Political Realism, Feasibility Wedges, and Opportunities for Collective Action on Climate Change

Mark Budolfson

Version 3.0

budolfson /AT/ princeton /DOT/ edu

1. Overview

According to an influential realist argument about what it is politically feasible to do about climate change, regimes are infeasible if they make current citizens of powerful nations worse off, and thus an intuitively unjust global response to the problem of climate change that involves compensating large emitters for reducing emissions is the best we can realistically hope for and is thus the solution that we should actively promote even from an ethical point of view.

The realist conclusion of this kind of argument has been endorsed by a wide range of commentators in philosophy, economics, law, and international affairs such as John Broome, Eric Posner, Cass Sunstein, David Weisbach, Richard Stewart, Jonathan Wiener, and many others. For example, Broome argues that the costs of mitigation must be shifted to future generations to ensure that no one now is made worse off by climate policy, whereas Posner, Weisbach, and Sunstein suggest that poor nations more vulnerable to climate change must make large transfer payments to rich nations who have less to gain to ensure that the rich nations are made no worse off.

In this paper I examine the substance of the realist conclusions that these authors reach, and I identify a number of worries that differ depending on whether the

implementation involves the kind of intergenerational transfers that Broome has in mind, or the intratemporal transfers that Posner, Sunstein, and Weisbach have in mind.

As a more general point, I also argue that the realist argument is invalid, and that the reason why it fails also points the way toward a more desirable realist response than these commentators endorse. The argument is invalid because it overlooks the fact that what it is in the interest of a nation to do can change depending on the actions of other nations. In particular, even if nations are invariably self-interested as the realist premise of the argument assumes, other nations can change what it is in their interest to do via threats, sanctions, and other measures, and can therefore make it in their interest to comply with international standards that it would not have been in their interest to comply with in the absence of those measures. As a concrete illustration of how this creates trouble for the argument in the particular case of climate change, I discuss an international climate treaty that requires signatories to impose globally optimal emissions taxes within their territory and empowers compliant nations to impose tariffs and other penalties on non-compliant nations as retaliation for non-compliance, somewhat akin to proposals by Joseph Stiglitz and William Nordhaus.

However, the positive proposal I offer focuses on adding a large number of feasibility enhancing features and complimentary measures that I call 'feasibility wedges', which focus attention on the diplomacy and political strategy that may be required in order to achieve anything like optimal policy outcomes. I identify a strategic dynamic that allows this 'meta-architecture' for agreement along with intra- and intertemporal transfers can be incorporated into the emerging post-Kyoto 'bottom-up' 'regime complex' for climate change, and how leveraging that dynamic can accelerate the evolution of a global response in the direction of optimal policy.

2. The Realist Argument and Efficiency Without Sacrifice

According to many, the primary problem standing in the way of an effective global response to climate change is a problem of political feasibility at the level of nations in

light of our circumstances. The basic problem is that it is simply not in the interest of the current citizens of many nations to make substantial emissions reductions, and so, realistically, we should not expect substantial emissions reductions unless those people are 'bribed' into making them.

This is a large part of a realist explanation why no meaningful action on climate change has been observed despite its dramatic threat to future generations, and why more substantial action will not happen unless something substantial is done to make it in the interest of current people to make such reductions. Realists often also cite the prisoner's dilemma-like nature of the situation, whereby if an individual nation makes large reductions unilaterally without the cooperation of the other nations – or, if even a large but insufficient number of nations join together to make such reductions – they will suffer significant economic losses, and the climate-related outcome will not be importantly different either for their own citizens or for others around the world who are vulnerable to the effects of climate change, partly because of the phenomenon of leakage, whereby 'unilateral' regulation to reduce emissions in one part of the world would create large incentives for industries there to simply relocate to the unregulated areas of the world and continue emitting in those unregulated areas – thereby increasing the costs to the unilaterally regulating regions, while undermining the realization of any global benefits due to the leakage, and thereby reducing the value of the already small sliver of the global benefits for regulating nations.[1]

From these premises about our circumstances together with the characteristic view of realism, the *Realist Argument* I will focus on here reaches a conclusion we can call *Efficiency Without Sacrifice*:

> *Realist Feasibility Constraint*: Nations act only in the interests of their current citizens, so a response to climate change is infeasible if it requires a nation to act contrary to the interests of its current citizens.

[1] EMF session on leakage, Rendall on leakage, Budolfson unilateral paper, Nordhaus and Boyer warming the world discussion.

Circumstances: An optimal response to climate change requires substantial emissions reductions from the no policy status quo from many rich nations, and the current citizens of those rich nations would be made much worse off relative to business as usual by such reductions unless they were compensated for making them.

Therefore, given the *Realist Feasibility Constraint* and our *Circumstances*,

The best feasible response is *Efficiency Without Sacrifice*, which involves reducing emissions to a level that is optimal,[2] while compensating current people, including citizens of rich nations, for the cost of making such reductions, thereby ensuring that no one including the rich has to make any sacrifices, and that the costs are instead borne by those who are vulnerable to climate damages, because only in this way can substantial emissions reductions become feasible.

This argument has been endorsed by a wide range of commentators in philosophy, economics, law, and international affairs such as John Broome, Eric Posner, Cass Sunstein, David Weisbach, Richard Stewart, Jonathan Wiener, and many others.[3] It depends crucially on the idea that there are transfers ('side payments') between people that are both feasible and yield a pareto improvement over the no policy 'business as usual' status quo.

---

[2] Following standard usage in emissions reductions contexts, the so-called 'optimal' level of emissions is one at which the global marginal cost of further reductions equals the global marginal damage of further emissions, typically under optimal policy. Although this can sound straightforward, and is usually presented as if it is straightforward, the idea that this is genuinely the *optimal* level of reductions is not straightforward. As just one example, in such calculations the cost of further reductions is usually assumed to be location-independent (with calculations based on global or at best regional average cost in dollars), when in fact a unit of reductions has dramatically higher welfare costs for the poor than the rich (which is not adequately accounted for by even regional averages). More generally, calculations of both marginal cost of emissions and abatement are based on global averages in a way that does not adequately (even if it does partially) account for the welfare impact of inequalities in the distribution of the costs at issue. So, in sum, the policy that has the lowest welfare cost or is welfare optimal might be quite different from the policy that is 'optimal' in this sense. I discuss some related issues in a section on optimal reductions further below.

[3] For example Eric Posner and David Weisbach, *Climate Change Justice*, pp. 6, 86, and 143, Eric Posner and Cass Sunstein, "Climate Change Justice", especially pp. 1569-1570 (but see also Cass Sunstein, "US Should Act Unilaterally on Climate Change", http://www.bloomberg.com/news/2013-01-23/u-s-should-act-unilaterally-on-climate-change.html), Richard Stewart and Jonathan Wiener, *Reconstructing Climate Policy: Beyond Kyoto*, pp. 102-103, Jonathan Wiener, "Incentives and meta-architecture", in Aldy and Stavins (eds.) *Architectures for Agreement*, pp. 75-76, John Broome, *Climate Matters*, pp. 44-7.

In what follows I first evaluate the different substantial versions of Efficiency Without Sacrifice that have been offered by these commentators. I then argue that the Realist Argument is invalid and depends on a mischaracterization of the feasible options. This leads to a substantive alternative proposal that also includes a smaller but nonetheless important place for transfers insofar as they are feasible.

