Philosophy & Technology
https://doi.org/10.1007/s13347-021-00452-9

RESEARCH ARTICLE

A Modest Defense of Geoengineering Research: a Case Study in the Cost of Learning

Eric Winsberg[1]

Received: 24 November 2020 / Accepted: 5 April 2021

Abstract
Recently, research into the possibilities of developing solar radiation management (SRM) and other geoengineering technologies has gained new momentum. Just last year, Cambridge University announced the opening of a "Centre for Climate Repair" as part of the university's Carbon Neutral Futures Initiative. Recent modeling work gives hope that SRM could confer more benefits than previously thought. But opposition to even conducting research into SRM remains strong. I use the case study of SRM to develop a framework, based on a theorem by I.J. Good, for thinking about the benefits and costs of acquiring new evidence and for thinking about the conditions under which new evidence could be harmful. I argue that the expected benefits of supporting public research in SRM technologies outweigh the expected costs and harms.

Keywords Climate · Geoengineering · Climate change · Solar radiation management · Stratospheric aerosol injection

1 Introduction

In 2011, an experiment at the University of Bristol to test technologies involving solar radiation management (SRM), a technology some hope could someday be used to partially offset the impact of greenhouse gases on the climate, was put on hold. This was due, at least in part, to the objections of environmental groups. It has never been resuscitated.[1] In a letter sent to Britain's climate minister, many environmental

[1] There is controversy surrounding why the experiment, which was part of the Stratospheric Particle Injection for Climate Engineering (SPICE) project, was terminated. It seems to have mostly been overdetermined. In addition to pressure from environmentalists, the project faced questions about intellectual property and about the safety of a part of its apparatus. But what is clear is that the project came under heavy fire from many environmental groups (Cressey, 2012; Kuo, 2012).

Eric Winsberg
winsberg@usf.eduwinsberg

[1] The University of South Florida, Tampa, FL, USA

Published online: 15 April 2021

groups argued that the experiment violated a decision not to undertake large-scale geoengineering tests made during the previous meeting of the Convention on Biological Diversity (CBD). The authors were appealing to a decision adopted in 2010 by the CBD that places a taboo on geoengineering (GE) implementations, and what they claimed was an extremely high bar on any research activities.[2,3] Perhaps more tellingly, they argued that the experiment would undermine Britain's capacity to negotiate effectively for global climate agreements.

More recently, research into the possibilities of developing SRM and other geoengineering technologies has gained new momentum. Just this year, Cambridge University announced the opening of a "Centre for Climate Repair" as part of the university's Carbon Neutral Futures Initiative. The center does research into a variety of so called "geoengineering" technologies including carbon capture and sequestration (CCS) and SRM. The response has not been overwhelmingly positive. Patrick Galey, the global science and environment correspondent for Agence France-Presse, tweeted that the center is "a tremendous waste of time and money [that] toes the fossil fuel lobby's line." He described SRM as "a bat@#$t crazy idea [that is] a bit like setting fire to your house then trying to put it out by turning on the air conditioning."[4] Some environmental groups believe this kind of research is being advanced on behalf of fossil fuel companies in an effort "to shift policy norms so that previously unthinkable notions and activities – like solar radiation management – start to become more mainstream and acceptable" (Geoengineering Monitor n.d.).[5] Explaining why we should oppose research into geoengineering generally, Geoengineering Monitor says, "This much we know: geoengineering techniques do nothing to address the root causes of climate change, and evidence points to a high likelihood that rather than improving climate, they will make things worse" (Monitor n.d.).

In general, many environmentalists and commenters on energy and climate policy argue that some of these technologies are dangerous because of the foreseeable and unforeseeable consequences their future implementation could produce. This might be right. What concerns me in this paper is the claim that even conducting research into such technologies should be discouraged (or opposed in more strenuous ways), because of those potential dangers as well as because of the moral hazard created by the mere act of investigating their potential, and

[2] The decision calls on members to "Ensure, in line and consistent with decision IX/16 C, on ocean fertilization and biodiversity and climate change, in the absence of science based, global, transparent and effective control and regulatory mechanisms for geoengineering, and in accordance with the precautionary approach and Article 14 of the Convention, that no climate-related geoengineering activities that may affect biodiversity take place, until there is an adequate scientific basis on which to justify such activities and appropriate consideration of the associated risks for the environment and biodiversity and associated social, economic and cultural impacts, with the exception of small scale scientific research studies that would be conducted in a controlled setting in accordance with Article 3 of the Convention, and only if they are justified by the need to gather specific scientific data and are subject to a thorough prior assessment of the potential impacts on the environment" https://www.cbd.int/decision/cop/?id=12299.

[3] Most legal scholars seem to agree is not a legally binding moratorium (Reynolds, 2019b).

[4] https://twitter.com/patrickgaley/status/1126809903492947968

[5] The claim that geoengineering research is promoted by the fossil fuel industry is highly contested. See (Reynolds et al., 2016; Reynolds, 2019a), for example. More on this later.

other similar considerations. This stronger conclusion (that even research ought to be opposed) has in fact been pursued by many environmental ethicists and political scientists.[6]

It is easy to understand why reasonable people would oppose the *implementation* of any sort of geoengineering strategy. The balance of risks presently looks very hard to assess, and the science is too immature to support any actions that could have wide direct consequences. We simply do not know enough at this time about how bad climate change will get in the absence of geoengineering interventions, and we do not know enough about the likely effectiveness of any geoengineering strategies. Most people would probably agree that it is simply too early to have this debate. But what could be the philosophical justification for opposing scientific *research* into *possible* strategies for using geoengineering to offset the harms caused by the emission of heat trapping gases? Most discussions of this issue assume one of two extreme positions: opponents argue that, because the prospect of geoengineering strategies being safe and effective look so poor, all research should be opposed, while proponents argue that, since it is only *research*, and there is no harm in *asking*, we should let the scientific chips fall where they may, and decide how to act accordingly if and when the time comes that we have better information. Many refer to this as "arming the future."

Groups most politically mobilized in favor of environmental policy tend to be averse to focusing on environmental adjustments that damage to the environment will require. This explains the attitude that all such research should be avoided, because these groups see that strategy as a distraction from the more pressing demand of avoiding the original cause of the damage. This is usually a sensible attitude, because past environmental challenges have been met without much need for adjustment and adaptation. Damage avoidance by eliminating the underlying cause, as in the cases of DDT and CFCs, has often been a success. But, as each year goes by, climate change is looking less and less like the kind of challenge where the damaging agent will be easily eliminated.[7]

The attitude can also partly be explained by the general distrust that many environmental activists direct at what they regard as "technofixes" to environmental problems. Such people are often mistrustful of cost–benefit analyses in environmental policy and prefer, instead, something akin to a precautionary principle. Such an attitude is reasonable in a number of contexts, but there are two reasons why it is inapt in the present case. The first reason has to do with a major theme of this paper. As the reader will see as we go, one of the key findings of this paper is that the central risks of geoengineering will, on balance, be mitigated, rather than aggravated, by supporting public research. So those who fear technofixes ought not necessarily oppose research into geoengineering. Some people argue for geoengineering research so that we can "arm the future." Opponents of "technofixes" arguably ought

[6] See, for example, Jamieson (1996), Jamieson (2013), Blomfield (2015), Gardiner (2011), McKinnon (2019), and Lin (2013).

[7] See Victor (2011) for some good discussion of why climate action has been harder to achieve than action on CFCs, for example.

to oppose that sort of argument. But that is emphatically *not* the approach of this paper.

The second reason has to do with the general applicability of precautionary principle reasoning in climate change mitigation and adaptation. Let us set aside worries about whether precautionary principle reasoning is even coherent. The problem, in the case of climate change in 2020, is that there is harm in every direction. Kevin Elliot argued this in his (2011), and the strength of this claim has only increased in the intervening decade. The time to avoid all harm probably passed us by in the 1980s. It is unavoidable, at this point, that we weigh the harms of all of our possible actions.

