

Embracing uncertainty in climate change policy

Friederike E.L. Otto^{1*}, David J. Frame², Alexander Otto¹, and Myles R. Allen^{1,3}

**friederike.otto@ouce.ox.ac.uk*

1 Environmental Change Institute, University of Oxford, Oxford, UK

2 NZ Climate Change Research Institute, Victoria University of Wellington, Wellington, NZ

3 Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, UK

The “pledge and review” approach to reducing greenhouse gas emissions presents an opportunity to link mitigation goals explicitly to the evolving climate response. This seems desirable because the progression from the IPCC 4th to 5th Assessments has seen little reduction in uncertainty. A common reaction to persistent uncertainties is to advocate mitigation policies that are robust even under worst-case scenarios, thereby focusing attention on upper extremes of both the climate response and the costs of impacts and mitigation, all of which are highly contestable. Here we ask whether climate policies can learn from “adaptive management” techniques. Recognising that long-lived greenhouse gas emissions have to be net zero by the time temperatures reach a target stabilisation level such as 2°C, and anchoring commitments to an agreed index of attributable anthropogenic warming, would provide a transparent approach to meeting such a temperature goal without prior consensus on the climate response.

Introduction

The primary reasons for slow progress in global mitigation policy are not scientific. They are strategic – economic and political barriers to action arising from weak incentives to mitigate and strong incentives to free-ride on the efforts of others, internationally and inter-generationally (1, 2, 3). To be successful, a climate change mitigation policy does not only have to overcome those economic and political barriers, but has to withstand and adapt to other external pressures that originate from shifts

in the economy (e.g. “austerity”), and political interests (e.g. “climate scepticism”). Attempts have been made to design policies that are more robust to these external pressures, e.g. by attempting to find ways for regulators to “credibly commit” both themselves and their successors in an environment of changing power structures (e.g., 4), by locking-in certain policies through institutional design (5), by capitalising on emergent government structures (e.g. 6) and self-reinforcing effects of certain policies (7). Connecting these lines of thought to those of adaptive management and the governance of complex systems (8), here we argue for a redesign of climate change mitigation policies to be “anti-fragile” (9) with respect to scientific uncertainty.

Anti-fragile means that uncertainty and changes in scientific knowledge make the policy more successful by allowing for trial and error at low societal costs. Hence anti-fragile re-design allows for incorporating a wider range of risks of concern to policy makers, potentially allowing more successful mitigation policies. Arguably, a key pre-requisite for an anti-fragile climate policy is an index not beholden to high scientific uncertainty. Here we suggest the “attributable anthropogenic warming” as an anti-fragile index against which pledges could be reviewed, independent of the details of individual countries’ mitigation policies.

Precautionary Mitigation Policies

The predominant approach to the design of climate mitigation policies refers to the precautionary principle, embedded in the UNFCCC (10, Article 2). As any climate policy has the joint goals of enabling continued human development while at the same time staying within the boundaries posed by the limitations of the climate system, the trade-off between these goals has to be struck. Two main methods for implementing the trade off are cost-benefit approaches and cost-effectiveness approaches. In the former, estimates of the costs and benefits of mitigation are evaluated, usually through the use of Integrated Assessment Models (IAMs). The target is set where the marginal costs of climate damages are equal to the marginal costs of climate mitigation. In the cost-effectiveness approach, the target is set by capping damages, with subsequent emphasis on finding the least-cost way of achieving the target. Damages themselves are uncertain functions of global mean surface

temperatures, but for simplicity we follow the usual practice of assuming a smooth and convex relationship between them. Discovering damages are a more convex function of temperature change than is usually supposed would simply reinforce the need for “anti-fragile” mitigation policies – mitigation paths which look attractive under moderate mitigation scenarios may turn out to be impossible to scale up if mitigation becomes more urgent.

