PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

rsta.royalsocietypublishing.org

Research

CrossMark click for updates

Cite this article: Morrow DR. 2014 Ethical aspects of the mitigation obstruction argument against climate engineering research. *Phil. Trans. R. Soc. A* **372**: 20140062. http://dx.doi.org/10.1098/rsta.2014.0062

One contribution of 15 to a Theme Issue 'Climate engineering: exploring nuances and consequences of deliberately altering the Earth's energy budget'.

Subject Areas:
climatology

Keywords:
ethics, moral hazard, mitigation obstruction, geoengineering, climate engineering, solar radiation management

Author for correspondence:
David R. Morrow
e-mail: dmorrow2@gmu.edu

Ethical aspects of the mitigation obstruction argument against climate engineering research

David R. Morrow

Department of Philosophy, University of Alabama at Birmingham, Birmingham, AL, USA

Many commentators fear that climate engineering research might lead policy-makers to reduce mitigation efforts. Most of the literature on this so-called 'moral hazard' problem focuses on the prediction that climate engineering research would reduce mitigation efforts. This paper focuses on a related ethical question: Why would it be a bad thing if climate engineering research obstructed mitigation? If climate engineering promises to be effective enough, it might justify some reduction in mitigation. Climate policy portfolios involving sufficiently large or poorly planned reductions in mitigation, however, could lead to an outcome that would be worse than the portfolio that would be chosen in the absence of further climate engineering research. This paper applies three ethical perspectives to describe the kinds of portfolios that would be worse than that 'baseline portfolio'. The literature on climate engineering identifies various mechanisms that might cause policy-makers to choose these inferior portfolios, but it is difficult to know in advance whether the existence of these mechanisms means that climate engineering research really would lead to a worse outcome. In the light of that uncertainty, a precautionary approach suggests that researchers should take measures to reduce the risk of mitigation obstruction. Several such measures are suggested.

1. Introduction

Safer race cars crash more often. Facing less danger if they crash, drivers take more risks. Compounded over many drivers and races, these extra risks lead to more frequent crashes [1]. The extra crashes partially

Royal Society Publishing

offset the gains from safety innovations. The drivers' behaviour exemplifies risk compensation, in which people respond to risk-reducing innovations by behaving less cautiously. Researchers have also claimed to detect the phenomenon in connection with ordinary drivers [2], public health measures to fight HIV [3], hurricane forecasts [4], government-financed rescue teams for mountaineers [5] and other contexts. Risk compensation can also endanger people besides the risk-takers, as when drivers injure pedestrians [2] or when rescuers die searching for stranded mountaineers. Thus, while rational from the risk-takers' perspective, risk compensation can reduce the gains from safety innovations and redistribute risks from risk-takers to bystanders.

Many commentators have worried that climate engineering research could have a similar effect by inducing society to slow its investment in mitigation [6–21]. The basic idea is as follows: greenhouse gas (GHG) emissions increase climate risk. Emissions abatement is a costly way to reduce that risk. Climate engineering might reduce climate risk more cheaply. Thus, developing climate engineering might make society willing to emit more GHGs—that is, to mitigate less—than it otherwise would. Following [7], commentators often express this concern as a 'moral hazard' problem. David Keith now argues that 'risk compensation' describes the phenomenon more accurately [22], since the effect depends not on voluntary risk-sharing between a risk-taker and an insurer (as moral hazard does), but by a technology-driven reduction in risk to the risk-takers. More generically, the effect is sometimes described as 'mitigation obstruction' [23].

Many commentators treat the possibility of mitigation obstruction as a reason to avoid climate engineering research [23]. Roughly, their argument is as follows: researching climate engineering would induce policy-makers to mitigate less than they otherwise would. The resulting outcome would be worse than the outcome that would result if there were no further serious study of climate engineering. Therefore, researching climate engineering would lead to a worse outcome than would result without that research. In order to avoid this worse outcome, the argument concludes, we ought to forgo further climate engineering research.

We can restate the argument more formally in terms of climate policy portfolios—that is, policies combining mitigation, adaptation and climate engineering [15]. We can represent a portfolio as an ordered quadruple, (m, a, c, s), denoting levels of mitigation, adaptation, carbon dioxide removal (CDR) and solar radiation management (SRM), respectively. Let (m_0, a_0, c_0, s_0) represent whatever portfolio policy-makers would choose in the absence of further climate engineering research. Call this the 'baseline portfolio'. Let $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ represent whatever portfolio policy-makers would choose if climate engineering research were to proceed in earnest.[1] The concern about mitigation obstruction is that climate engineering research would induce policy-makers to choose $m_{ce} < m_0$, and as a result, $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ would yield a worse outcome than would the baseline portfolio (m_0, a_0, c_0, s_0).

The argument from mitigation obstruction involves two very different claims. The first is a prediction that $m_{ce} < m_0$. The second is a normative claim about the value of the resulting outcome—namely, that $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ would yield a worse outcome than (m_0, a_0, c_0, s_0). Assessing the argument requires assessing both the prediction and the normative claim.

Most discussion in the literature has so far focused on the prediction that $m_{ce} < m_0$. Some commentators doubt that climate engineering research will cause significant reductions in mitigation [24–26]. Others describe mechanisms by which climate engineering research could obstruct mitigation [9,18,19,23]. This part of the literature addresses the as yet unsettled question, 'Would climate engineering research actually obstruct mitigation?'

The literature contains little discussion of the argument's normative premise. That is, discussions of the argument from mitigation obstruction have little to say about exactly why it would be bad if climate engineering research obstructed mitigation. Unless we pinpoint the ethical problems with reducing mitigation, it is unclear how to understand the argument from mitigation obstruction or how seriously to take it [27].

[1]There is significant uncertainty about each component of each portfolio, except c_0 and s_0, which are presumably zero. Furthermore, the components of the climate engineering portfolio depend heavily on how climate engineering research proceeds. While important, these uncertainties will not affect the presentation of argument.

