# Leveraging transfers to strengthen climate cooperation under the threat of solar geoengineering

Strategic Behavior and the Environment

David M. McEvoy
*Department of Economics, Appalachian State University, Boone, North Carolina, USA, and*

Matthew McGinty
*Department of Economics, University of Wisconsin – Milwaukee, Milwaukee, Wisconsin, USA*

Received 10 October 2025
Revised 14 February 2026
Accepted 31 May 2026

**Abstract**

The authors explore the role of transfers in a model of international environmental agreements (IEAs) on abatement under the threat of solar geoengineering (SGE) deployment. Transfers can expand the set of stable IEAs, reduce SGE deployment and lead to efficiency gains by closing more of the abatement gap. The authors show that transfers are part of an indirect governance approach to limit harmful SGE deployment by increasing global emissions abatement.

**Keywords** Solar geoengineering, Solar radiation management, International environmental agreements, Self-enforcing agreements, Global public goods, Transfers

**Paper type** Research paper

## 1. Introduction

As global greenhouse gas (GHG) emissions continue to rise there is growing interest in the potential of solar geoengineering (SGE) – technologies to reflect a fraction of sunlight back into space – to prevent impending climate damages. These techniques could cool the planet quickly and relatively cheaply. Despite their potential, SGE approaches are controversial for a number of reasons. One is that they are untested and could be inherently risky. Another is that their very availability may weaken the already lackluster collective resolve to mitigate GHG emissions; this is the "so-called" moral-hazard concern (Moreno-Cruz *et al.*, 2025). Another issue is that because SGE technologies are relatively inexpensive, it is possible they are deployed unilaterally, which brings rise to externalities imposed by a free driver (Wagner and Weitzman, 2012; Weitzman, 2015). The "free" part is an exaggeration of how cheap SGE deployment is and the "driver" part captures the ability of a single country to determine the global thermostat. This paper is motivated by these governance challenges and investigates how effective international climate agreements are in the face of a SGE free driver. Our contribution is investigating the role of financial transfers between IEA members to enhance climate cooperation in the face of a free driver.

We model the benefits and costs of abatement, and the impacts of SGE deployment following McEvoy *et al.* (2024). It is helpful to summarize a few important features of their setup. Deploying SGE can partially protect countries from impending climate damages they face because of inadequate emissions abatement. Countries are heterogeneous in their optimal SGE deployment levels, which is the only source of heterogeneity in the model. SGE

**JEL classification** – C7, D7, F5, H4

Strategic Behavior and the Environment

1944-012X
DOI 10.1108/SBE-10-2025-0001

is a good-or-bad (a “GoB”), in which deployment at or below a country’s optimal level is beneficial and beyond that level SGE turns costly (Weitzman, 2015). McEvoy *et al.* (2024) demonstrate that the threat of a menacing free driver creates additional incentives for some countries to abate emissions because the free driver’s optimal deployment level decreases as more emissions are abated. They refer to these as “anti-driver” incentives and if they are strong enough it is possible that the emergence of SGE can lead to more abatement compared to a world without SGE. However, consistent with the seminal work by Barrett (1994), McEvoy *et al.* (2024) show that stable agreements will tend to be small and aggregate abatement levels well below socially optimal levels. This paper introduces the possibility of transfers to the model of McEvoy *et al.* (2024).

The financial transfers we consider are not intended to directly stop the free driver. This is because the notion of the free driver is fluid, as it describes the country that prefers the lowest mean temperature and has the capability to deploy SGE. Even if a free driver could be stopped, it would simply give rise to the next one: the country with the second-lowest mean temperature preference. And so on. The financial transfers in our model can impact the free driver only indirectly, by potentially enhancing collective abatement levels within international environmental agreements (IEAs). In contrast, Bakalova and Belaia (2023) explore an IEA that directly governs SGE to analyze the effect of using side payments to curb the free driver’s deployment.

Following the self-enforcing IEA literature, we model a voluntary agreement between countries in which the members commit to choosing abatement levels to maximize their collective payoffs (Barrett, 1994; McGinty, 2007). Stable agreements are defined by being both *internally stable* – no member country wants to leave – and *externally stable* – no nonmember country wants to join. Without transfers, for an agreement to be internally stable, all members must each find it profitable to remain inside the agreement given their individual abatement responsibilities. Given that countries are heterogeneous, some IEA members may gain a large surplus from being in an agreement while others just break even. We model an IEA that leverages the positive surplus, where members agree to share the benefits of collective abatement to expand the set of stable agreements.

In principle, transfers or side payments, have the potential to increase the provision of global public goods (Chander and Tulkens, 1995; Barrett, 2001; McGinty, 2007; Sælen, 2016; Kapstein, 2025). Mechanisms to facilitate financial transfers are common in existing international agreements, including the Green Climate Fund (GCF) under the Paris Agreement. Tens of billions of dollars are transferred annually from Annex I countries (Global North) to non-Annex I countries (Global South) through the GCF to help fund mitigation and adaptation activities related to climate change. As SGE deployment becomes more likely, institutions like the GCF are well positioned to facilitate financial transfers linked to emissions abatement responsibilities.

Ghidoni *et al.* (2023) is the only paper we are aware of that directly explores transfers or side payments in the context of SGE. They conduct a series of laboratory experiments that test the effectiveness of using side payments to limit the deployment of SGE that has “GoB” properties. They find that side payments do very little to enhance cooperation above the baseline scenario. Our study is different from Ghidoni *et al.* (2023) in two fundamental ways. First, they do not consider mitigation in their study. The only decisions are on SGE deployment and therefore they cannot explore the interaction between SGE and mitigation, which underlies the moral hazard concern with SGE. Our study explicitly links emissions and SGE deployment. Second, the transfers in our model are tied to emissions decisions not SGE decisions. Following McEvoy *et al.* (2024), in our study, SGE technologies allow countries to protect themselves (imperfectly) from climate damages resulting from inadequate emissions control. Therefore, SGE deployment follows mitigation decisions and the free driver's optimal level of SGE decreases in mitigation.