In many settings, cost-benefit approaches might be expected to predominate, since this is a very familiar and effective policy approach. In the case of climate change, significant caveats arise from the incompleteness of IAMs: it is often claimed that these models significantly underestimate climate change damages (11). If this is so, and if there is no corresponding underestimate in the costs of mitigation, it introduces a significant bias to any cost-benefit analysis of climate change.

Partly for this reason, many have argued instead for a cost-effectiveness approach (12). In the context of climate change by far the most-discussed structure pertaining to the cost-effectiveness approach is the 2°C goal adopted by the United Nations Framework Convention on Climate Change (UNFCCC) in Cancun (13). A substantial body of research into how we might achieve successful management of the climate change problem has focussed on meeting the agreed 2°C target – an approach that is especially common among physical climate scientists researching the problem. One of the most cited publications in recent years in this field, Meinshausen et al. (14), updated by (15), provides limits to cumulative carbon emissions up to 2050 with corresponding limits to the probability of exceeding the 2°C target. Their analysis is based on the budget approach (16) considering an emittable CO₂ budget till 2050 under which “only a small amount of CO₂ may be emitted worldwide after 2050” (16, SPM). Their emittable budget is derived by limiting the probability of exceeding a 2°C target to levels at or below 25%. The UNFCCC has not formally adopted a position on the allowable probability of temperatures exceeding 2°C, although the focus on scenarios that are “likely” to achieve this goal (17, 18) might be interpreted as signalling a preference for a two-in-three odds or better of success.

The problem is that a policy that might be criticized for offering “only” a 3-in-4 or 2-in-3 odds of success in limiting impacts could also be criticized for offering similar odds of imposing excessive

mitigation costs. This focus on reducing the risk of unfavourable outcomes represents one possible interpretation of the precautionary principle, prominently employed by Weitzman (19). In this interpretation of precaution, actions should not only be taken to avoid an expected unfavourable outcome, but also to protect against unlikely but high-impact possibilities. Governments and other stakeholders care not only about the environmental costs of inaction, but also about the economic costs associated with (perhaps unnecessarily) aggressive action on climate change. Any approach that considers only environmental risks cannot capture the full set of risks to which governments may be averse, and the failure to incorporate these risks, and thus their fragility towards the consequences of overestimation of environmental costs, seems to be a major reason why stringent mitigation policies have not been embraced by governments, in spite of their willingness to sign up to a globally stringent temperature target.

Interpreting “precautionary” thinking in this way leads to two problems: firstly minimising the risk of high damages requires a huge and immediate mitigation effort that is too demanding of communities with multiple priorities. The stringency of the mitigation policies that appear to be needed to provide an acceptable chance of meeting the 2°C goal have led to calls for the goal itself to be abandoned as unachievable (20) in favour of alternate targets such as net radiative forcing in 2050. However, there are counter-arguments (21) and new targets present other challenges, not least that it could result in a cessation of mitigation activity on CO₂ emissions for the foreseeable future, since reducing these is a relatively expensive way of affecting near-term radiative forcing trends. Secondly, by focussing on the upper tail of the distribution of possible future warming, the required mitigation effort for meeting the climate target becomes very sensitive to the upper bound of the climate system response, which is badly-constrained by observations (22, 23), and hence easily contested by different interest groups. The addition or retraction of a couple of papers at the high end could materially affect policy goals, and hence investment, leading to an unacceptable level of political risk.

Policies invoking this interpretation of the precautionary principle can, therefore, lead to high and uncertain mitigation costs to guard against potentially high but equally uncertain impacts. They are, therefore, ‘fragile’ in the sense that uncertainty in both mitigation costs and impacts make it more

difficult for any policy to be adopted, providing a strong incentive to defer decisions until these uncertainties are resolved. Yet, as the history of progress on climate sensitivity has shown, this could mean a recipe for indefinite procrastination: some uncertainties, including the costs of mitigation and the speed with which temperatures respond to falling emissions, may only be resolved after substantial mitigation efforts are already under way. The potentially paralysing impact of uncertainty becomes particularly acute if rational fears of over-mitigating combine with the politics of special interests to create additional pressures on negotiations.