This paper addresses the question, 'What, exactly, would be wrong with obstructing mitigation?' The goal is to clarify the ethical aspects of the argument from mitigation obstruction. The paper considers several possible scenarios and identifies the key ethical assumptions on which those scenarios would be worse than whatever would happen without further climate engineering research. The analysis in this paper undermines the view that reducing mitigation *necessarily* leads to a worse outcome, but it supports the view that mitigation obstruction is still a serious risk. A precautionary approach justifies measures to reduce that risk. The penultimate section suggests specific precautionary measures that researchers could adopt.

2. Why some reduction in mitigation might be warranted

One natural but mistaken view about mitigation obstruction is that *any* reduction in mitigation would *necessarily* lead to a morally worse outcome. That is, one might think that if $m_{ce} < m_0$, then any portfolio involving m_{ce} would yield a worse outcome than one involving m_0. If climate engineering proves effective at reducing overall climate risk, however, this may not be true. This is not to say that climate engineering *would* justify reducing mitigation, only that it might.

To see why climate engineering research might justify some reduction in mitigation, consider that climate policy involves difficult trade-offs between morally important benefits and costs. The optimal climate policy portfolio would balance the morally important marginal social benefits from additional GHG emissions against the morally important marginal social costs of those emissions.[2] The most important benefits of GHG-emitting activities include human development for the global poor, though many lesser benefits to others are also morally important. The marginal costs come in the form of increased climate risk. Climate engineering, if it works and is used wisely, would reduce the climate risk from—and therefore the marginal social cost of—additional emissions. This would mean that the morally optimal level of mitigation, in the presence of effective, wisely used climate engineering, would be lower than the optimal level of mitigation in the absence of climate engineering. That is, the fact that $m_{ce} < m_0$ does not necessarily mean that $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ yields a worse outcome than would (m_0, a_0, c_0, s_0). Thus, if research led us to expect climate engineering could be used wisely and effectively, we might be justified in reducing mitigation efforts.

The wisdom of reducing mitigation depends partly on how much of the reduced climate risk comes from CDR and how much from SRM. If the relative benefits of $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ depend heavily on SRM, then two further ethical assumptions would be necessary to justify reducing mitigation. The most important assumption is that deploying SRM is not intrinsically wrong. If SRM is intrinsically wrong, then it would probably be wrong to deploy it for the purpose of reducing mitigation and reaping the benefits of additional emissions. The other necessary assumption is that, at least in the presence of climate engineering, the marginal social benefit of further GHG emissions is greater than their marginal social cost. The first assumption—that SRM is not intrinsically wrong—is contested among climate ethicists. The second assumption is more secure, as it rests on the relatively uncontroversial idea that the benefits of some additional emissions, such as 'subsistence emissions' by the global poor [28], are morally more important than their social costs.

This abstract discussion does *not* show that climate engineering research would actually justify any reduction in mitigation. Rather, the point is that the argument from mitigation obstruction cannot simply rest on the prediction that climate engineering research would lead to *some* mitigation obstruction. Showing that the argument does or does not work requires both predictions about climate policy and ethical assumptions about what counts as making the world worse off.

[2] There are, of course, tremendously difficult ethical and technical questions about the morally appropriate way to calculate marginal social benefit and marginal social cost for a global problem that will span many generations. We need not settle those questions to assert that the morally best policy balances morally important marginal benefits against morally important marginal costs.

3. In what sense might mitigation obstruction make things worse?

This section distinguishes three kinds of outcomes and identifies the ethical assumptions needed to describe each outcome as worse than whatever baseline outcome would result without further climate engineering research.

(a) Simple utilitarianism

From a simple utilitarian perspective, climate engineering research would cause a worse outcome if it induced a climate policy portfolio that yielded lower aggregate expected welfare than the baseline portfolio. Formally, simple utilitarianism compares outcomes based on an unweighted, additive social welfare function. Thus, to say that (m_{ce}, a_{ce}, c_{ce}, s_{ce}) yields a worse outcome than (m_0, a_0, c_0, s_0) from a simple utilitarian perspective is just to say that

$$\sum_{i \in S} Eu_i(m_{ce}, a_{ce}, c_{ce}, e_{ce}) < \sum_{i \in S} Eu_i(m_0, a_0, c_0, e_0), \tag{3.1}$$

where S denotes society and i denotes an arbitrary individual in society. That is, (m_{ce}, a_{ce}, c_{ce}, s_{ce}) is worse than (m_0, a_0, c_0, s_0) if and only if adding together the expected utility for each individual i in society under (m_{ce}, a_{ce}, c_{ce}, s_{ce}) yields a lower aggregate utility than adding together the expected utility for each individual i in society under (m_0, a_0, c_0, s_0).

This inequality would be satisfied if (m_{ce}, a_{ce}, c_{ce}, s_{ce}) imposed harms on some people that outweighed the benefits it provided to others, with harms and benefits measured relative to the baseline outcome. This could be true if the harms were concentrated but large while the benefits were diffuse but minor or vice versa.

Embracing simple utilitarianism requires assuming that the only thing that matters, morally speaking, is human (and perhaps animal) welfare. This contradicts a number of common ethical beliefs. For instance, it contradicts the idea that fairness matters for its own sake, that people have rights, and that intention affects the moral value of an action. Simple utilitarianism also seems to permit—or even require—actions that seem morally repugnant, such as killing socially useless people to harvest their organs for the benefit of socially useful people. Thus, despite its initial attractiveness, simple utilitarianism is a deeply controversial view.

(b) Fairness-adjusted utilitarianism

Supplementing simple utilitarianism with the assumption that fairness matters yields a view that we might call 'fairness-adjusted utilitarianism'. From a fairness-adjusted utilitarian perspective, climate engineering research would cause a worse outcome if it induced a portfolio that yields a lower *or less fairly distributed* aggregate expected welfare. Fairness-adjusted utilitarianism ranks outcomes according to the aggregate social welfare, after adjusting for the fairness of the distribution of welfare [29].

Inequality-adjusted social welfare functions, such as the inequality-weighted Human Development Index [30], provide one way to implement fairness-adjusted utilitarianism. These functions generally have a form like $f(q, w) = (1 - q)w$, where w is the sum total of welfare and q is a measurement of inequality in the distribution of welfare, with $q = 0$ under perfect equality and $q = 1$ under perfect inequality. Thus, when used to compare two outcomes with equal aggregate welfare, such a function ranks the one with a more equal distribution of welfare above the other.