Accordingly, my goal in this paper is to take a more careful look at the ethics of publicly supporting the *research itself*. I take a broadly consequentialist approach. My goal is to qualitatively estimate the expected utility of publicly supporting the research itself. I make no claim to being able to calculate the expected utility—only to qualitatively assess whether it is more likely positive or negative. I make no appeals to fundamental freedoms. I do not assume that we can unilaterally decide whether or not research, especially private research, will take place. We can only add our own voices for or against it. I reject the inference that supposes that if the expected utility of pursuing geoengineering implementations presently looks negative, then it follows that *the research itself* has negative expected utility. I reject even the inference that supposes that from the very strong assumption that the implementation is unlikely to look positive given *any* foreseeable new evidence and that it follows that the expected utility is negative. I ask, instead: are there in fact good reasons for opposing *mere research* into geoengineering strategies *even if* one has a high degree of skepticism in the likelihood of implementations being successful?

The IPCC Fifth Assessment Report defines geoengineering (GE) as "technological efforts to stabilize the climate system by direct intervention in the energy balance of the Earth for reducing *global warming*" (Stocker, 2014, spm21). In a Royal Society report (Ming, 2014), geoengineering is defined as the "*deliberate large-scale manipulation of the planetary environment to counteract anthropogenic climate change*." Many view GE as a middle road between *mitigation* (which primarily comprises methods of reducing emissions) and *adaptation* (the process of adapting our lives, habitats and technologies to a warmer world.) At present geoengineering strategies basically fall into 3 categories: carbon dioxide removal (CDR), earth radiation management (ERM), and solar radiation management (SRM). It has been speculated that SRM itself can be achieved via high albedo[8] crops, clearing the forests that block the albedo of fields of snow, marine cloud brightening, giant space mirrors, and stratospheric aerosol injection (SAI).[9]

SAI is imagined to be achieved by injecting very small sulfate particles into the stratosphere in a way designed to mimic the global cooling caused by volcanic eruptions like that of Mount Pinatubo in 1991. (Other similar proposals involve self-levitating particles made out of titanium or aluminum.)

[8] Albedo is the tendency to reflect solar radiation back into space.

[9] See Heyward (2013) for an excellent discussion differentiating different geoengineering strategies.

Since no single answer exists to the question "Should conducting research into geoengineering technologies be discouraged?", I will focus here on SAI. This choice is somewhat arbitrary, but it can be motivated by a number of considerations:

1. It is generally regarded to have one of the highest, if not the single highest, cooling capacity of all the strategies above.
2. It is expected to be relatively cheap to implement.
3. It appears to be reasonably achievable with existing technologies.
4. There is likely to remain substantial uncertainty about what its total effect will be (both on regional climate and hydrology, and in terms of non-climactic side effects).[10]
5. It has both a large number of undesirable foreseeable consequences and the potential for a large number of undesirable unforeseeable consequences.
6. Its impact on the planet is likely to be quite varied, with a considerable likelihood of causing net benefit to some regions alongside net harm to others.[11]

These considerations make SAI a good test case for the central philosophical question of this paper because they *both* make SAI the most tempting technology for advocates of geoengineering solutions to investigate *and* one of the most alarming to those who fear unGoodian costs of research.[12] Thus, if it turns out that SAI research should not be opposed based on the considerations canvased in this paper, then it is likely that most of the other strategies would come out the same way. In the end, of course, research into each strategy will need to be evaluated individually regarding whether it ought to be discouraged. Hopefully, however, the analysis here of research into SAI will provide a useful template.

Here is how the paper will proceed: in Section 2, I unpack how the appraisal of the expected utility of research ought to proceed in terms of Good's theorem. Good's theorem says that free learning is never bad. But of course the theorem assumes some preconditions. And looking at the circumstances in which those preconditions are violated reveals a great deal about what considerations need to be in place in order for the expected utility of research to be negative. In Section 3, I unpack how these considerations might apply to geoengineering. Section 4 surveys the standard objections to GE and SAI in the literature. Sections 5–8 evaluate four different possible lines of argument that SAI research could have negative expected utility: cost of learning; biased reporting of evidence; extremely low likelihood of success; and positive and negative externalities of SAI research. Sections 5–8 should also provide a template for evaluating the value of research into other possible GE strategies. In Section 9, I summarize the findings of Sections 6–9 and offer my own appraisal, which can be summarized in the following thesis: research into SAI

[10] See Section 4 for more details on items 4–6.

[11] See Lenton and Vaughan (2009) for support for all these claims.

[12] One possible exception here might be that research into CDR, especially in the form of carbon sequestration and capture [CSC] at the source of fossil fuel, might be argued to have the greatest degree of moral hazard (which I discuss below). I do not explore this possibility any further.

actually mitigates more of the dangers of SAI than it aggravates. On the basis of that thesis, I argue that a modest defense of *public* support for SAI research is warranted.

2 Opposing Research on Grounds of Expected Utility

Our goal is to evaluate whether the expected utility of supporting or opposing research into geoengineering strategies is positive or negative. Since it will be nearly impossible to block all privately funded research from happening anywhere in the world, we will focus on whether the expected utility of promoting public research is positive or negative.

How can the expected utility of conducting research—of simply trying to learn something new—be negative? To sharpen the question, we can remind ourselves of a theorem of I.J. Good. Good's theorem teaches us that under ordinary conditions of decision theory and learning,*free* information is *always good*.[13] More formally, getting new information, assuming the cost of the acquisition itself is free, can never reduce your expected utility. Of course, in the real world, getting new information is never free. And even if the information is free, waiting for it to arrive might have opportunity costs. So Good's theorem cannot show that getting new information will never have negative utility, all things considered. It only says that changing your utility calculation by changing your credences in the light of the new evidence (once that evidence has been collected and paid for) will never have negative utility.

But Good's theorem seems to imply that if I am not in a hurry to make a decision, and if someone else is willing to pay the cost of getting the new information (or if the cost is basically negligible), I can never increase my expected utility by preventing them from gathering the new information at their expense and handing it over to me. And even if I am going to bear some of the cost of the new information (say because the other person wants to spend my tax money on the research), it also suggests that if the cost to me of the information is modest compared to the differences in utilities of the various outcomes, I should be willing to spend the money. So why should anyone ever oppose research?

Presumably, those who oppose research into GE strategies are aware that the direct financial costs of the research programs they oppose are very small compared to the differences in the utilities we would associate with the outcomes of implementing (vs not implementing) any GE strategies. This is surely true regardless of whether it is a great success or a disastrous failure. The stakes will simply be high one way or the other, and the resources being expended on the research are small. Moreover, environmental groups often oppose this research even when it privately funded by others. (Indeed, it is often part of their rhetorical strategies to alert their audiences to the sources of this funding—that it is funded by the Gates foundation, or by companies whose fortunes are tied to the fossil fuel industry. More on this later.) Their objection is not just that the research is a waste of resources that could be used on mitigation efforts. They seem to regard it as intrinsically dangerous.

13 (Good, 1967). See Myrvold (2012) and references therein for further discussions.

We had better be careful not to think Good's theorem proves too much. In the real world, allowing research to proceed could have large negative consequences.[14]

But that is because in the real world, it is often the case that at least one of two assumptions required for Good's theorem are *violated*.

1. If there are negative externalities associated with gathering the information (ones that cannot be associated with the financial cost of collecting the evidence), this might be, for example, because by signaling your willingness to collect the evidence, you cause some unwanted action by others.
2. If a formal condition of Good's theorem does not apply.

The formal condition I have in mind is sometimes called "reflection." Most saliently in this context, in order for reflection to be in place, and for Good's theorem to hold, it must be that the person evaluating the expected utility of conducting the research is the one whose credences will be used when conditionalizing on the new information. More crudely, if present-me thinks that future-me might misinterpret the evidence, then present-me might judge the belief-revision that would occur in the light of the misinterpretation of the new evidence to have negative expected utility. Or if present me worries that future you will act on the available evidence using your credences, then I might have a similar worry. So in a group context, this means that if I am worried that the people with whom I need to coordinate my future decisions will not interpret new information in a manner with which I agree, then gathering the new information might reduce my expected utility more severely than its prima facie cost would suggest.

