Climatic Change (2023) 176:138
https://doi.org/10.1007/s10584-023-03612-2

Climate cooperation with risky solar geoengineering

Todd L. Cherry[1,2] · Stephan Kroll[3] · David M. McEvoy[4]

Received: 25 April 2023 / Accepted: 8 September 2023 / Published online: 2 October 2023

Abstract
Given the lack of progress on climate change mitigation, some scientists have proposed solar geoengineering as a means to manage climate change at least temporarily. One main concern with such a risky technological solution, however, is that it may create a "moral hazard" problem by crowding out efforts to reduce emissions. We investigate the potential for a risky technological solution to crowd out mitigation with theory and experiments. In a collective-risk social dilemma, players strategically act to cooperate when there is an opportunity to deploy a risky technology to help protect themselves from impending damages. In contrast to the moral hazard conjecture, the empirical results suggest that the threat of solar geoengineering can lead to an *increase* in cooperative behavior.

Keywords Collective risk · Social dilemma · Public goods · Experiments · Solar geoengineering

1 Introduction

Despite continued negotiations, global commitments to reduce greenhouse gases (GHGs) are still well below what is required to prevent dangerous temperature increases. For this reason, there is a growing public debate about whether the global community should consider the use of solar geoengineering—a method of introducing particles into the atmosphere to reflect sunlight away from Earth—to help avoid catastrophic climate change. While some scientists view solar geoengineering as a temporary bridge to keep temperatures below catastrophic levels until sufficient reductions in GHG emissions are realized (Wagner 2021; Keith 2021), others have called for an "International Non-Use Agreement" on solar geoengineering (Biermann et al. 2022). Concerns about the technology and particularly the risk of its deployment are manifold. In a summary of a report by a working group for the Forum for Climate Engineering Assessment (Chhetri et al. 2018), Jinnah

✉ Todd L. Cherry
tcherry@uwyo.edu

[1] Department of Economics, University of Wyoming, Laramie, WY, USA

[2] CICERO - Center for International Climate Research, Oslo, Norway

[3] Department of Agricultural and Resource Economics, Colorado State University, Fort Collins, CO, USA

[4] Department of Economics, Appalachian State University, Boone, NC, USA

et al. (2019) list the fear of "undesirable environmental impacts and inequitable distribution of burdens and benefits, as well as risk of geopolitical conflict and technological lock-in" and, in some circumstances, "termination shock"; additional risks from different poorly designed or implemented governance arrangements; and, not least, a "moral hazard" effect, in which solar geoengineering research and deployment "distracts from mitigation efforts."[1, 2] The flip side to the moral hazard concern is that the perceived risks of costly side effects associated with the technology may prompt more mitigation to diminish the incentive to deploy the technology (Reynolds 2019), something that Wagner and Zizzamia (2022) call "inverse moral hazard."

It is an open empirical question whether the threat of deploying risky geoengineering impacts mitigation efforts and how these efforts respond to changes in risk. Given that solar geoengineering has never been deployed and the threat of deployment is not yet salient, social scientists must turn to simulated environments to better understand behavioral responses to risk and risky technologies. We employ two methods suited to consider the strategic behavior associated with the introduction of new risky technological solution—i.e., moral hazard. First, following a robust literature (e.g., Millard-Ball 2012; Moreno-Cruz 2015; Heutel et al. 2016), we develop a game-theoretic model that captures the key features of SGE that underpin the strategic considerations of SGE. Second, following an emerging literature (e.g., Cherry et al. 2022; Andrews et al. 2022), we design a controlled laboratory decision experiment that mimics the SGE decision environment to test the theoretical predictions. Theory and experiments are simplifications of a more complex reality that allows researchers to isolate causal effects that explain strategic behavior with alternative technologies and policies (Falk and Heckman 2009). Experiments have been shown to be effective policy testbeds and tend to allow for generalizable inferences about behavior (Snowberg and Yariv 2021), which is particularly useful when field observations are unattainable.

In this paper, we report results from an economic experiment designed to investigate cooperation in a social dilemma with the introduction of a risky technological solution.[3] An emerging literature offers some evidence on how the threat of geoengineering impacts mitigation decisions. However, existing studies generally rely on survey responses to hypothetical scenarios (e.g., Fairbrother 2016; Cherry et al. 2021) or consumer responses to information provision (e.g., Merk et al. 2016; Austin and Converse 2021). These studies examine individual decisions made in isolation, which fails to capture the strategic considerations inherent in collective action problems like mitigation and solar geoengineering. To our knowledge, there are only three studies that use controlled lab experiments to capture

[1] The environmental risks associated with solar geoengineering include the potential for harmful effects to the ozone layer (Keith et al. 2016), greater ocean acidification (Williamson and Turley 2012), changes in global precipitation patterns (Irvine et al. 2019), and the long-term oscillations of natural climate systems (Gabriel and Robock 2015). See also the recent report by the National Academies of Science, Engineering and Medicine (NASEM, 2021).

[2] In a recent article, David Keith (2021) summarizes the issue: "Perhaps the central concern about solar geoengineering is that deployment, or even the credible possibility of deployment, will slow emissions cuts. This concern—moral hazard, or mitigation inhibition—arises from political links between decisions about solar geoengineering and emissions cuts in the face of climate risks, […]." Lin (2013) is among the first papers that examine whether geoengineering presents a moral hazard and how to ameliorate this moral hazard.