Even if fairness is not fundamentally about equality of welfare, inequality-adjusted welfare functions provide a decent descriptive model of fairness-adjusted utilitarianism, especially for the purposes of evaluating climate policies. (A 'descriptive model' of a moral theory is a decision-theoretic model that ranks choices in the same way as the moral theory would, even if the reasons for ranking choices differ from those articulated in the moral theory [31].) Adjusting for inequality reduces simple utilitarianism's tendency to justify the gross mistreatment of some people for the benefit of others, while retaining the utilitarian position that sufficiently large benefits for some can justify harms to others. Furthermore, in the case of climate policy, the equality of an outcome turns out to correlate with the degree to which different people get what they deserve, almost regardless of one's theory of desert: since the rich are contributing the most to climate change while the poor are suffering the most, more equal outcomes tend to shift the burdens of climate change towards those who are causing the problem and are best able to bear it. Thus, it makes sense to use inequality-adjusted welfare functions to model a fairness-adjusted utilitarian perspective on mitigation obstruction.

To say that $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ yields a worse outcome than (m_0, a_0, c_0, s_0) from a fairness-adjusted utilitarian perspective is just to say that

$$(1 - q_{ce}) \sum_{i \in S} Eu_i(m_{ce}, a_{ce}, c_{ce}, e_{ce}) < (1 - q_0) \sum_{i \in S} Eu_i(m_0, a_0, c_0, e_0), \tag{3.2}$$

where q_{ce} and q_0 measure the inequality of welfare under the two portfolios. That is, $(m_{ce}$, a_{ce}, c_{ce}, $s_{ce})$ is worse than $(m_0$, a_0, c_0, $s_0)$ if and only if the sum of expected utilities under the former, multiplied by a measure of inequality under that policy, is less than the sum of expected utilities under the latter, multiplied by a measure of inequality under that policy. Inequality (3.2) would be satisfied if $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ either decreased aggregate utility (relative to the baseline) without sufficiently improving the fairness of its distribution or aggravated existing unfairness in the distribution of utility without sufficiently increasing aggregate utility.

(c) Non-consequentialism

Non-consequentialism is a broad family of moral theories on which the moral value of an action does not depend solely on the consequences of performing it. For instance, most theories that stress human rights are non-consequentialist. If people have a right to freedom of religion, for instance, then it is wrong for governments to prohibit citizens from practicing some religion, even if doing so would improve social welfare. Human rights are therefore one species of 'moral constraint'—that is, constraints on what someone may do to bring about better consequences.

From a non-consequentialist perspective, climate engineering research could induce an *ex ante* worse outcome if it created too great a risk that someone would deploy climate engineering in violation of one or more 'moral constraints'. Ethicists have identified many possible moral constraints on climate engineering. These include violations of distributive, intergenerational or procedural justice [32]; a presumption against interfering with nature in certain ways [33]; the wrongness of exacerbating the relationship of domination between generations [34] and the stricter constraint on doing harm rather than allowing harm [35]. These constraints apply primarily to SRM but may also apply to specific forms of CDR. From some non-consequentialist perspectives, these constraints may be so strict that climate engineering—or, at least, SRM—is obviously wrong. But many non-consequentialist theories include 'catastrophe clauses', which permit the violation of moral constraints to prevent catastrophes. Such catastrophe clauses leave the door open for an important argument for climate engineering research.

The argument that we should research climate engineering now in case future generations need it to ward off a climate catastrophe is called the 'arm the future argument' [14]. Very roughly, the argument is that we should research climate engineering because humanity may someday face a climate catastrophe so massive that climate engineering is the 'lesser evil', and only by researching climate engineering in advance can society prepare itself for such a situation. But to call climate engineering the 'lesser evil' in such circumstances is to appeal to the catastrophe clause in non-consequentialist theories; if it were not an 'evil' of some kind—that is, a violation of some moral constraint—climate engineering would simply be the 'least bad' option in those circumstances.

The connection between non-consequentialism and the argument from mitigation obstruction comes out most clearly in Stephen Gardiner's criticisms of the 'arm the future argument'. Gardiner notes that the probability of a climate catastrophe depends partly on how aggressively humanity reduces its emissions. Thus, to the extent that climate engineering research obstructs mitigation, it increases the probability of such a catastrophe [14]. It also, therefore, increases the probability of violating any moral constraints on climate engineering. It may seem that enhancing humanity's ability to protect itself against such a catastrophe, should it occur, compensates for increasing the probability of a catastrophe. But from a non-consequentialist perspective, a world in which humanity protects itself from a catastrophe by violating some moral constraint is worse than one with a similar level and distribution of welfare in which humanity does not violate that moral constraint. Thus, even if climate engineering could completely protect us from any increased climate risk resulting from mitigation obstruction, the increased risk of violating moral constraints is its own sort of 'moral cost'.

One way to weigh the 'moral cost' of increasing the risk of wrongdoing against the potential benefits of climate engineering research is to model non-consequentialist theories in a decision-theoretic way. This is difficult and controversial but not impossible [31,36]. Any decision-theoretic representation of these concerns is necessarily a descriptive model; it aims only to identify the judgements that non-consequentialists would make about an action, not to replicate the reasoning behind those judgements. One approach to modelling non-consequentialist concerns, similar to the approach in [31], is to use a 'utility function' that yields lower utilities for outcomes that are brought about by the violation of moral constraints than for similar outcomes that are brought about in morally neutral ways. (The term 'utility function' is just a convention here. We need not suppose that it is measuring 'utility' in the economist's sense.) For instance, in calculating the value of a given portfolio, we might identify the possible outcomes from that portfolio in which society violates some moral constraint by deploying SRM. For those outcomes, we could calculate the benefits and risks from SRM separately from those derived from other parts of the portfolio, and then multiply the benefits by some discount factor $0 < d < 1$ and multiply the risks by some penalty term $p > 1$. Given such adjustments, the greater the chance of such deployment under a particular portfolio, the lower the value such an analysis would assign to that portfolio. Taking this approach, $(m_{ce}, a_{ce}, c_{ce}, s_{ce})$ might yield a greater 'expected utility' than (m_0, a_0, c_0, s_0), before adjusting for the risk of violating moral constraints, and yet yield a lesser 'expected utility' once those constraints were accounted for.

