (with 'High School' as the comparison group). Further, we also control for

political orientation via two individual dummies; conservatism (with '1 = Conservative') and liberalism (with '1 = Liberal'), with independents and the respective other orientation coded as the comparison group for both. Further, we also control for urbanicity/rurality (with '1 = Urban') as well as subjective financial well-being (which is coded as a 5-point Likert scale). Further, we add two further central control variables: prior knowledge of geoengineering (as a 5-point Likert scale where increasing scores denote increasing knowledge) and belief in anthropogenic climate change (as a 5-point Likert scale where increasing scores denote increasing belief in anthropogenic climate change).

We generated a total of five regression models. Model (1) has the choice to donate to a climate change charity as its dependent variable (with '1' if such a donation is made, and '0' if a donation is made to another type of charity or if no donation is made). Model (2) predicts behaviour with regard to the size of the donation to a climate change charity (donations can be up to £0.30, and donations to different charities as well as no donations at all are coded as '0'). Model (3) predicts link clicks to a climate change petitions (with '1' if the link has been clicked, and '0' if a link to a different type of petition has been clicked or if no link has been clicked at all). In Appendix A, we present preregistered logit models as robustness checks for our OLS regressions with binary dependent variables,[7] finding virtually identical results, suggesting that our main results are not sensitive to these model choices. In Model (4), support for a carbon tax is our dependent variable (ranging from '1 = Strongly oppose' to '5 = Strongly support'). For Model (5), the dependent variable is self-reported intention to reduce carbon emissions over the next 12 months (ranging from '1 = Very unlikely' to '5' = Very likely').

However, we also provide direct evidence in favour of the null. This is something that a regression analysis in a null-hypothesis testing framework technically cannot provide, which is why we present preregistered equivalence test results to not only provide evidence in favour of a failure to reject the null hypothesis, but instead in favour of the null itself (Lakens, McLatchie, Isager, Scheel, & Dienes, 2020). We use the TOST procedure as our equivalence test (Schuirmann, 1987). TOST stands for 'Two One-Sided Tests', and this procedure aims to statistically support the absence of a meaningful effect. It does this by defining a set of equivalence bounds, the lower bound in the negative direction and the upper bound in the positive direction, jointly setting what range of effect sizes are seen as 'null or negligible' results. These bounds are chosen by a set of plausible smallest effect sizes of interest. Then, this method tests the parameter of interest against both the upper and lower bound in one-tailed tests. If the test for both the upper and the lower bound are statistically significant, one can conclude that the effect is null or negligible, where negligibility is determined by the equivalence bounds.

3. Results

In Table 1, we display the demographics of our final sample. For age, ethnicity, and gender, we also list the corresponding values from the US Census Bureau (2022) in parentheses.

Our sample shows high representativeness with respect to age, ethnicity, and gender. However, we also find that with respect to education, our sample is considerably more educated than the US Census population, with more than twice as many undergraduate degree holders. We also observe a higher proportion of liberals participating (almost twice as many as conservatives). Analysing relationships between these participant characteristics as well as their views about geoengineering and anthropogenic climate change, we find results that

Table 1
Participant characteristics.

	%		%
Age		*Ethnicity*	
18-20	1.9 (4.5)	White	75.7 (75.5)
21-44	49.6 (38.9)	Black	12.3 (13.6)
45-64	34.6 (30.6)	Asian	6.7 (6.3)
65 and over	13.9 (20.9)	Mixed	3.2 (3.0)
		Other	2.1 (1.6)
Gender		*Political Affiliation*	
Male	47.2 (49.6)	Liberal	57.4
Female	51.1 (50.4)	Conservative	23.5
Other	1.7	Independent	19.1
Education		*Financial Wellbeing*	
High school	30.5 (62)	Finding it very difficult	8.6
Undergraduate	48.0 (23)	Finding it quite difficult	13.4
Graduate/Professional	21.5 (14)	Just about getting by	27.8
		Doing alright	36.1
Urbanicity/Rurality		Living comfortably	14.1
Urban	63.5		
Rural	36.5	*Anthrop. Climate Change*	
		Strongly agree	41.2
Knowl. Of Clim. Interventions		Agree	36.4
Strongly agree	18.8	Neither disagree nor agree	11.2
Agree	38.5	Disagree	7.0
Neither disagree nor agree	10.8	Strongly disagree	4.2
Disagree	21.2		
Strongly disagree	10.7		

Notes: Demographics for the full n = 2284 sample after exclusions. For age, gender, ethnicity, and education, we also report the US Census estimate for the relevant category in parentheses.

are very much in line with previous literature. For example, we find that conservatives show both lower belief in anthropogenic climate change, $r = -0.486$, $p < 0.001$, and lower trust in government, $r = -0.230$, $p < 0.001$, and science, $r = -0.388$, $p < 0.001$, while being older, $r = 0.144$, $p < 0.001$, and whiter, $r = 0.102$, $p < 0.001$. We also find that those who knew about SRM prior to this study showed higher trust in science, $r = 0.111$, $p < 0.001$, were younger, $r = -0.078$, $p < 0.001$, and were more likely to be male, $r = -0.089$, $p < 0.001$. We also observe a strong relationship between trust in government and trust in science, $r = 0.422$, $p < 0.001$. These results provide prima facie evidence that we do not have a sample that is fundamentally different from underlying populations that have been studied in previous work, strengthening the validity of our sample.