[3] We follow the literature and use the term "cooperation" to denote cooperative behavior that contributes (i.e., mitigation) to solving the collective action problem (i.e., climate change). Cooperative behavior may result from different motives, such as spite, self-interest, and altruism.

both individual and group decisions when investigating questions directly related to solar geoengineering.

Abatayo et al. (2020) conduct an experimental test of the "free-driver" hypothesis—the conjecture that the country that prefers the highest level of solar geoengineering will deploy the technology to a point that exceeds the social optimum. They find evidence in support of the free-driver hypothesis, but they do not explore the impact of geoengineering on mitigation or consider the potential for costly side effects from the technology. Andrews et al. (2022) and Cherry et al. (2022) both explicitly test the moral hazard conjecture and find that the threat of solar geoengineering either has no effect or increases investment in mitigation. Andrews et al. (2022) explore a world in which solar geoengineering may or may not be successful but can do no harm. They find that solar geoengineering opportunities did not reduce mitigation. Cherry et al. (2022) model solar geoengineering as a good or a bad (a "GoB") depending on the realized level. Contrary to the moral hazard conjecture, they find that the threat of the technology causes an increase in mitigation investment. Importantly, both studies do not consider uncertain and costly side effects from geoengineering (which is one of the major arguments against these technologies). Moreover, none of the existing experimental studies examine how heterogeneity in wealth may influence behavior surrounding solar geoengineering.

Our experimental study explores how the opportunity to deploy a risky technology (solar geoengineering) impacts contributions to a public good (mitigation) and how this behavior changes in response to changes in risk levels. We employ a collective-risk social dilemma in which players can avoid impending damages (climate change) by contributing to a public good. We consider a world with and without the possibility of using the quick, cheap, and risky technology, and we vary the probability that this technological solution imposes costly side effects. To illustrate the moral hazard argument, we present a simple game-theoretic model, which predicts that the availability of the technological solution either weakens the incentive for risk-neutral subjects to mitigate or does not change the incentives at all. Empirically, we find the opposite in most scenarios: the threat that somebody might deploy the technology when there is a substantial downside risk triggers an increase in mitigation efforts. We are able to map out the relationship between investments in mitigation and the expected net benefits of using the technological solution. The only instance in which we do not observe a significant increase in mitigation is when there is very little risk of costly side effects from solar geoengineering.[4] We also explore the impact of heterogeneity in wealth among players by varying the distribution of starting endowments. Higher endowment players tend to contribute more to mitigation, and heterogeneity in endowments does not negatively impact group-level mitigation.

In the next section, we introduce the mitigation and solar geoengineering games. In section 3, we present the testable predictions that are derived from a game-theoretic model presented in the Supplementary Information (SI).[5] The results are presented in Sect. 4, and Sect. 5 concludes.

[4] Note that another concern of geoengineering skeptics is that the risk of this technology is unevenly divided and falls mainly on poorer countries (Biermann et al. 2022). The focus of the current paper is on the impact of risk in general, so we assume homogeneity in downside risk. Related research (Cherry et al. 2022) considers that some countries are more (negatively) impacted by the deployment of solar geoengineering than others.

[5] All supplemental information (SI), including theoretical model and predictions, additional analysis, data and code, and experimental instructions, is available on the Open Science Framework (OSF) at https://osf.io/ve9kw/?view_only=3862b1e218a346939d9dc30cfbf73bfb.

2 The games

Consider a world in which a group of players face an impending costly disaster. In one situation, the only opportunity groups have to protect themselves is through marginal contributions to a group account that can protect all players equally. We call this the "mitigation game." In another situation, groups have an additional option to try to protect themselves from the impending disaster by deploying a technological solution. The technology is free to deploy but it is risky. We call this two-stage game the "mitigation and solar geoengineering game." We introduce the two games in turn.

2.1 Mitigation game

The *mitigation game* draws from the literature on collective-risk social dilemmas (e.g., Brown and Kroll 2017; Milinski et al. 2008). Players are endowed with a starting balance of tokens in a personal account, which changes depending on the decisions and outcomes in the game. Tokens are exchanged for money at the end of the game. Players are in groups, and each member faces the threat of losing some or all of the tokens in their personal account. To protect themselves from the potential losses, group members choose to keep their tokens or contribute them to a group account. Each token contributed to the group account does two things: it removes one token from the contributing member's personal account, and it protects 1% of the tokens in all group members' personal accounts. Thus, contributions benefit all members of the group, but it is costly to the individual making the contribution.

There are two features of the game to note. First, more contributions provide more benefits to the group, but only up to an upper limit. Once contributions reach a threshold, the risk of loss from the threat is eliminated. Second, as contributions approach the threshold, the potential loss from the threat is reduced proportionally.[6]

The prediction for contributions in the mitigation game is determined by the noncooperative Nash equilibrium (see SI). While the exact prediction depends on group size and endowment levels, contributions in our setting are predicted to be insufficient to avoid damages but greater than zero. The mitigation game presents players with a way to achieve the socially optimal outcome by collectively acting to avoid impending damages, but theory predicts a suboptimal outcome because individual members can free ride off the actions of others. Therefore, the game is a collective-risk social dilemma that captures the fundamental tension of the collective action needed to mitigate climate change and avoid impending damages.