Besides assuming that there are moral constraints on (certain kinds of) climate engineering, one must assume that those constraints apply even in the absence of better options. That is, one must hold that people sometimes find themselves in genuine moral dilemmas—that is, in situations in which each of their options requires acting wrongly. In the absence of this assumption, one's least bad option is always morally permissible, in which case there is no reason to discount the benefits derived from one's least bad option. While controversial, the assumption that there are genuine moral dilemmas is common among contemporary non-consequentialist ethicists [37] and explicitly endorsed by some climate ethicists in the context of climate engineering [14].

4. How might climate engineering research induce excessive reductions in emissions abatement?

The previous two sections argue that while reducing mitigation does not necessarily lead to a worse outcome, it is theoretically possible for mitigation obstruction to make things worse than it would be under the 'baseline portfolio'. Since the baseline portfolio, consisting only of mitigation and adaptation, would remain available to society even if it pursues climate engineering research, choosing an inferior portfolio involving climate engineering would be irrational. Critics of the argument from mitigation obstruction might doubt that society would choose inferior portfolios. Thus, some might argue, simply showing that portfolios involving climate engineering *could* be worse than the baseline portfolio does not show that mitigation obstruction is a serious risk, since the fact that they would be worse suggests that policy-makers would not choose them. The climate engineering literature, however, identifies several mechanisms that cause society to behave irrationally—that is, to choose a suboptimal policy portfolio. We can divide the mechanisms into three kinds: informational failures, cognitive failures and ethical failures.

Economic models of risk-taking typically assume that decision-makers have full information about the risks they run, meaning that they know the probability of each outcome and the magnitude of gain or loss in each outcome. This assumption does not hold in the case of climate engineering. For various reasons, policy-makers may remain ignorant or even develop false beliefs about the promise, efficacy and risk of climate engineering technologies. For instance, they may come to see SRM as a 'silver bullet' that can 'solve' the problem of climate change cheaply and easily, which could lead them to rely too heavily on SRM and invest too little in mitigation; such a portfolio might be worse than the baseline portfolio from any of the ethical perspectives discussed in the previous section. On the other hand, policy-makers may come to see SRM as an 'all-or-nothing' affair, in the sense that the only options involve offsetting all anthropogenic warming or none. Policy-makers would then overlook more moderate options, such as those proposed in [38] or [22], which may be the only way for a portfolio that includes climate engineering to outperform the baseline portfolio. Alternatively, policy-makers may conflate SRM and CDR in problematic ways. Such informational failures could lead to choosing irrational portfolios, even if policy-makers reasoned correctly from their false beliefs.

Even worse, the hope that policy-makers would reason correctly dims when we consider the range of cognitive biases that affect decisions about risk. A 'cognitive bias' is an unconscious psychological mechanism that tends to skew judgements in some domain (e.g. about risk), even when starting from accurate information. Albert Lin identifies a number of well-known biases that might 'foster unduly favourable perceptions of specific geoengineering options' [19]. Optimism bias and overconfidence bias could cause policy-makers to exaggerate the efficacy and underestimate the downside risks of climate engineering [19]. Lin mentions the importance of hyperbolic discounting, which causes an unduly large discounting of future costs and benefits [19]. Hyperbolic discounting may lead to an overly strong preference for SRM, with its immediate benefits and low short-term costs, over the delayed benefits and high short-term costs of mitigation. Lin suggests that the 'affect heuristic', by which people judge risks based on how they feel, rather than on objective analysis, might cause unduly rosy views of climate engineering in people who gain a comforting sense of control from climate engineering [19]. As other commentators have suggested, though, climate engineering evokes strongly negative feelings in many environmentalists [22] and political egalitarians [26]. Thus, the affect heuristic may generate unduly gloomy views of climate engineering in certain people. Such cognitive failures could cause even well-informed policy-makers to choose climate engineering-related portfolios that are worse than the baseline portfolio.

Another possible cognitive failure depends the 'cultural cognition thesis' about risk perception. This is the hypothesis that individuals process information about risks differently depending on how those risks interact with their value judgements about how society should be organized [26]. For example, if acknowledging anthropogenic climate risk threatens an individual's value judgements about government regulation, that individual will be more likely to reject evidence of climate change. This is understandable: if accepting the evidence for climate change means rejecting the values that his or her community prizes, an individual might suffer more from accepting the evidence than from rejecting it [26]. Such motivated reasoning, however, can lead to over- or underestimating the risks facing society. As applied to climate engineering, the fear is that because some people value humanity's technological mastery of environmental challenges, those people will downplay some risks of climate engineering and thus rely on it more heavily than they should [19].

Another kind of cognitive failure arises from what Gardiner calls 'moral corruption'. Though Gardiner does not explain it this way, moral corruption is a mixture of wishful thinking and confirmation bias: it is the unconscious tendency to concoct, seek out or more readily accept arguments that absolve us of some responsibility that we wish to escape [39]. Just as some scientists might feel more inclined to accept arguments that support their desired interpretation of the data, people might feel more inclined to accept arguments that would justify their avoidance of onerous responsibilities. But just as being a good scientist means resisting that inclination, so too does being a virtuous moral agent. Gardiner suspects that policy-makers and parts of the public have succumbed to moral corruption by accepting specious arguments against climate action [39]. By making people feel as if humanity has an 'easy way out', climate engineering research provides another tempting argument for irrationally low levels of mitigation [14]. Furthermore, those who stand to benefit from SRM might succumb to specious arguments that they are not responsible for the risks that it imposes on others, leading them to deploy it more readily than they ought.