In Table 2, we outline our main outcome variables split by our three conditions and report mean, frequency, and standard deviation.

We also report a correlation table of all five outcome variables pooled together in Table 3. We find a strong correlation between donation frequency and amount donated, while most other variables correlate at moderate levels. For our main analysis, we do not combine these measures but treat them individually as it is important to potentially capture distinct effects on policy preferences and actual behaviour. Importantly, though, all our dependent variables show some level of correlation, suggesting that they all measure roughly one type of overall behaviour, i.e., climate change mitigation behaviour/preferences.

Below, we present five preregistered regression models testing our two null hypotheses. Each model has one of the five outcome variables as the dependent variable. We enter our condition dummies for Treatment and Salience Control, as well as the full set of control variables. We

[7] We decided to report OLS models in the main document so that we can present more easily interpretable and comparable results, and because results are generally insensitive to model choice (Angrist & Pischke, 2008; Gomila, 2020; Hellevik, 2009).

Table 2
Outcome measures.

	Control	Salience Control	Treatment
Frequency of Donation to Climate Charity	14.8%	14.3%	15.4%
Link Clicks to Climate Petition	15.8%	16.2%	15.3%
Size of Donation to Climate Charity	3.16 (8.48)	3.32 (8.91)	3.40 (8.84)
Size of Donation (if donation)[a]	21.34 (9.93)	23.18 (9.73)	22.08 (9.79)
Support for Carbon Tax	3.80 (1.34)	3.81 (1.33)	3.71 (1.38)
Intention to Reduce Emissions	3.29 (1.42)	3.45 (1.36)	3.25 (1.45)
Sample Size	696	678	910

Notes: Outcome measures for all three conditions with frequency and mean (standard deviation).
[a] This includes only donations where a donation was made.

Table 3
Correlation table of outcome variables.

	Donation Freq.	Donation Amount	Link Click	Carbon Tax Support	Intention to Reduce Emissions
Donation Freq.					
Donation Amount	0.902				
Link Click	0.144	0.143			
Carbon Tax Support	0.232	0.222	0.154		
Intention to Reduce Emissions	0.222	0.209	0.156	0.462	

Notes: Correlation table for all five outcome variables.

did not enter the Content Control dummy as this serves as the reference group for the condition dummies. For interpretation, we focus primarily on the treatment effect of the Treatment condition. We also report our Salience Control treatment effect, which was intended to provide a test for whether any given effect is due to the salience of the treatment text's mention of climate change. For all analyses below, the preregistered threshold for significance is set to p = 0.01 to adjust for multiple comparisons following the Bonferroni method. Thus, we only interpret p-values below 0.01 as significant and will treat any values at or above 0.01 as unequivocally non-significant. See Table 4 for full regression results.

Table 4
OLS regression results for all five outcome variables.

Model	*Variable*	*B* (*SE*)	*β*	*p*
(1) Donation Frequency	Treatment	0.006 (0.017)	0.009	0.711
	Salience Control	−0.004 (0.019)	−0.006	0.817
(2) Donation Amount	Treatment	0.244 (0.427)	0.014	0.568
	Salience Control	0.144 (0.458)	0.008	0.753
(3) Link Click	Treatment	−0.004 (0.018)	−0.005	0.823
	Salience Control	0.007 (0.019)	0.009	0.706
(4) Carbon Tax Support	Treatment	−0.084 (0.049)	−0.030	0.083
	Salience Control	0.022 (0.052)	0.008	0.666
(5) Intention to Reduce Emissions	Treatment	−0.03 (0.063)	−0.011	0.628
	Salience Control	0.155 (0.067)	0.050	0.021

Notes: OLS regression results for five different outcome variables. All models control for all control variables not shown in the table. *p < 00.01, **p < 00.001.

