Governing Climate Geoengineering: Side Payments Are Not Enough

Riccardo Ghidoni, Anna Lou Abatayo, Valentina Bosetti, Marco Casari, Massimo Tavoni

Abstract: Climate geoengineering might reduce the economic and ecological impacts of global warming. However, its governance is challenging: since climate preferences vary across countries, excessive climate geoengineering relative to the socially optimal level is a likely risk. Through a laboratory experiment on a public good-or-bad game, we study to what extent side payments can curb geoengineering efforts and restore efficiency. Although the availability of side payments is theoretically effective, its impact is modest in the experiment, especially in a multilateral setup. Replacing unstructured bilateral agreements with a structured framework improves outcomes.

JEL Codes: C70, C90, H40, Q50

Keywords: climate treaty, public good or bad, free driving, transfers, experiment

GEOENGINEERING OFFERS A POSSIBLE WAY to cope with the climate emergency. However, its implementation at a planetary level remains uncharted territory for scientists and policy makers. The discussion on geoengineering has gained traction after the signing of the Paris Agreement, whose long-term goal is to limit global temperature

Riccardo Ghidoni is at the Department of Economics, University of Bologna, and Department of Economics, CentER, Tilburg University (riccardo.ghidoni@unibo.it). Anna Lou Abatayo is at the Environmental Economics and Natural Resources Group, Wageningen University and Research (anna.abatayo@wur.nl). Valentina Bosetti is at the European Institute on Economics and the Environment, Centro Euromediterraneo sui Cambiamenti Climatici, and Department of Economics, Bocconi University (valentina.bosetti@unibocconi.it). Marco Casari is in the Department of Economics, University of Bologna, and IZA (marco.casari@unibo.it). Massimo Tavoni is at the RFF-CMCC European Institute on Economics and the Environment, Centro Euromediterraneo sui Cambiamenti Climatici and Politecnico di Milano (massimo.tavoni@polimi.it). Bosetti, Casari, and Tavoni gratefully acknowledge the financial support provided by the PRIN-2017-201782J9R9, Abatayo and Bosetti the ERC-2013-StG 336703-RISICO, and Abatayo the NWO grant 406.XS.01.073. The funding agencies had no role in the study design, data collection and analysis, *Dataverse data:* https://doi.org/10.7910/DVN/B5V2QL

Received October 7, 2021; Accepted January 17, 2023; Published online July 13, 2023.

Journal of the Association of Environmental and Resource Economists, volume 10, number 5, September 2023.
 Published by The University of Chicago Press for The Association of Environmental and Resource Economists.
https://doi.org/10.1086/724286

increases to below +2.0°C. Participating governments made pledges to reduce emissions and contribute to a fund aimed at compensating for climate damage in the most vulnerable countries. The resulting policies have made emissions peak in several Organisation for Economic Co-operation and Development (OECD) countries, and efforts are underway to slow emissions in various non-OECD countries. Nevertheless, even if emission reductions occur at scale, the inertia in the climate and socioeconomic responses may call for additional strategies to deal with the rising climate risks that past and current emissions have already committed our planet to.

A strategy that has been researched and discussed for a while involves direct manipulation of the Earth's temperature through solar radiation management (SRM). Reducing incoming solar radiation can lower temperatures rapidly and, thus, limit climate impacts. This can be achieved, for example, by injecting reflecting aerosol particles into the stratosphere, mimicking the well-studied consequences of volcanic eruptions. SRM interventions can be climatically effective (Irvine et al. 2019; Kravitz et al. 2021) and economically appealing (Barrett 2008). This study employs a laboratory experiment to shed light on the strategic impacts of SRM, which are largely unknown yet critically important.

Climate geoengineering could substantially alter the strategic incentives of nations and raise significant governance challenges (Schelling 1996; Victor et al. 2009; Horton et al. 2018; Sugiyama et al. 2018; Rickels et al. 2020). In the model introduced by Weitzman (2015), nations have diverse ideal temperatures and can unilaterally use geoengineering to cool the planet. National diversities originate from heterogeneous exposure to climate damage resulting from different geographical locations. In equilibrium, this public good-or-bad setting is characterized by excessive levels of geoengineering. Given the low marginal cost of geoengineering, the country with the highest preference for cooling the planet uses this technology to a level that will hurt most other countries and, thus, decrease global welfare. The overprovision of effort is termed "free driving," in contrast to the well-known free-riding phenomenon characterizing public goods. Abatayo et al. (2020) provide experimental evidence of free driving and its welfare-deteriorating consequences. This study follows up by investigating a possible way to restore geoengineering toward a socially optimal level.

To curb free driving, Weitzman (2015) proposes a top-down approach involving a binding international treaty established through voting. Weitzman's approach follows early climate negotiation attempts aimed at an integrated global contract. In contrast, this study takes a decentralized and voluntary approach, which is more common to later

decision to publish, or preparation of the paper. We thank the editor Christian A. Vossler and two anonymous reviewers for their excellent comments. We also thank participants at the 2022 Symposium on Markets and Sustainability in Amsterdam, 2022 EAERE conference in Rimini, and 2022 Economic Science Association (ESA) European meeting in Bologna for insightful discussions.

climate negotiations, such as the voluntary national pledges framework of the Paris Agreement. We allow individual nations to offer side payments to others and make the transfer conditional to the receiver following a requested geoengineering effort.[1] In principle, side payments can promote international cooperation when countries are heterogeneous (Barrett 2005), observed, for instance, in freshwater international disputes (Dinar 2006). In our context, heterogeneity arises from the different regional preferences for geoengineering. Countries damaged by excessive geoengineering may hence find it beneficial to offer adequate compensations to those overexploiting such technology to induce them to limit their cooling efforts.

Side payments have already been proposed in climate change agreements as an effective policy instrument for making a small environmental treaty effective (Kverndokk 1993) and as a way to enlarge an existing climate club (Saelen 2016). The Green Climate Fund in the Paris Agreement can also be seen as a side payment arrangement where developed nations can help vulnerable societies to adapt to the impacts of climate change and finance their green transition. This solution has also been shown to foster efficient outcomes in various lab experiments (e.g., Hoffman and Spitzer 1982, 1985).[2]

We conducted a laboratory experiment to study the empirical effectiveness of side payments to limit free driving and restore efficiency. The experiment comprises two benchmark scenarios in the Baseline treatment and four scenarios where side payments are available. We manipulated the number of decision makers and the structure of the side-payment scheme. Economies included either two decision makers ($N = 2$) or six decision makers ($N = 6$). Side payments could take the form of bilateral agreements that any party can make across a wide range of possible transfers and geoengineering targets (Decentralized treatment). As documented in Hoffman and Spitzer (1982) and Dixit and Olson (2000), however, this setup can become increasingly complex as the action space widens and the number of decision makers increases.

Alternatively, side payments could be implemented in a structured setting where the action space is simplified to ease coordination (Structured treatment): offers are made jointly by the decision makers with lower preferences for geoengineering, and the resulting side payments are distributed equally among decision makers who limit their geoengineering efforts. This setup mirrors the creation of negotiation blocs of homogeneous countries, a typical process in UN Framework Convention on Climate Change (UNFCCC) negotiations, where countries that are similar in their vulnerability to climate change, such as the Alliance of Small Island Developing States, negotiate at smaller

1. Our experiment focuses on a scenario where the geoengineering technology is already available to all decision makers and no investment is necessary to develop it. We use the term “effort” to refer to the (costly) use of such a technology.

2. Allowing side payments is in line with the Coasian bargaining approach to externalities. We report experiments on the Coase theorem in appendix table D.6 (the appendix is available online) and discuss them in the concluding section.

tables and find common ground before engaging in broader negotiations. In the design, we first limit the set of side-payment actions and then define how side payments should be shared. In the setting with $N = 2$, the Structured treatment only restricts the offer space relative to the Decentralized treatment. In the multilateral setting of $N = 6$, the Structured treatment also implies an equal division of the received side payments within a negotiation bloc. With this staggered design, we can disentangle which feature of the structured negotiation may help increase coordination compared to Decentralized, if any.

Our experimental results show that the availability of side payments has limited effectiveness in curbing free driving, especially in the multilateral setting. While side payments could theoretically enable decision makers to achieve the socially optimal use of geoengineering, efficiency improves only by a modest amount in the Decentralized and Structured treatments compared to Baseline. A deeper analysis reveals the reasons for this failure. Participants tried to move away from the free-driving outcome through a flurry of side-payment offers. However, many offers were inadequate, as they were not in the mutual interest of both the sender and receiver. Furthermore, some adequate offers were not taken up. As a result, too few side payments took place. For instance, in Structured, the side payments' frequency was 48% in economies of two and only 28% in economies of six. Efficiency was higher in economies where more side payments occurred, as theory would suggest. Behavioral forces such as confusion, social preferences, or low stakes played a minor role in determining this outcome. As stated in the Conclusion, studying these patterns further may help design more effective institutions.

This study brings original contributions to the growing stream of experimental studies about climate change. The extant literature has mostly focused on emission reduction strategies. Experiments often involve modified versions of the voluntary contribution public good game and focus, for instance, on the role of tipping points (Tavoni et al. 2011; Barrett and Dannenberg 2017), uncertainty of impacts (Barrett and Dannenberg 2012; Ghidoni et al. 2017), dynamic externalities (Sherstyuk et al. 2016; Calzolari et al. 2018), or commitment devices (Dengler et al. 2018).[3] There is much less work about climate geoengineering. Abatayo et al. (2020) have investigated the public good-or-bad game by Weitzman (2015), which captures a situation where the stock of emissions has already reached a dangerous level, and climate damage can only be avoided through geoengineering. This setting alters in fundamental ways the strategic environment of a public good game: (1) decision makers differ in their goals about the global geoengineering level; (2) the provision cost is low, and, more precisely, it is lower than the potential benefit of climate geoengineering; (3) decision makers are characterized by single-peaked preferences over the global outcome, leading to penalties both in case of over- and underprovision.