We could call either one of these, that is, either negative externalities or the absence of the reflection condition, "unGoodian costs of learning."

For example, this is the best way to think about what Kicher argues in his (2003). When investigators conduct research into the biological basis of the differential success of members of economically or socially disadvantaged groups, they are gambling with the quality of life of the members of that group. And they are gambling in a way such that the deck is heavily stacked against the group—it is a loaded die, *and* the payout only occurs when the group in question loses the bet. We need to look and see if research into geoengineering has the same kind of expected utility.

3 Is the Expected Utility of Research into Geoengineering Negative?

Our task now is straightforward: we need to identify the unGoodian costs of geoengineering. Assessments of the prospects of various geoengineering projects will have to made, when the time comes, under substantial uncertainty, and in the absence of a scientific consensus of even what the relevant probabilities are. Thus, if

[14] For a good example of an argument that research can have negative consequences despite Good's theorem, see Kitcher (2003). He argues against the wisdom of conducting research into the biological origins of inequality. The basic framework I employ here is loosely structured around his work.

we are going to make a Goodian assessment of whether research into geoengineering strategies out to be promoted or opposed, we need to be able to estimate the probability that, as inconclusive evidence regarding the feasibility of various strategies comes in, harms or benefits will accrue. To do this, we need to answer the following questions: Are decision makers likely to overestimate or underestimate the evidence in favor of hypotheses regarding the prospects of various geoengineering projects? What asymmetries exist between how decision makers are likely to act when they believe that the evidence favors or disfavors hypotheses regarding the prospects of various geoengineering projects? What harms and benefits accrue when hypotheses regarding the prospects of various geoengineering projects being incorrectly and correctly appraised, respectively? And since we might also worry that geoengineering research has significant negative externalities, we also have to ask: What harms or benefits will accrue simply as a result of our tolerating, accepting, or promoting such research?

Arguably, however, such questions cannot be answered regarding a singular hypothesis about geoengineering. There are simply too many different possible geoengineering strategies, and it is unlikely the answers to the above questions will be the same for each one of them.

4 Opposition to Geoengineering Research

There are many different strategies that fall under the general umbrella of geoengineering. For the rest of the paper, I will focus on stratospheric aerosol injection (SAI). SAI is a method of reducing the amount of incoming solar radiation (so as to offset the decrease in outgoing thermal radiation) by injecting small sulfate particles into the stratosphere in a way designed to mimic the global cooling caused by volcanic eruptions like that of Mount Pinatubo in 1991. (Other similar proposals involve self-levitating particles made out of titanium or aluminum.) I focus on SAI both because it is likely to be one of the easiest geoengineering methods to achieve and because it among the riskiest and most controversial. But choosing SAI, I set the bar as high as it can be set for a defender of geoengineering research.

So let us begin in earnest by canvassing the standard consequentialist reasons people give for opposing research into SAI. They generally fall into three categories:

1. The moral hazard of *investigating* SAI
2. Harmful foreseeable consequences of *employing* SAI
3. Unforeseeable consequences of *employing* SAI

4.1 Moral Hazard

The idea of moral hazard first arose in the insurance industry and latter was applied in banking. The general idea is that if governments or central banks are perceived to be ready and willing to bail banks out, in the event that one of the bank's investments or loans causes the bank to become insolvent, then banks will stop managing their own risk carefully enough, and will engage in overly risky or "hazardous" behavior. The idea regarding geoengineering is similar: critics worry that if the public sees scientists conducting investigations into geoengineering strategies, they will conclude that climate change is less of a risk than they perceived it to be prior to seeing the scientists work on those strategies (Dembe & Boden, 2000; Lin, 2013). The public, in such a situation, might pursue mitigation strategies, such as decarbonizing their energy sources and agricultural practices, with less vigor. This, the critic maintains, is overly risky given the low probability (in the critics' opinion) that the geoengineering strategy will actually succeed in providing the same benefits that the mitigation strategy would have provided.

Moral hazard arguments can thus be understood as arguments that the cost of acquiring new information about SAI has much higher cost than merely the money and time spent on the scientific work—that SAI research has high negative externalities because of the signals it sends to other stakeholders. Thus, moral hazard arguments could by themselves be cogent arguments against SAI research even when the direct costs of the research are relatively insignificant and when Good's theorem applies.

4.2 Harmful Foreseeable Consequences

Some people seem to believe that simply because SAI has harmful foreseeable and unforeseeable consequences, we should oppose research into it.[15] So it is worth reviewing what people take each of these to be. Here is a list of some of the foreseeable ones:

1. Using SAI, or any form of SRM generally, to offset the impact of GHG, does nothing to alleviate any of the other harms of anthropogenic emissions—particularly the ocean acidification caused by CO_2. While this is not a consequence of SAI or SRM per se, many see it as a likely consequence of pursuing SRM insofar as SRM might license agents to pursue less vigorous mitigation efforts.[16]
2. The temperature of the earth is fixed by a balance between insolation (incoming solar radiation) and emissivity (escaping longwave radiation created by the black body radiation of the earth). Thus, the *net warming effect* of any decrease in emissivity can in principle be offset by a corresponding decrease in insolation. But a zero *net* change in mean temperature does not imply an absence of significant local changes. The pattern of local changes is likely going to have some relative local "winners and losers" (Robock, 2015).

[15] To be clear, arguments of this kind look prima facie like slippery slope arguments– "if we even do research into this it will inevitably be deployed, so reasons not to deploy are automatically reasons not to conduct research." In Section 8, I discuss more carefully worries that research inevitably leads to deployment. For now, I simply want to get on the table the reasons some have given for worrying about deployment.

[16] See Robock (2008).

3. SAI won't simply turn down the sun, and it won't simply decrease the *quantity* of solar radiation reaching the earth. It will also affect its character. SAI is predicted to whiten the sky and redden sunsets (Robock, 2015). Effects on crop production are uncertain. There is some evidence that corn production would be hurt (Tollenaar et al., 2017) but that other crops might benefit. Increases in CO_2, if they were accompanied by no other changes in climate (!!), would likely benefit most crops. So this might offset some of the possible harm of solar dimming.
4. SAI could cause ozone depletion by interacting with CFCs and other ozone destroying gases in polar stratospheric clouds (Tabazadeh et al., 2002). This could be mitigated either by using non-sulfate particles, like calcium carbonate, preventing the sulfates from drifting into polar stratospheric clouds, or by not engaging in SAI until the levels of CFCs fall to low enough levels, or by ensuring that the sulfates do not make it into polar stratospheric clouds *until* CFC levels have sufficiently fallen (Keith et al., 2016).
5. SAI effects have a very short lifespan. If a SAI system was compensating for what would otherwise be a significant degree of warming, and was suddenly shut down, the earth would suddenly undergo a very rapid change in climate. Many argue that this "termination shock" would be significantly worse than that had all the same GHGs been emitted and been allowed to have their damaging effects happen gradually. Some estimates suggest that termination shock could lead to climate warming at a pace twenty times greater than the warming evident today (Matthews & Caldeira, 2007; Trisos et al., 2018). Parker and Irvine (2018) claim that this risk is vastly overstated. Still, many argue that it would be an act of intergenerational injustice to employ such a system, since it would put a "Sword of Damocles" over the heads of future generations. Some argue that this would make any system designed to maintain SAI a tempting target for terrorists.

4.3 Harmful Unforeseeable Consequences

It is almost a performative contradiction to list unforeseeable consequences, but crudely, we can at least sort them into three categories.

1. Non-climatic consequences of conducting SAI. These could be anything from health effects on humans (especially respiratory, to direct damage to ecosystems, or any other similar effect of the particles themselves. See chapter 3 of (Callies, 2019) for a thorough discussion.)
2. Unforeseen climatic effects. It is reasonably foreseeable that SAI would produce "winners and losers," but insofar as the details of how this will play out are likely to remain uncertain up until the moment that SAI is deployed, the potential for unforeseen consequences here is large. Some argue that it would even be difficult to tell at first what the climatic consequences were, because the internal variabil-

ity of the weather makes detecting changes in the climate in real time difficult (Robock, 2015).