In view of these issues, we argue that an approach which (a) is less beholden to the contestable tails of climate distributions, (b) more fully accounts for the set of risks governments care about, and (c) less dependent on a globally binding mandate, may be a better way of preserving flexibility in climate mitigation.

There are multiple currents of thought associated with adaptive management (8), resilience and more recently “anti-fragility” that argue for a more iterative approach to the management of complex problems. Although these approaches are usually associated with environmental or natural resource management (24) or, when in the field of climate change, responses to climate change and adaptation strategies (e.g. Netherlands Ministry of Infrastructure and the Environment, 25), we argue that some of this thinking could be constructively employed in mitigation strategies, too.

Basing strategy on more robust statistical properties, such as median estimates of both climate impacts and mitigation costs, reduces dependence on contestable tails of these distributions. To be credible, however, such a policy must also adapt to new scientific findings in a predictable way that itself minimizes the risk of unacceptable outcomes such as a sudden and precipitate revision in mitigation pathway and avoids placing an intolerable burden on future decision-makers. Simply stating that policies will be revised in the light of new evidence is insufficient: some constraints are needed on the scale of these revisions if policies are to be used as a basis for investment.

Flexible policies have been advocated before that internalise costs of emission-externalities contingent on observed climate states (26) and thus adjust to new information about the uncertain climate response to emissions (e.g. see 27). Policies that automatically adjust expenditures or efforts on the basis of some numerical parameter (usually consumer prices) are commonplace (6). Indexing makes it easier for politicians to commit to long-term stability than might otherwise be the case if explicit assent were required for every policy adjustment. It can also help create a “normative aura” around policies if they are seen to reflect an underlying fairness in the indexing (6).

An index of anthropogenic warming

A number of features are desirable in the index variable: first, it should be clearly relevant to the overall policy goal; second, it should evolve predictably to minimise short-term policy volatility; and third, it should be simple to calculate and update regularly. Since governments have already adopted the goal of limiting global average warming above pre-industrial temperatures to 2°C, and recognising that the majority of climate impacts scale more closely with this than any other readily accessible variable, an index based on global average near-surface temperature is a logical starting point.

Global temperature itself, however, is subject to natural inter-annual and inter-decadal variability that would significantly increase the risks of indexing climate policy on this variable alone. Investments in energy infrastructure mature over timescales of decades. If a global carbon tax were anchored to global temperature, as proposed by (26), then a large volcano or an upward fluctuation in the Pacific Decadal Oscillation, although neither relevant to the long-term goal of limiting anthropogenic warming, could depress or inflate carbon prices for a decade or more, unnecessarily bankrupting investors in either renewable or fossil energy supplies respectively.

A more predictable variable that is also more closely tied to the overall policy goal would be an index of warming attributable to human influence: since the work of Hasselmann in the 1990s (28), this has been defined in terms of a weighted-least-squares-fit between observed temperatures on the one hand and the expected temperature responses to anthropogenic and natural factors on the other. Estimates

of attributable warming are traditionally updated in the scientific literature when new statistical methods or new simulations of anthropogenic and natural warming become available, and assessed every few years by the Intergovernmental Panel on Climate Change. This would be inappropriate for an index variable: the method of calculating the index should be subject to scientific scrutiny, but if the value of the index itself were directly dependent on scientific judgement, this would place undue pressure on the scientists making the assessment.

Fortunately, when the target is net anthropogenic warming, very simple approaches based solely on global mean temperature and radiative forcing time-series give results that are statistically indistinguishable to the most complex statistical and modelling tools available. Orange diamonds in Fig. 1 show how anthropogenic warming relative to the mid-19th-century has evolved since 1992, calculated every year using only data available up to that year. Observed temperatures (black diamonds) are decomposed into an anthropogenic (orange) and natural (blue) component, plus a noise residual (green), using an ordinary least-squares fit.