3. Posner, Sunstein, and Weisbach's Version of Efficiency Without Sacrifice: Intratemporal Transfers

In this section, I raise some concerns about the substance of Posner, Weisbach, and Sunstein's version of Efficiency Without Sacrifice. I ultimately conclude that their version of the view has to be mistaken.

Among commentators who endorse the realist argument for Efficiency Without Sacrifice above, Posner, Sunstein, and Weisbach have received much attention for claiming that the upshot of that argument is that a response to the problem of climate change that is intuitively highly unjust is the best that we can currently hope for, and is thus that such a response is the one we should aim for even from an ethical point of view. Specifically, their view is that given the realistic feasibility constraint outline above, the best outcome we can hope for involves poor nations such as Tuvalu making massive transfers to rich nations such as the United States so that the current citizens of the rich nations can be made better off vs. the no policy status quo by the combination of those transfers and incurring their share of the costs of a global harmonized carbon tax. (These authors do not see this arrangement as involving any substantial injustice, but I will set aside that aspect of their argument in what follows, emphasizing instead the uncontroversial fact that such an arrangement is 'intuitively unjust'.[4])

[4] Compare Broome: "Efficiency without sacrifice has the further, serious demerit that it is unjust. … Under efficiency without sacrifice emitters are paid to reduce their emissions by the receivers. Receivers in effect bribe emitters not to harm them. This benefits both emitters and receivers, but only relative to the initial unjust state of business as usual. Efficiency without sacrifice perpetuates the injustice" (46).

One problem with this view is that economic models suggest that it is impossible to have a pareto improvement of the sort that they assume is possible if the relevant transfers are supposed to involve simply redistributing national outputs over the coming decades. For example, William Nordhaus's multi-region RICE model implies that under optimal policy with a single global harmonized carbon tax all regions of the world have lower economic output than under the no policy status quo until nearly 2100. This means that even before making the transfers that Posner, Sunstein, and Weisbach have in mind, under optimal policy the current citizens of poor nations (along with everyone else) still have lower GDP than under business as usual.[5] If we add large transfer payments from these poor nations on top of that, then they would do even worse, and would certainly not be better off within these models than under the no policy status quo. So, the Posner, Sunstein, Weisbach view seems based on a basic misunderstanding of the relevant economic facts if it assumes that there is a possible reshuffling between nations of national economic outputs over the coming decades that in conjunction with optimal emissions reductions yields a pareto improvement for citizens of all nations over the no policy status quo. There is simply no such reshuffling that is even remotely possible, because everyone is a net loser in the coming decades under optimal emissions reductions, assuming there is no additional change (such as much lower savings rates) of the sort not contemplated in mainstream analyses of the sort that they cite with approval. (I consider the effect of changing these assumptions about savings rates etc. further below.)

To suggest otherwise is to misunderstand the most basic tradeoff in the mainstream economics of climate change, which is that *in all regions* near-term output must be sacrificed in our lifetimes to prevent larger sacrifices of output in the further future if there is a global harmonized carbon tax. Furthermore, these near-term sacrifices of GDP will be particularly large and painful for developing nations in the near-term if they are accomplished with a global harmonized carbon tax. So, it is mistake to suggest – as many philosophers glibly do – that optimal emissions reductions in practice would amount to nothing more than a small tax on the rich. The

[5] Nordhaus's RICE 2010 model is available on his website.

reality is that with a harmonized global carbon tax, the horrific tradeoff we face is to decide out how many millions of poor people to kill now with carbon taxes in order to protect more millions of people in the future. We should be deeply suspicious of anyone who claims to be a realist but suggests otherwise.[6]

If one endorses a global harmonized carbon tax – as do Posner, Weisbach, and Sunstein, and Broome (discussed later) – then this horrific tradeoff can seem inescapable. (I discuss alternatives to a global harmonized tax and to harmonized carbon prices further below.) However, even with a global harmonized carbon tax there is at least a theoretical way out of the problem that would allow us to protect the current poor while at the same time making optimal emissions reductions. This theoretically possible way out is to make transfers of future economic output from the future to the present (transfers of output across time), rather than merely transfers of economic output across nations within time as considered above.

However, Posner, Weisbach, and Sunstein's view falls apart if we endorse these intertemporal transfers as realistic. That is because if this way out is sufficiently realistic, then nations can redistribute their own output across time, which if this is sufficiently realistic means that nearly all individual nations would be able to secure a pareto improvement over the no policy status quo in conjunction with optimal emissions reductions by transferring *within their nation* from their future richer citizens to their current comparatively poorer citizens. This means that there is then no longer any good argument at all from the premises of the realist argument above for the distinctive conclusion of Posner, Sunstein, and Weisbach that the only way to achieve feasibility involves large transfers to rich nations from poor nations.

The background facts that support this are that there are only a few nations for which the no policy status quo does not ultimately lead to long-run GDP losses that far outweigh the current costs of their share of optimal mitigation under a global harmonized carbon tax. (Russia and Canada come to mind as possible examples of such 'invulnerable' nations, although some commentators would probably argue that even

[6] See Anthoff and Tol, Schelling.

those nations would ultimately face losses under no policy that overwhelm the cost of compliance with an optimal global tax.) If there are indeed a small number of these 'invulnerable' nations, we can then adopt the Posner, Weisbach, and Sunstein trick of side payments to bribe them into compliance – but because the number of these invulnerable nations is at least small and arguably nonexistent, there is at least no longer any good argument at all from the premises of the realist argument above for the distinctive conclusion that rich nations must in general be paid off by large transfers from poor nations.[7]

So, the upshot is that Posner, Weisbach, and Sunstein's version of Efficiency Without Sacrifice seems to face a crippling dilemma: we've just seen that their view that poor nations must make transfer payments to rich nations falls apart if we endorse large intertemporal transfers as realistic, because then rich nations can compensate their current citizens in a way that makes everyone better off than under business as usual, and we previously saw that if we don't endorse large intertemporal transfers as realistic, then their view is a nonstarter, because it then relies on the idea that optimal global emissions reductions can be conjoined with redistribution of national outputs over the coming decades in a way that leads to a pareto improvement over the no policy status quo – which is not even close to being possible if we do not allow intertemporal transfers, as all regions are worse off in coming decades under optimal policy than under the no policy status quo.

[7] Following all of the authors in this literature, I set aside the problem in this section and the next of how to realistically ensure universal compliance among nations even given transfer payments, given that many nations would have strategic reason to refuse the offer of a pareto-improving transfer in order to bargain for a much larger payment – as it would be common knowledge that the success of the entire emissions reductions scheme could in such a way be held hostage by one large nation or at least a small coalition of nations. A further problem is that unless a self-enforcing incentive structure is somehow created, it will be common knowledge that a coalition of nations can always scuttle the agreement at a future date by pulling out perhaps in a similar strategic move to bargain for even larger side payments. I return to this problem in a subsequent section, where I offer some substantive proposals for dealing with it.

4. Broome's Version of Efficiency Without Sacrifice: Intergenerational Transfers

The previous section suggests that Posner, Weisbach, and Sunstein's substantive view is not supported by a good argument even if one endorses the basic premises and logic of the Realistic Argument. In this section I discuss the more promising view of John Broome, who explicitly develops his version of Efficiency Without Sacrifice in terms of transfers across time. Broome's basic idea is to use intertemporal transfers as described in the previous section from the future to the present to ensure that current people are made no worse off than under business as usual even while making optimal emissions reductions.