We show that transfers can expand the size and effectiveness of stable IEAs that govern abatement under the threat of a free driver. We measure effectiveness by how much of the abatement gap between the Nash equilibrium and the social optimum is closed by the IEA. IEAs with transfers are able to close more of the gap compared to agreements without transfers. While McEvoy *et al.* (2024) showed that strong anti-driver incentives can lead to relatively high abatement levels, we show that transfers can help achieve similar outcomes without such extreme external damages imposed by the free driver. The important implication is that transfers can enhance the indirect-governance approach of collectively increasing abatement to contain the free driver and capture efficiency gains.

Our approach is to compare the outcomes and effectiveness measures of IEAs with and without transfers under the threat of an SGE free driver. The next section briefly summarizes the relationship between emissions abatement and SGE deployment. In section three, we consider an IEA with a transfer schemes and characterize the conditions for stable agreements. In section four we demonstrate the impact of transfers through examples and comparative statics. Then we conclude.

## 2. Emissions abatement and solar geoengineering

Following Moreno-Cruz (2015) and McEvoy *et al.* (2024), we assume that SGE-deployment decisions follow abatement decisions. Following Barrett (1994, 2003) GHG emissions abatement is modelled as a pure global public good, with $n$ countries, individual abatement denoted as $q_i$ and aggregate abatement is denoted as $Q$ where $Q = \sum_{i \in N} q_i$. The global benefit from abatement is $B(Q) = bQ$ and each country receives the same benefit share $\frac{1}{n}$ and therefore each country's benefit is $B_i(Q) = \frac{bQ}{n}$. Individual abatement costs are convex and specified as $C_i(q_i) = \frac{c(q_i)^2}{2}$. A country's payoff is $B_i(Q) - C_i(q_i)$ or:

$$\pi_i(q_i, Q) = \frac{bQ}{n} - \frac{cq_i^2}{2}, \tag{1}$$

and Nash equilibrium abatement levels are $q^* = \frac{b}{cn}$ and $Q^* = \frac{b}{c}$. The socially optimal abatement levels are $q^o = \frac{b}{c}$ and $Q^o = \frac{bn}{c}$ and the difference between the optimal and noncooperative aggregate abatement levels is:

$$Q^o - Q^* = \frac{b(n-1)}{c}. \tag{2}$$

If the noncooperative abatement level equals the optimal level, then there is no benefit from deploying SGE. When that difference is positive, countries can deploy SGE as an additional way to avoid climate damages. Thus, SGE deployment can lead to benefits in the same way that emissions abatement does. However, SGE cannot perfectly substitute for abatement because it does not address the source of the problem, nor can it offset all the damages caused by GHG emissions (e.g. continued ocean acidification) (Robock, 2008; Felgenhauer *et al.*, 2022). Following Weitzman (2015), Abatayo *et al.* (2020) and McEvoy *et al.* (2024), countries are heterogeneous in their payoff-maximizing level of SGE deployment.

Following McEvoy *et al.* (2024), we model a country's payoff-maximizing deployment of SGE as inversely related to the level of global emissions abatement. As mitigation efforts reduce climate damages, the incentive for any individual country to rely on SGE correspondingly declines. Cross-country variation in the benefits derived from SGE is captured by the parameter $\gamma_i$. Country $i$'s payoff-maximizing or "preferred" level of SGE is a function of the *abatement gap*, denoted as $G_i^p$ and takes the following form:

$$G_i^p = \gamma_i(Q^o - Q). \tag{3}$$

In this way, SGE can be both a "good" or a "bad" depending on the realized level, referred to as a "GoB" in the literature (Weitzman, 2015). We denote the realized level of aggregate SGE for all $n$ countries as $G$. Let $\beta(G)$ be global benefit from SGE, $\beta_i(G)$ be individual benefit and $\beta(G) = \sum_{i \in N} \beta_i(G)$. Country $i$'s benefit function from SGE takes the following form:

$$\begin{aligned} \beta_i(G) &= \theta G \quad \text{if} \quad G \le G_i^p \\ \beta_i(G) &= \theta G_i^p - \phi\left(G - G_i^p\right) \quad \text{if} \quad G > G_i^p, \end{aligned} \tag{4}$$

where $\theta$ is the marginal benefit to SGE when $G$ is less than country $i$'s preferred level and $\phi$ is the marginal loss in benefits (i.e. marginal cost) for additional deployment of SGE beyond the preferred level. Each country's benefit function is maximized at $G_i^p$ (see McEvoy *et al.*, 2024, p. 862). The parameter restriction $\theta < \frac{1}{n}$ implies the maximum individual marginal benefit from SGE is less than the individual marginal benefit from emissions abatement. As $\theta$ increases toward the upper bound $\frac{1}{n}$SGE becomes a more perfect substitute for abatement.

Given SGE is free to deploy, we know that the country with the highest optimal value of SGE will be the only country deploying SGE (i.e. the free driver). If we rank countries from 1 to $n$ in terms of their payoff-maximizing level, then $G_n^p = G^{\max} = G$ is the free-driver level SGE. From there we can solve for the Nash and optimal levels of abatement and SGE deployment for all countries (see McEvoy *et al.*, 2024 for the analytical expressions). Some important results are worth summarizing. In a noncooperative equilibrium, the availability of SGE causes the free driver to reduce their abatement levels relative to a world without SGE. The non free driver countries, on the other hand, may increase or decrease their abatement given the free driver's deployment of SGE. The direction depends on how strong the "anti-driver" incentives are; that is, if the externality imposed by the free driver ($\phi\left(G - G_i^p\right)$) is substantial, then a country may find it optimal to increase their abatement to reduce $G^{\max}$.