3. Geopolitical consequences. There are at least three possibilities (Victor, 2011):

   a. A state, or a collection of allied states, that is a net loser as a result of the employment of SAI could go to great lengths to try to bring an SAI program to a halt.
   b. A state or a collection of allied states that is not a net loser could, merely as a result of the natural variability of the weather, perceive itself to be a net loser, and go to great lengths to put an end to it.[17]
   c. A state with a non-climate-related grievance against another state could employ a well-researched SAI strategy as a weapon. (This is actually a possible consequence of merely researching SAI, and not necessarily of employing it.) Or, if SAI were already being employed, state A could launch an attack on state B by attempting to destroy the SAI system if state B was benefitting from the SAI much more than state A.[18]

5 Cost of Learning Considerations

Of course, the fact that SAI has harmful foreseeable and unforeseeable consequences is at least matched by the fact that climate change does too. And in any case, the question was not whether SAI presently looks promising. The question was whether the expected utility of supporting, or tolerating, research into SAI systems is obviously negative. As we have seen, to believe that SAI research has negative expected utility above and beyond its financial cost, you effectively have to believe one of two things (or both): That the results of research into SAI are very likely to be taken, by decision makers, to offer much stronger support for implementing SAI strategies than you yourself would take them to offer, or that research into SAI has intrinsic negative externalities.

Let us begin with the first of these: the absence of “reflection”—the worry that others will take the results of SAI research to be stronger evidence for the wisdom of its adoption than I would.

Why might someone have this worry? Here, I do not intend to canvas published or otherwise expressed reasons. Rather, I am concerned with the logical space of reasons that could give rise to this worry. I would argue that they fall into two general reasoning patterns. The first is based on the idea that others will have the wrong priors, and the second that they will have the wrong values.

The first line of reasoning goes something like this:

[17] The well-known climate scientist Gavin Schmidt has expressed something like this view, claiming that the first time the monsoon season failed in India, regardless of what caused the failure, they would go to great lengths to end an SAI program. https://twitter.com/ClimateOfGavin/status/1105848382000689152

[18] Given how hard it is to regionalize SAI, and the availability of other more powerful weapons to any country capable of developing such a system, this risk strikes some as farfetched.

1. My credence in the proposition that SAI strategies could possibly be beneficial is so low that I view it as nearly impossible for research ever to significantly raise it.
2. All scientific research has a non-trivial probability of being misinterpreted by decision makers in one direction or the other.
3. Therefore, scientific research into SAI has a non-trivial probability of being taken to offer stronger support for implementation than I would warrant and a nearly zero probability of doing the opposite.

The second line would go something like this:

> Those who would conduct such research, or those who would be responsible for disseminating its results, would not share my values, and therefore they would either conduct, or disseminate the results of, this research in a way that I took to be biased in favor of the proposition that it would be beneficial to implement an SAI strategy.

The second premise of the first argument is probably true. But what should we make of the first premise? And what about the central premise of the second argument? Let us evaluate them in turn in the next two sections. After that, we can evaluate the claim that the research itself has negative externalities.

6 Biased Reporting

One can indeed tell that biased evaluation or reporting due to conflicted interests is one of the major concerns of environmentalists who oppose GE research simply by looking at the language they use. Consider the following passages from *Geoengineering Monitor*'s SRM factsheet.

> **Key Players: ExxonMobil and Shell**
> There are large companies for whom 'saving the world' – exclusively through some sort of techno-fix – is increasingly becoming a structural prerequisite for continuing their business, particularly when those companies depend heavily on fossil fuels. They try to shift policy norms so that previously unthinkable notions and activities – like solar radiation management – start to become more mainstream and acceptable.
> Among them, ExxonMobil's Senior Scientific Advisor Dr. Haroon Kheshgi is the point person on geoengineering, recruited from the Lawrence Livermore National Laboratory.

As we have noted, the actual influence of fossil fuel companies, and other companies that depend on fossil fuels, on SAI research is at best controversial.[19] But let us be perfectly clear about this: General worries about bias in research due to financial

[19] (Reynolds et al., 2016; Reynolds, 2019a).

and similar conflicts of interest are *serious and well founded.* It is well-documented that when scientific research is sponsored or funded by industry players who have a financial interest in having the research come out in a particular way, it is more likely to come out the way that favors the industry players. This is a familiar result from toxicology. Studies concerning the toxic properties of low doses of bisphenol A, famously, seemed to be strongly affected by who was funding the research. Of the experiments on low-dose exposure to the substance, 90% of those that were funded by government concluded that it had significant toxic effects. But not one of the industry funded studies found any toxicity.[20] As Torsten Wilholt notes, moreover, most or all of this discrepancy can be explained without appeal to fraud in the industry sponsored studies (Wilholt, 2009). Rather, Wilholt attributes the differences to what he calls "preference bias" wherein the industry players systematically make methodological choices that accord with their "values" (which presumably place company profits ahead of consumer safety), but which are not, in any deep sense, objectively wrong. They are simply not the choices I would make given my priors and my values. But in *my* deliberations about the expected utility of research, it is my values and priors that matter. And so research that exhibits the preference bias of industry players is likely to have large unGoodian costs to me.

According to Wilholt, the solution to the problem of agents with different interests or values making methodological choices that favor their own values is the institution of what he calls "conventional methodological standards." These include standards such as publication rules and standards regarding experimental design.[21]

Environmentalists who worry about the dangers of SAI research leading to unjustified conclusions regarding its safety and efficacy should find this *extremely alarming*, especially since geoengineering research, being extremely novel and exploratory, cannot possibly have as rich a set of conventional research standards as toxicology does. The possibility, whether or not it presently exists of researchers who strongly represent the interests of companies that depend on fossil fuels to conduct research with high unGoodian costs, is extreme and should be taken seriously.

An obvious question, however, is what the best thing to do about this is. One option is the one being pursued by Geoengineering Monitor: Make them stop doing the research. It is an empirical question how effective this will be. It is interesting to note that geoengineering research seems to be divided along the following lines: Physical experiments like SPICE (discussed above) and the Stratospheric Controlled Perturbation Experiment (SCoPEx),[22] which are aimed at testing the physical effects of particulates and how they disperse, tend to be much more tentative and funded by unusual sources, whereas studies about the climatological impacts of possible SRM technologies (usually conducted with computer simulations) tend to be more

[20] See Wilholt (2009) and references therein for details.

[21] The discrepancy noted above in studies of bisphenol A turned out to arise from the fact that industry sponsored studies tended to choose less estrogen-sensitive rats in their studies, whereas it was well understood that the toxicity of bisphenol A acted through a channel that mimicked estrogen. This has led to the adoption of standard in choices of model organisms that are permitted in toxicity studies.

[22] https://projects.iq.harvard.edu/keutschgroup/scopex

mainstream, integrated into the broader scientific community, and funded in more mainstream ways. Indeed the climate modeling community interested in studying SRM has built up a rough analog to the highly community standard-governed climate model intercomparison project (CMIP)—the central body of modeling work that informs the projections of the IPCC. Theirs is called the GeoMIP,[23] and it is now officially integrated into CMIP. Arguably, the existence of inter-coordination projects like GeoMIP would *reduce* the likelihood of preference bias in research, since inter-coordination projects are the natural home of things like research standards. But they might also boost the legitimacy of such projects without eliminating bias entirely. I do not pretend to have a proof that the former effect is stronger, but I suspect it is. Let us at least give a name to the empirical premise that supporting inter-coordination research projects like the CMIP generally, and the GeoMIP particularly, will have the net effect of reducing the likelihood of public attitudes to SRM and SAI being influenced by research affected by preference bias. We can call it the bias reduction premise.