There is a third mechanism by which policy-makers might come to rely too heavily on climate engineering: politicians have, on occasion, acted unethically to promote their own careers. In general, the more heavily they plan for future generations to rely on climate engineering, the fewer costs politicians must bear in order to be seen as taking action on climate change. Thus, some politicians might prefer to rely too heavily on climate engineering, even if they fully understood that doing so was unethical. If the public is willing to accept such policies, then self-serving politicians can happily 'pass the buck' on to their successors. Since much of the public has so far accepted anemic mitigation policies, they may accept a climate policy that relies heavily also on climate engineering. Thus, even if everyone had full information about climate engineering and reasoned clearly about how they ought to respond to that information, ethical failures may lead society to choose an unjust path.

Taken together, these three kinds of potential failure—informational, cognitive and ethical—create a serious risk that climate engineering research would induce policy-makers to choose a portfolio that yields a worse outcome than the baseline portfolio would. For society to respond correctly to a climate engineering option, they would need to understand the risks and potential benefits of different policy portfolios, process information about those risks and benefits rationally, resist moral corruption and avoid being misled by unethical politicians. In short, it is not enough that climate engineering technologies work well. Society would also need to use them well, both in deploying them (if it ever comes to that) and in planning for their possible use.

Although the mechanisms described in this section make it plausible that policy-makers will choose an irrational portfolio, that does not entail that climate engineering research would lead them to choose one that is worse than the baseline portfolio. If the net benefits of climate engineering prove great enough, there might be some room for error in policy-makers' choice of portfolio. Furthermore, since policy-makers seem not to be choosing the best portfolio in the absence of climate engineering, even a far-from-optimal portfolio involving climate engineering might be better than the baseline portfolio. Thus, even if this section justifies doubts that policy-makers will choose the best portfolio involving climate engineering, that is not enough to show that climate engineering research would induce policy-makers to choose a worse portfolio than they would have chosen in the absence of that research.

5. Ethical decision-making under deep uncertainty

The preceding sections reveal deep uncertainty about the ethics of mitigation obstruction: limited reductions in mitigation might be unobjectionable—even desirable—as part of the right climate policy portfolio. But other portfolios could be worse than the baseline portfolio. And though we have good reason to doubt that policy-makers would choose optimal climate engineering policies, we cannot know in advance whether policy-makers would err so badly as to make things worse than they would be under the baseline portfolio—not least of all because we do not even know what the baseline portfolio would be. This uncertainty is 'deep' (or 'Knightian') because we cannot assign probabilities to the various outcomes. Climate engineering research, therefore, is a moral gamble. Doing the research might lead to better climate policy choices, but it might lead to worse choices, and we cannot know which in advance. How should researchers in particular respond to this dilemma?

One response would be to try to get a better grip on the likelihood of various outcomes. Game theoretic models can help identify likely deployment scenarios for climate engineering [40], but estimating the probabilities of various climate policy portfolios would be extremely challenging. Further climate modelling can better quantify the downside risks associated with SRM, placing upper bounds on the risks that climate engineering might create. For now, however, we must consider approaches to decision-making under deep uncertainty.

Precautionary principles provide the most salient such approach, especially in environmental contexts. International treaties, national laws and academics frequently recommend 'the precautionary principle' or a 'precautionary approach' in responding to environmental risks; but these sources express 'the' principle in different and often incompatible ways [41–43]. Although precautionary principles vary along several dimensions [41], they generally endorse some form of the 'Requirement of Precaution', which requires that '[a]ctivities [that] may bring great harm should not be ... undertaken unless they have been shown not to impose too serious risks' [44, p. 11]. Plausible precautionary principles will ignore risks below some *de minimis* threshold, however, lest they paralyse decision-making [44]. Thus, in what follows, we must assume that, given further climate engineering research, the risk of policy-makers choosing a policy portfolio worse than the baseline portfolio is large enough to trigger some plausible precautionary principle.

Philosophers and lawyers have reached mixed conclusions when applying precautionary approaches to climate engineering, depending on the specific precautionary principle invoked and the specific aspects or method(s) of climate engineering considered [45–48]. Precautionary principles generally support limited climate engineering research [45,47,48]. Existing analyses provide little insight, though, into the relevance of precautionary principles to mitigation obstruction.

In the light of previous sections, it seems probable that most precautionary principles would require precautions against mitigation obstruction. The key question is whether those precautions include a moratorium on climate engineering research. A moratorium would involve a risk–risk tradeoff: to research climate engineering is to run the risk of excessive mitigation obstruction. To impose a moratorium is to run the risk of high climate damages that could have been prevented through climate engineering [48]. Without knowing the magnitude of either risk, though, it seems we cannot know whether any plausible precautionary principles would require a moratorium.

We might turn to a precautionary principle that seems especially well-suited to cases of deep uncertainty: the maximin rule. The maximin rule says to choose the option whose worst possible outcome is no worse than that of any other option—that is, to maximize the minimum possible 'payoff'. While maximin is irrational as a general decision rule [49], some have argued for its applicability in specific circumstances, including those of climate change [50]. But even if it is applicable to climate change generally, it is not a suitable decision rule for climate engineering research because of a special feature of the present case. The worst possible outcome with climate engineering is the so-called 'double catastrophe' resulting from abrupt termination of SRM in a high-GHG world with high climate sensitivity [51]. Because it involves extremely rapid climatic change, it is worse than the worst possible outcome without climate engineering—namely, catastrophic climate change resulting from high emissions in the context of high climate sensitivity. A precautionary moratorium on climate engineering research would (presumably) avoid the double catastrophe. Maximin, therefore, would prescribe a moratorium.

The truly catastrophic nature of each worst-case scenario, however, makes maximin problematic for two reasons. First, even its defenders claim that maximin is only justified when, among other things, decision-makers 'care relatively little for potential gains that might be made above the minimum that can be guaranteed by the maximin approach' [50, p. 47]. The worst possible outcome achievable by forgoing climate engineering research, however, is still catastrophic. We care deeply about gains above that minimum. So, at least one of the special conditions needed to justify maximin does not apply. Relatedly, the worst outcomes of each choice are so horrendous that it seems reasonable to consider their relative likelihoods. The 'double catastrophe' requires (more or less) all of the conditions that the 'single catastrophe' does, *plus* the unlikely condition that no part of global society can rebuild the capacity for SRM within a decade or so of SRM's initial termination. Whatever the probability of the single catastrophe, then, the probability of the double catastrophe seems lower. Therefore, it is arguably irrational to forgo climate engineering research simply because the double catastrophe would be worse than the single catastrophe.