In more detail, Broome's idea is that in light of the Realist Feasibility Constraint above, the best feasible outcome is a welfare suboptimal outcome that is nonetheless 'efficient' in the sense used in this literature, which means achieving the desired (optimal) emissions reductions *at the least possible cost*, where the 'optimal' level is understood in the sense used in this literature, which is the level of emissions where the global marginal cost of emissions is equal to the global marginal cost of abatement.[8] Broome believes that the welfare optimal outcome is a particular version of what he calls Efficiency With Sacrifice that involves the optimal level of emissions reductions via something like a cap and trade scheme with permits allocated so as to maximize welfare – and such a welfare optimal distribution of permits would involve distributing them entirely to the world's poorest people, which would imply a massive transfer of wealth from rich to poor, which is why it would violate the Realist Feasibility Constraint. Further, even if we imagine an otherwise similar scheme but with permits allocated in proportion to economic output, current people would still be worse off than in the no policy status quo, and thus even this welfare inferior policy would violate the Realist Feasibility Constraint. In light of that, Broome's thought is that by using intergenerational transfers, we are able to compensate people now for the cost of

[8] Again, the generally accepted convention is to use 'optimal' in this way in connection to optimal emissions reductions. I discuss some problems with this in the footnote above, and other problems in a section below.

making optimal emissions reductions via transfers to them from future people (compatriots?) who will still be net beneficiaries due to their even larger benefit from reduced future climate damages, thereby making optimal level of emissions reductions feasible by the lights of the Realist Feasibility Constraint. (Note that implicit in this is the very thing that undermines the Posner, Sunstein, Weisbach argument that poor nations must compensate rich nations if it is assumed that such intergenerational transfers are feasible.)

In the rest of this section, I highlight a possible disagreement with Broome's assumptions about the implications of his view, which I do not think leads to any objection to the view itself, but merely makes the view much more attractive if Broome is incorrect. In the next section, I discuss more substantive issues about the sort of intertemporal transfers that Broome has in mind.

The assumption of Broome's that I find problematic is that Efficiency Without Sacrifice is welfare inferior to Efficiency With Sacrifice. Given how Broome defines those outcomes, I believe Efficiency Without Sacrifice could be welfare superior to Efficiency With Sacrifice – and could be dramatically welfare superior. To see why, note that Efficiency With Sacrifice assumes that transfers (in effect) happen via permit allocation and the resulting trading, and that these transfers are entirely between contemporaries (crucially, are *intra*temporal). Efficiency Without Sacrifice, in contrast, assumes that transfers are happening from future to the present – and it is fairly uncontroversial that future people will be richer than their contemporaries now. So, if it is really feasible to make those transfers across time in the way that Broome assumes, then the costs of climate change could be shifted to people in the future who are much richer than their counterparts now. There is at least in principle no reason why the result could not be a large welfare improvement over Efficiency With Sacrifice if the latter outcome only involves transfers (so to speak, via permit allocation) between contemporaries.

Broome explicitly considers and rejects the possibility that Efficiency Without Sacrifice could be a welfare improvement over Efficiency With Sacrifice in this way, but I think his analysis ignores many of the distributional issues that tell in favor of that

possibility, and so I find his analysis unconvincing. In his analysis, Broome quite rightly highlights the fact that if intergenerational transfer is accomplished with, say, a lower savings rate, this will lead to foregone consumption later, and that this *could* lead to welfare losses in the future that are larger than the welfare gains to people now who benefit from the transfer; Broome then goes on to claim that cost-benefit analyses show that it is actually the case that the resulting welfare losses would outweigh the gains. But I think that this is simply a matter of invoking inadequate cost-benefit analyses that ignore the distributional issues to answer a particular question that depends essentially on the distributional issues that are ignored by those analyses.

Although Broome might agree that there is a lack of adequate representation of inequalities in these models, he seems to believe that this does not raise problems for his argument here because he believes – falsely, by the lights of all standard analyses[9] – that emissions reductions only impose large costs on the rich. If that were true, then his argument would go through. To see where Broome makes this assumption, consider the key passage on these issues in *Climate Matters*:

> The difference between Efficiency Without Sacrifice and Efficiency With Sacrifice is the distribution of resources between people. Since emitters are mainly the current rich, whereas receivers are mainly the poor and future generations, the current rich are better off in Efficiency Without Sacrifice, whereas the poor and future generations are better off in Efficiency With Sacrifice.[10]

The crucial assumption here is that the current poor would not pay any important part of the cost of emissions reductions – this is crucial to the inference that "the [current] poor…are better off in Efficiency With Sacrifice". However, that assumption is simply false, at least if we pay any attention to standard cost benefit analyses, where, again, the entire essence of the problem is that both rich *and poor* in our lifetimes suffer losses relative to business as usual if we make large emissions reductions. Again, that is the

[9] I take leading standard analyses to be the results of eg DICE, FUND, PAGE, which are used in the US social cost of carbon estimates, where PAGE was used in the Stern Review, and where DICE and RICE are Nordhaus's models.

[10] pg. 45.

essence of the problem according to all standard models: as noted above, if we are going to rely on a single global price on emissions, then we have to choose the least bad among seriously regrettable tradeoffs between people now and people in the future, keeping in mind that many people now who will be seriously harmed by carbon prices are desperately poor. That is the essence of the problem according to the economists who are at the forefront of the standard modeling of climate change costs and benefits, and it is also the essence of the problem according to population-weighted representatives of the world's poor in our actual international climate negotiations. (The fact that representatives of a few nations like Tuvalu have a different view is not an objection to this, and neither is the fact that most philosophers have a different view. We saw in the last section that Posner, Sunstein, and Weisbach also seem to have a mistaken understanding of these issues, as their argument (on the most natural interpretation) also depends on the assumption that the current poor are net beneficiaries from optimal emissions reductions, which is the opposite of the truth according to all standard models.)

The upshot is that Broome faces a dilemma regarding his argument that Efficiency Without Sacrifice is worse than Efficiency With Sacrifice: on the one hand, he seems to want to argue directly for that ranking in the passage quoted above in a way that does not depend on the idiosyncrasies of existing cost-benefit analyses; but that argument relies on an assumption that is inconsistent with the essence of the climate change problem according to the standard literature that he quotes with approval. And if we set that aside and focus only on the existing calculations of cost-benefit analyses, which do agree with Broome's ranking of those policies, the problem is that it is clear that those ranking are likely to reverse under some plausible assumptions about the intratemporal distribution of mitigation cost, climate damage, and the incidence of Broome's intergenerational transfer. Those distributional facts are simply ignored by existing cost-benefit analysis models, which means that those models cannot answer this question.