Bias Reduction Public support for SRM and SAI research will bring it out into the open and promote the creation of inter-coordination projects that will establish conventional standards which reduce the possibility of preference bias. This effect would more than offset the extra legitimacy conferred on projects that did continue to be affected by preference bias.

This is one of the most important empirical premises that I discuss in this paper—and the evidence, from philosophy and sociology of sciences like toxicology, is quite strong: industry-funded research has high unGoodian costs for people who do not share industry's values.

7 Should We Assign an Extremely Low Credence to the Hypothesis that SAI Could Be Beneficial?

If you have an extremely low credence in some hypothesis, this can justify a belief that research surrounding that hypothesis has high unGoodian costs. Take the hypothesis that prayer has more powerful healing properties for cancer and heart disease than hospitals do. I would oppose research into this hypothesis because my credence in it is so low that I am apt to think that the only thing new data could do is lead people to erroneously accept the hypothesis, and this would lead to great harm.

Is the hypothesis that SAI could help the world deal with climate change like this? Two empirical premises suffice for mitigating the worry that it is:

Minimum Climate Pessimism It is reasonable to be nearly certain that SAI would invariably have some unavoidable harmful side effects—to human health, crops,

[23] http://climate.envsci.rutgers.edu/GeoMIP/index.html

ecosystems, etc. Climate pessimism is the claim humans are nearly certain not to achieve sufficient climate change mitigation such that the damage caused by climate change is significantly greater than the sum total of the *clearly foreseeable* harmful side effects of SAI.

Minimum SAI Optimism There is at least some non-trivial probability that SAI could in fact deliver some net benefit.

Without minimum climate pessimism, you would probably think that SAI would invariably bring about net harm. And without minimum SAI optimism, you might think no evidence could justifiably raise your credence to a high enough level to give any SAI strategy positive expected value.

So what is the right attitude to have to these two premises? Unfortunately, I think **minimum climate pessimism** has become, in recent years, nearly unassailable. It is hard to imagine that the *obviously foreseeable* side effects of SAI could rival the damage from climate change that is already baked into the cake. A recent letter to *Nature* argues that "The committed future CO2 emissions from *proposed and existing* energy infrastructure represent more than the entire remaining carbon budget if mean warming is to be limited to 1.5 degrees" (Tong et al., 2019). My own view is that keeping total GHG emissions below levels that, left to their devices, would eventually warm the planet by at least 2 °C is looking more and more like a goal in the rear view mirror. And it is hard to imagine that the damage caused by that much warming would not eclipse the negative impacts of an SAI program about which we could be nearly certain.

Regarding **minimum SAI optimism**, I should emphasize that this premise is less empirical than it might appear. That is because there is bound to be some fairly deep disagreement about what counts as overall net benefit. I can best explain what I mean by this by giving some crudely constructed examples. Suppose that we knew that an SAI scheme would accrue 10 positive utility points for each of the members of 90% of the world's population, but 40 negative utility points to 10% of that population (relative to some assumed emissions scenario with no geoengineering remedy). The coarsest utilitarians would automatically declare victory. Others would declare victory only conditional on a scheme of balance of payments designed to restore justice. Others might think that some of the damages accrued by the losers would be economically incommensurable. Or they might worry that some of the members of the 10% of the population live under political conditions of corruption too severe for economic aid to reach them. The last group might not see net benefit where the first group does. Thus, it is not a simple scientific or empirical matter to decide if a geoengineering scheme will deliver net benefits.

What about the empirical component of the premise? Recent work from the scientific community that models the regional effects of SRM interventions has become slightly more optimistic about the degree to which SRM and specifically SAI schemes could be made, with the right implementation, to reduce the degree of variation from winners to losers (Irvine et al., 2019; Tilmes et al., 2018, 2020).

Suffice it to say that one's attitude to the minimum SAI optimism premise will depend on how one views the present evidence, about how low of a credence one needs to have in a hypothesis before one should conclude that *any* future evidence gathering has negative expected utility, and about how much tradeoff between injustice and net utility one will tolerate. All three of these are moving parts that reasonable people can disagree about. My own view is that the first and third of these moving parts pull in the direction of accepting the premise, but the second pulls against it. We will revisit these considerations in the conclusion of the paper.

8 Negative (and Positive) Externalities of SAI Research

So far we have looked at the possible negative direct unGoodian costs of SAI research. These are the possible costs in expected utility that accrue as a result of our expectation that new evidence will be misinterpreted (in a way that is contrary to our own credences). But in addition to these costs, there are other unGoodian costs and benefits to such research that arise as externalities. Research takes money and intellectual labor, and it outputs evidence, but the doing of it also sends signals, and it also sometimes creates social structures that do not otherwise exist.

8.1 Creating Social Structures

We have already discussed the creation of social structures in the context of the creation of conventional epistemic standards and norms. But what about ethical norms and best practices regarding what are acceptable experiments and implementations? And perhaps more importantly, what about the ethical norms that can form the foundation of international governance?

Regardless of what we come to learn about the benefits and perils of SAI—indeed regardless of whether we do or do not learn anything—international governance of SAI will be soon become crucial.

One thing many opponents of SAI fear is unilateral state action on SAI.[24] This is both something that is, ceteris paribus, likely to occur as well as something likely to be dangerous. To see why it is likely to occur, consider a state like Saudi Arabia. Saudi Arabia has the wealth to do it, the military power to stave off moderate geopolitical pressure, and a geological situation that would make it a likely winner from SAI. Their only concern vis a vis climate change, moreover, is probably sea level rise. It is plausible, therefore, that Saudi Arabia, just for example, would face a strong temptation to take unilateral action with SAI. And for reasons we will see below, I think SAI research is likely to lower the risk a nation going rogue, not raise it. For now, let us encode two basic assumptions underlying the fear of unilateral action into two named premises:

Rogue Nation SAI is sufficiently cheap, and sufficiently potent (in the sense of producing dramatic effects, irrespective of how net good they are in the aggregate), and

[24] See Rabitz (2016) and references therein.

climate change will eventually put enough states under enough stress, that, conditional on what we presently know about SAI, the probability is **significant** that some state with both the financial and technical capacity to implement SAI and to defend itself militarily, will eventually try to implement it unilaterally.[25] A rogue nation is much more likely to do a bad job of implementing SAI and produce an outcome that is bad for the world as a whole.[26]

An **addendum to Rogue Nation** is that, unless SAI research was to surprisingly discover that SAI technology was impossible to implement even very badly (which I think we are unlikely to discover), the probability of a state going Rogue will likely remain significant, regardless of what we learn about SAI. The reason that the addendum is probably true is that if SAI research goes badly, it will likely be in the form of showing that it is impossible to do a reasonably decent job of smoothing out the winners and losers, or of attenuating harmful side effects. But such findings are only likely to reshuffle the deck regarding which states are likely to go rogue. There is no predictable way in which we should expect the results of SAI research will alter our probabilities of *some* state going rogue.

The first premise tells us that *the risk of unilateral action is already quite high*. The second tells us that *SAI research is unlikely to raise that risk*. If these premises are true, then one of the most important externalities of SAI research would be if its pursuit increased or decreased the probability of the emergence of effective international governance structures.

Indeed, Blomfield (2015) argues that conducting SAI research would actually diminish the probability of successful international governance structures taking shape. She argues, on the explicit analogy to a Rawlsian theory of justice, that hammering out an international governance body is a bit like choosing a just society: it is best conducted under the veil of ignorance. In other words, just as Rawls argues that individuals would be best able to deliberate about what the most just social structures would be if they do not yet know what position they will occupy in society, states are most likely to reach agreements underpinning SAI governance bodies if they do not yet know whether they will be SAI winners or losers. On her account, as SAI research reveals new facts, some states will begin to see themselves as likely winners, and will be less likely to enter into governance agreements that might inhibit their capacity to use SAI to their own benefit. It is thus crucial, according to her, that we defer research for at least as long as is necessary to reach such agreements.