The results suggest that the treatment showed no statistically significant effect on either behavioural outcomes, as evidenced by Models (1) to (3), or stated preferences, as shown by Models (4) and (5), see Table 4. In terms of behavioural outcomes, the coefficients for treatment in relation to both charity decisions and petition link clicks were not statistically significant. Likewise, for stated preferences, neither the support for a carbon tax nor self-reported intentions to reduce emissions were significantly affected by the treatment. The Salience Control condition also showed no statistically significant effects across all five outcomes. These findings collectively point to a failure to detect moral hazard. For a visual representation of all standardised coefficients alongside their standard errors for the Treatment condition, see Fig. 3.[8]

These results for Models (1)–(5) suggest that we did not find evidence that would allow us to reject either null hypothesis as we fail to find that the SRM treatment significantly impacts either behaviour (on any of the three models) or stated preferences (on either of the two models). This is true both before and after adjusting for multiple comparisons and after conducting robustness checks (like logit models for binary dependent variables).

In exploratory analyses, we also test for an effect on a variable we did not preregister as an outcome variable but which might be similarly valid: self-reported intention to reduce eating meat. This exploratory result replicates our preregistered results in that we find a null effect, with the regression results of the same model as above for the intention to reduce meat consumption variable being $B = -0.001$ (0.064), $\beta < -0.001$, $p = 0.989$. As such, we also do not find evidence of a moral hazard effect on the intention to reduce one's meat consumption.[9]

As preregistered, we conduct equivalence tests to provide evidence in favour of the null. Following the method set out by Alter and Counsell (2021), we conduct these tests on the standardised coefficients from Models (1) to (5) and test them against a range of plausible upper and lower equivalence bounds, all in standardised coefficients to allow for

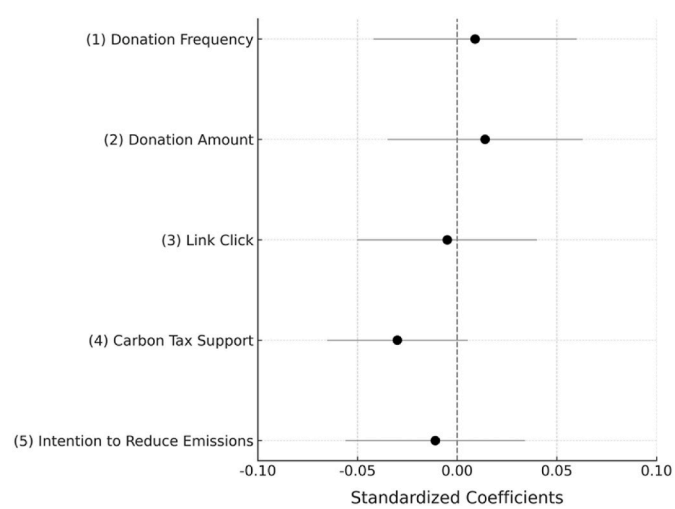

Fig. 3. Forest plot. *Notes:* Standardised coefficients of treatment effect and standard errors for all five outcome variables.

[8] One may worry that given our large number of control variables, that our estimates may be subject to overcontrol bias (Li, 2021). To show that our results are robust to our inclusion of that many (preregistered) control variables, we report the same set of main regressions in Appendix B without any control variables, finding no difference in magnitude, size, or significance of the treatment effect coefficients.

[9] We thank an anonymous reviewer for suggesting including this analysis.

easier cross-model comparison. This approach enables scrutiny of our results across a variety of potentially interesting levels while also allowing us to delineate where we do not have enough evidence in favour of a null or negligibly small effect. This approach is akin to a sensitivity analysis across potential equivalence bounds, which we set to show the effects of this test on a range of plausible bounds. Specifically, we report results for equivalence bounds of (−).01, (−).05, and (−).075. We do not report results of larger equivalence bounds as we already have strong evidence in favour of a negligible effect at the present parameters, even after adjusting for multiple comparisons by putting the level of significance at the 1% level, see Table 5.

Our results suggest that while we do not have evidence in favour of a null (or a negligibly small effect) at the tight equivalence bounds of 0.01, we do already provide such evidence at an 5% alpha-level at the 0.05 equivalence bounds for standardised treatment effects of Model (1), Model (3), and Model (5). Most centrally, all our standardised treatment effect estimates fall within the equivalence bounds of standardised coefficients −0.075 and 0.075 at the adjusted 1% level of significance, indicating strong evidence in favour of a null effect as big as 0.075.

In unstandardised and more easily interpretable terms, this means that the treatment effect in Model (1) is smaller than a 5-percentage point increase or decrease in the probability of donating to a climate-related charity and amounts to an effect smaller than a donation increase or decrease of 1.3 pence in Model (2) (out of a maximum donation of 30 pence). In Model (3), the treatment effect is smaller than a 6-percentage point increase or decrease in the probability of clicking on a link to a climate-related petition. For Model (4), our treatment effect is smaller than a 0.21-point increase or decrease in support with a carbon tax on a 5-point Likert scale. For Model (5), the treatment effect is smaller than a 0.20-point increase or decrease in self-reported intention to reduce greenhouse gas emissions over the next twelve months on a 5-point Likert scale. These results suggest that the effect of being provided with information about SRM does not impact behaviour or preferences to an extent exceeding these estimates. While estimations of policy-relevance are always difficult to make, we argue that these results provide robust evidence in favour of a null at these specified bounds.