3. Another strand of the experimental literature investigates the efficacy of emission trading schemes (e.g., Cason 1995; Jakob et al. 2017).

The remainder of the paper is organized as follows: section 1 provides a description of the experiment, including the illustration of the public good-or-bad game and its treatments, as well as details on the experimental procedures; section 2 presents the theoretical benchmarks; section 3 presents the main experimental results and the underlying the mechanisms; and section 4 concludes.

1. THE EXPERIMENT

1.1. Public Good-or-Bad Game

The theoretical model follows the public good-or-bad (GoB) game by Weitzman (2015), which considers an economy of $N \geq 2$ decision makers choosing their geoengineering efforts independently. The interaction lasts for a finite number of periods, T. At the beginning of every period, each decision maker i receives an endowment E, which can be used to exert a geoengineering effort, g_i, bounded below by zero and above by $\bar{g}$.[4] The marginal effort cost is $\alpha > 0$. Decision makers choose their effort level simultaneously and observe all effort decisions in the economy at the end of each period. The sum of the efforts in the economy determines the global geoengineering level, $G = \Sigma_{i=1}^{N} g_i$. Decision makers have different ideal levels of global geoengineering, G_i^* (henceforth, ideal point). Furthermore, preferences for global geoengineering are single peaked: decision maker i incurs losses when G is either above or below G_i^*. For simplicity, we assume that i suffers the same loss if G undershoots or overshoots G_i^* by the same amount. The marginal loss of deviations from G_i^* is larger than the marginal effort cost, $\lambda > \alpha$. Hence, decision maker i's stage-game payoff is:

$$\pi_i = E - \alpha g_i - \lambda |G - G_i^*|. \tag{1}$$

In the experiment, $E = 150, \alpha = 4, \lambda = 10$, and g_i is an integer number with $\bar{g} = 15$. All parameters are public information.

This characterization of geoengineering entails four simplifying assumptions that make the experiment tractable. First, geoengineering is the only climate strategy available (other climate strategies, such as mitigation and adaptation, are ruled out).[5] Second, geoengineering outcomes are fully deterministic, and we abstract from indirect risks like its effects on ozone levels or rainfall.[6] Third, the GoB game assumes a repeated

4. For simplicity, we rule out counter-geoengineering, i.e., the possibility to undo other decision makers' efforts by choosing $g_i < 0$ (see Heyen et al. [2019] and Abatayo et al. [2020], for discussions on the effects of counter-geoengineering).

5. While mitigation and adaptation require cumulative investments over several years, geoengineering is a short-term strategy. Therefore, we believe that analyzing this climate strategy in isolation makes sense.

6. There is much scientific uncertainty regarding the link between the injection of aerosols and global cooling. To our knowledge, only Proctor et al. (2018) have investigated this link in the field. When Mount Pinatubo in the Philippines erupted in 1991, the eruption injected

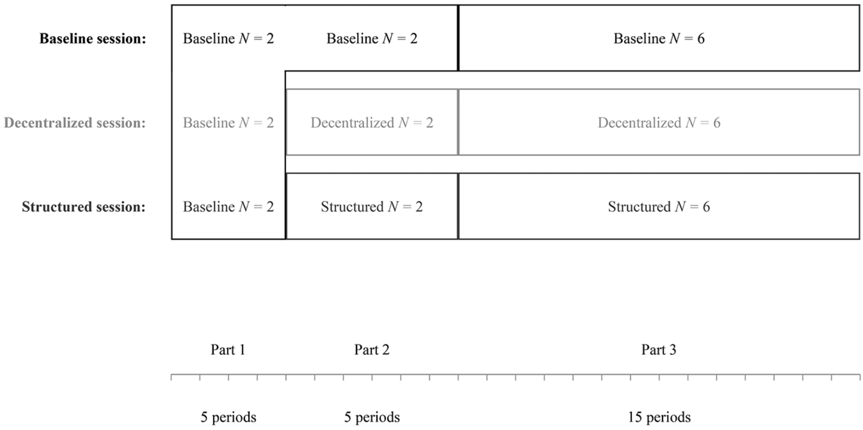

Figure 1. Experimental sessions. Number of participants (no. sessions) were 72 (3) in Baseline, 72 (3) in Decentralized, and 120 (5) in Structured. Session dates: May 21, 2018, May 23, 2018, May 24, 2018, June 11, 2018, June 18, 2018, June 20, 2018, and June 28, 2018. Data from the Baseline sessions have already been used by Abatayo et al. (2020).

but static interaction: geoengineering efforts in a given period do not affect the temperature in subsequent periods. Fourth, decision makers only differ in their ideal points G_i^*, while they have an identical initial endowment, action space, effort cost, and losses from deviations from G_i^*.

1.2. Treatments

The experiment comprises three treatments—Baseline, Decentralized, and Structured—differing in the availability of side payments and the way to negotiate them. At the beginning of each session, decision makers were randomly assigned their ideal point, which remained fixed throughout the session. Within each economy, we matched decision makers with different ideal points to generate a conflict over the optimal level of global geoengineering. All sessions included both bilateral and multilateral interactions. Decision makers interacted in fixed pairs of $N = 2$ in the first 10 periods (parts 1 and 2). We then pooled together three economies of two to form an economy of six, and the same decision makers interacted in a fixed group of $N = 6$ in the following 15 periods (part 3).

While the multilateral setting is closer to reality, the bilateral setting is simpler to analyze and easier to understand for the participants. By comparing $N = 2$ and $N = 6$ outcomes, we can shed some light on how the complexities due to multilateralism

20 Mt of SO_2 into the atmosphere and cooled the planet by 0.5°C. However, under different wind, weather, and temperature conditions, the resulting global cooling may have been different.

can affect geoengineering governance. In the Decentralized and Structured treatments, participants faced the same rules as in the Baseline treatment during part 1, while the rules were treatment specific in parts 2 and 3 (fig. 1). This structure enabled participants in all treatments to familiarize themselves with the simplest version of the GoB game and provided us with a common performance measure to assess the treatment effects.[7] We now describe each treatment in detail.

1.2.1. Baseline

The stage game in this treatment corresponds to the GoB game described in section 1.1. In each session, four decision makers received a low ideal point $G_i^* = 2$ (L_1, L_2, L_3, L_4), four received a medium ideal point $G_i^* = 6$ (M_1, M_2, M_3, M_4), and four received a high ideal point $G_i^* = 10$ (H_1, H_2, H_3, H_4). We adopted the following matching protocol to generate six independent economies of two in parts 1 and 2 and two independent economies of six in part 3. Economies of two combined (L_1, M_1), (L_2, H_2), (M_2, H_1), (L_3, M_3), (L_4, H_4), and (M_4, H_3), respectively. Economies of six combined (L_1, L_2, M_1, M_2, H_1, H_2) and (L_3, L_4, M_3, M_4, H_3, H_4), respectively, which ensured the same number of low, medium, and high ideal points within every economy of six.[8]

1.2.2. Decentralized

In this treatment, side payments take the form of bilateral agreements between two decision makers. Each period has two stages. First, each decision maker can offer a side payment to any other decision maker in the economy in exchange for a requested level of geoengineering effort. Next, all offers are publicly displayed, and decision makers choose their effort (with the same GoB game as in Baseline). Under both $N = 2$ and $N = 6$, every decision maker i can make at most one offer to a decision maker j for an amount $T_{i,j}$ between 1 and 150. The offer is binding as the side payment automatically occurs if the receiver satisfies the sender's request. Requests must specify an effort lower than, greater than, or equal to an arbitrary level between 0 and 15.[9] The per-period payoff of decision maker i is:

$$\pi_i = E - \alpha g_i - \lambda |G - G_i^*| - I_{i,j} \times T_{i,j} + \sum_{j \neq i} (J_{j,i} \times T_{j,i}). \quad (2)$$

7. A drawback of this design is the lack of control for potential order effects, which could complicate the identification of the impact of $N = 6$. We quantified the magnitude of order effect through an econometric analysis that exploits the variations in parts 1 and 2 on outcomes in part 3. Our findings suggest that order effects have limited significance (see n. 26).

8. This matching protocol ensures that decision makers who have interacted with each other in the economies of two (parts 1 and 2) will interact with each other again in the economies of six (part 3).

9. In the experiment, the offer was framed as follows: "I will transfer XXX tokens if you produce an amount lower than/greater than/equal to YYY units."

where $I_{i,j} = 1$ if i made an offer to j and j satisfied i's request and $I_{i,j} = 0$ otherwise, and $J_{j,i} = 1$ if i received an offer from j and satisfied j's request and $J_{j,i} = 0$ otherwise.