Expected responses to external forcing, normally obtained from complex Atmosphere-Ocean general circulation models, are here calculated from global mean radiative forcing series using the simple two-component impulse-response model already used for metric calculations by the IPCC and UNFCCC (29). Crucially, because attributable warming is based on a least-squares-fit to observed temperatures, it does not require an estimate of either the Equilibrium Climate Sensitivity (ECS) or Transient Climate Response (TCR), and is only very weakly dependent on their ratio (the Realised Warming Fraction, RWF) and the adjustment times in the impulse-response model (the figure uses the model settings documented in (30)). These give an attributable warming in 2014 of 0.91°C, which varies by only one hundredth of a degree as the RWF is varied from 0.2 to 0.8 and the shorter thermal adjustment time from 2 to 10 years. Uncertainties in forcing have more impact, but are still small: setting anthropogenic aerosol forcing to zero, for example, only reduces estimated anthropogenic warming in 2014 to 0.88°C.

This index of anthropogenic warming requires no complex model calculations and can be updated as soon as new figures for annual mean temperatures and radiative forcing are released. It would have been assigned a value of 0.54°C in 1992, and has since increased monotonically by 0.37°C . The rate of increase slowed slightly after 2000 in response to the so-called “hiatus” in observed warming, showing how this index responds to evolving observations, but it does so sufficiently slowly that it would not compromise its use as a policy index. A plot of regression residuals shows nothing unprecedented about the past two decades (Fig. 1 (left panel, green)).

Anti-fragile policies

Given the burgeoning uptake of adaptive management techniques in the climate adaptation and natural resource management domains, their absence from mitigation discussions is striking. Using the index described above (or a variant on it) a range of automatically indexed policies could be explored: we simply outline some illustrative examples here reflecting the goal of limiting anthropogenic warming to 2°C and the recognition that net emissions of long-lived greenhouse gases, including carbon dioxide and nitrous oxide, have to reach zero to stabilize temperatures.

The simplest policy would be indexed emission reductions: countries could commit to reduce their emissions from a predetermined baseline by a fraction proportional to anthropogenic warming from the time the policy is adopted, rising to 100% when this warming reaches 2°C . Had such a policy been enacted in 1992, global emissions would now be 25% below baseline (anthropogenic warming since 1992 divided by the difference between 2°C and anthropogenic warming in 1992). That would be 16% above 1992 emissions if the global baseline emissions had been assumed to increase at 2% per year, or about 10% below current emissions, but still rising. Countries would argue for different rates of baseline growth depending on their stages of development, leading to different times of peak national emissions, but all would converge to zero by the time anthropogenic warming reaches 2°C . There is, of course, a risk that feedbacks in the climate system could lead to a further temperature increase even after net emissions reach zero, but in the absence of concrete evidence for this, reducing

emissions of long-lived gases to zero by the time anthropogenic warming reaches 2 °C is a defensible and realistic goal.

Although attractively direct, an indexed emission cap, supported by a permit auction, would lead to potentially destabilising price volatility should a national or regional economy over- or under-perform the expectations used in setting the baseline. An alternative would be an indexed carbon tax, but to be effective in achieving net zero, this would have to rise to a level that would discourage all further emissions by the time anthropogenic warming reaches 2°C. This, in effect, means predicting the cost of Direct Air Capture.

This has been recognised in an emerging literature on planning for climate change which emphasises the “stickiness” – the lack of immediate policy reversibility – of options as a way of embedding low carbon paths, as well as focussing on the durability and scalability of climate mitigation initiatives (5). Part of this approach involves attempts at pre-commitment, which is a potentially central feature of a problem with inter-temporal incentives to free-ride (31). Hence, arguably the simplest of all policies would be an indexed sequestration mandate: a regulation on fossil fuel extractors to sequester, or pay for the sequestration of, a monotonically increasing fraction of the carbon they extract, rising to 100% by the time attributable warming reaches 2°C (32).