## 3. IEAs with transfers under the threat of a free driver

The IEA has three stages. In stage 1 (participation), countries decide independently and simultaneously whether to join the IEA. In stage 2 (abatement), the IEA members choose abatement levels to maximize their joint (i.e. collective) payoffs. Meanwhile, nonmembers choose their abatement levels unilaterally (noncooperatively). Finally, in stage 3 (SGE), countries choose their SGE levels.

The game is solved by backward induction and the third stage is trivial given that the free driver – the country with the highest optimal SGE level – deploys $G_n^p = G^{\max}$ while all other countries deploy $G_i = 0$ so that $G = G^{\max}$. The precise level of $G$ is a decreasing function of aggregate abatement determined in stage 2.

### *3.1 Stage 2: abatement*

Let the members of an agreement be a coalition $S$ and the set of nonmember countries outside an agreement be denoted $T$ with $S \cup T = N$, where $N$ is the set of countries and $|N| = n$. The number of agreement members is denoted as $|S| = s$ and indexed by $i = 1, 2, \ldots, s$. The number of nonmembers is $|T| = n - s$ and indexed by $j = 1, 2, \ldots, (n - s)$.

In stage 2, the $n - s$ nonmembers have a dominant abatement strategy denoted as $q_j^t(G^{\max})$. Meanwhile, the $s$ agreement members determined from stage 1 choose $q_i^s$ to maximize joint payoffs, which is denoted as $q_i^s(G^{\max})$ (these two equations can be found in the Appendix).

The novel feature of the IEA is that the payoffs to the members are determined by the transfer mechanism, which has two important features (McGinty, 2011):

(1) *Incentive compatibility*: Each member is guaranteed an amount equal to their outside payoff from leaving the agreement, $\pi_i^t(S\ i)$.

(2) *Share of surplus*: Any positive surplus $\sigma(S)$ beyond the incentive compatibility payments is allocated equally among the $s$ IEA members.

The payoff after the transfer under an equal share is:

$$\pi_i^s(S, \tau_i) = \pi_i^t(S\ i) + \frac{\sigma(S)}{s}, \tag{5}$$

and the transfer $\tau_i$ is the difference between (5) and the payoff under an agreement without transfers:

$$\tau_i = \pi_i^s(S, \tau_i) - \pi_i^s(S).$$

Note that transfers are budget-balanced and purely redistributive.

### *3.2 Stage 1: participation*

In the first stage all countries decide independently and simultaneously whether to join the IEA. An equilibrium IEA is one that is both *internally* and *externally* stable. Internal stability (IS) requires that no member could increase their payoff by leaving the agreement and external stability requires that no nonmember could increase their payoff by joining. Without transfers, the internal stability condition for each IEA member is:

$$IS_i(S) = \pi_i^s(S) - \pi_i^t(S\ i) \geq 0. \tag{6}$$

Without transfers, the IS condition in (6) must be satisfied for all $s$ members of the agreement [1]. Introducing transfers changes the internal stability condition. With transfers, a coalition $S$ is internally stable if the *sum* of the individual internal stability conditions is positive and is expressed as:

$$\sigma(S) = \sum_{i \in S} IS_i(S) = \sum_{i \in S} \left(\pi_i^s(S) - \pi_i^t(S\ i)\right) \geq 0. \tag{7}$$

Clearly, if (6) is satisfied for all members of a given ($S$) then (7) must also be satisfied, which leads to the result that the set of stable IEAs with transfers is weakly greater than the set of stable IEAs without transfers. This was originally proved, without the threat of SGE, in McGinty (2007).

The intuition on why financial transfers may expand the size and effectiveness of stable IEAs is straightforward. With heterogeneous preferences for optimal SGE deployment, each member country faces different external costs imposed by the free driver's level of SGE.

Those with low $\gamma_i$s suffer more from high deployment and therefore they gain more from increasing abatement relative to those with $\gamma_i$s closer to the free driver's. Therefore, even though each member country takes on the same level of abatement, members have different payoffs. One can easily imagine an IEA member to a stable agreement enjoying a relative large surplus from being in the agreement. That surplus, in effect, is used to entice additional countries to join the agreement by sharing it among members.

## 4. Transfers in practice

To illustrate the impact of transfers, we start with the parameterized example from McEvoy *et al.* (2024) (example 1, page 19). In this case $n = 5$, which is collection of countries small enough to easily test stability conditions and calculate abatement levels and payoffs, but large enough to illustrate the impact of transfers on stable agreements and effectiveness. The distribution of SGE preferences is $\gamma_1 = 0.2$, $\gamma_2 = 0.4$, $\gamma_3 = 0.5$, $\gamma_4 = 0.6$ and $\gamma_5 = \gamma^{max} = 0.8$. Note that $\bar{\gamma} = 0.5$ and, like McEvoy *et al.* (2024) we set $b = c = 1$ and $\phi = \theta = 0.1$ [2].

Table 1 illustrates stable agreements both with and without transfers. The first column shows which countries are part of the agreement where the numbers 1, 2,[…], $n$ index countries depending on their relative payoff-maximizing level of SGE (where $n$ is the free-driver country). The top section of the table reproduces parts from McEvoy *et al.* (2024, page 876) and the bottom section contains additional stable agreements under an IEA with transfers. Note that there are six stable agreements without transfers and twenty stable agreements with transfers (all possible two and three-country agreements are stable with transfers).