We also run exploratory Bayesian analyses aimed at testing the sensitivity of our results by not only relying upon the frequentist approach. In Appendix C, we report Bayes factor model odds for the null models on a number of different priors (uniform, beta binomial, and Wilson), showing strong evidence in favour of the null for Models (1)–(4). The model averaged coefficients similarly replicate our frequentist regression results of showing a null effect. For further details, please see Appendix C.

Table 5
TOST for standardised treatment effects.

	−0.01	0.01	−0.05	0.05	−0.075	0.075
Model (1) Std. Treatment Effect	0.73	0.04	2.27**	1.58*	3.23***	2.54**
Model (2) Std. Treatment Effect	0.96	−0.16	2.56***	1.44	3.56***	2.44**
Model (3) Std. Treatment Effect	0.22	0.65	1.96*	2.39**	3.03***	3.48***
Model (4) Std. Treatment Effect	−1.11	2.22*	1.11	4.44***	2.50**	5.83***
Model (5) Std. Treatment Effect	−0.04	0.91	1.70*	2.65**	2.78**	3.74***

Notes: All t-statistics for TOST procedures on a variety of lower and upper equivalence bounds (in standardized coefficients) of treatment effects from Models (1)–(5). *p < 00.05, **p < 00.01, ***p < 00.001.

4. Discussion

The data collected here do not allow for the rejection of either null hypothesis. Based on preregistered equivalence tests, however, we are able to provide strong evidence in favour of the null that information about SRM does not lead to a reduction (or increase) in climate change mitigation behaviours or stated preferences, and thus as such does not constitute a moral hazard. Moreover, we are able to concretely specify the bounds of these null effects, suggesting that the treatment effect of being informed about SRM is either null or rather small and thus unlikely to be relevant to public policy. These results are corroborated by exploratory Bayesian analyses, providing evidence in favour of the null from a non-frequentist framework.

Our work thus adds to the existing literature which has so far found mixed results, with Raimi, Maki, Dana, & Vandenbergh (2019) finding a reduction in mitigation support, Cherry et al. (2021) finding the converse, and Fairbrother (2016) finding no effect. As argued earlier, our work is the first to jointly implement a number of methodological choices that were present in much of the previous work, those being opaque study design, salience control, and behavioural measures. As such, we argue that our study is among the most rigorous conducted in this field so far, with our data being most in line with Fairbrother (2016) and Andrews et al. (2022). Importantly, both of these papers featured two of the three design features we outlined before, compared to those with fewer than two for the other three papers, suggesting that the more rigorously designed the study, the more likely it is to find no evidence of a moral hazard.

Notably, our results also show a number of secondary relationships that strengthen the validity of our outcome variables and thus our results. For example, we find that those who self-identify as conservative (in the US-political sense) are significantly less likely to support a carbon tax or report intentions to reduce their own emissions. Similarly, higher trust of government also predicts both of these outcomes while higher trust in science only predicts support for a carbon tax. Furthermore, belief in anthropogenic climate change stands in a positive relationship to donating to a climate change charity and to the size of that donation, as well as both support for a carbon tax and intention to reduce emissions. Lastly, previous knowledge of geoengineering only impacts self-reported intentions to reduce emissions positively. These secondary relationships are in line with the extant literature.

In conclusion, our study was designed to mimic an environment as realistically as possible, where a variety of types of information was provided, and a plurality of outcome measures were collected. Overall, we did not find evidence that being provided with information about SRM significantly impacts either stated preferences or actual behaviour. Specifically, we argue that our paper substantially contributes to this literature by presenting a robust null effect. Because our design had several methodological strengths—e.g., usage of behavioural measures (like donations to climate change charities and clicking on links to climate change related petitions), incentivising preference measures (via the Bayesian Truth Serum), keeping the topic of the study opaque (by including a large number of additional questions and texts), and by controlling for salience (by including an intervention only providing information about climate change generally), and because the design was highly powered, preregistered, and used a representative sample alongside appropriate statistical methods to draw conclusions regarding a null effect—we believe that the work presented here represents the best evidence available regarding the question whether SRM information poses a moral hazard. Our answer is that it does not (or that its effect is negligibly small). We hope that this will contribute to a more nuanced discussion where, instead of talking about moral hazards of particular climate measures in isolation, we discuss risks in terms of packages of climate measures (Markusson, McLaren, & Tyfield, 2018; Jebari et al., 2021).