I agree with Blomfield that this would be a compelling reason to defer research. The Rogue Nation premise is very likely true, and hence, governance structures are vital. If deferring research would facilitate governance, then that would be a strong

[25] In fact Fruh and Hedahl (2019) have argued that, under considerations of just war theory, some nations would be justified in implementing rogue SRM strategies, even if those strategies harmed other nations. I take no position on that claim other than to note that it does suggest the outcome is not entirely unlikely, whether or not it would be justified.

[26] See Victor (2011) for more details.

incentive to defer research. But there are two considerations that weigh against her basic premise.

1. In the first place, it is somewhat misleading to view the current situation as being analogous to the situation of being behind the veil of ignorance.

   It is true that it is difficult for some states to determine, at the present time, whether they will be SAI winners or losers. But other states do not presently occupy this position. As we noted above, rich, powerful, and well-armed gulf oil states are almost guaranteed winners from SAI if GHG forcings start to drive sea levels beyond those that are dangerous to them. The USA is arguably a likely winner from SAI under similar circumstances, since we probably have better options for mitigating precipitation changes than we do sea level rise (and "category 6" hurricanes) in cities like New York and Miami. Unsurprisingly, when a group of nations, led by Switzerland, recently tried to raise the issue of SRM governance at the UN environment assembly in Nairobi, the motion was blocked by opposition promoted by the USA and Saudi Arabia.[27]
2. The second consideration is that there is evidence from political science that international scientific research which creates a body of accepted shared knowledge can lead to shared and accepted ethical norms and best practices regarding what are acceptable experiments and implementations, which can in turn form the foundation of international governance (Victor, 2011, 196).

The underpinnings of this have been studied in both cases of economic uncertainty (Ascher, 1983) and environmental uncertainty (Haas, 1990); the underlying logic is the same in each case. If international scientific research creates a body of accepted or shared knowledge, it creates the grounds for the formation of what Haas calls an epistemic community (Haas, 1990). Such groups not only share knowledge but also share the same values and beliefs as to the proper way to obtain such knowledge. This enables the community to evaluate and adapt to new evidence more swiftly as the beliefs and values that govern the assimilation of new evidence are shared. This can easily lead to a more firm foundation for international governance.

David Victor argues, in fact, that in the absence of an open epistemic community, international governance treaties are nearly impossible to achieve. The basic idea is this: only when states have a clear and shared understanding of what they are giving up and what they are getting in return do useful governance agreements emerge. Even if states agree to a taboo, in the absence of shared knowledge, suspicion will eventually lead to secret research (as it did in the contrasting case of secret military programs to develop battlefield weather modification techniques.)

The key insight that Victor, Haas, and Ascher share is one that comes from public choice theory. Blomfield's argument assumes that states act like individuals, looking to minimize the harm of the worst-case scenario for themselves as the Rawlsian subject does. But states are governed by politicians and bureaucrats, who have complex

[27] (Jinnah & Nicholson, 2019).

sets of motives that rarely match those of the people they govern. What is minimax for the policy maker is not necessarily what is minimax for the state itself.

Thus, research communities need not even reach consensus, or eliminate uncertainty, in order to create a moral order. Ascher argues that "once it is established that the international economic regime is not a straightforwardly determinable vector of nations' interests and power…uncertainty gives greater power to those (whether individuals or subunits) who "absorb uncertainty". In such circumstances, politicians and bureaucrats, in other words, can protect themselves from future criticism by deferring to experts. Being able to point to experts and say "we just did what they said" is the best way for policy makers to minimax. This leads, Ascher argues, to situations in which expert bodies, whether or not they have in fact reduced uncertainty, get a life of their own, and an attendant set of powers.

Whatever the evidence in its favor, this brings us to another important empirical premise:

Research Builds Norms Public support for SRM and SAI research will bring it out into the open and promote the creation of scientific communities that will establish shared ethical norms and best practices regarding what are acceptable experiments and implementations. This raises the probability of the emergence of international governance structures that could reduce the probability of a rogue state implementing an SAI or other SRM program.

If **research builds norms** is true, then it suggests that SAI research has a significant positive externality.

8.2 Signaling

We now come to one of the most common[28] complaints about all forms of geoengineering research: that they create moral hazard. *Geoengineering Monitor* put the point thusly:

> SRM, and geoengineering more broadly, is a "perfect excuse" for climate deniers and governments seeking to avoid the political costs of carbon reductions. For those looking to stall meaningful climate action, the active development of geoengineering tools and experiments will be presented as a preferred pathway to address climate change and be used as an argument to ease restrictions on high carbon emitting industries.[29]
>
> Environmental law professor Lin (2013) argues as follows:
>
> Geoengineering presents a strong economic, political, and psychological temptation to defer difficult and costly actions to future generations. This tempta-

[28] "One of the main ethical objections to geoengineering" is "moral hazard" according to the United Kingdom's Royal Society Report Geoengineering the Climate (Ming et al., 2014, 39).

[29] http://www.geoengineeringmonitor.org/2018/06/stratosphericaerosolinjection/

tion, whether characterized as moral hazard, risk compensation, or political opportunism, is a serious concern because geoengineering is widely acknowledged to be an inferior, problematic, and at best temporary option for responding to climate risks. (p. 711)

But there is a palpable tension in Lin's quotation. In making the case for moral hazard, he acknowledges that geoengineering is "*widely acknowledged* to be an inferior, problematic, and at best temporary option." Some commenters on the topic have pointed out that it is hard to see how an option "widely acknowledged to be an inferior, problematic, and at best temporary" could be seen as a "perfect excuse" to avoid real climate action. Martin Bunzl writes that moral hazard arguments "seem far-fetched since, at least among policy makers, nobody believes that geoengineering offers anything but a relatively short stopgap to buy time for other action" (Bunzl, 2009, 2).[30]

There is a large literature on moral hazard arguments concerning geoengineering, and I won't review it all here. Hale (2012) offers an extensive catalog of different varieties of moral hazard arguments, all of which he claims fail to establish that geoengineering strategies should not be pursued. Here, I limit myself to the observation that there are two empirical questions: one needs to answer if we are interested in knowing the signaling value of doing SAI research.

1. How likely is it that the "difficult and costly actions" that need to be taken to achieve the carbon reductions that would make geoengineering a luxury will be taken in time to ensure that it remains one?
2. What actually is the signaling value of SAI research. Does it, as moral hazard proponents argue, send the signal that "all is well, technology will save us and we are free to ease restrictions on high carbon emitting industries?" Or does it, rather, signal that climate change isn't just a conspiracy promoted by luddites and haters of capitalism who just want to stop economic growth for its own sake?[31]

Regarding the first question, I think I have already made my view known. I think 1.5 °C of warming is already baked into the cake, and staying below 2 °C would require changes in our energy and agricultural systems at a speed that is difficult to imagine. Some moral hazard arguments strike me as grounded in a belief in the idea that a solution to all our problems just requires a commitment to urgent action that could be right around the corner. I am less optimistic.

Regarding the second question, there is in fact some empirical research.[32] Evidence seems to suggest that the following are true:

[30] See Neuber and Ott (2020) and citations therein for more on the prospects of "buying time" with SAI.

[31] One prominent libertarian philosopher recently posted a story on Facebook about SAI technology and claimed that if "climate alarmists" really believed their projections, they would be furiously pursuing such technology.

[32] See (Corner & Pidgeon, 2014; Merk, 2018; Merk et al., 2016; Raimi et al., 2019) and references therein for more details.