When distributional considerations are taken into account, I believe the sort of Efficiency Without Sacrifice that Broome has in mind would likely lead to a welfare

improvement over Efficiency With Sacrifice – but that is merely a conjecture. Getting to the bottom of this requires figuring out the answer to a very complicated economic/utilitarian question, and one that requires *careful* economic modeling to answer, and is not something that can be decided on the basis of existing economic modeling that is clearly *inadequate* for the task, as existing models ignore many of the distributional impacts of both climate damages and mitigation cost.[11]

To illustrate why I have this intuition, suppose that the distribution of mitigation cost is regressive, where mitigation cost is understood to include all near-term foregone consumption under optimal emissions reductions. If so, then existing cost-benefit analysis models underestimate the welfare loss imposed by emissions reductions paid for by current people (because they implicitly assume that mitigation cost is distributed proportional to consumption), which means that they underestimate the welfare gain that would result from having some future richer people pick up the tab instead as would be the case under Efficiency Without Sacrifice as compared to Efficiency With Sacrifice. Insofar as the intergenerational transfers that Broome has in mind can also be made so that their incidence is progressive among the future people who pick up the tab, then that would be a further welfare-improving dimension that is not taken into account in existing cost-benefit analyses. When I speculate about the actual distributional facts, including under Broome's intended intergenerational transfer scheme, I speculate that both of the distributional considerations here would add welfare to Efficiency Without Sacrifice over the estimates of standard models – and would add enough welfare to make it a welfare improvement over Efficiency With Sacrifice; thus my conjecture.

I take it that there is no downside for Broome's view on this particular point, since if what I've conjectured is correct, that would mean that the view has even more desirable properties than has previously been acknowledged, since might allow for a welfare improvement over even the policy that is welfare optimal assuming no transfers. It would also at least mitigate the worry that Efficiency Without Sacrifice amounts to an

[11] For work that highlights the importance of the distribution of damages and mitigation cost, see Dennig et al 2015, and Budolfson et al under review.

injustice – and, depending on how the costs and benefits play out, it could entirely remove that worry.[12]

Let me take a step back and make a 'big picture' comment about Broome's view, all of this, and how it relates to the existing literature. In some sense, what is going on here is that by assuming the possibility of intergenerational transfers, Broome has implicitly moved outside the box of the standard literature on climate economics, which assumes that there can be no transfers except via the instruments of climate policy. (I.e. the standard assumption is that the only transfers allowed are the small-scale transfers that happen implicitly via carbon prices, and perhaps also – although this is almost never investigated in the standard literature – via permit allocation; the assumption is also that savings rates are determined by a rule that is not a matter of independent policy choice). Once we are outside of that box of the standard literature, we can then consider non-climate-instrument-related transfers from rich to poor that are welfare improving, and if we are able to assume such transfers in large measure, then it is *easy* to describe conjunctions of those transfers with a climate policy achieving optimal emissions reductions that amount to a vast welfare improvement over any climate policy considered in isolation, including the climate policy that is welfare optimal assuming that no such transfers are possible. And the climate policy that is welfare optimal assuming that no such transfers are possible is simply Efficiency With Sacrifice, as described above (if we follow, as Broome does, other standard modeling conventions in describing that view, such as a harmonized global carbon price). So, if we allow large transfers, we can achieve a large welfare improvement over Efficiency With Sacrifice; and since Broome's Efficiency Without Sacrifice assumes that some large transfers are feasible, it is then easy to see how it *could* be a welfare improvement over Efficiency With Sacrifice.

In the last section of this paper, I argue that if we assume that these intergenerational transfers are possible, and if we believe that a political strategy I

[12] See Broome on the injustice of efficiency without sacrifice, quoted several footnotes above.

describe in a later section is feasible, then a much better and more fair response to climate change is feasible.

5. Is Broome's Proposal Feasible? How Realistic is It?

All of the preceding emphasizes why it is very important whether the intergenerational transfers that Broome has in mind are genuinely feasible. If they are not, then this entire discussion is irrelevant to what we should actually do assuming the Realist Feasibility Constraint.

I leave it to others to analyze the feasibility of intergenerational transfers. To be honest, I don't really understand how exactly they are supposed to work in practice in this context. But I look forward to seeing the results, and I am hopeful they could work – it is an important and exciting idea. In the meantime, in the rest of this section I'll focus on a number of further problems that seem to emerge, especially from a realist perspective.

Most importantly, even if some forms of intergenerational transfers are feasible, it is important to wonder what the distributional implications are of that subset of transfers that are feasible. The big worry here is that insofar as intergenerational transfers are feasible, it may be that only regressive transfers are feasible (regressive on future people who 'pay' the transfer), with the consequence that the non-climate welfare cost of making those transfers could be larger than the climate-related welfare gain from making them. For example, if the only feasible way to make large intergenerational transfers is via a political bargain that somehow bakes in funding cuts to schools, pensions, and investment for the future poor and vulnerable, then this would be a major problem.

If one adopts a realist stance, then presumably one should take this worry particularly seriously, as it seems all-too realistic to imagine a public policy initiative that aspires to be a progressive version of Efficiency Without Sacrifice being hijacked and devolving into a welfare-destroying compromise deal that achieves a carbon tax at

the price of cuts in effective entitlement programs. Or more darkly, one can imagine that the bargain that emerges is a fraudulent carbon pricing scheme that only serves to enrich sophisticated investment banks and thus accomplishes no good, but succeeds in harming the economy for non-elites, and is purchased at the price of very large sacrifices of entitlement programs – and also kills millions of poor people in the coming decades through food price spikes and other more indirect ways of killing poor people, and so on.

The positive 'flip side' of this worry is that insofar as progressive intergenerational transfers are indeed possible, then there is an opportunity to promote welfare by analyzing them, and then choosing the form of intergenerational transfers that is welfare optimal.

What needs to happen next, then, is presumably something similar to what is happening now with the current 'hot topic' in the climate economics modeling community of estimating the effects of combining a carbon tax with (intratemporal) revenue recycling and tax reform.[13] The point of this current modeling exercise is, first, to evaluate how close we can get to a 'double dividend' whereby a carbon tax combined with other structural reforms immediately also yields a net gain for aggregate economic output, and, second, to evaluate what the net distributional effect of such policy combinations would be for different socio-economic groups (e.g. income quintiles). The results suggest that we probably cannot get a double dividend, but we can significantly reduce the cost to the economy with such policy combinations (vs. a carbon tax only, or (even worse) a command and control approach such as embodied in the US EPA Clean Power Plan).[14] Perhaps more importantly for considerations of both wellbeing and feasibility, the results indicate that some of these policy combinations such as equal per

[13] See current EMF exercise, RFF reports, other citations.

[14] citations Metcalf, RFF, Marisa Beck, RFF commentators report that the EPA plan is much worse along dimensions of both cost per level of emissions reductions and regressivity of the policy than any of the other possibilities studied. Some people claim to get a double dividend, although this is a minority result at this point.

capita rebate of the revenues from a carbon tax immediately make over 80% of citizens net beneficiaries of the combined policy.

In light of these results, one idea is for Broome to add these *intra*temporal transfers to the overall package of measures he recommends. The view of many of those who are doing and promoting this modeling in policy circles is that these intratemporal transfer measures are indeed a crucial part of the best way forward.

In any event, this provides perhaps an example of how the notion of intergenerational transfers could be investigated in a way that would be perceived as rigorous and policy relevant. Given that distributional issues are crucial, I don't think that e.g. the DICE model is ultimately going to be an adequate tool for fully answering these questions, even if it is the best tool for getting an initial sense of some of the relevant facts.

The preceding indicates that it is crucial to Broome's argument that the intergenerational transfers he has in mind are both feasible and welfare-improving over business as usual, even when their effects outside the box of the climate problem are taken properly into account.