While we have outlined the strengths of our design above, we also want to draw attention to the limitations and downsides of our

approach. In most cases, optimising for one parameter (like experimenter demand) may lead to trade-offs with other worthwhile design goals. For example, one limitation of this design is inherent in its focus on reducing experimenter demand. Specifically, by having aimed so heavily on ensuring that experimenter demand concerns are minimised, the current design may in fact be biasing the results towards the null by making the stimuli themselves too subtle. While we argue that the treatments themselves provided ample reason to think that there was a relatively high chance of detecting very small effects (recall that the a priori power analysis indicated enough power to detect small global effect of $f^2 = 0.01$), we want to point out this trade-off so that readers understand this limitation.

A further, similar, limitation of our design is that by putting so much effort on minimising experimenter demand, this may have induced participant fatigue, which may have driven the results towards a null. While this is certainly possible, we argue that our low failure rates at the comprehension quizzes (at 3.96%) provides some evidence against this worry. One further limitation of our study is that it was conducted entirely in the (online) lab and did not include field experiment aspects. While this is in line with the literature as a whole (e.g., Cherry et al., 2021; Fairbrother, 2016), it is a weakness worth noting. Furthermore, one may worry that more compelling and evocative treatment texts may have led to significant effects. While this may be true, we argue that our choice of treatment text was motivated by having it be neutral and similar to informative media one may encounter in the 'real world'. While we do acknowledge this limitation, we argue that our approach was justified on these grounds.

Another issue is about the target population. We were interested in the general public or citizens; one might think that, instead, what is relevant for climate behaviour is predominantly—or even exclusively—decision-makers. Of course, evidence for the general population *may* apply to decision-makers, but this is uncertain. Furthermore, in many political systems, decision-makers are sensitive to public preferences to greater or lesser degrees, so even if moral hazard is (more) observable amongst decision-makers, they may still be subject to the preferences of publics which are not subject to moral hazard. However, if the population of interest for some readers is predominantly or exclusively decision-makers, this research may be less relevant for them.

Additionally, we want to raise a potential confound in our Treatment and control texts.[10] Since our Treatment text and our Salience Control text present participants with two distinct types of solutions to climate change, this may differentially prime participations about mitigation across these two conditions. If this is true, responses to our outcome variables may thus be, at least in part, due to one's beliefs about the response efficacy of each type of solution (SRM and mitigation). If this is true, this might mean that data patterns in favour or against moral hazard may be explainable to some extent due to this heterogeneity in beliefs. We hope that further research responds to this concern and subjects it to empirical scrutiny.

5. Conclusion

The present paper investigated the risk of a moral hazard in the context of solar radiation management. We studied this in a large US sample, concluding that providing US Americans with information about SRM does not meaningfully impact their behaviour or their stated preferences regarding climate change mitigation behaviour and, therefore, our results indicate that information about solar radiation management does not constitute a moral hazard.

Ethical approval

Granted by the University Teaching and Research Ethics Committee (UTREC) at the University of St Andrews (SA15687).

Funding

We acknowledge funding from the Forethought Foundation and the Center for Effective Altruism. However, nobody except the authors had an impact on study design, analysis, or write-up. Gold Open Access costs are covered by IReL Ireland or KEMÖ Austria. There are no further potential sources of bias to disclose.

Availability of data and materials

Full data are available as supplementary information and materials are enclosed in the appendices.

CRediT authorship contribution statement

Philipp Schoenegger: Data curation, Formal analysis, Funding acquisition, Writing – original draft, Writing – review & editing. **Kian Mintz-Woo:** Funding acquisition, Writing – original draft, Writing – review & editing.

Conflicts of interest

We do not disclose any further potential sources of bias except the funding disclosed above.

Appendix F. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvp.2024.102288.

Appendix A. Logit Robustness Check

Here, we report two preregistered robustness checks, i.e. logit models of Model (1) and Model (3) respectively. These are reported to show that our results are not sensitive to model choice as the outcome variables are binary and as such both approaches would be valid. The results indicate that there is no difference in estimation of treatment effect.

Appendix Table 1
Logistic Regression Robustness Checks

	(6)	(7)
Treatment	0.053 (0.148)	−0.030 (0.142)

(*continued on next page*)

[10] We thank an anonymous reviewer for this helpful suggestion.

Appendix Table 1 (*continued*)

	(6)	(7)
Salience Control	−0.008 (0.160)	0.057 (0.150)
Cox & Snell R^2	0.089	0.035
Sample size	2284	2284

Notes: Log odds and standard errors. *p < 00.01, **p < 00.001.