1. Most people (in the USA, the UK, and Canada) say, about themselves, that they would not pursue mitigation less vigorously when they learn about the possibility of geoengineering. More people said they become more motivated to pursue mitigation efforts when they learn about geoengineering research than the contrary.
2. Many people do believe that *others* would become less motivated to pursue mitigation efforts when *they* learn about geoengineering results.
3. The "galvanizing" effect of learning about geoengineering (wherein when the person learns that governments are pursuing geoengineering efforts, the learner claims that they will become more motivated to pursue mitigation efforts) is less pronounced among people who describe themselves as skeptical about anthropogenic climate change. On the other hand people who are skeptical of climate change are more likely to have their estimates of climate change risks increase. (A common explanation of the last fact is that many people underestimate the risk of climate change precisely because they believe that if the risks are real, they will have to modify their behavior. Learning about geoengineering reduces the degree of belief in that conditional.)
4. Fear of the implementation of SAI can cause learning about the possibility of SAI research to boost desire to mitigate.
5. How information is framed in questionnaires has a very large impact on how people answer.[33]

The empirical evidence, in other words, is hardly unequivocal, and reasonable people can still disagree about whether the signaling value of geoengineering research generally, and SAI research in particular, is negative, neutral, or positive.[34] But reasonable people might also disagree about whether any of this matters much, given the pace of mitigation efforts in a world in which there is little geoengineering research taking place, and even less public awareness of it. Rather than naming a premise with a binary truth value here, let us give a name to a spectrum of views:

Signaling Value One's views on the signaling value of geoengineering and SAI research can range from highly negative (mitigation efforts will be seriously harmed by the existence of GE research) to neutral, to highly positive. The right view to have on this is not well fixed by the evidence. One can also have diverging views on how successful mitigation would be in the absence of SAI research. If one is pessimistic about mitigation, the political impact of people learning about SAI research might not seem to be terribly significant.

8.3 Lock in

A final alleged negative externality of various kinds of technological research is so-called lock in. Though Catriona McKinnon does not argue against SRM research

[33] See especially Raimi et al. (2019) and Merk et al. (2016).

[34] In addition to the specific findings about framing effects in SAI-related surveys, there are general reasons to doubt the results of this kind of research. See, for example, Bullock and Lenz (2019).

tout court, she does warn that "If we govern to enable and stimulate SRM research so that it aims at delivering a deployable technology as soon as possible, we could lock in a pathway that commits us to deployment" (McKinnon, 2019, 444). The idea seems to have come from an influential piece by Jamieson (1996), but it has been echoed in many places. The Royal Society report on geoengineering governance warns, for example, that "Scientific momentum and technological and political 'lock-in' may increase the potential for research on a particular method to make subsequent deployment more likely, and for reversibility in practice to be difficult even when technically possible" (The Royal Society, 2009). Stephen Gardiner has argued that, "It is not clear that geoengineering activities can really be limited to scientific research…. In our culture, big projects that are started tend to get done. This is partly because people like to justify their sunk costs; but it is also because starting usually creates a set of institutions whose mission it is to promote such projects" (Gardiner, 2011). And Albert Lin argued, similarly, that "[e]ven very basic and safe research … could be a first step onto a 'slippery slope,' creating momentum and a scientific lobbying constituency for development and eventual deployment" (Lin, 2013).[35]

The empirical premise here seems to be:

Lock in Research into technologies like SAI will inevitably lead to their deployment.

But it is not clear what the evidence for this premise is meant to be. As Callies (2018) argues, there seem to be more counter-examples to this than examples. Drug companies and medical device companies invest billions of dollars into researching candidate therapies, and the majority of these candidates is abandoned. We have, of course, the example of the atom bomb at the end of World War II. But most countries that have conducted the research necessary to build nuclear bombs have never *actually* deployed one, and some countries, like Brazil and South Africa, invested large sums into researching nuclear weapons but subsequently refrained from arming themselves or have disarmed themselves. For a strange and awkward technology like SAI, the worry seems poorly motivated.[36]

9 Conclusion

Where does this leave us? The most important premises surveyed in this paper are **bias reduction** and **research builds norms**. The evidence for these, from philosophy and sociology of science and from public choice theory, respectively, is quite strong. And their impact is likely to be substantial. The greatest danger of scientific evidence being misappraised arises when the process of gathering evidence is captured and monopolized by agents with the economic interest to engage in preference bias. And the greatest general danger we face with regard to

[35] The above three quotations are assembled in Callies (2018).

[36] But see Callies (2018) for a more detailed discussion.

the implementation of SAI is that someone will implement it unilaterally. That is, **Rogue Nation** is likely true, and so any process that lowers the attendant probability is highly desirable. If all of this is right, it means that greatest risks of SAI risks specifically, and of geoengineering in general, are *mitigated*, rather than aggravated, when we support public research into them. I take these to be the two most important positive findings of the paper.

If the central question regarding SAI, SRM, or other GE technologies is how to prevent people funded by the carbon-emitting industries or interested in patenting and profiting from technologies from becoming the experts we need to rely on to assess the relevant hypotheses, then we should support SAI research, and we should in any case support better integration of the communities that conduct modeling research and that conduct physical implementation research. If the central question is how to ensure that, when the time comes, there are good international governance structures in place so that whether or not to implement SAI is decided via a global process, then we *probably* ought to support the development of an international community of experts on SAI. If both of these considerations are important, then, we *very probably* ought to support SAI research.

The strongest negative finding of the paper is that SAI research is unlikely to provide misleading evidence simply because our rational priors ought to be so low to begin with that it can help. One only has to believe extremely weak versions of **minimun climate pessimism** and **minimum SIA optimism** in order to dismiss the worry that unbiased research, governed by a sensible group of researchers with a representative set of values, would be much more likely to reach a false positive conclusion regarding the wisdom of pursuing SAI than we would if we were evaluating the evidence yourself.

Regarding the remaining externalities of SAI research, the evidence for **lock in** is poor. Indeed, there is moderately strong evidence that it is false. After surveying the empirical evidence from opinion surveys, it is difficult to remain anything other than agnostic about the *signaling value* of SAI research and where it lies on the spectrum from strongly negative to strongly positive. At best (in the sense of the most epistemically strong claim one ought to endorse), the research weakly suggests that the signaling value is positive, but probably not by much. More importantly, this is largely irrelevant. Even if the signaling value of SAI research is negative, it will not impact many people. More significantly, no matter how the political winds shift with regard to mitigation effort in a few select countries, we should, at this stage in the game, expect fairly serious and dangerous climate outcomes. These mitigation efforts are unlikely to get much more urgent in the USA and Australia, have probably reached maximum urgency in most of the rest of the developed world, and will remain low in developing economies.

Given all of this, my conclusion is that modest support of SAI research is warranted, and I hope to have provided a useful template for thinking about other forms of geoengineering research.

Acknowledgement Special thanks to Judah Kreinbrook, who worked as a research assistant on this paper.

References

Ascher, W. (1983). New development approaches and the adaptability of international agencies: The case of the World Bank. *International Organization, 37*(3), 415–439.

Blomfield, M. (2015). "Geoengineering in a climate of uncertainty." *Climate Change and Justice*, 39–58.

Bullock, J. G., & Lenz, G. (2019). Partisan bias in surveys. *Annual Review of Political Science, 22*(1), 325–342. https://doi.org/10.1146/annurev-polisci-051117-050904.

Bunzl, M. (2009). Researching geoengineering: Should not or could not? *Environmental Research Letters, 4*(4), 045104.

Callies, D. E. (2018). "The slippery slope argument against geoengineering research". *Journal of Applied Philosophy*.

Callies, D. E. (2019). *Climate Engineering: A Normative Perspective*. Lexington Books.

Corner, A., and Pidgeon, N. (2014). "geoengineering, climate change scepticism and the 'moral hazard' argument: An experimental study of UK public perceptions." *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 372* (2031). https://doi.org/10.1098/rsta.2014.0063.

Cressey, D. (2012). Geoengineering experiment cancelled amid patent row. *Nature News*. https://doi.org/10.1038/nature.2012.10645.

Dembe, A. E., & Boden, L. I. (2000). Moral hazard: A question of morality? *NEW SOLUTIONS: A Journal of Environmental and Occupational Health Policy, 10*(3), 257–279. https://doi.org/10.2190/1GU8-EQN8-02J6-2RXK.

Elliot, K. C. (2011). Direct and indirect roles for values in science. *Philosophy of Science, 78*(2), 303–324. https://doi.org/10.1086/659222.