2.2 Mitigation and solar geoengineering game

In the *mitigation and solar geoengineering game*, there are potentially two stages. Stage 1 is identical to the previously described mitigation game. However, when making their decisions in stage 1, group members know that if the contributions to the group account are insufficient to avoid impending damages, they will have an opportunity in stage 2 to deploy a risky technology (solar geoengineering) that may prevent damages. Note that if group

[6] Potential losses decline proportionally as contributions approach the threshold, which corresponds to more mitigation leading to reduced potential damage from climate change.

contributions in stage 1 reach the threshold to avoid damages, there is no stage 2—mitigation is sufficient to avoid damages from climate change, and therefore, there is no need for solar geoengineering.

In stage 2, group members decide simultaneously and independently whether or not to deploy a risky but costless technology that, if successful, protects all members from damages. Thus, if only one member of the group deploys the technology, the effects are imposed on everyone.[7] Setting the cost to deploy the technology at zero is the simplest and most salient way to capture how relatively cheap solar geoengineering is (Barrett 2008; Weitzman 2015). The risk of deployment is that, while it can succeed and therefore result in a good outcome, there is a chance it will fail and result in a bad outcome that is worse than the damages it was supposed to prevent. A good outcome avoids any damage and group members keep the tokens in their private accounts. A bad outcome causes all group members to lose all the tokens in their private account. It is all or nothing.[8] The probability of a good outcome (denoted as π in the theoretical model presented in the SI) is common for all players in a group. It is not dependent on how many group members deploy the technology. Specifically, if any player(s) in the group deploys the technology, there is a probability (π) of a good outcome in which all group members keep the tokens in their private accounts after their contributions in stage 1, and there is a probability ($1 - \pi$) of a bad outcome in which all group members lose the tokens in their private account. Importantly, when players make mitigation decisions in stage 1, they know the probability of a good outcome from the technological option in stage 2.

To illustrate the trade-offs that group members face in stages 1 and 2, we use a benchmark game-theoretic model of the *mitigation and solar geoengineering game* with risk-neutral agents (see SI), which we solve for the subgame-perfect equilibrium with backward induction that starts with outcomes in stage 2. A risk-neutral group member will deploy the technology in stage 2 if the expected payoff from doing so is larger than the certain payoff from the contribution in stage 1. If in equilibrium the technology is deployed in stage 2, then group members contribute nothing to the group account in stage 1. If in equilibrium the technology is not deployed in stage 2, then group members make the same contributions to the group account as they would in the mitigation game. Thus, if deployment is expected, the model predicts mitigation to be lower than expected in the mitigation game, and if deployment is not expected, the model predicts mitigation to be the same as in the mitigation game. In the next section, we provide the exact subgame-perfect Nash equilibria for the parameters used in the experiment.

[7] This reflects the often-cited concern by geoengineering skeptics that "[g]iven the anticipated low monetary costs of some of these technologies, such as stratospheric aerosols injection, a few countries could engage in solar geoengineering unilaterally or in small coalitions even when other countries oppose such deployment" (Biermann 2022).

[8] In our model and experiments, solar geoengineering is a binary choice and is either universally good or universally bad depending on the outcome from deploying the technology. This approach differs from other models and experiments that specify a distribution of "preferred" levels of solar geoengineering and continuous choices (e.g., Weitzman 2015; Abatayo et al. 2020; Cherry et al. 2022). Our simplified approach allows us to focus on the impact of potentially costly side effects on mitigation decisions. We vary risk by varying the probability of a bad outcome while holding the severity constant. Future research should consider varying the size of losses while holding the probability constant, as well as interacting probability and severity.

Table 1 Equilibrium predictions

Treatment	Stage-decision	Endowment: Homogeneous No spread [50, 50, 50, 50]	Endowment: Heterogeneous Low spread [67, 67, 33, 33]	Endowment: Heterogeneous High spread [75, 75, 25, 25]
Mitigation	1-contribution 2-deploy?	10, 10, 10, 10 (40) n/a	22, 22, 0, 0 (44) n/a	25, 25, 0, 0 (50) n/a
SGE				
π = 0.9	1-contribution 2-deploy?	0, 0, 0, 0 (0) Yes	0, 0, 0, 0 (0) Yes	0, 0, 0, 0 (0) Yes
π = 0.5	1-contribution 2-seploy?	0, 0, 0, 0 (0) Yes	0, 0, 0, 0 (0) Yes	Ambiguous
π = 0.3	1-contribution 2-deploy?	10, 10, 10, 10 (40) No	22, 22, 0, 0 (44) No	25, 25, 0, 0 (50) No
π = 0.1	1-contribution 2-deploy?	10, 10, 10, 10 (40) No	22, 22, 0, 0 (44) No	25, 25, 0, 0 (50) No

Predicted individual contributions are listed in the same order as individual endowment levels; predicted group contributions are in parentheses.

3 Experimental design and predictions

The experimental design consists of two treatments that correspond to the two games just described. The baseline *mitigation treatment* is the mitigation game that has one stage of decision making (mitigation), and the *SGE treatment* is the mitigation and solar geoengineering game that has the potential for two stages (mitigation and deployment). See the SI for the experimental instructions.