Appendix B. Robustness Check for Main Regressions without Controls

We also investigate the results of our central regressions without any control variables to test the robustness of our result to a potential overcontrol bias (Li, 2021). We find that coefficients show the same directionality throughout all five models, and the same significance level throughout four models, with Model (11) no longer being significant at the 10% level (which was not interpreted either way due to our move to a 1% level because of adjustment for multiple comparisons). Further, the magnitude difference of all five coefficients is also negligibly small, with 0.009 in the original regression turning into 0.008 when all controls are dropped, and 0.014 into 0.013, −0.005 into −0.007, −0.30 into −0.034, and −0.011 into −0.016 for the four other models respectively. This suggests that our results are not influenced by overcontrol bias.

Appendix C. Bayesian Analyses

We also conduct additional exploratory Bayesian analyses (Rouder & Morey, 2012) using Bayesian linear regression analyses that draw on Bayesian model averaging (e.g., Hinne, Gronau, van den Bergh, & Wagenmakers, 2020; Maier, Bartoš, & Wagenmakers, 2023) to provide further evidence that does not rely on a frequentist framework. We report results for our five basic models that include our treatment dummies as covariates. Below, for simplicity's sake, we report only the null model's results. (It should be noted that each analysis for each outcome variable actually includes four models: the null model (only intercept and error term), a model with only the treatment dummy, one with only the control salience dummy, and one with both.)

We report results with three sets of priors to show the sensitivity of our results to different model prior choices. First, we report the Bayes factor model odds for null model results with a uniform model prior at 0.25. Second, we report the same analyses using a beta binomial model prior at 0.33 which is not biased against sparse and dense models. Third, to provide an even harsher test we also report results with a Wilson prior that is a variant of the beta binomial prior that assigns more mass to models with fewer predictors (Van den Bergh et al., 2020). For the Wilson prior, we set $\alpha = 1$ and $\lambda = 2$ (with $\beta = \lambda$*predictors), and the model prior for the null model is thus set at 0.667. As before, (1) refers to frequency of donation to a climate charity, (2) to the amount of that donation, (3) to clicks on a link to a climate petition, (4) to support for a carbon tax, and (5) to the self-reported intentions to reduce emissions.

APPENDIX TABLE 2
Bayesian Linear Regression Bayes Factor Model Odds for Null Model

	Uniform Prior	Beta Binomial Prior	Wilson Prior
Freq. of Donation to Charity (1)	26.799	34.459	22.753
Donation to Climate Charity (2)	29.241	37.523	24.852
Petition Link Clicks (3)	27.766	35.715	23.570
Support for Carbon Tax (4)	10.823	13.759	9.252
Intention to Reduce Emissions (5)	0.915	1.138	0.791

Notes: Bayes Factor Model Odds for the Null Model with Uniform Prior, Beta Binomial Prior, and Wilson Prior.

The results in Appendix Table 2 suggest that in four of our five models (1)–(4), the odds in favour of the model being the null model after observation of the data have increased by a factor of between 9.252 for Model (4) on a Wilson prior to 37.523 for Model (2) on a beta binomial prior. Model (5), predicting self-reported intentions to reduce emissions, is markedly different in that the Bayes factor model odds to not suggest that the data fit the null model, with a Bayesian model odds factor of between 0.791 at the Wilson prior and 1.138 at the beta binomial prior.

Below we report the model averaged coefficients that allow us to deal with uncertainty over the estimates as well as uncertainty over model choice. We report coefficients as well as 95% credible intervals that represent a weighted average (weighted by the posterior probability of predictor inclusion). We use a JZS parameter prior with the default r scale of 0.354 (Liang, Paulo, Molina, Clyde, & Berger, 2008) and use the uniform model prior to compute the model averaged results.

APPENDIX TABLE 3
Bayesian Linear Regression Coefficients and 95% Credible Intervals

	Treatment	Salience Control
Freq. of Donation to Charity (1)	0.0004 [-0.0015, 0.0021]	−0.0004 [-0.0020, 0.0000]
Donation to Climate Charity (2)	0.0083 [-0.0639, 0.0099]	0.0015 [0.0000, 0.0244]
Petition Link Clicks (3)	−0.0004 [-0.0047, 0.0000]	0.0004 [-0.0001, 0.0011]
Support for Carbon Tax (4)	−0.0154 [-0.1208, 0.0000]	0.0034 [0.0000, 0.0550]
Intention to Reduce Emissions (5)	−0.0138 [-0.1401, 0.0024]	0.1224 [0.0000, 0.2803]

Notes: Model averaged coefficients and 95% credible intervals.

The results in Appendix Table 3 provide additional evidence in favour of a null effect for the Treatment condition across all five outcome variables, though note that for the self-reported intentions to reduce emissions, these results suggest a notable influence of the Salience Control condition, which we did not observe in our preregistered null-hypothesis testing results reported in the main text (and which is also captured in the Bayes factor model odds above, explaining the divergence of results in Model (5)). However, the central estimate of interest is the treatment effect, which is why we take our exploratory Bayesian analyses to provide strong evidence in favour of the claim that there is no moral hazard with regard to being presented with information about climate interventions.