1.2.3. Structured

In this treatment, side payments are designed to facilitate coordination by mimicking the existence of negotiating blocs that pre-agree on a set of conditions. Compared to Decentralized, the Structured treatment exhibits three differences. First, not everyone is allowed to make an offer: the decision makers with the highest ideal point in the economy can only receive offers but not make them. Hence, there can be at most one offer under $N = 2$ and four offers under $N = 6$. Second, the offer space is simplified. The offered amount T_i and the requested effort can only be chosen from a subset of elements including those leading to the efficient equilibrium: the requested effort could be 2 or 6 and the offered amount 34 or 70. Third, under $N = 6$, side-payments negotiations are across negotiation blocs. For one thing, side payments are only enacted if both high-ideal-point decision makers jointly satisfy the request by splitting geoengineering efforts equally $(1 + 1$ or $3 + 3)$. Moreover, when the two high-ideal-point decision makers satisfy the requests for receiving one or more side payments, the offered amounts are pooled and equally split between them.[10] Thus, under both $N = 2$ and $N = 6$, the per-period payoff of a decision maker i who can make offers is:

$$\pi_i = E - \alpha g_i - \lambda|G - G_i^*| - I_i \times T_i. \tag{3}$$

The stage-game payoff of a decision maker j who can only receive offers is:

$$\pi_j = E - \alpha g_j - \lambda|G - G_j^*| + I_i \times T_i \text{ if } N = 2, \tag{4}$$

$$\pi_j = E - \alpha g_j - \lambda|G - G_j^*| + \sum_{i \neq j}(I_i \times T_i)/2 \text{ if } N = 6. \tag{5}$$

Under $N = 2$, $I_i = 1$ if i made an offer and j satisfied i's request, and $I_i = 0$ otherwise. Under $N = 6$, $I_i = 1$ if i made an offer and the request was satisfied by both high-ideal-point decision makers, and $I_i = 0$ otherwise.

1.3. Experimental Procedures

A team of two participants constituted the basic decision-making unit in all treatments. We used teams to capture the collective processes behind national choices generated by countries. Moreover, teams are generally considered more rational than individuals making decisions in isolation (see, e.g., Charness and Sutter 2012). The team composition was random and kept constant throughout the session. The team members

10. This outcome is viable in the Decentralized treatment as well, but substantial coordination is necessary to achieve it.

could communicate via chat for up to one or two minutes in each period and had to reach unanimous decisions. No eye contact was possible among participants.

Although we ruled out communication among teams, climate negotiations involve this type of communication.[11] To assess the robustness of our results, it would be interesting to study the impact of communication among teams.[12] While pre-play communication can boost cooperation in social dilemma experiments (see Balliet [2010] for a meta-analysis), a behavioral prediction about its effect in the Baseline GoB game is not clear cut. Communication may ease coordination on offer decisions in the Decentralized and Structured treatments, given the multiplicity of equilibria. Yet, the experimental evidence from coordination games suggests that the specific type of game and communication protocol matter considerably for the outcome (Ellingsen and Östling 2010). For these reasons—and for the additional hurdle of having here teams as decision makers—a dedicated study of geoengineering with communication among decision makers could be a fascinating follow-up.

The experiment was neutrally framed: we never mentioned climate change; decision makers were referred to as "teams," economies as "groups," and geoengineering as "production." At the beginning of each part, the experimenter handed out the relevant instructions and read them aloud (see app. B.1; apps. A–H are available online). Participants completed a quiz on the instructions (see app. B.2) and were asked to write on paper the results for each period to ensure that they paid attention. The software featured a built-in calculator that participants could use to simulate their earnings by entering hypothetical efforts.[13] Before leaving the lab, participants completed a questionnaire (see app. B.3).

A total of 264 students participated in the experiment, divided into sessions of exactly 24 participants; they were recruited via ORSEE (Greiner 2015). All sessions were conducted at the BLESS laboratory of the University of Bologna.[14] On average, participants earned 21 EUR.[15] The experiment was programmed with zTree (Fischbacher 2007).

---

11. One could argue that the screen displaying all offers made in the economy, coupled with the many periods of repeated interaction, is a form of structured communication among teams.

12. We thank the editor and an anonymous referee for this suggestion.

13. In Structured, participants were also allowed to simulate the payoff impact of paying or receiving side payments.

14. We conducted two additional sessions of Structured because we detected a bug in the software after the first three sessions. The bug, which affected five out of six economies in part 3, concerned the visualization of the results in the final screen of some periods. Recall that a side payment in Structured is only enacted if both H teams satisfy the request. The bug occurred when only one H team's effort matched the request: the left side of the screen showed the correct information (i.e., that H teams did not receive the side payment), while the right side showed otherwise. Payoffs were computed correctly. Since robustness checks show that the bug did not have an impact on any outcome (see app. E), we decided to use data from all sessions for our analyses.

15. Including a show-up fee of 8 EUR; conversion rate 0.01 EUR = 2 tokens.

Table 1. Theoretical Benchmarks

	Economies of Two ($N = 2$)			Economy of Six ($N = 6$)	
	Economy (L, M)	Economy (L, H)	Economy (M, H)	Economy (L, L, M, M, H, H)	
SO	G = 2	G = 2	G = 6	G = 6	
SPNE Baseline	$g_L = 0$, $g_M = 6$	$g_L = 0$, $g_H = 10$	$g_M = 0$, $g_H = 10$	$g_{H1} + g_{H2} = 10$	
SPNE Decentralized:					
Efforts	$g_L = 0$, $g_M = 2$	$g_L = 0$, $g_H = 2$	$g_M = 0$, $g_H = 6$	$g_{H1} + g_{H2} = 10$	$g_{H1} + g_{H2} = 6$
Side payments	$T_L = -25$	$T_L = -49$	$T_M = -25$	None	$T_{L1} + T_{L2} + T_{M1} + T_{M2} = -66$
SPNE Structured:					
Efforts	$g_L = 0$, $g_M = 2$	$g_L = 0$, $g_H = 2$	$g_M = 0$, $g_H = 6$	$g_{H1} + g_{H2} = 10$	$g_{H1} + g_{H2} = 6$
Side payments	$T_L = -34$	$T_L = -70$	$T_M = -34$	None	$T_{L1} + T_{L2} + T_{M1} + T_{M2} = -68$
Gini index (SO)	.10	.21	.14	.05	.05
Gini index (SPNE)	.04	.12	.01	.13	.04

Note. For $N = 6$, only strictly positive efforts are reported (efforts equal to 0 are omitted). The range of Pareto-improving side payments T varies among economy types, with intervals of [24, 40], [48, 80], and [24, 40] in (L, M), (L, H), and (M, H) economies, respectively, and [64, 80] in (L, L, M, M, H, H) economy. In Structured, where only two side-payment levels are available, optimal side payments fall within these intervals. In Decentralized, the reported side-payment level is the minimum that makes receivers strictly prefer it over their Baseline earnings. For $N = 2$ the Gini index was computed using cumulative equilibrium earnings over five periods. For $N = 6$ the Gini index was computed using cumulative equilibrium earnings over 15 periods. SPNE = subgame perfect Nash equilibria; SO = social optimum.

2. THEORETICAL BENCHMARKS

To evaluate the empirical results, we present the theoretical benchmarks of the social optimum (SO) and the subgame perfect Nash equilibria (SPNE). Our measure of social optimality is the total surplus, that is, the sum of all decision makers' payoffs in the economy in a period. We first analyze economies of two and then economies of six. Table 1 provides an overview of the benchmarks across conditions. Note that, in the GoB game, the repeated-game equilibria correspond to the (finite) repetition of the one-shot game equilibria.[16]

16. In this section, we report benchmarks for the specific parameters used in the experiment. The formal proofs for a general N are in app. A.

2.1. Economies of Two

We present the social optimum and the SPNE for $N = 2$.

Proposition 1 (Social optimum with $N = 2$): A level of global geoengineering equal to the lowest ideal point in the economy is socially optimal in all treatments.

It follows from proposition 1 that the socially optimal level of global geoengineering is 2 in the economies (L, M) and (L, H) and 6 in the economy (M, H). The intuition behind this proposition rests on the linearity and symmetry of payoff losses when deviating from ideal points, $\lambda > 0$, as well as on the need to save on effort costs, which make a level of global geoengineering equal to the lowest ideal point the most efficient outcome. Any combination of individual efforts that sums up to that level is socially optimal because, by design, the marginal effort cost α is positive and the same for all decision makers.[17]

Proposition 2 (SPNE with $N = 2$): The SPNE is unique in all treatments.

- In Baseline, the highest ideal point among decision makers determines the level of global geoengineering. The decision maker with the highest ideal point puts in all the effort, and the other puts in zero.
- In Decentralized, the lowest ideal point among decision makers determines the level of global geoengineering. The decision maker with the highest ideal point puts in all the effort, while the other puts in zero. A side payment of 25 occurs in the economies (L, M) and (M, H), and a side payment of 49 occurs in the economy (L, H).
- In Structured, the lowest ideal point of decision makers determines the level of global geoengineering. The decision maker with the highest ideal point puts in all the effort, while the other puts in zero. A side payment of 34 occurs in the economies (L, M) and (M, H), and a side payment of 70 occurs in the economy (L, H).