This leaves us, then, with less demanding precautionary principles—principles that require *some* kind of precautions against objectionable mitigation obstruction, but not a moratorium. In the short term, the most important precautionary measures are those available to climate engineering researchers themselves.

6. What precautions can researchers take against mitigation obstruction?

The first of several things that researchers can do is to *research relevant alternatives* to ensure that policy-makers have information about their various options. This involves structuring research projects to include a relevant range of possibilities with respect to any given technology. The goal is to provide policy-makers with the information they need about relevant options. As an example, consider two recent studies that model the abrupt termination of SRM [52,53]. Both studies model termination in the context of the highest representative concentration pathway (RCP) from the Intergovernmental Panel of Climate Change, but only one models termination in the context of the more optimistic RCP 2.6 scenario [52]. The highest pathway, RCP 8.5, constitutes a rather pessimistic view of future emissions. Abrupt termination in that context would lead to large, rapid warming: although their methodologies differ slightly, each paper predicts warming of roughly 2°C over two decades if SRM were stopped suddenly after several decades of deployment [52,53]. In the context of RCP 2.6, however, abrupt termination results in less than half that rate of warming [52]. Thus, the study that focuses only on RCP 8.5 omits a crucial piece of information: in the context of aggressive mitigation, the abrupt termination of SRM poses a far less serious problem than it does under a high-emissions scenario. Conversely, a study that only considered RCP 2.6 would omit the even more important fact that abrupt termination would be catastrophic under a high-emissions scenario. Similarly, studies that focus only on a single method of climate engineering will omit information about alternative methods; studies that focus only one level of climate engineering (e.g. the level needed to offset all anthropogenic warming) will omit information about alternative levels (e.g. the level needed to partially offset additional warming) and so on. It is impossible, of course, to consider all alternatives in a single study. Large coordinated research projects, such as GeoMIP [54], are helping to fill this gap, and researchers can always highlight the limitations of their studies by explaining available alternatives and why those alternatives might give different results.

The second thing that researchers can do is to *be mindful of messaging*. That is, researchers can think carefully about how various audiences will understand and react not just to *what* information the researchers convey, but also to *how* they convey it. Researchers studying emerging technologies can educate themselves about the science of science communication [55,56], especially in its application to climate engineering [26,57–61]. With respect to mitigation obstruction, two specific points deserve emphasis: first, researchers can highlight the differences between various climate engineering technologies and stress both the different benefits and the limitations of each technology. For instance, reminding readers that SRM would not address ocean acidification would help prevent the impression that SRM provides a 'silver bullet' solution to climate change. Conversely, researchers can help prevent unduly pessimistic beliefs about climate engineering by emphasizing the potential benefits of SRM, rather than focusing exclusively on risks or on especially intense deployments. Second, researchers can avoid framing their conclusions in ways that are liable to be distorted, exaggerated or misinterpreted as they pass from research papers to public relations offices to journalists to editors to the public. Each step of that process involves people with their own biases and agendas. Easily misunderstood or sensationalized conclusions are especially likely to be picked up and broadcast widely, often to the detriment of public understanding.

The third thing that researchers can do is to *engage the public and policy-makers* [61]. This may take the form of testimony to policy-makers; delivering popular lectures or engaging in debates on university campuses; writing books, op-eds and so on for the mass market; or doing interviews with science journalists and the popular press. The advantages of such direct engagement include getting the opportunity to present one's work in a way that is more easily understood (and therefore less subject to misinterpretation) by the public and policy-makers; learning first-hand what the public understands about climate engineering and correcting misinformation and misinterpretations. Without such direct engagement, researchers effectively cede public discussion of climate engineering to activists and entrepreneurs whose views will strike many researchers as unduly pessimistic or optimistic. Not all researchers are comfortable with public engagement, of course, but various institutions offer courses and other opportunities to train scientists to communicate with the public.

Some researchers may fear that taking these steps blurs the line between science and advocacy, but this is not necessarily so. Researchers can pursue each of these options—even public engagement—with the goal of delivering accurate and easy-to-understand information, rather than advocating for or against particular uses of climate engineering. The purpose of taking these precautionary measures is, after all, not to induce society to choose a particular portfolio, but to reduce the chance that one's own research induces policy-makers to choose irrationally.

7. Conclusion

The so-called 'moral hazard' problem for climate engineering is that climate engineering research might obstruct mitigation in ways that lead to a worse outcome than would result without further climate engineering research. The notion of a 'worse outcome' can be interpreted in various ways: on a simple utilitarian interpretation, a climate policy portfolio involving climate engineering is worse than one that does not if and only if the former yields lower expected aggregate social welfare than the latter. On a fairness-adjusted utilitarian interpretation, the former portfolio is worse than the latter just in case it yields lower expected aggregate social welfare after adjusting for the fairness of its distribution. On a non-consequentialist interpretation, the former portfolio is worse than the latter if it runs too high a risk of forcing society to deploy climate engineering in morally objectionable ways. None of these interpretations imply that reducing mitigation *necessarily* yields a worse outcome; if climate engineering proves sufficiently cost-effective and is used wisely, it might justify some reduction in mitigation.

The fact that climate engineering research might justify some reductions in mitigation makes it much harder to know whether the argument from mitigation obstruction provides a morally compelling reason not to pursue climate engineering research. Various commentators have described plausible mechanisms by which climate engineering research might induce *some* reductions in mitigations. But it is very difficult to know whether those reductions would be so large or so poorly planned that they would lead to a worse outcome than would result without climate engineering research. While that uncertainty does not currently justify a moratorium on climate engineering research, it does provide a strong reason for climate engineering researchers to take precautions against objectionable mitigation obstruction. These precautions involve structuring research, communicating results, and engaging the public and policy-makers in ways that reduce the chance of irrational climate policy portfolios.