For any of these three policy approaches, indexing to attributable anthropogenic warming allows a transparent link between the policy instrument and the policy goal. It renders the policy “anti-fragile”, or “adaptive” in the sense that disputes over the climate response are no longer an impediment to policy adoption. In fact, such disputes make the policy easier to adopt, since stakeholders who are convinced that future anthropogenic warming will be slower than current models predict will be reassured that the policy will “bite” correspondingly slower, while the converse is also true for those concerned about unexpectedly rapid warming in future.

Even if climate policies directly indexed to attributable anthropogenic warming are not adopted formally, this concept provides a simple and natural way of monitoring the overall consistency

between the evolving climate change signal, individual countries' emission 'pledges' and the overall goal of achieving net zero emissions of long-lived greenhouse gases by the time anthropogenic warming reaches 2°C. Annual updates of anthropogenic warming, based on a simple and transparent algorithm, should be as much part of a full suite of climate services as an annual update of global temperature.

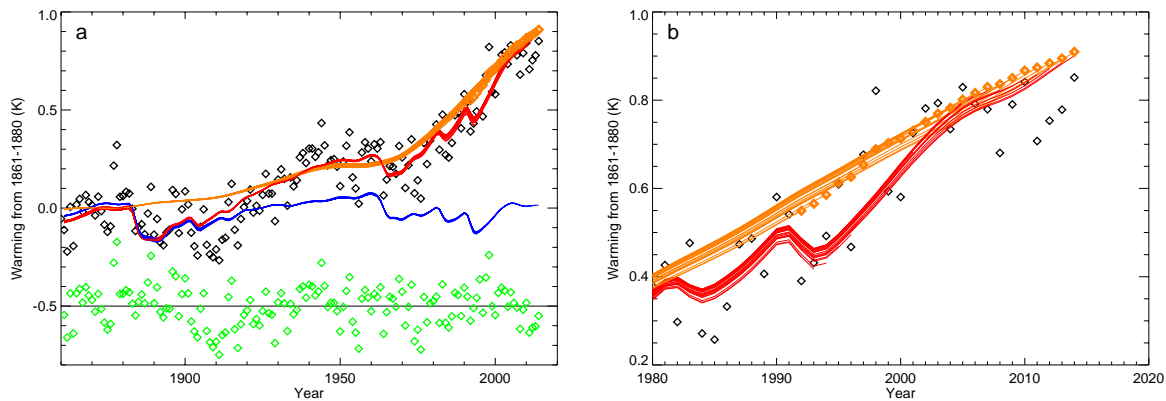


Figure 1: An index of anthropogenic warming. On both panels orange diamonds show anthropogenic warming estimated by a least-squares fit between observed temperatures (black diamonds) and the response to anthropogenic (orange) and natural (blue) forcing and their combination (red). Panel (a) shows the whole historical period and panel (b) zooms in on the anthropogenic warming since 1980. . Green diamonds in panel (a) show regression residuals for data to 2014, offset by 0.5°C.

References

1. Schelling, T. (1992) Some economics of global warming. *American Economic Review*, **82**, 1-14.
2. Barrett, S. (2005) The theory of international environmental agreements. *Handbook of Environmental Economics*, **3**, 1457-1516.
3. Victor, D.G. (2009) Plan B for Copenhagen. *Nature*, **461**, 342-344.
4. Pierson, P. (2000) The limits of design: Explaining institutional origins and change. *Governance*, **13**, 475-499.
5. Levin, K., B. Cashore, S. Bernstein and G. Auld (2012) Overcoming the tragedy of super wicked problems: constraining our future selves to ameliorate global climate change *Policy Science*, **45**, 123-152.
6. Weaver, K. (1986) *Automatic Government*. Brookings Institute
7. Patashnik, P. (2008) *Reforms at Risk*. Princeton University Press.
8. Holling, C. S. (1978). *Adaptive Environmental Assessment and Management*. Wiley, London. Reprinted by Blackburn Press in 2005.
9. Taleb, N. N. (2012) *Antifragile: Things that gain from disorder*. Penguin
10. UNFCCC (1992) United Nations Framework on Climate Change Convention.
https://unfccc.int/files/essential_background/background_publications_htmlpdf/application/pdf/conveng.pdf