The SPNE outcome outlined in proposition 2 is socially optimal in Decentralized and Structured but not in Baseline. In Baseline, there is free driving because the preference of the decision maker with the highest ideal point prevails as each decision maker has a unilateral incentive to exert effort until its ideal point is reached. This imposes an excessive level of geoengineering on the lowest-ideal-point decision maker. Off-equilibrium, if the latter exerts some effort, it will just induce an equivalent effort reduction by the

17. If α were equal to zero, then any global geoengineering level between the two ideal points would be socially optimal.

decision maker with the highest ideal point (who will save on the effort cost) without any change in the final global geoengineering outcome. Hence, since efforts are substitutes, the best response of the team with the lowest ideal point is to exert zero effort. Finally, it is worth noticing that inequality is less under the SPNE outcome than under the social optimum in all economy types.[18]

To better understand the GoB platform, it may be helpful to consider some analogies between the Baseline treatment with $N = 2$ and a canonical dictator game. First, the decision maker with a high ideal point is similar to a dictator, and, in equilibrium, the decision maker with a low ideal point gets a payoff set by the other, like the recipient in a dictator game. Second, assigning ideal points in the GoB game is like assigning roles in a dictator game. Third, equilibrium payoffs are highly unequal in both games. However, other design differences suggest caution in comparing experimental results between the two games. In a GoB game, all decision makers are active players, as they can exert a geoengineering effort. A positive effort by a low-ideal-point decision maker would be an off-equilibrium behavior but, in principle, possible.[19] Another subtle difference concerns the framing of the distributional conflict. In the dictator game, the framing is clear: there is a fixed endowment, and one party unilaterally decides how much to give to the other, with a unique equilibrium at the corner. Instead, in the GoB game, the total surplus can vary, and the equilibrium solution is an interior point of the action space.

In the Decentralized and Structured treatments, the low-ideal-point decision maker is better off compensating the high-ideal-point decision maker for reducing its effort. In equilibrium, the offered amount represents the minimum feasible side payment that receivers will accept in exchange for reducing their effort. The offered amounts differ between the Structured and Decentralized treatments as a result of variations in the width of the offer space in each treatment. Besides receiving the side payment, a decision maker who meets the request also saves on the effort cost. Hence, in equilibrium, the total surplus in Decentralized and Structured surpasses that in Baseline and corresponds to the socially optimal one. Crucially, all economies can achieve this outcome because offers are binding.

---

18. The effort cost sustained by the highest-ideal-point decision maker is higher at the SPNE than at the social optimum. This cost reduces the difference between the two decision makers' payoffs. Conversely, the effort cost at the social optimum is cheaper than at the SPNE, resulting in higher payoff differences.

19. This GoB game's feature echoes experiments where a punishment technology is available to the disadvantaged player. Consider an ultimatum game where, for instance, the proposer keeps \$9 and offers \$1 to the responder. By rejecting, the responder gives up \$1 and inflicts damage of \$9 to the proposer. In a GoB game, this fee-to-damage ratio is around 1:1 instead of 1:9, which is less appealing because it does not lower earnings inequality in the pair and requires much stronger other-regarding preferences to be attractive.

2.2. Economies of Six

Under $N = 6$, the social optimum and SPNE change due to the different composition of the economy.

> **Proposition 3** (Social optimum with $N = 6$): A level of global geoengineering equal to G = 6, the medium ideal point in the economy, is socially optimal in all treatments.

The driving forces behind proposition 3 are the presence of three diverse, equally spaced ideal points in the economy and the symmetric linear losses for deviations from these ideal points.

> **Proposition 4** (SPNE with $N = 6$ in Baseline): The highest ideal point among decision makers, $G^* = 10$, determines the SPNE level of global geoengineering. There exist multiple equilibria because the two decision makers with the highest ideal point can share geoengineering efforts as they like. The others always put in zero effort.

The intuition behind the inefficient free-driving outcome with $N = 6$ described in proposition 4 follows a similar logic to that of proposition 2 for $N = 2$. In economies of six, however, the SPNE is characterized by more inequality than the social optimum. Furthermore, decision makers with the high ideal point need to coordinate their efforts to achieve a global geoengineering level equal to 10.

> **Proposition 5** (SPNE with $N = 6$ in Decentralized and Structured): There exist multiple SPNE:
>
> - Free-driving equilibria in Decentralized and Structured: as in Baseline, there exist multiple equilibria without side payments, leading to the inefficient outcome G = 10.
> - Side payments equilibria in Decentralized: there exist multiple equilibria characterized by the socially optimal outcome G = 6 and involving side payments. These equilibria prescribe two out of four decision makers with low or medium ideal points to offer a side payment of 33 to each of the high-ideal-point decision makers in exchange for a total effort of 6. The high-ideal-point decision makers satisfy the requests.
> - Side payments equilibria in Structured: there exist six equilibria characterized by the socially optimal outcome G = 6 and involving side payments. These equilibria prescribe two out of four decision makers with low or medium ideal points to offer a side payment of 34 to the

high-ideal-point decision makers in exchange for an effort of 3 each. The high-ideal-point decision makers satisfy the requests.

Proposition 5 highlights how the equilibrium sets of Decentralized and Structured under $N = 6$ are particularly rich. In the free-driving equilibria, no offer is made, and the outcome is identical to that of Baseline. Free driving remains a possible equilibrium outcome in Decentralized and Structured because, by design, a single decision maker is incapable of offering an adequate side payment to compensate both high-ideal-point decision makers for a substantial effort reduction. The free-driving equilibria are Pareto inferior to the ones with side payments. Moreover, they are not coalition stable, as the joint offer of two decision makers can move the economy to an equilibrium with side payments.[20]

The socially optimal outcome of 6 is sustainable if the high-ideal-point decision makers are compensated through adequate side payments by two decision makers with lower ideal points. In both Decentralized and Structured, any combination of two decision makers with low or medium ideal points has the incentive to offer a sufficient amount to achieve the social optimum: (L_1, L_2), (L_1, M_1), (L_1, M_2), (L_2, M_1), (L_2, M_2), and (M_1, M_2). This raises a coordination problem among the four decision makers with lower ideal points. Structured simplifies this problem by limiting the action space and giving more structure to the admissible offers.[21]

3. RESULTS

3.1. Aggregate Results

Before presenting the results on side payments, we report key statistics for the Baseline treatment.[22] The Baseline treatment data are the same used in Abatayo et al. (2020) to address a different research question and strongly support the free-driving hypothesis. The modal level of global geoengineering corresponded to the inefficient SPNE outcome both in economies of two and six (fig. C.1 and fig. 2, respectively; figs. C.1–C.5, E.1, F.1, G.1, H.2 are available online). Moreover, teams with high ideal points

20. While the system of side payments that ensures the coalition stability is zero-sum as in McGinty et al. (2012), our strategic context is different since we focus on a public good-or-bad game instead of a public good game.

21. In Structured, levels of global geoengineering below the socially optimal one cannot be reached in equilibrium. Consider, for example, a case where the two low ideal point decision makers offer an amount of 70 each in exchange for a global geoengineering level of 2. The high-ideal-point decision makers would find it profitable to adjust their efforts to satisfy such a request. However, either one of the medium ideal point decision makers would have an individual incentive to increase its effort to reach a global geoengineering level of 6.

22. In all our analyses, we consider teams as the decision-making unit. Hence, we collapse our subject-level data at the team level. We examine within-team disagreements on effort and offer decisions in app. F.

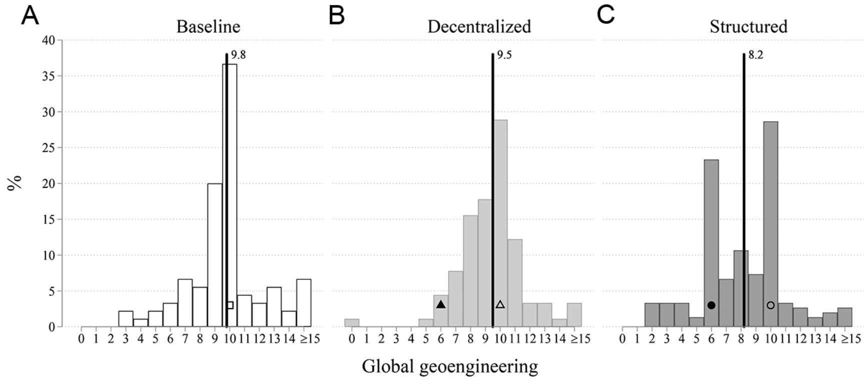

Figure 2. Global geoengineering in economies of six. *Vertical lines* mark average global geoengineering levels by treatment. *Solid markers* indicate the equilibrium predictions with side payments, which coincide with the socially optimal level; *hollow markers* indicate equilibrium predictions without side payments (free driving). The unit of observation is one economy in a period of part 3.

in economies of two provided on average 91% of the total effort in the economy. The analogous figure for economies of six is 86% (fig. C.4). Next, we report on the aggregate effects of side payments' availability on global geoengineering, total surplus, and inequality.

Result 1: While Decentralized did not reduce free driving, Structured partially did and improved welfare, albeit moderately.

Support for result 1 is in figure 2 and table 2. We first focus on economies of six. On average, global geoengineering was 9.8 in Baseline, 9.5 in Decentralized, and 8.2 in Structured.[23] According to Wilcoxon-Mann-Whitney tests (table 2), the reduction in global geoengineering compared to Baseline is statistically significant for Structured (p = .020, 16 observations), but not for Decentralized (p = .228, 12 observations). As in Baseline, the modal outcome corresponded to the free-driving outcome of 10 in both Decentralized and Structured (fig. 2). However, global geoengineering distributions in these two treatments were slightly more skewed toward lower levels than in Baseline. This pattern is especially visible in Structured, where the second most frequent global geoengineering outcome was 6 (i.e., the socially optimal level). In fact, the global geoengineering distribution in Structured is roughly bimodal, with 10 and 6 almost on par as the most likely outcomes.

23. In all treatments, average global geoengineering was relatively stable across periods. Outcomes were lower in Structured than in Baseline in all periods (fig. C.2b).