But at this point we should wonder, darkly, whether in practice those transfers would even be welfare improving over business as usual. Putting on my realist hat, nothwithstanding everything above, I don't really see why we should expect *any complex* climate interventions to be welfare improving *in practice*, even if we agree that their effects would be wonderful *in theory* were we to assume, completely unrealistically, that there would be perfect implementation by perfectly benevolent agents. Realists are apt to have a very dark view according to which any complex grand scheme by welfare economists is certain to be hijacked by special interests, who are much more sophisticated and powerful than welfare economists. So, from that perspective, even if the proposal would result in a large improvement in theory, the actual expected effect of this kind of grand-level, complex policy would be to make things worse than the current trajectory of increasingly enlightened self-interested

action by nations (which is not the same thing as business as usual).[15] Given the pervasive phenomenon of *government failure* (on a dark view, this is typically the other horn of the regulation dilemma to *market failure*), it is difficult to understand why this sort of problem is not even discussed by anyone in this literature (which is not to say that Broome has any special obligation to discuss it himself). Presumably, any climate policy discussion that purports to be discussing *realist* climate policy must explicitly engage with these dark issues.

With these dark considerations in mind, like many realist economists and other commentators, in response to climate change I see a lot of virtue in a simple transparent policy: perhaps the best current contender is a carbon tax with all revenues rebated on an equal per capita basis, of the sort that is currently the focus of the many modeling exercises noted above.[16] We might say that this is an example of a maximally realistic

[15] It is worth distinguishing between the scenario called 'business as usual' in climate change discussions, which generally assumes no substantial greenhouse gas policy, except as implicit in existing policies, vs. what might seem the more plausible result of nations not really doing anything special about climate change, which is a scenario in which nations do whatever they perceive to be in their self-interest on climate change, including given an evolving more 'enlightened' recognition of what is in their self-interest, a Kuznets curve-like evolution of demand for environmental protection and air pollution reductions etc as people become richer, and so on. To mark this distinction, we might call the first 'no meaningful GHG policy', and the second 'enlightened self-interest GHG policy'. Within the latter enlightened self-interest GHG policy category, there is a further useful distinction between three scenarios: first, one in which nations act independently as best they can to promote self-interest without nations engaging in any further cooperation or international policy to influence the outcome, second, the more realistic scenario in which nations also agree to coordinate with each other insofar as it is in their perceived self-interest to do so, but without nations engaging in any further international policy beyond mere coordination to influence the outcome, and third, a scenario is like the last, but perhaps even more realistic, where nations also engage in further non-coercive international policy (either individually or in clubs) in whatever way is in their self-interest, which might include border tax adjustments, self-interested transfer payments conditional on joining and complying with the rules of the 'emissions reduction club', and the like. We can then ask, first, how far toward optimal policy we should expect to get in each scenario of these scenarios, second, how likely it is that each scenario would evolve on its own, without need for any particularly savvy climate diplomacy or political and economic analysis. Finally, we can ask whether there is any feasible scenario that is an improvement over these, perhaps involving 'coerced cooperation' among nations, unilateral geoengineering, or some other such thing. Ultimately, we can must ask what kind of diplomacy would be required to get us to the best outcome (savvy or otherwise), and how likely it is that such diplomacy (or diplomacy by other means) could succeed. We must also ask what the comparative downsides are if various strategies not succeed.

[16] Should we set the carbon price at the 'optimal' level? Probably not nearly, since 'we' refers to the policymakers of a particular country, who can anticipate less than perfect global coordination on that

policy portfolio that includes *intra*temporal transfers and aims at optimal policy: it allows us to combine an optimal carbon price with 'outside the standard box' transfers in a way that appears both feasible[17] and not-easily-subject-to-capture, and along both of those dimensions appears to do about as well as we can hope if we are going to add intratemporal transfers to climate policy. The key practical question for Broome is whether he has a similarly realistic *inter*generational transfer policy to add to this – i.e., something as simple and transparent that could avoid being perverted by special interests while accomplishing the intertemporal transfer he has in mind. If (and only if) he does, then realists should enthusiastically endorse it as central to the maximally realistic policy portfolio that aims at an optimal policy response to climate change.

6. What is the Optimal Level of Emissions Reductions, and Should it be Achieved with a Harmonized Global Price on Emissions?

Broome's discussion follows the standard literature in assuming that (a) there is an optimal level of emissions reductions, (b) efficiency demands a single harmonized global carbon price that yields that level of reductions, and thus (c) the welfare optimal response involves that harmonized global carbon price that follows from (a) and (b) (but welfare optimality is not necessarily implied by that price, depending on how eg permits are allocated), where (d) the optimal level of emissions reductions can be calculated as by existing standard models that impose a harmonized global carbon price as a constraint on optimization. I think these assumptions are reasonable for Broome's

price. A better target is probably the enlightened self-interest level, which might arguably be ~$10/tCO2 in the US. Even if 'we' were perfectly benevolent, such a level might be wise in practice as a near-term starting point, not least of which because it allows for valuable learning from a large-scale carbon tax, while limiting the potentially catastrophic downsides, which are myriad, but include at the most welfare extreme food price spikes starving lots of people elsewhere in the world, etc, etc, etc. I discuss some of the issues that arise from the prospect of unilateral emissions reductions by nations further in my paper "Should the US and Other Nations Make Unilateral Greenhouse Gas Emissions Reductions?".

[17] I.e. it makes over 80% of citizens net beneficiaries, in contrast to a carbon tax only, which harms most citizens (in many modeling exercises), or at least makes a much smaller percentage of citizens net beneficiaries.

purposes in his work, but it is important to stress that they are probably not ultimately the correct way of thinking about things.

Following the economics literature,[18] the main problem with all this is that assuming as is standard that there are no transfers beyond those achievable with carbon prices alone (and thus assuming a standard harmonized tax rather than an unrealistic extremely progressive permit allocation system), harmonized carbon price outcomes might all be welfare inferior to some outcomes that allow different carbon prices in different locations in the world, for example with lower carbon prices in poorer regions that would sufferer much larger near-term welfare losses from higher carbon prices. To put it in terms of Broome's discussion, if we consider Efficiency With Sacrifice implemented by a harmonized global carbon tax that achieves optimal emissions reductions, that may be welfare inferior to (and may lead to less emissions reductions than) an alternative arrangement in which there are different carbon prices in different locations at each moment in time.

More importantly, recent modeling suggests that *in fact* Efficiency With Sacrifice is welfare inferior to other policies involving differential regional prices (that also involve greater emissions reductions than Efficiency With Sacrifice).[19] So, Efficiency With Sacrifice is not even welfare optimal in actual fact, assuming no non-climate-instrument-related transfers, as there is some other policy that does even better along both welfare and climate dimensions by allowing differential regional prices. The intuitive idea, again, is that differential prices allow one to prevent large welfare losses in the near term in developing regions by having only a tiny carbon price there now, where that comes at a smaller loss of welfare to the richer future by way of increased future damages; this is a welfare improvement over even the optimum with a harmonized price constraint that is ruled out by the imposition of the standard constraint of a harmonized global carbon price.

[18] ie Heal and chichilnisky 1994; Sandmo

[19] Anthoff provided the initial modeling demonstration of this in FUND; see also Budolfson and Dennig for a demonstration in RICE; both papers provide further discussion and references to the theoretical literature.