The shaded rows in Table 1 highlight the most effective stable agreements both with and without transfers between IEA members. The stable agreement containing the three countries with the lowest payoff-maximizing SGE level (1, 2 and 3) achieves the highest level of abatement closing about 38% of the abatement gap and 45% of the payoff gap between the social optimum and the Nash equilibrium. Note that without transfers the most effective stable agreement consists of two countries (1 and 2) and they are only able to close about 14% of the abatement gap and 19% of the payoff gap.

The example demonstrates that transfers can expand the set of stable agreements, reduce the amount of SGE deployed by the free driver, close more of the abatement gap and lead to higher efficiency gains. However, there are no stable agreements larger than three countries even with transfers for this example.

### *4.1 The free driver as a cooperative IEA member*

Note from Table 1 that there are 10 stable agreements that include the free driver (country 5) as a member. The free drivers, even when members of an IEA, continue to unilaterally determine aggregate SGE deployment. Without transfers, there is only one and that single agreement consists of the free driver and the next-in-line to be the free driver (country 4). They both take on the least amount of abatement of all possible two-member agreements. The new transfer system adds nine additional stable agreements that include the free driver and the most effective one consists of $S = 1, 2, 5$. The intuition why this particular agreement has relatively high collective abatement is straightforward. Countries 1 and 2 suffer the most from the free driver's deployment of SGE and thus have the strongest anti-driver incentives. A cooperative agreement of three members provides enough surplus to these two nondriving countries that they are able to pay the free driver to join and still be better off. Of course, countries 1 and 2 would prefer a three-member agreement that includes another non free driver (e.g. $S = 1, 2, 3$ or $S = 1, 2, 4$) because those agreements would take on relatively more abatement, result in lower SGE and increase their payoffs.

**Table 1.** Stable IEAs under the threat of a solar geoengineering free driver

| Agreement | Member abatement | Avg nonmember abatement | Total abatement ($Q$) | Total payoffs ($\Pi$) | SGE deployed ($G$) | Abatement gap closed (%) | Payoff gap closed (%) |
|---|---|---|---|---|---|---|---|
| *Stable IEAs without transfers* | | | | | | | |
| 1,2 | 0.44 | 0.15 | 1.34 | 1.48 | 2.93 | 13.5 | 18.9 |
| **1,3** | 0.42 | 0.16 | 1.32 | 1.47 | 2.94 | 12.9 | 18.2 |
| **2,3** | 0.38 | 0.17 | 1.28 | 1.46 | 2.98 | 11.7 | 17.4 |
| **2,4** | 0.36 | 0.18 | 1.26 | 1.45 | 2.99 | 11.1 | 16.7 |
| **3,4** | 0.34 | 0.19 | 1.24 | 1.44 | 3.01 | 10.5 | 16.0 |
| **4,5** | 0.28 | 0.21 | 1.18 | 1.42 | 3.06 | 8.6 | 14.5 |
| *Additional stable IEAs with transfers* | | | | | | | |
| **1,4** | 0.40 | 0.17 | 1.30 | 1.47 | 2.96 | 12.3 | 18.2 |
| **1,5** | 0.36 | 0.18 | 1.26 | 1.46 | 2.99 | 11.1 | 17.4 |
| **2,5** | 0.32 | 0.19 | 1.22 | 1.44 | 3.02 | 9.8 | 16.0 |
| **3,5** | 0.30 | 0.20 | 1.20 | 1.43 | 3.04 | 9.2 | 15.2 |
| **1,2,3** | 0.62 | 0.14 | 2.14 | 1.83 | 2.29 | 38.2 | 44.9 |
| **1,2,4** | 0.60 | 0.15 | 2.10 | 1.83 | 2.32 | 36.9 | 44.9 |
| **1,2,5** | 0.56 | 0.17 | 2.02 | 1.82 | 2.38 | 34.5 | 44.2 |
| **1,3,4** | 0.58 | 0.16 | 2.06 | 1.82 | 2.35 | 35.7 | 44.2 |
| **1,3,5** | 0.54 | 0.18 | 1.98 | 1.81 | 2.42 | 33.2 | 43.4 |
| **1,4,5** | 0.52 | 0.19 | 1.94 | 1.80 | 2.45 | 32.0 | 42.7 |
| **2,3,4** | 0.54 | 0.18 | 1.98 | 1.81 | 2.42 | 33.2 | 43.4 |
| **2,3,5** | 0.50 | 0.21 | 1.90 | 1.79 | 2.48 | 30.8 | 41.9 |
| **2,4,5** | 0.48 | 0.21 | 1.86 | 1.78 | 2.51 | 29.5 | 41.2 |
| **3,4,5** | 0.46 | 0.22 | 1.82 | 1.77 | 2.54 | 28.3 | 40.5 |

This leads to the interesting result that a cooperative agreement can be reached that includes the free driver which leads to a reduction of SGE deployment, but only indirectly, through the emissions abatement channel.