Table 2. Descriptive Statistics and Tests of Economy-Level Variables

	Baseline	Decentralized		Structured	
	Mean (SD)	Mean (SD)	*p*-value	Mean (SD)	*p*-value
Part 1:					
Global geoengineering	8.133 (1.922)	8.078 (1.577)	.812	8.200 (2.037)	.848
Total surplus	205.467 (26.803)	207.244 (24.353)	.800	203.867 (29.274)	.741
Gini index	.051 (.040)	.051 (.031)	.704	.046 (.043)	.594
Observations	18	18		30	
Part 2:					
Global geoengineering	8.289 (2.229)	8.033 (2.252)	.506	6.327 (2.683)	.020
Total surplus	203.067 (35.624)	197.867 (38.602)	.635	215.627 (34.192)	.166
Gini index	.051 (.044)	.055 (.028)	.359	.046 (.045)	.565
Observations	18	18		30	
Part 3:					
Global geoengineering	9.844 (.986)	9.500 (.530)	.228	8.240 (1.419)	.020
Total surplus	586.400 (41.675)	605.333 (30.198)	.522	632.640 (42.948)	.030
Gini index	.088 (.017)	.092 (.013)	.749	.078 (.017)	.129
Observations	6	6		10	

Note. *p*-values refer to Wilcoxon-Mann-Whitney exact tests of Baseline vs. Decentralized and Baseline vs. Structured; the null hypothesis is that the samples come from the same population. The unit of observation corresponds to an economy of six in part 3; the unit of observation is the average of the three economy types of two in parts 1 and 2. Global geoengineering and Total surplus are computed as the average outcome in an economy in a given part of the experiment; the Gini index is computed on cumulative earnings in the last period of each part.

In Structured, levels of global geoengineering closer to the social optimum positively reflected on the total surplus, computed as the sum of all teams' earnings in an economy in a period, which was 8 percentage points higher than that in Baseline and statistically significantly so ($p = .030$, 16 observations). Instead, the total surplus in Decentralized was only 3 percentage points higher than that in Baseline, and the difference is not statistically significant ($p = .522$, 12 observations). Neither Decentralized nor Structured had a significant impact on inequality, as measured by the Gini index calculated on cumulative earnings in an economy at the end of part 3 ($p \geq .129$, 12 and 16 observations, respectively).

Patterns in economies of two confirm the evidence from economies of six. When pooling data from all types of economies in part 2, a series of Wilcoxon-Mann-Whitney tests fail to detect a treatment effect of Decentralized on any of the three outcomes of interest (table 2; $p = .506$ for global geoengineering, $p = .635$ for total surplus,

$p = .359$ for inequality).[24] Instead, we identify a negative and statistically significant effect of Structured on global geoengineering ($p = .020$) but not on the other two dimensions ($p = .166$ for total surplus, $p = .565$ for inequality).[25]

Overall, Structured was more effective than Decentralized in curbing global geoengineering. Under $N = 6$, two aspects set the designs of Decentralized and Structured apart. On the one hand, the institutional setting is different: in Decentralized, only bilateral agreements are available, while in Structured, there are negotiating blocs. On the other hand, the offer space is wider in Decentralized than in Structured. Our design allows us to test to which extent these aspects contributed to achieving more efficient outcomes in Structured. A first insight comes from the observation that in part 2 with $N = 2$, where Decentralized and Structured only differ in the offer space width, Structured already outperformed Decentralized in a statistically significant way ($p = .005$, see Wald test in table D.3).

A difference-in-difference approach can provide a more formal test of the relative importance of the institutional setting versus the offer space width. If the more efficient outcomes of Structured in part 3 are due to its institutional setting, we should expect the difference in global geoengineering between Structured and Decentralized to increase from part 2 to part 3. Instead, if the offer space width is crucial, the observed difference when moving from part 2 to part 3 will be roughly the same for Decentralized and Structured. The evidence supports the latter interpretation: the prime determinant of the more efficient outcomes observed in Structured is the smaller offer space rather than the different institutional setting. Indeed, the estimated difference-in-difference coefficient is small and statistically insignificant ($\beta = -0.468$ and $p = .535$ according to a wild cluster bootstrap linear regression).[26] Note that this evidence does not imply

24. As a robustness check, we implemented one session of Decentralized where instead of playing part 3 in economies of six, teams played additional 15 periods in economies of two. These data suggest that even after substantial learning in the simple setting of economies of two, decentralized side payments were unsuccessful at mitigating free driving (see fig. C.3).

25. Wild cluster bootstrap p-values from linear regressions corroborate the results from these nonparametric tests (see tables D.2, D.3). We use the wild cluster bootstrap procedure because of the small number of clusters (Cameron et al. 2008). Similar results are obtained when clustering standard errors at the economy level.

26. Order effects can complicate the comparison of economies with $N = 6$ vs. $N = 2$. To assess the magnitude of possible order effects, we exploit between-economies variation in parts 1 and 2 in global geoengineering outcomes to study their impacts on part 3 global geoengineering outcomes. More precisely, we employ treatment-specific regressions, where the dependent variable is the global geoengineering level of one economy in a period of part 3 (table D.1). In Baseline, the independent variable is the average global geoengineering in parts 1 and 2 by the three economies that will form the economy of six in part 3. For Decentralized and Structured, we separately include the average global geoengineering level of parts 1 and 2. All estimated coefficients are statistically indistinguishable from zero ($p \geq .277$), suggesting that outcomes under $N = 2$ of the experiment played a minor role in shaping outcomes under $N = 6$.

Table 3. Frequency of Offers and Side Payments

Outcome	Frequency				Offer Pattern
	Decentralized		Structured		
$N = 2$:					
No side payments	71 %	13 (14 %)	52 %	48 (32 %)	No offer
		51 (57 %)		30 (20 %)	Offer not taken
With side payments	29 %	19 (21 %)	48 %	72 (48 %)	From team w/low ideal point
		7 (8 %)		–	From team w/high ideal point
	90 (100%)		150 (100%)		
$N = 6$:					
No side payments	32 %	3 (3 %)	72 %	19 (13 %)	No offer
		26 (29 %)		88 (59 %)	Offer not taken
With side payments	68 %	36 (40 %)	28 %	8 (5 %)	From one team
		20 (22 %)		29 (19 %)	From two teams
		5 (6 %)		6 (4 %)	From more than two teams
		90 (100%)		150 (100%)	

Note. One observation corresponds to one economy in a period. Teams with the highest ideal point could make offers only in the Decentralized treatment. When two offers were made under $N = 6$, usually the senders were the two M teams in Structured and one L team and one M team in Decentralized. Under $N = 6$, 93 side payments occurred in Decentralized and 84 side payments occurred in Structured.

that multilateralism had no impact on the results. In fact, efficiency was low, especially under $N = 6$ in both Decentralized and Structured.

Result 2: Economies with more side payments generally reached higher total surplus.

Evidence for result 2 comes from table 3 and figure 3. Side payments were much less frequent than predicted. According to theory, side payments should always be enacted in both Decentralized and Structured (propositions 2, 4, and 5).[27] Table 3 takes an economy in a period as a unit of observation and shows that, under $N = 2$, a side payment occurred in 29% of the observations in Decentralized and 48% of the observations in Structured. These patterns are almost flipped under $N = 6$, where the

27. For $N = 6$, we refer to the side payment equilibria.

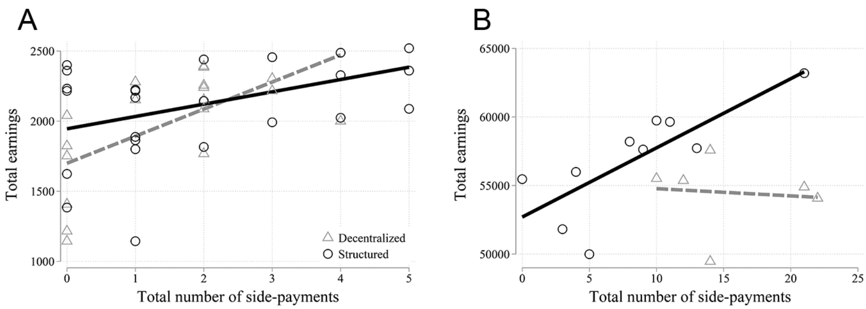

Figure 3. Higher total earnings with more side payments. The *dashed* (*solid*) line shows the linear prediction in Decentralized (Structured). One observation corresponds to an economy at the end of part 2 in panel *A* and part 3 in panel *B*: for $N = 2$, 18 observations in Decentralized and 30 observations for Structured; for $N = 6$, 6 observations in Decentralized and 10 observations in Structured.

percentage of observations with at least one side payment rises to 68% in Decentralized, while it shrinks to 28% in Structured.

Next, we examine the correlation between the frequency of side payments and the total surplus in the economy within each treatment condition. Figure 3 illustrates a scatter plot of the cumulative number of side payments versus total earnings, including a line of best fit. For $N = 2$, there is a statistically significant and positive relationship in both Decentralized and Structured treatments ($p = .006$ and $p = .007$, respectively). For $N = 6$, the correlation is significant and positive in Structured ($p = .008$) but not in Decentralized ($p = .858$). Therefore, under $N = 6$, there were fewer side payments in Structured than in Decentralized, but they were more effective. This treatment difference can be traced back to the coordination complexity characterizing the Decentralized treatment with $N = 6$. Recall that two side payments are necessary to compensate both high-ideal-point decision makers and reach the social optimum. While every single offer in Structured automatically targeted both *H* teams, in Decentralized, the offer senders had to coordinate on whom to target. Of the 53 cases where at least two offers to *H* teams were made in Decentralized, in 36% of them, only one of the two *H* teams was targeted. This evidence highlights the lack of coordination among offer senders and can explain why global geoengineering was on average higher in Decentralized than in Structured.