Fruh, K., & Hedahl, M. (2019). Climate change is unjust war: Geoengineering and the rising tides of war. *The Southern Journal of Philosophy, 57*(3), 378–401.

Gardiner, S. M. (2011). *A perfect moral storm: The ethical tragedy of climate change*. Oxford University Press.

Good, I. J. (1967). On the principle of total evidence. *The British Journal for the Philosophy of Science, 17*(4), 319–321.

Haas, P. M. (1990). *Saving the Mediterranean: The politics of International Environmental Cooperation*. Columbia University Press.

Hale, B. (2012). "The world that would have been: Moral hazard arguments against geoengineering." In *Reflecting Sunlight: The Ethics of Solar Radiation Management*. Rowman Littlefield.

Heyward, C. (2013). Situating and abandoning geoengineering: A typology of five responses to dangerous climate change. *PS: Political Science & Politics, 46*(1), 23–27. https://doi.org/10.1017/S1049096512001436.

Irvine, P., Emanuel, K., He, J., Horowitz, L. W., Vecchi, G., & Keith, D. (2019). Halving warming with idealized solar geoengineering moderates key climate hazards. *Nature Climate Change, 9*(4), 295. https://doi.org/10.1038/s41558-019-0398-8.

Jamieson, D. (1996). Ethics and intentional climate change. *Climatic Change, 33*(3), 323–336.

Jamieson, D. (2013). some whats, whys and worries of geoengineering. *Climatic Change, 121*(3), 527–537.

Jinnah, S., & Nicholson, S. (2019). The hidden politics of climate engineering. *Nature Geoscience, 12*(11), 876–879. https://doi.org/10.1038/s41561-019-0483-7.

Keith, D. W., Weisenstein, D. K., Dykema, J. A., & Keutsch, F. N. (2016). Stratospheric solar geoengineering without ozone loss. *Proceedings of the National Academy of Sciences, 113*(52), 14910–14914. https://doi.org/10.1073/pnas.1615572113.

Kitcher, P. (2003). *Science, truth, and democracy*. Oxford University Press.

Kuo, K. (2012). "Geoengineering trial cancelled: More regulation needed". *The Conversation*. http://theconversation.com/geoengineering-trial-cancelled-more-regulation-needed-7297.

Lenton, T. M., and Vaughan, N. E. (2009). "The radiative forcing potential of different climate geoengineering options." *Atmospheric Chemistry and Physics* 23.

Lin, A. C. (2013). "Does geoengineering present a moral hazard." *Ecology Law Quarterly* 40: 673. https://heinonline.org/HOL/Page?handle=hein.journals/eclawq40&id=697&div=&collection=.

Matthews, H. D., & Caldeira, K. (2007). Transient climate–carbon simulations of planetary geoengineering. *Proceedings of the National Academy of Sciences, 104*(24), 9949–9954.

McKinnon, C. (2019). Sleepwalking into lock-in? Avoiding wrongs to future people in the governance of solar radiation management research. *Environmental Politics, 28*(3), 441–459.

Merk, C. (2018). Do climate engineering experts display moral-hazard behaviour? *Climate Policy, 19*(2), 231–243. https://doi.org/10.1080/14693062.2018.1494534.

Merk, C., Pönitzsch, G., & Rehdanz, K. (2016). Knowledge about aerosol injection does not reduce individual mitigation efforts. *Environmental Research Letters, 11*(5), 054009. https://doi.org/10.1088/1748-9326/11/5/054009.

Ming, T., de Richter, R., Liu, W., and Caillol, S. (2014). Fighting global warming by climate engineering: Is the earth radiation management and the solar radiation management any option for fighting climate change?. *Renewable and Sustainable Energy Reviews, 31*(March): 792–834. https://doi.org/10.1016/j.rser.2013.12.032.

Geoengineering Monitor. (n.d.). "Geoengineering monitor – hands off mother earth." Accessed Sept 14, 2019. http://www.geoengineeringmonitor.org/.

Myrvold, W. C. (2012). Epistemic values and the value of learning. *Synthese, 187*(2), 547–568.

Neuber, F., & Ott, K. (2020). The buying time argument within the solar radiation management discourse. *Applied Sciences, 10*(13), 4637. https://doi.org/10.3390/app10134637.

Parker, A., & Irvine, P. J. (2018). The risk of termination shock from solar geoengineering. *Earth's Future, 6*(3), 456–467. https://doi.org/10.1002/2017EF000735.

Rabitz, F. (2016). Going rogue? Scenarios for unilateral geoengineering. *Futures, 84*(November), 98–107. https://doi.org/10.1016/j.futures.2016.11.001.

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

Reynolds, J. L. (2019a). "Does the fossil fuel industry support geoengineering?". https://geoengineering.environment.harvard.edu/blog/does-fossil-fuel-industry-support-geoengineering.

Reynolds, J. L. (2019b). *The Governance of Solar Geoengineering: Managing Climate Change in the Anthropocene*. Cambridge University Press.

Reynolds, J. L., Parker, A., & Irvine, P. (2016). Five solar geoengineering tropes that have outstayed their welcome. *Earth's Future, 4*(12), 562–568. https://doi.org/10.1002/2016EF000416.

Robock, A. (2008). 20 reasons why geoengineering may be a bad idea. *Bulletin of the Atomic Scientists, 64*(2), 14–18. https://doi.org/10.2968/064002006.

Robock, A. (2015). Stratospheric aerosol geoengineering. *AIP Conference Proceedings, 1652*(1), 183–197. https://doi.org/10.1063/1.4916181.

Stocker, T. (2014). *Climate change 2013: The physical science basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Tabazadeh, A., Drdla, K., Schoeberl, M. R., Hamill, P., & Toon, O. B. (2002). Arctic 'ozone hole' in a cold volcanic stratosphere. *Proceedings of the National Academy of Sciences of the United States of America, 99*(5), 2609–2612. https://doi.org/10.1073/pnas.052518199.

The Royal Society. (2009). *Geoengineering the climate: Science, governance and uncertainty*. The Royal Society.

Tilmes, S., Richter, J. H., Kravitz, B., MacMartin, D. G., Mills, M. J., Simpson, I. R., & Glanville, A. S. (2018). Cesm1(WACCM) stratospheric aerosol geoengineering large ensemble project. *Bulletin of the American Meteorological Society, 99*(11), 2361–2371. https://doi.org/10.1175/BAMS-D-17-0267.1.

Tilmes, S., MacMartin, D. G., Lenaerts, J. T. M., van Kampenhout, L., Muntjewerf, L., Xia, L., Harrison, C. S., et al. (2020). "Reaching 1.5 and 2.0 °C Global surface temperature targets using stratospheric aerosol geoengineering". *Earth System Dynamics, 11*(3): 579–601. https://doi.org/10.5194/esd-11-579-2020.

Tollenaar, M., Fridgen, J., Tyagi, P., Stackhouse Jr, P. W., & Kumudini, S. (2017). The contribution of solar brightening to the US maize yield trend. *Nature Climate Change, 7*(4), 275–278. https://doi.org/10.1038/nclimate3234.

Tong, D., Zhang, Q., Zheng, Y., Caldeira, K., Shearer, C., Hong, C., Qin, Y., and Davis, S. J. (2019). "Committed emissions from existing energy infrastructure jeopardize 1.5 °c climate target". *Nature*, July, 1. https://doi.org/10.1038/s41586-019-1364-3.

Trisos, C. H., Amatulli, G., Gurevitch, J., Robock, A., Xia, L., & Zambri, B. (2018). potentially dangerous consequences for biodiversity of solar geoengineering implementation and termination. *Nature Ecology & Evolution, 2*(3), 475–482.

Victor, D. G. (2011). *Global warming gridlock: creating more effective strategies for protecting the planet*. Cambridge University Press.
Wilholt, T. (2009). "Bias and values in scientific research." *Studies in History and Philosophy of Science Part A, 40*(1): 92–101. http://www.sciencedirect.com/science/article/pii/S0039368108001155.