In both treatments, participants are randomly assigned to groups of four ($n = 4$), and groups have either *homogenous endowments* or *heterogeneous endowments*. Previous experimental studies use homogeneous endowments with different desired levels of geoengineering. We consider heterogeneity in endowments to reflect the reality of countries having vastly different resources to deal with climate change.[9] In the homogeneous cases, each player starts with 50 tokens in their private account. In the heterogenous cases, two players are “high-endowment” members, and two are “low-endowment” members. Following the literature (e.g., Brown and Kroll 2017), we consider two different distributions of heterogeneous endowments: in the *low-spread heterogeneity* case, the high endowment is twice the size of the low endowment $(67, 67, 33, 33)$, and in the *high-spread heterogeneity* case, the high endowment is three times the size of the low endowment $(75, 75, 25, 25)$. Note that the sum of token endowments in the group is always 200.

To summarize, both mitigation and SGE treatments have three endowment distributions—homogeneous, low-spread heterogeneous, and high-spread heterogeneous. Given these general parameters, we use the game-theoretic model to solve for the equilibrium predictions in our two treatments. Predictions are derived in the SI, summarized below, and presented in Table 1.

[9] Previous experimental studies consider homogeneous endowments while introducing heterogeneity in the desired levels of solar geoengineering without costly side effects (Cherry et al.,2022; Abatayo et al. 2020).

In the baseline *mitigation treatment*, participants independently and simultaneously decide how many tokens from their private account to contribute to the group account. In case of homogeneous endowments in which all group members start with 50 tokens, the equilibrium prediction is that each player will contribute 10 tokens, for a total group contribution of 40 tokens. In equilibrium, group members keep 40% of their remaining token endowment. In the heterogeneous endowment cases, theory predicts that group members with relatively low endowments will contribute 0 tokens to the group account. For members with relatively high endowments, equilibrium contributions depend on the distribution of endowments. As reported in Table 1, theory predicts that high-endowment members will contribute 22 and 25 tokens in the low- and high-spread heterogeneous cases, respectively. This translates to predicted group contributions of 44 and 50 tokens in the low- and high-spread cases, respectively.

In the *SGE treatment*, there is the potential for two stages—mitigation and SGE deployment. Players contribute to mitigation in stage 1 knowing the probability of a good outcome if deployment occurs in stage 2. Our experimental design considers the risk of SGE by varying the probability of a good outcome (π) which is either 0.9, 0.5, 0.3, or 0.1.[10] Given the parameters, we derive subgame-perfect equilibrium predictions for risk-neutral players (see SI), which are summarized in Table 1.[11]

When a good outcome from SGE deployment is highly likely ($\pi = 0.9$), the benchmark theoretical model predicts that group members will contribute nothing to the group account in stage 1 and the SGE technology will be deployed in stage 2, independent from the endowment conditions. Predictions largely remain the same when the probability of a good outcome is 50–50 ($\pi = 0.5$), with the one exception being an ambiguous result in the high-spread heterogeneous endowment case. When the likelihood of a good outcome is relatively low ($\pi = 0.3$ and $\pi = 0.1$), theory predicts positive contributions by the group in stage 1 and no deployment of SG technology in stage 2. Also, individual and total contributions depend on the distribution of endowments—greater inequality increases total contributions via higher individual contributions from high-endowment members.

The key takeaway from the benchmark theoretical model is that, with the assumption of risk-neutral players, the availability of the technological solution might reduce contributions to mitigation when a good outcome from the technology is more likely ($\pi = 0.5$ or $\pi = 0.9$) and has no effect on contributions when a good outcome is less likely ($\pi = 0.1$ or $\pi = 0.3$). Thus, the introduction of SGE might crowd out mitigation or, at best, have no impact on mitigation. Theoretical predictions from such a model therefore are suggestive of a “moral hazard” response. However, behavioral research provides ample evidence that observed behavior often deviates from theory in systematic ways, so we turn to controlled experiments to gain behavioral insights.

We close this section with a summary of the key features of the experiment. Table 2 summarizes the design, and the experimental instructions are provided in the SI. The experiment used a program designed specifically for this study on z-Tree (Fischbacher 2007). Subjects were recruited using ORSEE (Greiner 2015) from a pool of registered participants to participate in-person in one of six sessions. Subjects did not have

[10] Given the ongoing debate about the potentially high risk of deploying solar geoengineering technologies (e.g., Biermann et al. 2022), we parameterized the experiment to investigate both extremes ($\pi = 0.1$ and $\pi = 0.9$), a central value ($\pi = 0.5$), and we explored one additional relatively high-risk option ($\pi = 0.3$).

[11] At the end of the experiment, we measured the level of risk aversion, using the Eckel and Grossman (2008) elicitation exercise (see Dave et al. 2010).