3.2. Why Were Side Payments Infrequent?

There are three possible obstacles behind the small number of side payments. First, a small number of offers were made in the first place. Second, offers were not mutually beneficial or were otherwise unattractive. Third, even profitable offers were turned down by their receivers. In this section, we explore the empirical relevance of each of these obstacles.

> **Result 3**: In Decentralized and Structured, many offers were made under both $N = 2$ and $N = 6$.

Evidence for result 3 comes from table 3 and figure 4. The absence of offers under $N = 6$ could be an indication of economies coordinating on the free-driving equilibrium that survives even in Decentralized and Structured (see proposition 5). However, our data show that teams tried to avoid this inefficient equilibrium: only in 3% of cases no offer was made in Decentralized and 13% of cases in Structured (table 3, $N = 6$). Usually, one or more offers were made each period. Yet, receivers often did not satisfy the requests (59% in Structured and 29% in Decentralized). Hence, for $N = 6$, a lack of offers was not a primary reason for the low number of side payments. Similar patterns emerge under $N = 2$ where the cases without any offer were 14% in Decentralized and 32% in Structured. Among the offers made, 63% were not taken in Decentralized and 20% were not taken in Structured.

Figure 4 provides more details by showing the absolute number of offers and side payments over time in an average economy of Structured and Decentralized. The figure shows that there were more offers in economies of six (periods 11–25) than in economies of two (periods 6–10). The reason for this pattern is most likely a mechanical one: more teams can make offers under $N = 6$. The figure also shows that the average numbers of side payments (i.e., offers that were taken) remained well below

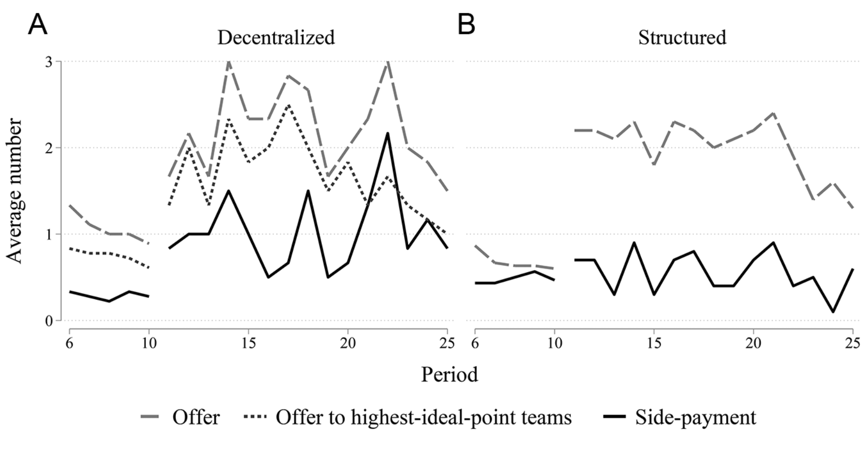

Figure 4. Offers and side payments over time. Economies of two in periods 6–10 (part 2) and economies of six in periods 11–25 (part 3). One observation corresponds to one economy in a period.

the predictions of one and two side payments for $N = 2$ and $N = 6$, respectively, with a relatively stable trend. The trend in the number of offers is approximately flat or declining in all treatment conditions. To facilitate treatment comparisons in figure 4, we added an extra line for the Decentralized treatment representing the subset of offers made by decision makers with the lowest ideal point toward those with the highest ideal point.[28]

To understand what went wrong, we analyze the type of offers that were made with particular attention to their incentive compatibility. The analysis focuses on $N = 2$ because this condition provides the cleanest setting to evaluate offers. Assessing the profitability of an offer under $N = 6$ is complex for three reasons. First, the ex post evaluation of an offer should take into account also the offers made by others. Second, given that offers were made simultaneously, ex ante, there exists considerable strategic uncertainty on the other teams' strategies. Third, the presence of multiple teams with the same ideal point raises an issue of coordination. These complexities are absent or less relevant under $N = 2$.

Result 4: While offers were generally profitable for the senders, many were unprofitable for the receivers, especially in Decentralized.

Support for result 4 is in figure 5. We define an offer as profitable when it could yield a higher or equal payoff to the free-driving payoff.[29] Under $N = 2$, the overwhelming majority of offers targeting teams with the highest ideal point were profitable for the sender (88% in Decentralized and 89% in Structured). Were these offers also profitable for their receivers? There is a substantial treatment difference: only 28% of offers were beneficial for the receiver in Decentralized, while 79% of the offers were beneficial for the receiver in Structured. Mutually beneficial offers represented 21% of the total offers in Decentralized and 76% of the total offers in Structured.[30] Hence, there was no shortage of offers in both treatments, but their quality was a key issue in Decentralized and much less so in Structured.

28. We consider only offers from one decision maker to another under $N = 2$ and from four decision makers to the other two under $N = 6$. This restriction mimics the rules of Structured, where high-ideal-point teams were not allowed to make offers to low-ideal-point teams.

29. When evaluating offers' profitability in Decentralized, to account for the fact that teams could also request an effort greater or smaller than a certain level, we make the following assumptions. If an offer to a high-ideal-point team is conditional on a maximum effort, we assume that the receiver will exert exactly that maximum effort level. When the offer is conditional to a minimum effort, we assume that the high-ideal-point team's effort will equal its ideal point. See the fig. 5 caption for more details.

30. With experience, the share of mutually beneficial offers tends to increase in the Structured treatment and decrease in the Decentralized treatment. However, as shown in fig. C.5, both trends are of modest magnitude.

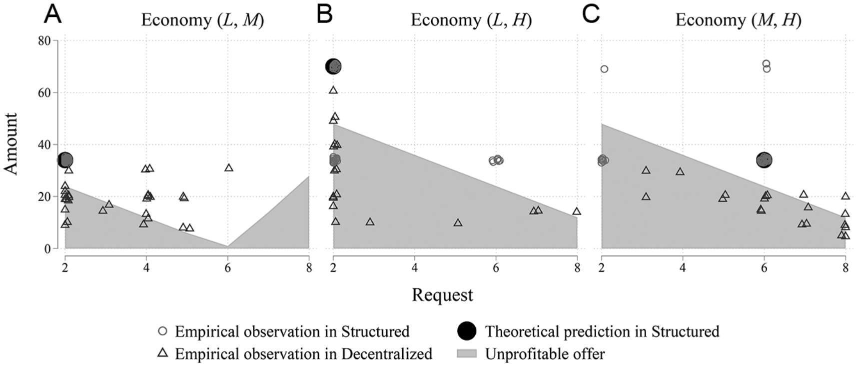

Figure 5. Profitable and unprofitable offers for the receivers. One observation corresponds to an offer made by a low-ideal-point team in a period to a high-ideal-point team. In Decentralized, if the low-ideal-point team requested was an effort lower than a certain level (32 observations), it was re-coded as an effort request equal to that level minus 1 (i.e., the most profitable choice for the high-ideal-point team wanting to meet the request); if the low-ideal-point team requested an effort greater than a certain level (1 observation), it was re-coded as an effort request equal to the ideal point of the receiver. Observations were jittered to display them better.

Figure 5 provides disaggregated evidence on the offers in each type of $N = 2$ economy. Each panel shows a scatter plot of the offered amount versus the requested effort. The shaded area indicates offers that were unprofitable for the receiver. In the following result, we investigate which offers were less likely to be taken.

Result 5: Many profitable offers were not taken in Decentralized and only some in Structured.

In the Decentralized treatment, 58% of the profitable offers to the highest-ideal-point teams were taken (11 out of 19). In the Structured treatment, instead, 82% of the profitable offers were taken (66 out of 81). Hence, offers' receivers were also partially responsible for the low number of side payments. If we focus on mutually profitable offers, patterns remain similar: 43% of these offers were taken in Decentralized (6 out of 14) and 81% in Structured (62 out of 77). Finally, receivers rarely took offers that were unprofitable for them (17% in Decentralized and 29% in Structured).

3.3. Evidence from the Chat

Using within-team chat dialogues, we explore the relevance of some behavioral mechanisms that might have led to the empirical results reported above.[31] Chat activity can shed light on reasoning in decision making, and we study both the volume and content of chats.