So, contrary to (d), if 'optimal' is understood as welfare optimal, then the optimal level of emissions reductions cannot be computed by standard models, which impose the constraint of harmonized global prices. This means that the substantial conception of (a) relied upon by commentators is mistaken. In light of that, it is important to keep in mind that 'efficiency' in (b) means only 'least cost to aggregate output' relative to a particular level of emissions reductions, and thus imposing the constraint of a harmonized carbon price calculated in way (d) may well amount to the imposition of a welfare loss vs. other possible outcomes.

In sum, when (a) is calculated in way (d), then (c) does not follow from (a) and (b) – and existing modeling indicate that in fact (c) is false under those conditions. The practical upshot is that we should be open to the idea that the best feasible policy response does not involve a single harmonized price on carbon. This should perhaps be welcome news, including to commentators like Broome, because the recent pivot toward a more 'bottom-up' strategy suggests that negotiators now think that differential regional prices are the only feasible thing in practice.[20]

7. Why the Realist Argument for Efficiency Without Sacrifice is Invalid

In the preceding discussion, I accepted the Realist Argument on its own terms, and directly evaluated the substance of the conclusions that it is taken to support. But now it is time to take a step back and note that there is a fundamental problem with Realist Argument: it is invalid, and the nature of its invalidity points the way towards what seems to be a more ethical response than *Efficiency Without Sacrifice* that is consistent with both the *Realist Feasibility Constraint* and the premise about our Circumstances. Here again is the *Realist Argument*:

[20] However, presumably the first-best differential prices optimum is not feasible. But the current point is that there is some possible welfare gain from having low initial prices in developing nations, which is part of the current trajectory.

> *Realist Feasibility Constraint*: Nations act only in the interests of their current citizens, so a response to climate change is infeasible if it requires a nation to act contrary to the interests of its current citizens.
>
> *Circumstances*: An optimal response to climate change requires substantial emissions reductions from the no policy status quo from many rich nations, and the current citizens of those rich nations would be made much worse off relative to business as usual by such reductions unless they were compensated for making them.
>
> Therefore, given the *Realist Feasibility Constraint* and our *Circumstances*,
>
> The best feasible response is *Efficiency Without Sacrifice*, which involves reducing emissions to a level that is optimal, while compensating current people, including citizens of rich nations, for the cost of making such reductions, thereby ensuring that no one including the rich has to make any sacrifices, and that the costs are instead borne by those who are vulnerable to climate damages, because only in this way can substantial emissions reductions become feasible.

The argument is invalid because it overlooks the fact that what it is in the interest of a nation to do can change depending on the actions of other nations – in particular, even if nations are as self-interested as the *Realist Feasibility Constraint* assumes, other nations can change what it is in their interest to do via threats, sanctions, and other measures, and can therefore make it in their interest to comply with international standards that it would not have been in their interest to comply with in the absence of those measures.

As a simple illustration of how this creates trouble for the argument, imagine an international climate treaty that requires signatories to impose globally optimal emissions taxes within their territory and empowers compliant nations to impose tariffs and other penalties on non-compliant nations as retaliation for non-compliance. Given the existence of such a treaty, even if high-emitting nations would be made worse off than business as usual by compliance, they might be made even worse off by non-compliance if there are many signatories to the treaty the collective sanctions of which make the penalties for non-compliance larger than the costs of compliance. If that is the case, then the *Realist Feasibility Constraint* entails that high-emitting nations would comply and reduce emissions even though doing so would make them worse off relative

to business as usual.[21] Because this is consistent with the truth of *Circumstances*, this shows, first, that the conclusion that *Efficiency Without Sacrifice* is the best feasible option does not obviously follow from the *Realist Feasibility Constraint* and *Circumstances*, and, second, that that conclusion is false if there is any response analogous to the simplistic example just described that is both feasible and ethically superior to *Efficiency Without Sacrifice*.

Of course, in the real world the simplistic response just described is arguably not feasible. That is because if sanctions are threatened in the simplistic way just described, it is common knowledge that they are likely to be met with a retaliatory trade sanctions and/or non-trade measures that would make the costs of following through on them unacceptably high, thereby preventing the threat of such sanctions from being credible in the first place, thereby preventing such a response from getting off the ground. This is the problem with simple proposals involving trade sanctions from commentators such as Joseph Stiglitz and William Nordhaus.

8. Feasibility Wedges and a Meta-Architecture for Global Agreement

Nonetheless, despite the problems with the simple proposals involving trade sanctions, there are a number of feasibility enhancing measures that enable the creation of a feasible analogous regime that could realistically succeed in changing what it is in the interest of nations to do over time, ultimately making it in the interest of rich nations to reduce emissions even beyond the point at which they are made worse off than the business as usual status quo. Although none of these 'feasibility wedges' are individually sufficient to guarantee the success of such an ethically superior treaty, in conjunction they make the prospect of success sufficiently high to make such a response a better bet

[21] Such a structure is implicit in the proposal of Joseph Stiglitz, "A New Agenda for Global Warming", and William Nordhaus, "Climate Clubs: Overcoming Free-riding in International Climate Policy". Nordhaus provides an empirically informed model of how tariff sanctions, if they could be credibly threatened and if they would have no further effects outside the modeled interactions, could change what carbon price it is in the interest of nations to impose.

for humanity than the response represented by *Efficiency Without Sacrifice*.[22] In what follows, I sketch an example of such a response that incorporates such feasibility enhancing features both in its architecture and strategy.

So, imagine a climate treaty architecture that has three somewhat familiar components: first, a cap and trade scheme among signatories to the treaty, where the cap decreases each year along a path that is insensitive to the number of signatories to the treaty; second, an undemanding initial cap that ensures that when the treaty initially enters into force, no nation has to make costly emissions reductions in the short run; third, a right granted to compliant signatory nations to impose a duty on imports from non-compliant nations.[23] Ideally, the magnitude of this duty would be large, as could be feasible if one imagines the WTO appellate body faced with the actual choice of deciding whether a reasonable regime of this kind with a large duty was permissible under, say, Article XX of the GATT (which can be read as a catch-all exemption for any tariff that is inherently reasonable given the values that a nation endorses if it doesn't otherwise discriminate between particular nations for any of the familiar more technical reasons). In that scenario, the appellate body would be under enormous pressure to decide that it was indeed permissible. This would be for many reasons, including that deciding otherwise could threaten the long term survival of the WTO in its current form and the careers of the bureaucrats on the appellate body.[24]

---

[22] My notion of feasibility wedges is inspired to some extent by the notion of stabilization wedges in Stephen Pacala and Robert Socolow, "Stabilization Wedges: Solving the Climate Problem for the Next 50 Years with Current Technologies".

[23] Because my goal in this paper is merely to identify a mechanism by which various emissions reductions architectures could realistically be implemented in as ethical a way as is practically possible given the constraint of feasibility, I wish to remain agnostic on the question of what specific metrics, cost benefit analysis, and architecture ought to be used. Depending on those factors, such a treaty would also include an allocation scheme for emissions permits, any number of which would be consistent with the suggestions in this paper. Such a treaty might also include a mechanism to redistribute to developing nations some of the duties collected, in order to secure the universal compliance of developing nations and to offset the more ethically significant costs of the treaty to developing nations.