### *4.2 Sensitivity to parameters*

We conclude this section by exploring how coalition stability under a transfer system is impacted by changes in the anti-driver incentives, which requires an analytical solution to the coalition surplus $\sigma(S)$ in equation (7). To simplify the expression slightly we follow McEvoy *et al.* (2024) and normalize $b = c = 1$. After a significant amount of algebra, the coalition stability condition simplifies to the following (derivation in Appendix):

$$\sigma(S) = \frac{(s-1)(s-3)(1+n\phi\gamma^{\max})}{2n}\left[\frac{-s(1+n\phi\gamma^{\max})}{n} + 2(\theta+\phi)\sum_{j\in S}\gamma_j\right] \tag{8}$$

$$+\frac{(\theta+\phi)^2}{2}\left[(2s-3)\sum_{j\in S}(\gamma_j)^2 - (s-2)\left(\sum_{j\in S}\gamma_j\right)^2\right].$$

The comparative statics show that global monotonicity cannot be established analytically. However, all else equal, increases in $\phi$ strengthen anti-driver incentives and expand the surplus available for transfers, which tends to enlarge the parameter region under which equation (8) holds. The intuition is that as anti-driver incentives increase through a heightened marginal negative externality from SGE, the benefits of cooperation for non free drivers increases. In turn, that surplus can be used to entice additional members to join under the transfer system. Similarly, greater dispersion in SGE preferences (i.e. $\gamma^{\max} - \overline{\gamma}$) strengthens anti-driver incentives and enlarges the potential surplus. That said, we find the stability results for our example – that only coalitions of size two and three are stable – are robust to substantial changes to $\phi$ a the distribution of $\gamma$s.

Note that under a limited set of parameters, it is possible that the transfer scheme leads to a stable grand coalition. For example, a relatively large $\phi$, small $n$ and wide distribution of $\gamma$'s can lead to a stable grand coalition with transfers [3].

## 5. Conclusion

A major concern with SGE is the risk of excessive deployment, leading to overcooling and economic losses (Abatayo *et al.*, 2020). One of the key governance challenges is that these technologies can be implemented unilaterally (the free-driver problem), making direct regulation or restriction difficult. To address this, we examine how strengthening international agreements on emissions abatement through transfers can serve as an indirect strategy to reduce reliance on SGE.

Our theoretical analysis and examples show that a transfer scheme among IEA members can expand the set of stable agreements, increase overall abatement levels, reduce SGE deployment and improve collective payoffs. The most effective agreements involve countries that are most negatively impacted by SGE – those with the strongest anti-driver incentives. We demonstrate that with appropriate transfers, even a free driver can be incentivized to join an IEA, indirectly curbing their reliance on SGE through enhanced cooperation on mitigation. In limited cases, this approach can lead to a stable grand coalition.

The results highlight an important insight for SGE governance. Rather than attempting to regulate deployment directly – a task complicated by the free-driver incentives – policy efforts may be more effective when focused on strengthening mitigation incentives and expanding transfer mechanisms within existing climate agreements. By increasing the gains from cooperation on abatement, the international community can reduce the incentives for unilateral SGE deployment without requiring a separate system of governance.

## Notes

[1.] See equation (31) in McEvoy *et al.* (2024) for an expanded version of equation (6).

[2.] In a world without SGE, or in a world with SGE and homogenous countries, the largest stable agreement is $s = 3$ as long as $n \geq 3$, regardless of the parameter values for $b$ or $c$ (Barrett 1994).

[3.] For example, the following parameter choices lead to a stable grand coalition: $n = 4$, $\theta = 0.249$, $\phi = 0.254$ and $\gamma = [0.01, 0.01, 0.99, 0.99]$.

## References


Abatayo, A.L., Bosetti, V., Casari, M., Ghidoni, R. and Tavoni, M. (2020), “Solar geoengineering may lead to excessive cooling and high strategic uncertainty”, *Proceedings of the National Academy of Sciences*, Vol. 117 No. 24, pp. 13393-13398.

Bakalova, I. and Belaia, M. (2023), “Stability of efficient international agreements on solar geoengineering”, *Environmental and Resource Economics*, Vol. 86 No. 3, pp. 673-712.

Barrett, S. (1994), “Self-enforcing international environmental agreements”, *Oxford Economic Papers*, Vol. 46 No. Supplement_1, pp. 878-894.

Barrett, S. (2001), “International cooperation for sale”, *European Economic Review*, Vol. 45 No. 10, pp. 1835-1850.

Barrett, S. (2003), *Environment and Statecraft: The Strategy of Environmental Treaty-Making: The Strategy of Environmental Treaty-Making*, Oxford University Press, Oxford, England.

Chander, P. and Tulkens, H. (1995), “A core-theoretic solution for the design of cooperative agreements on transfrontier pollution”, *International Tax and Public Finance*, Vol. 2 No. 2, pp. 279-293.

Felgenhauer, T., Bala, G., Borsuk, M., Brune, M., Camilloni, I., Wiener, J.B., and Xu, J. (2022), “Solar radiation modification: a risk-risk analysis”, *Carnegie Climate Governance Initiative (C2G)*, March, New York, NY.

Ghidoni, R., Abatayo, A.L., Bosetti, V., Casari, M. and Tavoni, M. (2023), “Governing climate geoengineering: side payments are not enough”, *Journal of the Association of Environmental and Resource Economists*, Vol. 10 No. 5, pp. 1149-1177.

Kapstein, E.B. (2025), “Side payments in world politics: theory and practice”, *Global Policy*, Vol. 16 No. 2, pp. 258-265.

McEvoy, D.M., McGinty, M., Cherry, T.L. and Kroll, S. (2024), “International climate agreements under the threat of solar geoengineering”, *Journal of the Association of Environmental and Resource Economists*, Vol. 11 No. 4, pp. 853-885.

McGinty, M. (2007), “International environmental agreements among asymmetric nations”, *Oxford Economic Papers*, Vol. 59 No. 1, pp. 45-62.

McGinty, M. (2011), “A risk-dominant allocation: maximizing coalition stability”, *Journal of Public Economic Theory*, Vol. 13 No. 2, pp. 311-325.