31. Abatayo et al. (2020) did not analyze the chat data.

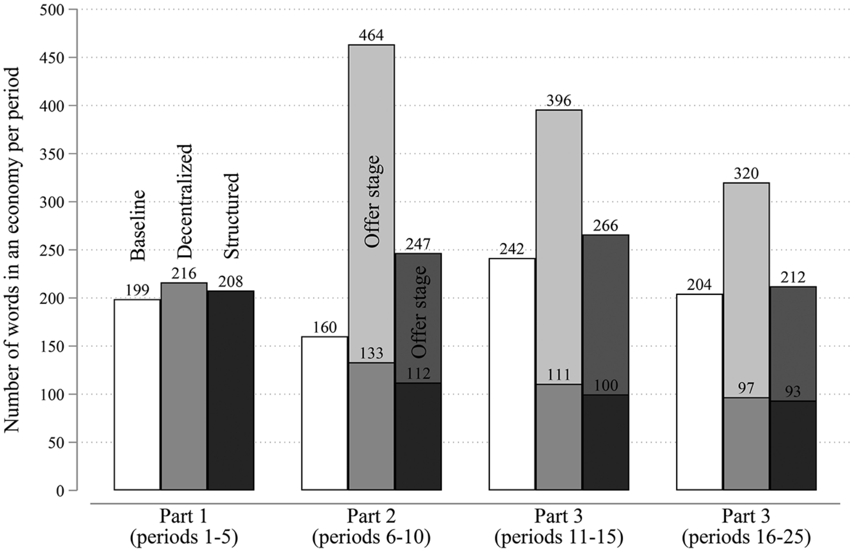

Figure 6. Chat activity by treatment and experience. Statistics are normalized for comparability across parts. The figure always reports the chat volume of six teams. To account for differences in economy size (two vs. six teams), in parts 1 and 2, we summed up the chat volume of three economies of two. Moreover, the time for chatting decreased in later periods. When chat time was less than two minutes, we normalized the chat volume using the share of words in the previous period that were recorded during the (now) unavailable time.

We report three main insights based on the evolution of chat volume (fig. 6). First, in part 1, where all teams played Baseline with $N = 2$, the chat volume is similar across treatments (three bilateral Wilcoxon-Mann-Whitney tests: $p \geq .359$, 18 observations for Baseline and Decentralized, 30 observations for Structured). This pattern is consistent with the evidence on other outcomes, and it reassures us that the behavior of participants in all three treatments was highly comparable under Baseline with $N = 2$. Second, in Baseline, where teams played under $N = 2$ for five additional rounds compared to the other treatments, the chat activity significantly declined with experience (Wilcoxon signed-rank test of part 1 vs. part 2: $p < .001$, 18 observations). This pattern suggests in learning that team members converge on an agreement and a reduction.[32] Third, under $N = 6$, we observe that the chat volume decreases with experience in all treatments (three Wilcoxon sign-rank tests of periods 11–15 vs. 16–25:

32. In Decentralized and Structured, adding the offer stage in part 2 makes teams' decisions richer and doubles the opportunities for chatting. Consequently, we observe an increase in chat volume in part 2, especially for Decentralized (two Wilcoxon signed-rank tests of part 1 vs. part 2: $p < .001$ with 6 observations for Decentralized, $p = .004$ with 10 observations for Structured). A

$p = .156$ for Baseline with 6 observations, $p = .031$ for Decentralized with 6 observations, $p = .006$ for Structured with 10 observations).[33] Therefore, this evidence indirectly suggests that participants achieved, over time, a good comprehension of the experimental settings.

To investigate the chat content, we instructed two research assistants to classify messages according to nine prespecified categories. Here we discuss the categories we consider to be the most insightful: confusion, coordination, social preferences, and the usefulness of offers. Furthermore, since we are primarily interested in the drivers behind offers, we focus on the chat in the offer stage of Decentralized and Structured.[34] Messages related to confusion, either within the team or from other teams, are few in both Decentralized and Structured (7% and 3%, respectively, when pooling parts 2 and 3). The frequency is higher in economies of six than in economies of two, but it does not seem a major concern (6% vs. 9% in Decentralized and 1% vs. 5% in Structured). More prevalent are the difficulties of predicting other teams' behavior and discussions of postponing offers to study the others' behavior (coordination category: 13% in Decentralized and 6% in Structured when pooling parts 1 and 2). These concerns are similar in economies of two and six (14% vs. 13% in Decentralized and 4% vs. 8% in Structured). Considerations originating from prosocial and antisocial preferences are uncommon (3% and 1%, respectively, when pooling parts 1 and 2). What is central in the conversations is instead the uselessness of offers. This topic represents 33% of the chat in Decentralized and 30% in Structured.

---

clean comparison between Decentralized and Structured requires excluding the highest-ideal-point teams in the economy because, by design, they could not participate in the offer stage of Structured. When doing so, we still detect significantly higher volumes in Decentralized than in Structured when pooling parts 1 and 2 data (two Wilcoxon-Mann-Whitney tests: $p = .042$ in both offer and effort stages, 6 observations for Decentralized, 10 observations for Structured).

33. Decentralized sticks out as having more chatting than the other treatments (e.g., two Wilcoxon-Mann-Whitney tests of Decentralized vs. Structured in periods 11–15 or 16–25: $p < .001$, 6 observations for Decentralized, 10 observations for Structured). When side payments become available, most chat activity happens during the offer stage, as one would expect from backward induction. When pooling parts 2 and 3, the chat volume in the effort stage is similar across treatments (Wilcoxon-Mann-Whitney test, $p = .313$). Moreover, teams with the highest ideal point debate less, as expected, since they should not make side payments to others (Wilcoxon sign-rank tests pooling parts 2 and 3, highest G^* vs. others in Decentralized, $p = .031$ for both offer and effort stages).

34. Appendix H provides a description of the coding procedure and the definition of chat categories. We pool the classifications of both research assistants. A category breakdown of chat dialogues is in fig. H.2, which also includes the "smooth" category. This category captures a large share of dialogues and includes comments on other teams' decisions, sometimes highlighting frustration.

We complement the chat evidence with a final set of analyses. Here we summarize the main findings (see app. G for a detailed discussion). To corroborate the evidence on (the lack of) confusion, we perform three exercises using the number of mistakes in the quiz on instructions as a proxy of teams' confusion.[35] First, we study treatment differences in the distribution of mistakes and find that, while mistakes are few in all treatments (the median team made no mistakes), slightly more mistakes were made in Structured than Decentralized (on average 0.36 vs. 0.72, Wilcoxon-Mann-Whitney test: $p = .148$, 36 observations for Decentralized, 60 observations for Structured). If confusion (as proxied by mistakes) was key, we would have expected the opposite result since Structured was more effective than Decentralized. Second, we reestimate treatment effects controlling for the number of quiz mistakes and find that results are robust for both $N = 2$ and $N = 6$ (see tables D.2 and D.3; tables D.1–D.6, E.1, E.2, G.1, G.2 are available online).[36] Third, along with other covariates, we correlate mistakes with the decision to make an offer, make a mutually beneficial offer, or take an offer (see table G.2).[37] Teams with a poorer understanding of the instructions were weakly less likely to make offers in general ($p = .141$ in Decentralized and $p = .033$ in Structured), but not so much mutually beneficial ones. Moreover, confusion negatively correlates with the decision to take an offer in Decentralized ($p = .033$) but not in Structured ($p = .775$). All in all, the empirical evidence suggests that confusion was not the main driver of our results.

The same set of regressions brings insights into another possible mechanism: the size of the stake might be perceived as too low by participants to take side payments seriously. The stake to offer a mutually beneficial offer was highest for low-ideal-point teams in (L, H) economies, where the distance between the two ideal points is maximal. Thus, if the stake size mattered, we should expect a positive and significant coefficient for the (L, H) dummy. However, the estimated coefficient is negative.[38] Finally, within the same regression framework, coefficients related to altruism and negative reciprocity—elicited using the method proposed in Falk et al. (2018)—are mostly small and statistically insignificant, confirming the chat evidence.

---

35. Although we collected mistakes at the participant level, we prefer to focus on teams as they are the decision-making unit in the experiment. We coded a mistake as such if both team members gave the wrong answer to a given control question. Results are similar when adopting alternative definitions of mistakes. While estimated treatment effects are robust to the inclusion of this control, on average, economies of six in which teams were more confused achieved higher global geoengineering levels with negative repercussions on efficiency. This was not the case for economies of two.

36. Confusion does not appear to influence Baseline results in the first place. Teams' average efforts were close to predicted levels (see table G.2). Teams also correctly modified their effort levels when moving from economies of two to economies of six.

37. We focus on economies of two because this simple setting allows a cleaner test.

38. A rationale for this result is discussed in app. G.

4. CONCLUSION

As global temperatures increase and climate risks intensify, geoengineering might gain popularity as a strategy to cope with global warming, even in the presence of aggressive emission reductions. Yet, solar geoengineering can reconfigure the strategic relations among countries regarding the economics and politics of climate change. Although any consideration is speculative at this stage because of the absence of any large-scale implementation of solar radiation management, we have focused on some of the governance challenges that might emerge and empirically tested a possible solution. Specifically, we conducted a laboratory experiment testing a widely used economic mechanism: side payments. In situations with low transaction costs and perfect information, like the one studied here, voluntary agreements à la Coase are expected to be welfare improving. We investigate the empirical effectiveness of using side payments to contain excessive geoengineering.

Our findings based on the public good-or-bad game do not support a Coasian prediction. Although excessive geoengineering was reduced under structured side payments, efficiency gains were just 3%–8% of the total surplus compared to the Baseline treatment without side payments. This pattern appears across all treatments examined, wherein we varied the economy size and the institutional rules for side payments. Why was allowing side payments not sufficient to overcome free driving? Not enough side payments took place: depending on the treatment, their frequency ranged from 28% to 64% instead of 100%. Even when side payments were enacted, they were not effective enough in curbing free driving. Our findings are in line with previous experiments on the Coase theorem in which free-form communication was not available (Jacques 1992; Spencer and Shogren 2000), as in our setting.[39]

Three main conclusions emerge from our study. First, the availability of side payments had a smaller-than-expected effect, even in simple economies of two, where the theoretical prediction is sharp. We document this result in both the Decentralized and Structured treatments. According to exploratory analyses, behavioral drivers such as confusion, social preferences, and low stakes cannot fully explain this empirical finding.[40]

Second, coordination failure over side payments was a major issue. In theory, limiting global geoengineering requires decision makers with lower ideal points to coordinate their side-payment offers. In most of the experimental economies, instead, only a single decision maker offered side payments. The lack of coordination made side payments largely ineffective in limiting geoengineering efforts.