Appendix D. Additional Analyses

We also report the following non-preregistered analysis. In our preregistered regression models, we compare the Treatment and Salience Control to the Content Control. However, in exploratory analyses, we do find a statistically significant effect when we use the Salience Control as the comparison group. Running the same specifications for all five Models (1)–(5), we find that for Model (5) – Intention to Reduce Emissions, the standardised coefficient is significant at the adjusted significance level with $B = -0.063$ ($SE = 0.022$). However, this effect does not in itself constitute evidence for a moral hazard because it actually captures the fact that, empirically, those in Salience Control $M = 3.45$ ($SD = 1.358$) show a higher intention to reduce emissions than both the Treatment $M = 3.25$ ($SD = 1.452$) and the Control $M = 3.29$ ($SD = 1.416$). Because we do not find that the Treatment is lower than the Control, we do not take this as evidence for a moral hazard, but wanted to outline this pattern of results nonetheless.

References

Alter, U., & Counsell, A. (2021). Equivalence testing for multiple regression. Available at: https://psyarxiv.com/ugc9e/.

Andrews, T. M., Delton, A. W., & Kline, R. (2022). Anticipating moral hazard undermines climate mitigation in an experimental geoengineering game. *Ecological Economics, 196*, Article 107421. https://doi.org/10.1016/j.ecolecon.2022.107421

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics*. Princeton University Press.

Austin, M. M. K., & Converse, B. A. (2021). In search of weakened resolve: Does climate-engineering awareness decrease individuals' commitment to mitigation? *Journal of Environmental Psychology, 78*, Article 101690. https://doi.org/10.1016/j.jenvp.2021.101690

Baker, T. (1996). On the genealogy of moral hazard. *Texas Law Review, 75*(2), 237–292.

Brañas-Garza, P., Jorrat, D., Kovářík, J., & López, M. C. (2021). Hyper-altruistic behavior vanishes with high stakes. *PLoS One, 16*(8), Article e0255668.

Butler, R. J., & Worrall, J. D. (1991). Claims reporting and risk bearing moral hazard in workers' compensation. *Journal of Risk & Insurance, 58*(2), 191–204.

Campbell-Arvai, V., Hart, P. S., Raimi, K. T., & Wolske, K. S. (2017). The influence of learning about carbon dioxide removal (CDR) on support for mitigation policies. *Climatic Change, 143*(3–4), 321–336. https://doi.org/10.1007/s10584-017-2005-1

Cherry, T. L., Kallbekken, S., Kroll, S., & McEvoy, D. M. (2021). Does solar geoengineering crowd out climate change mitigation efforts? Evidence from a stated preference referendum on a carbon tax. *Climatic Change, 165*(1–2), 6. https://doi.org/10.1007/s10584-021-03009-z

Fairbrother, M. (2016). Geoengineering, moral hazard, and trust in climate science: Evidence from a survey experiment in britain. *Climatic Change, 139*(3–4), 477–489. https://doi.org/10.1007/s10584-016-1818-7

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160.

Frank, M. R., Cebrian, M., Pickard, G., & Rahwan, I. (2017). Validating Bayesian truth serum in large-scale online human experiments. *PLoS One, 12*(5), Article e0177385.

Gardiner, S. M. (2017). Geoengineering: Ethical questions for deliberate climate manipulators. In S. M. Gardiner, & A. Thompson (Eds.), *The oxford handbook of environmental ethics*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199941339.013.44.

Gomila, R. (2020). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*.

Hale, B. (2012). The world that would have been: Moral hazard arguments against geoengineering. In C. Preston (Ed.), *Engineering the climate: The ethics of solar radiation management* (pp. 113–132). Lexington: Lanham.

Hart, P. S., Campbell-Arvai, V., Wolske, K. S., & Raimi, K. T. (2022). Moral hazard or not? The effects of learning about carbon dioxide removal on perceptions of climate mitigation in the United States. *Energy Research & Social Science, 89*, Article 102656.

Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality and Quantity, 43*(1), 59–74.

Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E. J. (2020). A conceptual introduction to Bayesian model averaging. *Advances in Methods and Practices in Psychological Science, 3*(2), 200–215.

Hudson, P., Wouter Botzen, W. J., Czajkowski, J., & Kreibich, H. (2017). Moral Hazard in Natural Disaster Insurance Markets: Empirical Evidence from Germany and the United States. *Land Economics, 93*(2), 179–208. https://doi.org/10.3368/le.93.2.179

Hulme, M. (2012). Climate change: Climate engineering through stratospheric aerosol injection. *Progress in Physical Geography, 36*(5), 694–705. https://doi.org/10.1177/0309133312456414

Jebari, J., Táíwò, O. O., Andrews, T. M., Aquila, V., Beckage, B., et al. (2021). From moral hazard to risk-response feedback. *Climate Risk Management, 33*, Article 100324. https://doi.org/10.1016/j.crm.2021.100324

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B, 75*(1), 45–57.