Table 2 Experimental design

	Mitigation	*SGE*
Subjects	56	48
Groups	14	12
Sessions	3	3
Group-level obs	168	144
Individual-level obs	672	576
Endowments	50, 50, 50, 50—rounds 3, 6, 8, 11 67, 67, 33, 33—rounds 1, 9, 10, 12 75, 75, 25, 25—rounds 2, 4, 5, 7	
Probability of good outcome (π)	– – – –	0.9—rounds 4, 8, 12 0.5—rounds 2, 10, 11 0.3—rounds 1, 6, 7 0.1—rounds 3, 5, 9

experience with the game and did not communicate during the game. All decisions were made in private. Subjects played twelve rounds of the game, with groups being reshuffled each round to avoid reputation effects. The distribution of endowments varied each round, but to maintain saliency, subjects had either a high or low endowment for all rounds with heterogeneity. Individual endowments were announced at the start of each round. If SGE is deployed in stage 2, a computerized random draw determined whether the good or bad outcome was realized.[12]

With three different endowment combinations and four probabilities, there are twelve endowment-probability combinations. In the twelve rounds of the *mitigation treatment* without stage 2, subjects experience each endowment combination four times. In the twelve rounds of the *SGE treatment* with the possibility of stage 2, subjects experience each endowment-probability combination once. The order of the combinations was randomly determined for the first session and then copied for all following sessions—the "pseudo-random" approach (see Cox et al. 2001 and Kroll and Shafran 2018). To incentivize subjects to pay close attention in each round, it was announced that final earnings were determined from one randomly selected round. With a total of 104 subjects, we observe 1248 individual-level and 312 group-level observations. Sessions lasted less than 1 h, and subjects earned an average of $20.

4 Results

We first examine group-level contributions to mitigation. Consistent with other collective-risk social dilemma experiments (e.g., Brown and Kroll 2017), mitigation contributions are higher than equilibrium predictions in all scenarios. Table 3 reports the average group contributions by treatment, endowment distribution, and probability of success. Qualitatively, for three of the four levels of likely success ($\pi = 0.5$,0.3, and 0.1), the numbers show that the availability of SGE *increases* cooperative behavior in the first stage. This is

[12] Outcomes from the random draw were not announced until the end of the session because studies show the realization of a random outcome, good or bad, in one round may have an impact on behavior in following rounds, even though round-specific payoffs are independent (e.g., Kroll and Shafran 2018).

Table 3 Average group contributions to mitigation

	All endowments	$E=[50, 50, 50, 50]$	$E=[67, 67, 33, 33]$	$E=[75, 75, 25, 25]$
Mitigation	77.88	80.00	75.55	78.09
SGE, pooled	86.49	87.21	84.10	88.17
$\pi=0.9$	65.39	63.75	56.08	76.33
$\pi=0.5$	93.22	89.83	93.83	96.00
$\pi=0.3$	92.33	95.92	93.58	87.5
$\pi=0.1$	95.03	99.33	92.92	92.83

not consistent with the moral hazard conjecture. The one exception is the highest probability of success ($\pi = 0.9$), when the availability of SGE led to a *decrease* in contributions to mitigation. Also, from Table 3, we note these aggregate findings are not particularly sensitive to whether endowments are homogeneous or heterogeneous. We scrutinize this finding with a conditional analysis of group contributions that yield estimates conditional on the likelihood of success, endowment distribution, round-specific effects, and observational dependence with robust standard errors (see SI). The conditional estimates corroborate the finding that the availability of SGE *significantly increases* group contributions to mitigation for three of the four SGE success rates ($\pi = 0.5$, 0.3, and 0.1). The one exception is the highest level of success ($\pi = 0.9$), and in this case, SGE has no significant effect on mitigation by groups.

Figure 1 shows the impact of SGE on group contributions to mitigation by illustrating the average group contributions over the twelve rounds. The round number is on the horizontal axis and contributions on the vertical axis. The solid diamonds and trend line are from the *mitigation treatment*. The other data points are averages from the *SGE treatment* and differ by the probability of a good outcome from deploying the technological solution. Different colors represent the three cases of heterogeneity. Consistent with the summary in Table 3, the only rounds of the *SGE treatment* that have lower contributions compared to the *mitigation treatment* are those that have a 0.9 probability of a good outcome. This result holds across the different cases of heterogeneity.

Next, we consider the ability of groups to successfully contribute enough to mitigation to completely protect themselves from losses. Recall, meeting (or exceeding) the contribution threshold allowed participants to keep their entire remaining endowments and therefore avoiding the need to rely on SGE. The results in Table 4 show a dramatic increase in the likelihood of groups reaching the threshold when they had the opportunity to deploy SGE. Apart from the scenario with a high probability of success ($\pi = 0.9$), the percentages in the *SGE treatment* are notably higher than the *mitigation treatment*. When pooling the *mitigation treatment* over all endowment distributions, only 13 out of 168 groups (7.74%) made it to the threshold (100 tokens). This is markedly lower than the pooled success rate in the *SGE treatment* of 29.17% (42 out of 144 groups). We again scrutinize this finding with a conditional analysis that estimates the effect of SGE and endowment distribution has on the likelihood of a group mitigating enough to avoid losses (see SI). The conditional estimates follow the aggregate numbers. SGE has a significant positive impact on the probability that a group will contribute to mitigation enough to avoid losses. The one exception is when SGE has the highest level of likely success ($\pi = 0.9$), in which SGE has no significant effect. This finding is consistent across endowment distributions.