39. Most experiments on the Coase theorem employ a free-form communication protocol and typically find that negotiating parties achieve efficient results. A complete overview of the experiments on the Coase theorem is reported in table D.6.

40. Other behavioral factors, like the endowment effect, could have played a role, but the current experiment was not designed to test for it. For instance, high-ideal-point decision makers might develop a sort of entitlement of a level of geoengineering equal to their ideal point, and hence they require a side payment higher than predicted to reduce their effort.

Third, the institutional framework of a structured negotiation, mimicking negotiation blocs with agreed-upon compensation schemes, performed better than that of decentralized agreements. One lesson from the experiment is the necessity of providing structure to the negotiation, as Pareto-dominant equilibria may not be achieved unless proper institutions channel decision makers toward them. Economic theory is silent on this point, and our results warn that the wide strategy spaces and the high number of parties characterizing international negotiations may pose problems for effective side-payment schemes.

There exists, of course, an issue with external validity because the study is an unframed laboratory experiment conducted using students as participants. The goal of our experimental setup is not to mimic all aspects of reality but to provide the groundwork for future research. The inclusion of economies of six is an attempt to bring the experiment closer to the reality of complex negotiations. As such, our results carry implications for sovereign countries' strategic behavior. In particular, expanding the negotiations' space to include compensatory side payments may not necessarily be a panacea. Although this might be disappointing from the point of view of standard economic theory, the experience of negotiations, including the complexities around Article 6 of the Paris Agreement, confirms that economic transfers are one of the many tools available in a negotiation arsenal and not necessarily the most important one. As one of the biggest obstacles to negotiations seems to be miscoordination, institutional capacity needs to be built up to deal with additional complexities such as those brought about by climate engineering.

In the field, including that of climate negotiations, complexity is arguably higher than the one simulated in this experiment. For instance, we ruled out imperfect information about the counterparts' ideal points and secret offers. Furthermore, the maximum number of parties involved was six rather than hundreds. Climate negotiations involve years of meetings of parties that provide abundant space for communication. In addition, the institution of negotiation blocs, involving smaller groups agreeing on a subset of acceptable strategies and any other governance mechanism that might reduce miscoordination, should all be factors increasing the success of international negotiations. Leaving out geoengineering from climate negotiations might lead to unilateral and fragmented actions to cool the planet, which would be far worse in terms of global welfare. The increasing likelihood of disruptive strategies such as geoengineering thus calls for strengthening and, possibly, restructuring the international institutions dealing with climate change.

REFERENCES

Abatayo, Anna Lou, Valentina Bosetti, Marco Casari, Riccardo Ghidoni, and Massimo Tavoni. 2020. Solar geoengineering may lead to excessive cooling and high strategic uncertainty. *Proceedings of the National Academy of Sciences* 117 (24): 13393–98.

Balliet, Daniel. 2010. Communication and cooperation in social dilemmas: A metaanalytic review. *Journal of Conflict Resolution* 54 (1): 39–57.

Barrett, Scott. 2005. The side payments game. In *Environment and statecraft: The strategy of environmental treaty-making*, 1457–1516. Oxford: Oxford University Press.

———. 2008. The incredible economics of geoengineering. *Environmental and Resource Economics* 39 (1): 45–54.

Barrett, Scott, and Astrid Dannenberg. 2012. Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences* 109 (43): 17372–76.

———. 2017. Tipping versus cooperating to supply a public good. *Journal of the European Economic Association* 15 (4): 910–41.

Calzolari, Giacomo, Marco Casari, and Riccardo Ghidoni. 2018. Carbon is forever: A climate change experiment on cooperation. *Journal of Environmental Economics and Management* 92:169–84.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics* 90 (3): 414–27.

Cason, Timothy. 1995. An experimental investigation of the seller incentives in the EPA's emission trading auction. *American Economic Review* 85 (4): 905–22.

Charness, Gary, and Matthias Sutter. 2012. Groups make better self-interested decisions. *Journal of Economic Perspectives* 26 (3): 157–76.

Dengler, Sebastian, Reyer Gerlagh, Stefan T. Trautmann, and Gijs Van De Kuilen. 2018. Climate policy commitment devices. *Journal of Environmental Economics and Management* 92:331–43.

Dinar, Shlomi. 2006. Assessing side-payment and cost-sharing patterns in international water agreements: The geographic and economic connection. *Political Geography* 25 (4): 412–37.

Dixit, Avinash, and Mancur Olson. 2000. Does voluntary participation undermine the Coase theorem? *Journal of Public Economics* 76 (3): 309–35.

Ellingsen, Tore, and Robert Östling. 2010. When does communication improve coordination? *American Economic Review* 100 (4): 1695–1724.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. Global evidence on economic preferences. *Quarterly Journal of Economics* 133 (4): 1645–92.

Fischbacher, Urs. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10:171–78.

Ghidoni, Riccardo, Giacomo Calzolari, and Marco Casari. 2017. Climate change: Behavioral responses from extreme events and delayed damages. *Energy Economics* 68:103–15.

Greiner, Ben. 2015. Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association* 1 (1): 114–25.

Heyen, Daniel, Joshua Horton, and Juan Moreno-Cruz. 2019. Strategic implications of counter-geoengineering: Clash or cooperation? *Journal of Environmental Economics and Management* 95:153–77.

Hoffman, Elizabeth, and Matthew L. Spitzer. 1982. The Coase theorem: Some experimental tests. *Journal of Law and Economics* 25 (1): 73–98.

———. 1985. Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *Journal of Legal Studies* 14 (2): 259–97.

Horton, Joshua B., Jesse L. Reynolds, Holly Jean Buck, Daniel Callies, Stefan Schäfer, David W. Keith, and Steve Rayner. 2018. Solar geoengineering and democracy. *Global Environmental Politics* 18 (3): 5–24.

Irvine, Peter, Kerry Emanuel, Jie He, Larry W. Horowitz, Gabriel Vecchi, and David Keith. 2019. Halving warming with idealized solar geoengineering moderates key climate hazards. *Nature Climate Change* 9 (4): 295–99.

Jacques, Stephanie. 1992. The endowment effect and the Coase theorem. *American Journal of Agricultural Economics* 74 (5): 1316–23.

Jakob, Michael, Dorothea Kübler, Jan Christoph Steckel, and Roel van Veldhuizen. 2017. Clean up your own mess: An experimental study of moral responsibility and efficiency. *Journal of Public Economics* 155:138–46.

Kravitz, Ben, Douglas G. MacMartin, Daniele Visioni, Olivier Boucher, Jason N. S. Cole, Jim Haywood, Andy Jones, et al. 2021. Comparing different generations of idealized solar geoengineering simulations in the Geoengineering Model Intercomparison Project (GeoMIP). *Atmospheric Chemistry and Physics* 21 (6): 4231–47.

Kverndokk, Snorre. 1993. Coalitions and side payments in international CO_2 treaties. Discussion Papers Research Department, Central Bureau of Statistics, Norway 97.

McGinty, Matthew, Garrett Milam, and Alejandro Gelves. 2012. Coalition stability in public goods provision: Testing an optimal allocation rule. *Environmental and Resource Economics* 52 (3): 327–45.

Proctor, Jonathan, Solomon Hsiang, Jennifer Burney, Marshall Burke, and Wolfram Schlenker. 2018. Estimating global agricultural effects of geoengineering using volcanic eruptions. *Nature* 560 (7719): 480–83.

Rickels, Wilfried, Martin F. Quaas, Katharine Ricke, Johannes Quaas, Juan Moreno-Cruz, and Sjak Smulders. 2020. Who turns the global thermostat and by how much? *Energy Economics* 91 (C): 104852.

Saelen, Håkon. 2016. Side-payments: An effective instrument for building climate clubs? *International Environmental Agreements: Politics, Law and Economics* 16 (6): 909–32.

Schelling, Thomas C. 1996. The economic diplomacy of geoengineering. *Climatic Change* 33 (3): 303–7.

Sherstyuk, Katerina, Nori Tarui, Majah-Leah V. Ravago, and Tatsuyoshi Saijo. 2016. Intergenerational games with dynamic externalities and climate change experiments. *Journal of the Association of Environmental and Resource Economists* 3 (2): 247–81.

Spencer, Michael A., and Jason F. Shogren. 2000. Protocol for inexperienced Coasean bargainers confronting delay costs. *Resource and Energy Economics* 22 (1): 79–90.

Sugiyama, Masahiro, Yosuke Arino, Takanobu Kosugi, Atsushi Kurosawa, and Shingo Watanabe. 2018. Next steps in geoengineering scenario research: Limited deployment scenarios and beyond. *Climate Policy* 18 (6): 681–89.

Tavoni, Alessandro, Astrid Dannenberg, Giorgos Kallis, and Andreas Löschel. 2011. Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences* 108 (29): 11825–29.

Victor, David G., M. Granger Morgan, Fay Apt, and John Steinbruner. 2009. The geoengineering option: A last resort against global warming. *Foreign Affairs* 88:64.

Weitzman, Martin L. 2015. A voting architecture for the governance of free-driver externalities, with application to geoengineering. *Scandinavian Journal of Economics* 117 (4): 1049–68.