Latham, J., Bower, K., Choularton, T., Coe, H., Connolly, P., Cooper, G., et al. (2012). Marine cloud brightening. *Philosophical Transactions of the Royal Society A, 370*, 4217–4262. https://doi.org/10.1098/rsta.2012.0086

Li, M. (2021). Uses and abuses of statistical control variables: Ruling out or creating alternative explanations? *Journal of Business Research, 126*, 472–488.

Liang, F., Paulo, R., Molina, G., Clyde, M. A., & Berger, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association, 103*(481), 410–423.

Lin, A. C. (2013). Does geoengineering present a moral hazard? *Ecology Law Quarterly, 40*(3), 673–712.

Mahajan, A., Tingley, D., & Wagner, G. (2019). Fast, cheap, and imperfect? US public opinion about solar geoengineering. *Environmental Politics, 28*(3), 523–543. https://doi.org/10.1080/09644016.2018.1479101

Maier, M., Bartoš, F., & Wagenmakers, E. J. (2023). Robust Bayesian meta-analysis: Addressing publication bias with model-averaging. *Psychological Methods, 28*(1), 107–122. https://doi.org/10.1037/met0000405

Markusson, N., McLaren, D., & Tyfield, D. (2018). Towards a cultural political economy of mitigation deterrence by negative emissions technologies (NETs). *Global Sustainability, 1*, Article e10. https://doi.org/10.1017/sus.2018.10

Martin, A. (2006). Liquidity provision vs. deposit insurance: Preventing bank panics without moral hazard. *Economic Theory, 28*(1), 197–211. https://doi.org/10.1007/s00199-005-0613-x

Merk, C., Pönitzsch, G., Kniebes, C., Rehdanz, K., & Schmidt, U. (2015). Exploring public perceptions of stratospheric sulfate injection. *Climatic Change, 130*(2), Article 299312. https://doi.org/10.1007/s10584-014-1317-7

Pamplany, A., Gordijn, B., & Brereton, P. (2020). The ethics of geoengineering: A literature review. *Science and Engineering Ethics, 26*(6), 3069–3119. https://doi.org/10.1007/s11948-020-00258-6

Prelec, D. (2004). A Bayesian truth serum for subjective data. *Science, 306*(5695), 462–466.

Quiggin, J. C., Karagiannis, G., & Stanton, J. (1993). Crop insurance and crop production: An empirical study of moral hazard and adverse selection. *Australian Journal of Agricultural Economics, 37*(2), 95–113. https://doi.org/10.1111/j.1467-8489.1993.tb00531.x

Rami, K. T., Maki, A., Dana, D., & Vandenbergh, M. P. (2019). Framing of geoengineering affects support for climate change mitigation. *Environmental Communication, 13*(3), 300–319. https://doi.org/10.1080/17524032.2019.1575258

Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research, 47*(6), 877–903.

Schoenegger, P. (2023). Experimental philosophy and the incentivisation challenge: A proposed application of the bayesian truth serum. *Review of Philosophy and Psychology, 14*(1), 295–320.

Schoenegger, P., & Verheyen, S. (2022). Taking a closer look at the Bayesian truth serum: A registered report. *Experimental Psychology, 69*(4), 226–239.

Schuirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics, 15*, 657–680.

Svoboda, T. (2017). *The Ethics of Climate Engineering: Solar radiation management and non-ideal justice*. New York: Routledge. https://doi.org/10.4324/9781315468532

Van den Bergh, D., Clyde, M. A., Gupta, A. R. K. N., de Jong, T., Gronau, Q. F., Marsman, M., … Wagenmakers, E. J. (2020). A tutorial on Bayesian multi-model linear regression with BAS and JASP. *Behavior Research Methods, 53*(6), 2351–2371. https://doi.org/10.3758/s13428-021-01552-2

Weaver, R., & Prelec, D. (2013). Creating truth-telling incentives with the Bayesian truth serum. *Journal of Marketing Research, 50*(3), 289–302.

Zhou, F., Page, L., Perrons, R. K., Zheng, Z., & Washington, S. (2019). Long-term forecasts for energy commodities price: What the experts think. *Energy Economics, 84*, Article 104484.

Zweifel, P., & Manning, W. G. (2000). Moral hazard and consumer incentives in health care. In A. J. Culyer, & J. P. Newhouse (Eds.), *Handbook of health economics*. Boston: Elsevier.