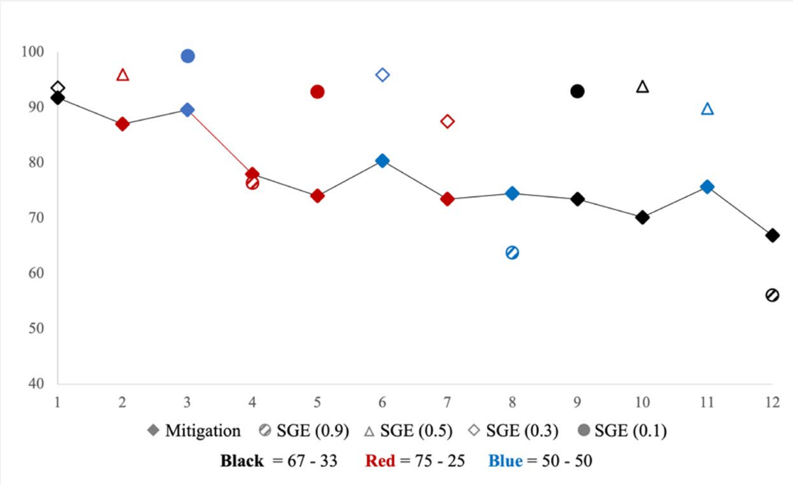

Fig. 1 Average group contributions by treatment and probability of success

Table 4 Percentage of rounds that groups met the threshold

	All endowments	$E=[50, 50, 50, 50]$	$E=[67, 67, 33, 33]$	$E=[75, 75, 25, 25]$
Mitigation	7.74	7.14	7.14	8.93
SGE, pooled	29.17	33.33	22.92	31.25
$\pi=0.9$	8.33	8.33	0.00	16.67
$\pi=0.5$	36.11	25.00	25.00	58.83
$\pi=0.3$	33.33	50.00	33.33	16.67
$\pi=0.1$	38.89	50.00	33.33	33.33

We now turn to individual-level contributions to mitigation. To examine mitigation decisions over the twelve rounds, we use generalized least-squares panel models to estimate the level of individual contributions as a function of the likelihood of SGE success and individual risk preferences. Estimates control for round-specific effects and use robust standard errors (see SI for details). Table 5 reports results for all endowments and by specific endowment distribution.

Overall, estimates indicate the effect that SGE has on individual contributions varies across the likelihood of success and the distribution of endowments. Three key results emerge across all models. First, except when SGE is very likely to succeed ($\pi=0.9$), 16 of the 18 estimated coefficients are positive, indicating that the availability of SGE tends to increase individual contributions to mitigation. Ten of the 16 estimated positive effects are significant, while neither of the two negative estimates are significant. Second, when SGE is most likely to succeed ($\pi = 0.9$), contributions weakly decrease (one of six estimated coefficients is significant). Estimates offer some indication that SGE may crowd out mitigation only at high probabilities of success. Third, with heterogeneous endowments, the positive impact that SGE has on individual contributions is observed among those with relatively high endowments. As expected, these findings correspond closely to the group-level results, and we find little evidence that the threat of SGE will decrease

Table 5 Regression analysis on individual contribution levels

	All endowments	50–50 endowments	67–33 endowments		75–25 endowments	
			High	Low	High	Low
SGE, 0.9	−2.03 (2.00)	−2.88 (1.98)	−1.69 (4.10)	−3.96** (1.73)	−0.018 (3.90)	−1.06 (1.69)
SGE, 0.5	3.80** (1.80)	3.33** (1.65)	10.71*** (3.01)	0.863 (2.06)	6.14* (3.61)	−1.92 (1.58)
SGE, 0.3	2.51 (1.62)	3.68*** (1.32)	3.53 (2.50)	−2.90 (1.82)	6.51** (2.87)	0.250 (1.40)
SGE, 0.1	3.90** (1.84)	2.23 (1.53)	6.78** (2.99)	2.70* (1.60)	7.41** (3.47)	1.73 (1.60)
Risk preference	−0.34 (0.58)	−0.703* (0.414)	−0.424 (0.691)	−0.248 (0.468)	−0.438 (0.675)	−0.255 (0.430)
Constant	23.47*** (2.72)	21.72*** (2.06)	28.63*** (3.27)	20.28*** (2.22)	29.80*** (3.50)	9.71*** (1.94)
N	1,248	416	208	208	208	208

All regressions control for round fixed effects (not included in table) and use robust standard errors.

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

individual and group contributions to mitigation. To the contrary, in three of the four risk scenarios, we observe an increase in contributions in response to the availability of the risky technology.

One question arises: why do individuals increase their mitigation efforts, contrary to theoretical predictions? One conjecture would be that they are afraid of “rogue actors” to deploy SGE unilaterally if the threshold is not met. To assess whether such a fear is warranted, we conclude our analysis by examining the individual-level decisions to deploy SGE. Table 6 shows the proportion of cases that individuals deployed SGE by endowment type and probability of SGE success. When the probability of success is the highest ($\pi = 0.9$), we observe a high deployment rate across endowment types, ranging from 70.8 to 85%. Unsurprisingly, when the probability of success drops, so too does