[24] For some reason, the literature on the practicalities of WTO permissibility seems to completely ignore this sort of 'extra-legal' consideration, which from a realist perspective might be the most important consideration at all in making a judgment of feasibility.

If well-designed, an initial set of nations would initially join such a treaty for a variety of reasons: in many cases, because the treaty would not require emissions reductions early on, but would only require reductions gradually later as the cap decreased and the largest emitters joined the treaty, at which point the cost to early signatories would in some cases tend to be offset by payments from later-joining nations in exchange for emissions permits, and in other cases those later costs would not be relevant to the short-run political calculations of politicians in democratic nations making the initial decision whether to join; in other cases, because it might be to the advantage of some initial signatories to have the power to impose unreciprocated duties on some imports from initial non-signatories; and in other cases, for the sort of complex reasons that have led many regions such as the EU, Australia, British Columbia, California, and others to independently enact such policies.

As a result, the initial set of signatories might be comparable to the set of nations that ratified the Kyoto Protocol. From this starting point, the idea is that the sanctions imposed by that initial critical mass of nations would set in motion a chain reaction that would make it in the interest of an increasing number of nations to join over time. Such a chain reaction would be driven by the fact that for each nation that joined the treaty, the cost would increase for each remaining nation that had not joined, because each additional cooperating nation means an additional nation imposing duties on imports from non-cooperating nations. In addition, as more nations joined the treaty and the cap decreased over time, the cap and trade scheme would become increasingly effective, thereby increasing the emissions differential between signatories and non-signatories, thereby increasing the magnitude of each individual duty imposed.

The idea is that over several decades, more and more nations would gradually join, until eventually the costs of not joining would be so high that even the most recalcitrant nations would ultimately find it in their interest to join rather than continue to hold out.

Of even greater importance, once nations joined such a treaty, it would never be in their interest to pull out, at least if the treaty is successfully designed with the self-

enforcing structure imagined above, because given that structure the costs of pulling out always outweigh the costs of staying in after the point at which it is initially in a particular nation's interest to join. This would solve the most serious problem for climate treaties, which is the problem of securing not merely initial ratification of the treaty, but long-run compliance.[25]

Of course, the realistic worry remains that strategic retaliatory measures by hostile powerful nations could scuttle even this sort of more sophisticated proposal. With that in mind, such a treaty should be introduced under only the most favorable conditions that can be realistically expected in order to further raise its probability of success. Toward that end, additional feasibility wedges should be identified that represent the most favorable conditions that can realistically be expected. There are a great many of these wedges, which represent an especially important place for contributions from (among other) climate negotiators and actual practitioners of international negotiation and political maneuvering. As a few examples, consider:

> (1) The treaty could be introduced early in the term of a US President who supports the treaty, enabling him or her to sign the treaty and publicly endorse its permissibility under all international laws and treaties, and to use his or her power and influence to ensure that international court decisions establish as *precedent* that the treaty is permissible. This will *institutionalize* the permissibility of the treaty in a way that cannot be reversed by anything short of dramatic (and hence unlikely) power politics. The US President can do all of this unilaterally even if the US Senate is initially disposed to reject the treaty unanimously, and even if powerful nations such as China oppose the treaty.
>
> (2) The treaty could be designed so that upon introduction, it is quickly ratified a substantial proportion of developed nations, as well as a substantial proportion of developing nations, thus enhancing its perceived legitimacy and, more generally, further institutionalizing it.

[25] For real-world examples and game-theoretic analysis of international environmental treaties that have a structure analogous to the *cascade to universal self-enforcement* structure described here, see Scott Barrett, *Environment and Statecraft*, chapter 9. A treaty is self-enforcing in the relevant sense relative to a set of nations if and only if it is both individually and collectively rational to maintain agreement to the treaty from the point of view of all of those nations.

(3) The treaty could be only one part of a policy portfolio that also includes all emissions reduction measures that are in the interest of individual nations to impose unilaterally, such as technical regulations to realize negative net cost emissions reductions, subsidies for research and development, and perhaps even geoengineering. This will reduce the magnitude of the costs that must be imposed by the treaty in order for the overall policy portfolio to be effective, thereby reducing incentive for hostile nations to invest in strategic retaliatory measures to scuttle the treaty.

(4) The cascade to self-enforcing structure described here is also consistent with other complementary incentive schemes to encourage compliance. For example, one promising addition would be for duties from all nations to be held by a single global administrator until the end of each year, at which point each nation's proceeds would be disbursed only if that nation complied with the treaty's provisions in the previous year; duties could then be subtracted from the accounts of non-compliant signatory nations based on their degree of non-compliance, with the proceeds distributed to compliant signatories. Other ideas might include a bonus for initial signatories, and many others.

(5) The treaty could be designed so that upon introduction, it tends to require only emissions reductions that nations would find feasible even without the treaty. For example, consider the 'tax and dividend' approach described several sections above that conjoins a carbon tax with an intragenerational transfer equal to equal per capita refund of all of the proceeds of the carbon tax. Perhaps such a policy is not generally feasible if it involves the carbon tax level that would be globally optimal; nonetheless, for some x% of the optimal tax level, it would be feasible. The treaty could be designed for that upon introduction, it requires on average only a carbon price equal to n% of the globally optimal level, for whatever the largest n is such that n tends to be feasible even without such a treaty. Then, sympathetic nations (or national leaders) could feasibly enact the n% tax and dividend policy, and then sell compliance with the treaty as in the self-interest of their nation, on the grounds that joining the treaty would be an improvement over a situation in which they are making the same emissions reductions but are not members of the treaty (which would be true only in the short run, but perhaps their audience tends to perceive only with the short run, in which case joining the treaty would be in their perceived self-interest even if not in their long-run self-interest).

The preceding are intended merely as initial examples of feasibility wedges that can be added to the meta-architecture for a global agreement described here in order to make an ethical collective action more feasible. For example, a treaty with the basic structure

described above makes it easier to satisfy conditions (1) and (2), because the structure of such a treaty makes it likely that it would actually be effective, which means that nations will recognize that the costs of the treaty are likely to be non-futile, including any political costs to leaders who support the treaty – which are already mitigated by the fact that the treaty shifts the costs of emissions reductions into the future, beyond the short-term time horizon of most political leaders. This increases the likelihood that (1) will be satisfied relative to any particular US President, and increases the likelihood of (2) being satisfied for the various reasons described above. In this context, it is worth noting that even the transparently flawed and ineffective Kyoto Protocol was ratified by an impressive set of nations and signed by US President Bill Clinton. As a result, it is realistic to think that a US President might well be willing to expend the political capital necessary to satisfy condition (1) relative to a treaty far superior to Kyoto, such as the sort of treaty aimed at here involving a cascade to universal self-enforcement that would also predictably involve satisfaction of (2).

For example, if (1) is satisfied by a supportive US President and (2) is satisfied by an initial coalition that includes roughly the nations that ratified Kyoto, then it is unrealistic to think that other nations hostile to the treaty would have the power and influence to convince international courts to rule against both the US administration and that coalition regarding the permissibility of the treaty under WTO rules, which together with the resulting institutionalized legitimacy of the treaty would substantially reduce the probability that subsequent retaliatory measures could scuttle the treaty.