11. Stern, N. (2013) The structure of economic modeling of the potential impacts of climate change: grafting gross underestimation of risk onto already narrow science models. *Journal of Economic Literature*, **51**, 838-859.
12. Schneider, S.H. and M.D. Mastrandrea (2005) Probabilistic assessment of “dangerous” climate change and emissions pathways. *Proc. Natl. Acad. Sci.*, **102**, 15728-15735.
13. UNFCCC, The Cancun Agreements (2010) <http://cancun.unfccc.int/>
14. Meinshausen, M., *et al.* (2009) Greenhouse-gas emission targets for limiting global warming to 2 °C. *Nature*, **458**, 1158-1162.
15. Rogeli, J., *et al.* (2011) Emission pathways consistent with a 2 °C global temperature limit. *Nature Climate Change* **1**, 413–418.
16. German Advisory Council on Global Change (2009) *Solving the Climate Dilemma: The Budget Approach*
17. Stocker *et al.* (eds.) (2013) *The IPCC Fifth Assessment Report: The Physical Science Basis*. Available at ipcc.ch.
18. UNFCCC (2015) The Lima decision
https://unfccc.int/files/meetings/lima_dec_2014/application/pdf/auv_cop20_lima_call_for_climate_action.pdf
19. Weizmann, M. (2012) GHG targets as insurance against catastrophic climate damages. *Journal of public economic theory*, 221-244.
20. Victor, D.G. and C.F. Kennel (2014) Climate policy: Ditch the 2°C warming goal. *Nature*, **514**, 30-31.
21. Jordan, A., T. Rayner, H. Schroeder, N. Adger, K. Anderson, A. Bows, C. Le Quere, M. Joshi, S. Mander, N. Vaughan and L. Whitmarsh (2013) Going beyond two degrees? The risks and opportunities of alternative options. *Climate Policy*, **13**, 751-769.
22. Aldrin, M. *et al.* (2012) Bayesian estimation of climate sensitivity based on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content. *Environmetrics* **23**, 253-271.
23. Otto, A. *et al.* (2013) Energy budget constraints on climate response. *Nature Geoscience*, **6**, 415-416.
24. Williams, B.K., (2011) Adaptive management of natural resources – framework and issues. *J. of Environmental Management*, 1346-1353.
25. Delta Programme 2015: Working on the Delta — The decisions to keep the Netherlands safe and liveable, The Ministry of Infrastructure and the Environment, September 2014
26. McKittrick, R. (2011) A simple state-contingent pricing rule for complex intertemporal externalities. *Energy Economics*, **33**, 111-120.
27. Leach, A. J. (2007) The climate change learning curve. *Journal of Economic Dynamics and Control*, **31**, 1728-1752.
28. Hasselmann, K. (1997) Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dynamics*, **13**, 601-611.
29. Boucher, O., and M. Reddy, 2008: Climate trade-off between black carbon and carbon dioxide emissions. *Energy Policy*, **36**, 193–200.
30. Myhre, G., *et al.* (2013) Anthropogenic and Natural Radiative Forcing. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
31. Lazarus, R. (2009) Super wicked problems and climate change: Restraining the present to liberate the future. *Cornell Law Review*, **94**, 1153-1233.
32. Allen, M. R., D.J. Frame and C.F. Mason (2009) The case for mandatory sequestration. *Nature Geoscience*, **2**, 813-814.

Author contributions

All authors contributed extensively to the writing of the paper.