Table 6 Proportion deploying SGE by endowment type and probability of success

	All endowments	*E*=75	*E*=67	*E*=50	*E*=33	*E*=25
SGE, 0.9	0.795 (0.035) [132]	0.85 (0.082) [20]	0.708 (0.095) [24]	0.773 (0.064) [44]	0.833 (0.078) [24]	0.85 (0.082) [20]
SGE, 0.5	0.153 (0.038) [92]	0.30 (0.153) [10]	0.056 (0.056) [18]	0.083 (0.047) [36]	0.111 (0.076) [18]	0.50 (0.167) [10]
SGE, 0.3	0.031 (0.018) [96]	0.05 (0.05) [20]	0.063 (0.063) [16]	0.00 (0.00) [24]	0.063 (0.063) [16]	0.00 (0.00) [20]
SGE, 0.1	0.068 (0.027) [88]	0.00 (0.00) [16]	0.125 (0.085) [16]	0.083 (0.058) [24]	0.063 (0.063) [16]	0.063 (0.063) [16]
N	408	66	74	128	74	66

Standard errors in parentheses and number of observations in brackets the proportion of players that deploy SGE, but, importantly, not to 0. Given $\pi = 0.5$, the percentage of deployment ranges from 5.6 to 50% with an average of 15.3%. Overall, the percentage of deployment drops from 79.5 to 15.3% when the likelihood of success drops from 0.9 to 0.5 (t-test, $p=0.000$). Moving from a 0.5 to a 0.3 success rate causes deployment to drop from 15.3 to 3.1% ($p=0.004$). The deployment percentage is statistically equivalent under the 0.3 and 0.1 success rates ($p=0.248$). Overall, results indicate that the likelihood of SGE success matters only when success becomes highly likely and that endowment levels have little impact on this relationship.

5 Conclusion

Solar geoengineering has been described as a "fast, cheap, and imperfect" solution to the climate problem (Wagner 2021). One of the arguments against considering solar geoengineering as part of a strategy to manage climate change is the "moral hazard" conjecture, which contends that serious consideration or deployment of solar geoengineering can crowd out incentives to reduce emissions. To help better understand cooperative behavior under the threat of a risky technological solution, our study uses controlled experimental methods. Through experiments, we are able to move beyond previous individual-level survey studies to consider important strategic interactions that define the collective action nature of climate policy while observing how the threat of deploying a cheap but risky technology (solar geoengineering) impacts participants' willingness to cooperate in a collective-risk social dilemma (mitigation).

In contrast to the moral hazard conjecture, we find that the threat of solar geoengineering tends to lead to an *increase* in cooperative behavior. This finding is consistent with other empirical social science studies exploring the moral hazard conjecture in other contexts (e.g., Cherry et al. 2022; Merk et al. 2016). The one exception to this finding is when deploying solar geoengineering comes with very little risk (i.e., 90% chance of being effective). In this case, the threat of solar geoengineering, on average, has a negative but insignificant impact on cooperation. Compared to a baseline treatment without the SGE option, groups were three to four times more likely to avoid impending damages through cooperation under the threat of SGE deployment. To summarize, in contrast to the moral hazard conjecture, we present empirical evidence that the availability of a risky technology is a clarion call to increase cooperation. The threat that one actor may deploy the risky technology appears to trigger greater mitigation to lower the chances of deployment. Future research is needed to examine the robustness of this finding, including how behavior may change due to varying the size of potential losses from a negative outcome.

The laboratory setting, like theory, is a dramatic simplification of a more complicated reality, but it has been shown to be a useful method to consider behavioral influences and to testbed policies (Falk and Heckman 2009). Our results provide suggestive empirical evidence on the strategic responses to SGE in the absence of field observations. For many, the thought of deploying solar geoengineering causes great trepidation (Wagner 2021). The technology is inherently risky and poses many difficult governance challenges (Barrett 2014). However, failing to adequately mitigate GHG emissions is also risky. This study, and the emerging social science research on this topic (Aldy et al. 2021, NASEM 2021), will help inform the ongoing global debate on how to best manage climate change.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10584-023-03612-2.

Acknowledgements We thank Mark Borsuk, Tyler Felgenhauer, Khara Grieger, Alex James, Jennifer Kuzma, Juan Moreno-Cruz, Billy Pizer, and Jonathan Wiener for valuable comments.

Author contribution All authors contributed to the conception and design, data analysis, and writing first, subsequent, and final drafts. All authors read and approved the final manuscript.

Funding This material is based upon work supported by the National Science Foundation under Grant No. 2033855.

Data availability Data and code are available via the Supplementary information.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abatayo AL, Bosetti V, Casari M, Ghidoni R, Tavoni M (2020) Solar geoengineering may lead to excessive cooling and high strategic uncertainty. Proc Natl Acad Sci 117(24):13393–13398

Aldy J, Felgenhauer T, Pizer WA, Tavoni M, Belaia M, Borsuk ME, Ghosh A, Heutel G, Heyen D, Horton J, Keith D, Merk C, Cruz JM, Reynolds JL, Ricke K, Rickels W, Shayegh S, Smith W, Tilmes S, Wagner G, Wiener JB (2021) Social science research to inform solar geoengineering. Science 375(6569):815–818

Andrews TM, Delton AW, Kline R (2022) Anticipating moral hazard undermines climate mitigation in an experimental geoengineering game. Ecol Econ 196:107421

Austin MMK, Converse BA (2021) In search of search of weakened resolve: does climate-engineering awareness decrease individuals' commitment to mitigation? J Environ Psychol 78:101690

Barrett S (2008) The incredible economics of geoengineering. Environ Resource Econ 39:45–54

Barrett S (2014) Solar geoengineering's brave new world: thoughts on the governance of an unprecedented technology. Rev Environ Econ Policy 8(2):249–269