With all of these pieces of the puzzle in hand – feasibility wedges together with the basic meta-architecture described above – it is important to see how they fit into a strategic dynamic at both the national and international level that can be leveraged for faster progress toward optimal policy, and that is emerging in the wake of the pivot away from trying to achieve a Kyoto-like 'grand global agreement' directly, and instead aiming to begin with a more 'bottom-up' strategy for addressing climate change. One key to this dynamic is that the biggest moves in the right direction are now the result of nations learning more about co-benefits and their own vulnerabilities, even within our lifetimes, and adopting unilateral emissions reductions that are in their perceived self-interest,

where that self-interest is becoming more enlightened and in favor of deeper emissions cuts all the time. The second key to this dynamic is that, in light of the first, nations increasingly see it in their interest to join 'clubs' of like-minded nations seeking to maximize the benefit of the emissions reductions that they have self-interested reason to make unilaterally. The third key to this dynamic is that the resulting reductions are often most easily achieved initially via much-less-than-efficient policy that is for various reasons initially more feasible, such as the US EPA command and control Clean Power Plan, which is much more regressive and inefficient than any policy for achieving those reductions that would be considered by an economist, primarily because of the constraints on policy imposed by the existing regulatory authority.

In light of these three key features of the situation, the strategic dynamic that emerges is that leaders of nations can increasingly credibly threaten some of their climate *opponents* (most importantly, other nations, and powerful industrial interests within their own nations) that they have the power to and will in fact make some substantial emissions reductions, and that those reductions will be made more costly to their opponents if they do not cooperate. For example, within the US, powerful industrial interests perceive that unless they are able to defeat the EPA Clean Power Plan in court, they will have to deal with a policy that is more costly to them than a tax and dividend policy would be that achieves the same level of emissions reductions. This creates incentive for them to favor the more efficient policy, even though that policy makes them worse off than the no policy status quo (as the no policy alternative is no longer available, and they now face the choice between a more efficient policy and the more costly policy default). Then, once the more efficient policy is in place, it becomes easier in the longer run to ratchet emissions down even further, partly because the cost of each unit of emissions reductions is then much lower. As this strategic dynamic plays out repeatedly in the coming decade, it will be increasingly clear to opponents that emissions reductions are coming in one form or another, and that they are better off

getting ahead of the curve with support of efficient policy, cooperation, and more low-carbon investment decisions.[26]

By recognizing this emerging strategic dynamic and leveraging it as fully as possible, including by using the tools of feasibility wedges and intra- and intertemporal transfers described above, we can move things in the right direction as quickly as possible even within the constraints of realism.

With the preceding outline in hand, a climate treaty that exploits the feasibility wedges above and the particular cascade to universal self-enforcement design structure outlined above, together with additional tools such as intra- and intergenerational transfers is a promising route to an ethically superior climate treaty that is feasible even in light of the *Realist Feasibility Constraint* and *Circumstances*. The challenge of crafting such a treaty is then a more well-defined engineering problem that can be addressed by experts on the relevant legal issues, economic mechanisms, dispositions of nations, and political strategy. The primary constraint on such a treaty is that it must predictably lead to universal ratification and compliance in the long run via the cascade of economic incentives described above. Subject to that constraint, the treaty should be designed to distribute the costs of emissions reductions as fairly as possible, which will include much more fair distributions than would be available if there were a genuine feasibility constraint that current citizens of rich nations can be made no worse off than under the no policy status quo. If this engineering project succeeds, it will result in a treaty-based response that is a far more ethical bet for humanity than the response intended under *Efficiency Without Sacrifice* and other contemporary 'realist' proposals.

Of course, even when all of the pieces of the puzzle above are combined, there is no way of establishing that this kind of more ethical response would ultimately succeed with anything near 100% certainty – but as noted above, there are analogous doubts

[26] As we are already see with substantial carbon pricing within the internal accounting of energy companies. See also world bank embargo on coal power financing, http://www.reuters.com/article/2013/07/24/eu-coal-finance-idUSL6N0FU32R20130724

http://www.nytimes.com/2013/10/30/us/us-says-it-wont-back-new-international-coal-fired-power-plants.html

about the success of other realist proposals including even the best versions of *Efficiency Without Sacrifice*. But in general, it would be surprising if the best feasible way forward was *obviously* feasible. Instead, we should expect the best feasible way forward to appear infeasible to many, as did the best feasible way forward at the height of US-Soviet tensions, and at many other points in world history after which a better solution than seemed feasible at the outset emerged from a combination of skilled diplomacy, good fortune, and bold leadership.

[END]

Bibliography

Joseph Aldy and Robert Stavins (eds.), *Architectures for Agreement*

aldy and stavins, 2 other volumes

paper from preceding on BTAs

Anthoff, job market paper

Anthoff and Tol, on Schellings conjecture in Climate Change and Common Sense

Scott Barrett, *Environment and Statecraft*

Scott Barrett, *Why Cooperate?*

Marisa Beck, paper on revenue recycling

Geoff Brennan and other Geoff Sayre McCord on Broome

John Broome, *Climate Matters: Ethics in a Warming World*

John Broome, "The Public and Private Morality of Climate Change", Tanner Lectures

Mark Bryant Budolfson, Collective Action, Climate Change, and the Ethical Significance of Futility, phd dissertation

Mark Budolfson, "Should the US and other Nations Make Unilateral Emissions Reductions? How to be a Leader Without being a Sucker"

Budolfson and Dennig, paper on Optimal carbon prices allowing differential regional prices

Budolfson, Dennig, Fleurbaey, Siebert, Socolow, under review, The Comparative Importance for Optimal Climate Policy of Discounting, Inequalities, and Catastrophes

Dennig, Budolfson, Fleurbaey, Siebert, Socolow, 2015, Inequalities, Carbon Prices, and the Future Poor.

Dena Gromet, Howard Kunreuther, and Richard Larrick, "Political ideology affects energy-efficiency attitudes and choices", 2013 *Proceedings of the National Academy of Sciences*

EMF (Energy Modeling Forum) Session on Leakage, published in their journal

EMF (Energy Modeling Forum) Session on Revenue Recycling [intratemporal transfers], forthcoming

Heal and Chichilnisky

Metcalf, several papers on revenue recycling

Nordhaus and Boyer, Warming the world

Nordhaus, Climate Casino

Nordhaus, Climate Clubs

Weitzman and , book

david keith

victor and keohane, regime complex for climate change paper

victor and keohane, new paper

David Victor, new climate clubs paper

David victor, gridlock book

RFF several sessions on revenue recycling

Sabel and Victor, climatic change paper

Stephen Pacala and Robert Socolow, "Stabilization Wedges: Solving the Climate Problem for the Next 50 Years with Current Technologies"

Joost Pauwelyn, "U.S. Federal Climate Policy and Competitiveness Concerns: The Limits and Options of International Trade Law"

Eric Posner and David Weisbach, *Climate Change Justice*

Eric Posner and Cass Sunstein, "Climate Change Justice"

Rendall, leakage paper

Thomas Schelling, *The Strategy of Conflict*

Schelling 1995 etc

Nicholas Stern et. al., *The Economics of Climate Change: The Stern Review*

Richard Stewart and Jonathan Wiener, *Reconstructing Climate Policy: Beyond Kyoto*

Joseph Stiglitz, "A New Agenda for Global Warming"