Moreno-Cruz, J.B. (2015), “Mitigation and the geoengineering threat”, *Resource and Energy Economics*, Vol. 41, pp. 248-263.

Moreno-Cruz, J.B., McEvoy, D.M., McGinty, M. and Cherry, T.L. (2025), "The economics and governance of solar geoengineering", *Review of Environmental Economics and Policy*, Vol. 19 No. 1.

Robock, A. (2008), "20 Reasons why geoengineering may be a bad idea", *Bulletin of the Atomic Scientists*, Vol. 64 No. 2, pp. 14-18.

Sælen, H. (2016), "Side-payments: an effective instrument for building climate clubs?", *International Environmental Agreements: Politics, Law and Economics*, Vol. 16 No. 6, pp. 909-932.

Wagner, G. and Weitzman, M. (2012), "Playing god", Foreign Policy (October 24).

Weitzman, M. (2015), "A voting architecture for the governance of free-driver externalities", *Scandinavian Journal of Economics*, Vol. 117, pp. 1049-1068.

**Corresponding author**
David M. McEvoy can be contacted at: mcevoydm@appstate.edu

**Appendix. Stage 2 abatement equations**

The abatement equations (p. 4) for nonmembers and members are, respectively:

$$q_j^t(G^{\max}) = \frac{b}{cn} - \frac{\theta\gamma_j}{c} + \frac{\phi\left(\gamma^{\max} - \gamma_j\right)}{c},$$

and:

$$q_i^s(G^{\max}) = \frac{bs}{nc} - \frac{1}{c}\left[(\theta+\phi)\sum_{i=1}^{s}\gamma_i - s\phi\gamma^{\max}\right].$$

Derivation of the internal stability condition with transfers (equation (8))

Total abatement is $Q(S) = sq^s + \sum_{j\in N\ S} q_j^t$:

$$Q(S) = sq^s + \sum_{j\in N\ S} q_j^t$$

$$Q(S) = s\left[\frac{s + sn\phi\gamma^{\max}}{n} - (\theta+\phi)\sum_{i\in S}\gamma_i\right] + \sum_{j\in N\ S}\left[\frac{1 + n\phi\gamma^{\max}}{n} - \gamma_j(\theta+\phi)\right]$$

$$Q(S) = \frac{s^2(1 + n\phi\gamma^{\max})}{n} - s(\theta+\phi)\sum_{i\in S}\gamma_i + (n-s)\left[\frac{1 + n\phi\gamma^{\max}}{n}\right] - (\theta+\phi)\sum_{j\in N\ S}\gamma_j$$

$$Q(S) = \frac{(s^2 + n - s)(1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\left[s\sum_{i\in S}\gamma_i + \sum_{j\in N\ S}\gamma_j\right]$$

If member $i$ leaves coalition $S$ abatement becomes $Q(S\ i)$:

$$Q(S\ i) = (s-1)q^s(S\ i) + q_i^t + \sum_{j\in N\ S} q_j^t$$

$$Q(S\ i) = (s-1)\left[\frac{(s-1)(1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\left(\sum_{j\in S}\gamma_j - \gamma_i\right)\right]$$
$$+ \frac{1 + n\phi\gamma^{\max}}{n} - \gamma_i(\theta+\phi) + \sum_{N\ S} q_j^t$$

$$Q(S\ i) = (s-1)\left[\frac{(s-1)(1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\left(\sum_{j\in S}\gamma_j - \gamma_i\right)\right]$$
$$+ \frac{1 + n\phi\gamma^{\max}}{n} - \gamma_i(\theta+\phi) + (n-s)\left[\frac{1 + n\phi\gamma^{\max}}{n}\right] - (\theta+\phi)\sum_{j\in N\ S}\gamma_j$$

$$Q(S\ i) = \frac{\left[(s-1)^2 + n - s + 1\right](1 + n\phi\gamma^{\max})}{n} - (s-1)(\theta+\phi)\left(\sum_{j\in S}\gamma_j - \gamma_i\right)$$
$$- \gamma_i(\theta+\phi) - (\theta+\phi)\sum_{j\in N\ S}\gamma_j$$

$$Q(S\ i) = \frac{\left[s^2 - 2s + 1 + n - s + 1\right](1 + n\phi\gamma^{\max})}{n} - (s-1)(\theta+\phi)\sum_{j\in S}\gamma_j$$

$$+ \gamma_i(s-1)(\theta+\phi) - \gamma_i(\theta+\phi) - (\theta+\phi)\sum_{j\in N\ S}\gamma_j$$

$$Q(S\ i) = \frac{\left[s^2 - 3s + n + 2\right](1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\left[(s-1)\sum_{j\in S}\gamma_j + \sum_{j\in N\ S}\gamma_j\right]$$

$$+ \gamma_i(s-2)(\theta+\phi)$$

$$Q(S\ i) = \frac{\left[s^2 - 3s + n + 2\right](1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\left[-\gamma_i(s-2) + (s-1)\sum_{j\in S}\gamma_j + \sum_{j\in N\ S}\gamma_j\right]$$

The reduction in abatement when $i$ leaves $S$ is $Q(S) - Q(S\ i)$:

$$Q(S) - Q(S\ i) = \frac{(s^2 + n - s)(1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\left[s\sum_{i\in S}\gamma_i + \sum_{j\in N\ S}\gamma_j\right]$$

$$- \frac{\left[s^2 - 3s + n + 2\right](1 + n\phi\gamma^{\max})}{n}$$

$$+ (\theta+\phi)\left[-\gamma_i(s-2) + (s-1)\sum_{j\in S}\gamma_j + \sum_{j\in N\ S}\gamma_j\right]$$

$$Q(S) - Q(S\ i) = \frac{2(s-1)(1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\left[\gamma_i(s-2) + \sum_{j\in S}\gamma_j\right]$$