Biermann F, Oomen J, Gupta A, Ali SH, Conca K, Hajer MA, Kashwan P, Kotzé LJ, Leach M, Messner D, Okereke C, Persson A, Potocnik J, Schlosberg D, Scobie M, VanDeveer SD (2022) Solar geoengineering: the case for an international non-use agreement. WIREs Clim Chang 13(3):754. https://doi.org/10.1002/wcc.754

Brown TC, Kroll S (2017) Avoiding an uncertain catastrophe: climate change mitigation under risk and wealth heterogeneity. Clim Change 141:155–166

Cherry TL, Kroll S, McEvoy D, Campoverde D, Cruz JM (2022) Climate cooperation in the shadow of solar geoengineering: an experimental investigation of the moral hazard conjecture. Environ Polit. https://doi.org/10.1080/09644016.2022.2066285

Cherry TL, Kallbekken S, Kroll S, McEvoy DM (2021) "Does solar geoengineering crowd out climate change mitigation efforts? Evidence from a stated preference referendum on a carbon tax". Climatic Change 165(6): https://doi.org/10.1007/s10584-021-03009-z.

Chhetri N, Chong D, Conca K, Falk R, Gillespie A, Gupta A, Jinnah S, Kashwan P, Lahsen M, Light A, McKinnon C, Thiele LP, Valdivia W, Wapner P, Morrow D, Turkaly C, Nicholson S (2018) Governing solar radiation management. Forum for Climate Engineering Assessment, American University, Washington, DC

Cox JC, Dinkin S, Swarthout JT (2001) Endogeneous entry and exit in common value auctions. Exp Econ 4:163–181

Dave C, Eckel CC, Johnson CA, Rojas C (2010) Eliciting risk preferences: when is simple better? J Risk Uncertain 41:219–243

Eckel C, Grossman P (2008) Forecasting risk attitudes: an experimental study using actual and forecast gamble choices. J Econ Behav Organ 68(1):1–17

Fairbrother M (2016) Geoengineering, moral hazard, and trust in climate science: evidence from a survey experiment in Britain. Clim Change 139(3–4):477–489

Falk A, Heckman JJ (2009) Lab experiments are a major source of knowledge in the social sciences. Science 326:535–538

Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. Exp Econ 10(2):171–178

Gabriel CJ, Robock A (2015) Stratospheric geoengineering impacts on El Niño/Southern oscillation. Atmos Chem Phys 15(20):11949–11966

Greiner B (2015) Subject pool recruitment procedures: organizing experiments with ORSEE. J Econ Sci Assoc 1(1):114–125

Heutel G, Moreno-Cruz J, Shayegh S (2016) Climate tipping points and solar geoengineering. J Econ Behav Organ 132:19–45

Irvine P, Emanuel K, He J, Horowitz LW, Vecchi G, Keith D (2019) Halving warming with idealized solar geoengineering moderates key climate hazards. Nat Clim Chang 9:295–299

Jinnah S, Nicholson S, Morrow DR, Dove Z, Wapner P, Valdivia W, Thiele LP, McKinnon C, Light A, Lahsen M, Kashwan P, Gupta A, Gillespie A, Falk R, Conca K, Chong D, Chhetri N (2019) Governing climate engineering: a proposal for immediate governance of solar radiation management. Sustainability 11:3954

Keith DW (2021) Toward constructive disagreement about geoengineering. Science 374(6569):812–815

Keith DW, Weisenstein DK, Dykema JA, Keutsch FN (2016) Stratospheric solar geoengineering without ozone loss. Proc Natl Acad Sci 113(52):14910–14914

Kroll S, Shafran AP (2018) Spatial externalities and risk in interdependent security games. J Risk Uncertain 56(3):237–257

Lin AC (2013) Does geoengineering present a moral hazard? Ecol Law Quart 40:673–712

Merk C, Pönitzsch G, Rehdanz K (2016) Knowledge about aerosol injection does not reduce individual mitigation efforts. Environ Res Lett 11(5):1–6

Milinski M, Sommerfeld RD, Krambeck H-J, Reed FA, Marotzke J (2008) “The collective-risk social dilemma and the prevention of simulated dangerous climate change. Proc Natl Acad Sci 105:2291–2294

Millard-Ball A (2012) The Tuvalu syndrome. Clim Change 110(3):1047–1066

Moreno-Cruz JB (2015) Mitigation and the geoengineering threat. Resourc Energy Econ 41:248–263

National Academies of Sciences, Engineering, and Medicine (NASEM) (2021) Reflecting sunlight: recommendations for solar geoengineering research and research governance. The National Academies Press, Washington, DC. https://doi.org/10.17226/25762

Reynolds J (2019) The governance of solar geoengineering. Cambridge University Press, Cambridge

Snowberg E, Yariv L (2021) Testing the waters: behavior across participant pools. Am Econ Rev 111(2):687–719

Wagner G (2021) Geoengineering: the gamble. Polity Press

Wagner G, Zizzamia D (2022) Green moral hazards. Ethics, Policy Environ 25(3):264–280

Weitzman M (2015) A voting architecture for the governance of free-driver externalities. Scandinavian J Econ 117:1049–1068

Williamson P, Turley C (2012) Ocean acidification in a geoengineering context. Phil Trans R Soc A 370:4317–4342