Acknowledgements. Thanks to Christopher Preston, Toby Svoboda and two reviewers for helpful comments on an earlier draft of this paper. All errors remain my own.

References

1. Sobel RS, Nesbit TM. 2007 Automobile safety regulation and the incentive to drive recklessly: evidence from NASCAR. *South. Econ. J.* **74**, 71–84.
2. Peltzman S. 1975 The effects of automobile safety regulation. *J. Polit. Econ.* **83**, 677–726. (doi:10.1086/260352)
3. Cassell MM, Halperin DT, Shelton JD, Stanton D. 2006 Risk compensation: the Achilles' heel of innovations in HIV prevention? *Br. Med. J* **332**, 605–607. (doi:10.1136/bmj.332.7541.605)

4. Sadowski NC, Sutter D. 2005 Hurricane fatalities and hurricane damages: are safer hurricanes more damaging? *South. Econ. J.* **72**, 422–432. (doi:10.2307/20062119)
5. Clark JR, Lee DR. 1997 Too safe to be safe: some implications of short- and long-run rescue Laffer curves. *East. Econ. J.* **23**, 127–137.
6. Schneider SH. 1996 Geoengineering: could—or should—we do it? *Clim. Chang.* **33**, 291–302. (doi:10.1007/BF00142577)
7. Keith DW. 2000 Geoengineering the climate: history and prospect. *Annu. Rev. Energy Environ.* **25**, 245–284. (doi:10.1146/annurev.energy.25.1.245)
8. Lawrence MG. 2006 The geoengineering dilemma: to speak or not to speak. *Clim. Change* **77**, 245–248. (doi:10.1007/s10584-006-9131-5)
9. Parson EA. 2006 Reflections on air capture: the political economy of active intervention in the global environment. *Clim. Change* **74**, 5–15. (doi:10.1007/s10584-005-9032-z)
10. Robock A. 2008 20 reasons why geoengineering may be a bad idea. *Bull. At. Sci.* **64**, 14–18. (doi:10.2968/064002006)
11. Shepherd JG *et al.* 2009 *Geoengineering the climate: science, governance and uncertainty*. London, UK: Royal Society.
12. Morrow DR, Kopp RE, Oppenheimer M. 2009 Toward ethical norms and institutions for climate engineering research. *Environ. Res. Lett.* **4**, 045106. (doi:10.1088/1748-9326/4/4/045106)
13. Virgoe J. 2008 International governance of a possible geoengineering intervention to combat climate change. *Clim. Change* **95**, 103–119. (doi:10.1007/s10584-008-9523-9)
14. Gardiner SM. 2010 Is 'arming the future' with geoengineering really the lesser evil? some doubts about the ethics of intentionally manipulating the climate system. In *Climate ethics* (eds SM Gardiner, S Caney, D Jamieson, H Shue), pp. 284–312. New York, NY: Oxford University Press.
15. Wagner G, Zeckhauser RJ. 2012 Climate policy: hard problem, soft thinking. *Clim. Chang.* **110**, 507–521. (doi:10.1007/s10584-011-0067-z)
16. Burns WCG. 2011 Climate geoengineering: solar radiation management and its implications for intergenerational equity. *Stanf. J. Law Sci. Policy.* **4**, 37–55.
17. Preston CJ. 2013 Ethics and geoengineering: reviewing the moral issues raised by solar radiation management and carbon dioxide removal. *Wiley Interdiscip. Rev. Clim. Chang.* **4**, 23–37. (doi:10.1002/wcc.198)
18. Hamilton C. 2013 *Earthmasters: the dawn of the age of climate engineering*. New Haven, CT: Yale University Press.
19. Lin AC. 2013 Does geoengineering present a moral hazard? *Ecol. Law Q.* **40**, 673–712.
20. Zürn M, Schäfer S. 2013 The paradox of climate engineering. *Glob. Policy* **4**, 1–12. (doi:10.1111/gpol.12004)
21. Rayner S, Heyward C, Kruger T, Pidgeon N, Redgwell C, Savulescu J. 2013 The Oxford principles. *Clim. Change* **121**, 499–512. (doi:10.1007/s10584-012-0675-2)
22. Keith D. 2013 *A case for climate engineering*. Cambridge, MA: MIT Press.
23. Betz G, Cacean S. 2012 *Ethical aspects of climate engineering*. Karlsruhe, Germany: Kalrsuher Institut für Technologie.
24. Bunzl M. 2009 Researching geoengineering: should not or could not? *Environ. Res. Lett.* **4**, 045104. (doi:10.1088/1748-9326/4/4/045104)
25. Goeschl T, Heyen D, Moreno-Cruz J. 2013 The intergenerational transfer of solar radiation management capabilities and atmospheric carbon stocks. *Environ. Resource Econ.* **56**, 85–104. (doi:10.1007/s10640-013-9647-x)
26. Kahan DM, Jenkins-Smith HC, Tarantola T, Silva CL, Braman D. In press. Geoengineering and climate change polarization: testing a two-channel model of science communication. *Ann. Am. Acad. Polit. Soc. Sci.* **658**.
27. Hale B. 2012 The world that would have been: moral hazard arguments against geoengineering. In *Engineering the climate: the ethics of solar radiation management* (ed. CJ Preston), pp. 113–132. Plymouth, UK: Lexington Books.
28. Shue H. 1993 Subsistence emissions and luxury emissions. *Law Policy* **15**, 39–60. (doi:10.1111/j.1467-9930.1993.tb00093.x)
29. Hooker B. 2000 *Ideal code, real world: a rule-consequentialist theory of morality*. Oxford, UK: Oxford University Press.
30. United Nations Development Programme. 2010 *Human development report 2010*. New York, NY: Palgrave Macmillan.