The signatory abatement cost is $\frac{(q_i^s)^2}{2}$:

$$\frac{(q_i^s)^2}{2} = \frac{1}{2}\left[\frac{s(1 + n\phi\gamma^{\max})}{n} - (\theta+\phi)\sum_{j\in S}\gamma_j\right]^2$$

$$\frac{(q_i^s)^2}{2} = \frac{s^2(1 + n\phi\gamma^{\max})^2}{2n^2} - \frac{(\theta+\phi)s(1 + n\phi\gamma^{\max})\sum_{j\in S}\gamma_j}{n} + \frac{(\theta+\phi)^2}{2}\left(\sum_{j\in S}\gamma_j\right)^2$$

The dominant strategy abatement cost is $\frac{(q_i^t)^2}{2}$:

$$\frac{(q_i^t)^2}{2} = \frac{1}{2}\left[\frac{1 + n\phi\gamma^{\max}}{n} - \gamma_i(\theta+\phi)\right]^2$$

$$\frac{(q_i^t)^2}{2} = \frac{(1 + n\phi\gamma^{\max})^2}{2n^2} - \frac{\gamma_i(\theta+\phi)(1 + n\phi\gamma^{\max})}{n} + \frac{(\gamma_i)^2(\theta+\phi)^2}{2}$$

The gain from leaving is the reduction in abatement cost $\frac{\left[(q_i^s)^2 - (q_i^t)^2\right]}{2}$:

$$\frac{\left(q_i^s\right)^2}{2}-\frac{\left(q_i^t\right)^2}{2}=\frac{1}{2}\left[\begin{array}{c}\frac{(s^2-1)(1+n\phi\gamma^{\max})^2}{n^2}-\frac{2(\theta+\phi)(1+n\phi\gamma^{\max})\left(s\sum_{j\in S}\gamma_j-\gamma_i\right)}{n}\\ +(\theta+\phi)^2\left[\left(\sum_{j\in S}\gamma_j\right)^2-(\gamma_i)^2\right]\end{array}\right]$$

Putting these results into (6), then simplifying results in the individual internal stability condition, $IS_i(S)$:

$$IS_i(S)=\left(\frac{1+n\phi\gamma^{\max}-n(\theta+\phi)\gamma_i}{n}\right)\left[Q(S)-Q(S\;i)\right]-\frac{\left[(q_i^s)^2-\left(q_i^t\right)^2\right]}{2}$$

$$IS_i(S)=\left(\frac{1+n\phi\gamma^{\max}-n(\theta+\phi)\gamma_i}{n}\right)\left[\frac{(2s-2)(1+n\phi\gamma^{\max})}{n}-(\theta+\phi)\left[\gamma_i(s-2)+\sum_{j\in S}\gamma_j\right]\right]$$
$$-\frac{\left[(q_i^s)^2-\left(q_i^t\right)^2\right]}{2}$$

$$IS_i(S)=\left(\frac{(2s-2)(1+n\phi\gamma^{\max})^2-n(\theta+\phi)(2s-2)(1+n\phi\gamma^{\max})\gamma_i}{n^2}\right)$$
$$-\left(\frac{1+n\phi\gamma^{\max}-n(\theta+\phi)\gamma_i}{n}\right)\left[(\theta+\phi)\left[\gamma_i(s-2)+\sum_{j\in S}\gamma_j\right]\right]-\frac{\left[(q_i^s)^2-\left(q_i^t\right)^2\right]}{2}$$

Simplifying the first two terms results in the following:

$$IS_i(S)=\frac{(2s-2)(1+n\phi\gamma^{max})^2}{n^2}-\frac{(2s-2)(\theta+\phi)(1+n\phi\gamma^{max})\gamma_i}{n}$$
$$-\frac{(\theta+\phi)(1+n\phi\gamma^{max})}{n}\left[\gamma_i(s-2)+\sum_{j\in S}\gamma_j\right]+(\theta+\phi)^2\gamma_i\left[\gamma_i(s-2)+\sum_{j\in S}\gamma_j\right]$$
$$-\frac{[(q_i^s)^2-(q_i^t)^2]}{2}$$

$$IS_i(S)=\frac{(2s-2)(1+n\phi\gamma^{max})^2}{n^2}-\frac{(2s-2)(\theta+\phi)(1+n\phi\gamma^{max})\gamma_i}{n}$$
$$-\frac{(s-2)(\theta+\phi)(1+n\phi\gamma^{max})\gamma_i}{n}-\frac{(\theta+\phi)(1+n\phi\gamma^{max})}{n}\sum_{j\in S}\gamma_j$$
$$+(s-2)(\theta+\phi)^2(\gamma_i)^2+(\theta+\phi)^2\gamma_i\sum_{j\in S}\gamma_j-\frac{[(q_i^s)^2-(q_i^t)^2]}{2}$$

$$IS_i(S)=\frac{(2s-2)(1+n\phi\gamma^{max})^2}{n^2}-\frac{(3s-4)(\theta+\phi)(1+n\phi\gamma^{max})\gamma_i}{n}$$
$$-\frac{(\theta+\phi)(1+n\phi\gamma^{max})}{n}\sum_{j\in S}\gamma_j+(s-2)(\theta+\phi)^2(\gamma_i)^2+(\theta+\phi)^2\gamma_i\sum_{j\in S}\gamma_j$$
$$-\frac{[(q_i^s)^2-(q_i^t)^2]}{2}$$