31. Colyvan M, Cox D, Steele K. 2010 Modelling the moral dimension of decisions. *Noûs* **44**, 503–529. (doi:10.1111/j.1468-0068.2010.00754.x)
32. Svoboda T, Keller K, Goes M, Tuana N. 2011 Sulfate aerosol geoengineering: the question of justice. *Public Aff. Q.* **25**, 1–42.
33. Preston CJ. 2011 Re-thinking the unthinkable: environmental ethics and the presumptive argument against geoengineering. *Environ. Values* **20**, 457–479. (doi:10.3197/096327111X13150367351212)
34. Smith PT. 2012 Domination and the ethics of solar radiation management. In *Engineering the climate: the ethics of solar radiation management* (ed. CJ Preston), pp. 43–61. Plymouth, UK: Lexington Books.
35. Morrow DR. 2014 Starting a flood to stop a fire: some moral constraints on solar radiation management. *Ethics Policy Environ.* **17**, 123–138. (doi:10.1080/21550085.2014.926056)
36. Portmore DW. 2007 Consequentializing moral theories. *Pacific Philos. Q.* **88**, 39–73. (doi:10.1111/j.1468-0114.2007.00280.x)
37. Sinnott-Armstrong W. 1998 *Moral dilemmas*. Oxford, UK: Oxford University Press.
38. Wigley TML. 2006 A combined mitigation/geoengineering approach to climate stabilization. *Science* **314**, 452–454. (doi:10.1126/science.1131728)
39. Gardiner SM. 2011 *A perfect moral storm: the ethical tragedy of climate change*. New York, NY: Oxford University Press.
40. Ricke KL, Moreno-Cruz JB, Caldeira K. 2013 Strategic incentives for climate geoengineering coalitions to exclude broad participation. *Environ. Res. Lett.* **8**, 014021. (doi:10.1088/1748-9326/8/1/014021)
41. Sandin P. 1999 Dimensions of the precautionary principle. *Hum. Ecol. Risk Assess.* **5**, 889–907. (doi:10.1080/10807039991289185)
42. Ahteensuu M, Sandin P. 2012 The precautionary principle. In *Handbook of risk theory* (eds S Roeser, R Hillerbrand, P Sandin, M Peterson), pp. 961–978. Dordrecht, The Netherlands: Springer.
43. Hartzell-Nichols L. 2013 From 'the' precautionary principle to precautionary principles. *Ethics Policy Environ.* **16**, 308–320. (doi:10.1080/21550085.2013.844569)
44. Munthe C. 2011 *The price of precaution and the ethics of risk*. Dordrecht, The Netherlands: Springer.
45. Elliott K. 2010 Geoengineering and the precautionary principle. *Int. J. Appl. Philos.* **24**, 237–253. (doi:10.5840/ijap201024221)
46. Tedsen E, Homann G. 2013 Implementing the precautionary principle for climate engineering in international law. *Carbon Clim. Law Rev.* **2**, 90–101.
47. Hartzell-Nichols L. 2012 Precaution and solar radiation management. *Ethics Policy Environ.* **15**, 158–171. (doi:10.1080/21550085.2012.685561)
48. Reynolds JL, Fleurke F. 2013 Climate engineering research: a precautionary response to climate change? *Carbon Clim. Law Rev.* **2**, 101–107.
49. Harsanyi JC. 1975 Can the maximin principle serve as a basis for morality? A critique of John Rawls's theory. *Am. Polit. Sci. Rev.* **69**, 594–606. (doi:10.2307/1959090)
50. Gardiner SM. 2006 A core precautionary principle. *J. Polit. Philos.* **14**, 33–60. (doi:10.1111/j.1467-9760.2006.00237.x)
51. Baum SD, Maher TM, Haqq-Misra J. 2013 Double catastrophe: intermittent stratospheric geoengineering induced by societal collapse. *Environ. Syst. Decis.* **33**, 168–180. (doi:10.1007/s10669-012-9429-y)
52. McCusker KE, Armour KC, Bitz CM, Battisti DS. 2014 Rapid and extensive warming following cessation of solar radiation management. *Environ. Res. Lett.* **9**, 024005. (doi:10.1088/1748-9326/9/2/024005)
53. Keller DP, Feng EY, Oschlies A. 2014 Potential climate engineering effectiveness and side effects during a high carbon dioxide-emission scenario. *Nat. Commun.* **5**, 1–11. (doi:10.1038/ncomms4304)
54. Kravitz B, Robock A, Boucher O, Schmidt H, Taylor KE, Stenchikov G, Schulz M. 2011 The geoengineering model intercomparison project (GeoMIP). *Atmos. Sci. Lett.* **12**, 162–167. (doi:10.1002/asl.316)
55. Nisbet MC, Scheufele DA. 2009 What's next for science communication? Promising directions and lingering distractions. *Am. J. Bot.* **96**, 1767–1778. (doi:10.3732/ajb.0900041)
56. Cheng D, Claessens M, Gascoigne T, Metcalfe J, Schiele B, Shi S. 2008 *Communicating science in social contexts: new models, new practices*. Dordrecht, The Netherlands: Springer.

57. Pidgeon N, Corner A, Parkhill K, Spence A, Butler C, Poortinga W. 2012 Exploring early public responses to geoengineering. *Phil. Trans. R. Soc. A* **370**, 4176–4196. (doi:10.1098/rsta.2012.0099)
58. Corner A, Parkhill K, Pidgeon N, Vaughan NE. 2013 Messing with nature? Exploring public perceptions of geoengineering in the UK. *Glob. Environ. Chang.* **23**, 938–947. (doi:10.1016/j.gloenvcha.2013.06.002)
59. Pidgeon N, Parkhill K, Corner A, Vaughan N. 2013 Deliberating stratospheric aerosols for climate geoengineering and the SPICE project. *Nat. Clim. Chang.* **3**, 451–457. (doi:10.1038/nclimate1807)
60. Mercer AM, Keith DW, Sharp JD. 2011 Public understanding of solar radiation management. *Environ. Res. Lett.* **6**, 044006. (doi:10.1088/1748-9326/6/4/044006)
61. Carr WA, Preston CJ, Yung L, Szerszynski B, Keith DW, Mercer AM. 2013 Public engagement on solar radiation management and why it needs to happen now. *Clim. Chang.* **121**, 567–577. (doi:10.1007/s10584-013-0763-y)