Then add the cost term from above $-\frac{\left[(q_i^s)^2-(q_i^t)^2\right]}{2}$:

$$
\begin{aligned}
IS_i(S) = {} & \frac{(2s-2)(1+n\phi\gamma^{\max})^2}{n^2} - \frac{(3s-4)(\theta+\phi)(1+n\phi\gamma^{\max})\gamma_i}{n} \\
& - \frac{(1+n\phi\gamma^{\max})(\theta+\phi)\sum_{j\in S}\gamma_j}{n} - (s-2)(\theta+\phi)^2(\gamma_i)^2 - (\theta+\phi)^2\gamma_i\sum_{j\in S}\gamma_j \\
& - \frac{1}{2}\left[\begin{array}{c} \frac{(s^2-1)(1+n\phi\gamma^{\max})^2}{n^2} - \frac{2(\theta+\phi)(1+n\phi\gamma^{\max})\left(s\sum_{j\in S}\gamma_j - \gamma_i\right)}{n} \\ + (\theta+\phi)^2\left[\left(\sum_{j\in S}\gamma_j\right)^2 - (\gamma_i)^2\right] \end{array}\right]
\end{aligned}
$$

Combining terms results in the following:

$$
\begin{aligned}
IS_i(S) = {} & \frac{\left[2(2s-2)-(s^2-1)\right](1+n\phi\gamma^{\max})^2}{2n^2} - \frac{(3s-4)(\theta+\phi)(1+n\phi\gamma^{\max})\gamma_i}{n} \\
& - \frac{(\theta+\phi)(1+n\phi\gamma^{\max})\sum_{j\in S}\gamma_j}{n} + \frac{(\theta+\phi)(1+n\phi\gamma^{\max})\left(s\sum_{j\in S}\gamma_j-\gamma_i\right)}{n} \\
& + (s-2)(\theta+\phi)^2(\gamma_i)^2 + (\theta+\phi)^2\gamma_i\sum_{j\in S}\gamma_j - \frac{(\theta+\phi)^2}{2}\left[\left(\sum_{j\in S}\gamma_j\right)^2 - (\gamma_i)^2\right]
\end{aligned}
$$

$$
\begin{aligned}
IS_i(S) = {} & \frac{(-s^2+4s-3)(1+n\phi\gamma^{\max})^2}{2n^2} - \frac{(3s-3)(\theta+\phi)(1+n\phi\gamma^{\max})\gamma_i}{n} \\
& + \frac{(s-1)(\theta+\phi)(1+n\phi\gamma^{\max})}{n}\sum_{j\in S}\gamma_j + (\theta+\phi)^2\left[(s-2)(\gamma_i)^2 + \gamma_i\sum_{j\in S}\gamma_j\right] \\
& - \frac{(\theta+\phi)^2}{2}\left[\left(\sum_{j\in S}\gamma_j\right)^2 - (\gamma_i)^2\right]
\end{aligned}
$$

$$
\begin{aligned}
IS_i(S) = {} & \frac{-(s-1)(s-3)(1+n\phi\gamma^{\max})^2}{2n^2} - \frac{3(s-1)(\theta+\phi)(1+n\phi\gamma^{\max})\gamma_i}{n} \\
& + \frac{(s-1)(\theta+\phi)(1+n\phi\gamma^{\max})}{n}\sum_{j\in S}\gamma_j \\
& + \frac{(\theta+\phi)^2}{2}\left[2(s-2)(\gamma_i)^2 + 2\gamma_i\sum_{j\in S}\gamma_j - \left(\sum_{j\in S}\gamma_j\right)^2 + (\gamma_i)^2\right]
\end{aligned}
$$

$$
\begin{aligned}
IS_i(S) = {} & \frac{-(s-1)(s-3)(1+n\phi\gamma^{\max})^2}{2n^2} + \frac{(s-1)(\theta+\phi)(1+n\phi\gamma^{\max})}{n}\left(-3\gamma_i + \sum_{j\in S}\gamma_j\right) \\
& + \frac{(\theta+\phi)^2}{2}\left[(2s-3)(\gamma_i)^2 + 2\gamma_i\sum_{j\in S}\gamma_j - \left(\sum_{j\in S}\gamma_j\right)^2\right]
\end{aligned}
$$

The final step involves summing this up across all $i \in S$ to get a coalition stability condition, $\sigma(S) \equiv \sum_{i \in S} IS_i(S) = \sum_{i \in S}[\pi^s(S) - \pi^t(S\ i)] \geq 0$:

$$\sigma(S) \equiv \sum_{i \in S} IS_i(S) = \frac{-s(s-1)(s-3)(1+n\phi\gamma^{\max})^2}{2n^2} + \frac{(s-1)(\theta+\phi)(1+n\phi\gamma^{\max})}{n}\left(-3\sum_{j \in S}\gamma_j + s\sum_{j \in S}\gamma_j\right) + \frac{(\theta+\phi)^2}{2}\left[(2s-3)\sum_{j \in S}(\gamma_j)^2 + 2\sum_{j \in S}\gamma_j\sum_{j \in S}\gamma_j - s\left(\sum_{j \in S}\gamma_j\right)^2\right]$$

Which simplifies to the following (equation (8)):

$$\sigma(S) = \frac{(s-1)(s-3)(1+n\phi\gamma^{\max})}{2n}\left[\frac{-s(1+n\phi\gamma^{\max})}{n} + 2(\theta+\phi)\sum_{j \in S}\gamma_j\right] + \frac{(\theta+\phi)^2}{2}\left[(2s-3)\sum_{j \in S}(\gamma_j)^2 - (s-2)\left(\sum_{j \in S}\gamma_j\right)^2\right].$$