

Collective Action, Climate Change, and the Ethical Significance of Futility

Dissertation Summary

Mark Bryant Budolfson

Some collective action problems are easy. For example, if a serious tragedy of the commons arises within a nation's borders, and if it is clear that the problem can be solved only by a particular kind of coercive governmental response, and if it is clear that every citizen would want the government to adopt such a response, then the government should adopt such a response, and such a response is justified. This is an easy collective action problem because it is obvious *what* course of action should be taken, and it is also clear enough *why* that course of action would be justified.¹

In this dissertation I answer questions that arise from the most difficult collective action problems, such as our situation with respect to climate change, with respect to which neither economists nor philosophers have made any real progress. For example, in the first part of the dissertation I answer the following question:

What should we do when we face a collective action problem that is impervious to our standard tools for achieving cooperation and avoiding a tragic outcome? In other words, what should we do when we are unable to rely on benevolence, natural 'free-market' coordination, bargaining, governmental coercion, and other familiar mechanisms to avoid an undesirable outcome?

My answer is that even when our standard tools are useless, we can often escape a tragic dilemma by turning the force of such a dilemma against itself. The trick is to intentionally place ourselves into a new dilemma of our own design, the predictable outcome of which is a desirable solution to the initial, otherwise insoluble dilemma. For example, with respect to anthropogenic climate change, where any effective response will be contrary to the interests of the current citizens of many nations, the trick is to engineer and introduce an international climate treaty that creates a multi-player prisoner's dilemma at the level of nations, the predictable outcome of which is that all nations will agree to the treaty, thereby making the current citizens of some of those nations worse off relative to the no-treaty status quo as a side-effect of a mechanism within the treaty that effectively combats climate change. I show that such a treaty is the only attractive

¹ For example, consider taxation to provide a public good that is more valuable to each taxpayer than the tax required to provide it, where that good cannot plausibly be provided in any other way. In many nations, law enforcement, national defense, and national parks are approximate examples of such goods.

option if – as seems true – we must secure the cooperation of many influential nations whose current citizens will be made worse off by any effective response, and if such nations are disposed to reject treaties whenever they are contrary to the interests of their current citizens. I also explain why such a treaty is the key to engineering the most *ethical* climate treaty possible, and to making genuine progress beyond the unpromising *incrementalist*, *idealist*, and *hard-headed realist* approaches to climate treaties that dominate the literature in philosophy, politics, economics, and law.

In the second part of the dissertation, I show that nations can be justified in acting in ways that make their citizens worse off, are unfair to their citizens, and that would be reasonably rejected by their citizens – which provides a counterexample to many political theories. Such situations arise most clearly when the citizens of a nation are, through no fault of their own, morally required to favor a national course of action that is unfair to them, contrary to their interests, and contrary to the interests of future generations of their descendants and fellow citizens. Arguably, the United States is in such a situation with respect to climate change, given that inaction on the part of the US may well ensure a catastrophic outcome for untold billions of future people, and given that any real solution to the problem arguably requires the US to agree to a climate treaty that is unfair to its citizens and contrary to their interests.²

At the same time, even given worst-case-scenario assumptions about the effects of climate change, I argue that *individual citizens* are not required to reduce their emissions by a significant amount, because without an effective global response such reductions are both costly and futile in a way that makes such reductions not required. In fact, I argue that such futility and lack of requirement at the individual level explains why individuals are required *to favor* a coercive intergovernmental response to climate change, and why nations are justified in adopting such a response even if it is unfair to their citizens and contrary to their interests. (Here it might be useful to recall that in the first part I explain why an effective intergovernmental solution is realistic even given pessimistic assumptions about the collective action problems that arise at the level of nations.)

In the course of arguing for these conclusions, I make progress on a number of issues at the foundations of normative ethics and political philosophy. Among other things, I analyze the ethical significance of futility, distinguishing between, on the one hand, cases in which actions are required despite being futile in some intuitive sense, and, on the other hand, cases in which actions are not required because they are futile in a more complete sense that makes actions genuinely not required. Futility has not been adequately investigated by philosophers and other theorists, and is of great practical and theoretical importance, because worries about futility arise

² My own view is that an effective and ethical response need not be unfair to the United States or contrary to its long-run interests. Such a view of the empirical facts does not undermine my argument that a nation *could* be justified in acting in a way that is unfair to its citizens and contrary to their long-run interests if the empirical facts were different.

in almost every area of practical concern, and threaten to undermine most normative theories. For example, I show that existent normative theories are unable to offer a plausible account of what individuals are required to do in the kind of collective action situations that are common in a market-based society, and are therefore unable to explain many of the most important facts about modern moral life. In particular, straightforward mathematical and empirical considerations show that an appeal to expected consequences cannot possibly deliver the verdicts on such cases that consequentialists themselves assume, and I show that alternative theories that appeal to ‘universalizability’, ‘direct harm’, and other notions also cannot deliver plausible verdicts on such cases. For these and other reasons, I argue that a plausible account of what individuals are required to do in a large market-based society must invoke a distinction between activities that are *essential* to a product or to the actual production of a product, and activities that are not. These and other results related to futility have important practical implications, because many serious practical issues are tied to situations in which the action of individuals is in some sense futile – for example, whether it is permissible to consume animal products given the terrible way that many animals are treated, whether it is permissible to consume products that are produced in a way that harms or exploits other people, whether individuals are required to contribute to charities that help innocent people who are suffering from easily curable afflictions, and, of course, whether individuals are required to reduce their greenhouse gas emissions.

In the final chapters, I show that morality and other forms of normativity are sometimes *dramatically directly collectively self-defeating*, where a normative system has that property when it requires everyone to act in a way that everyone can see is certain to be dramatically worse along the relevant normative dimension than if everyone did not follow the requirements of that normative system instead. This shows that a wide range of normative theories are either false, or at least don’t have the consequences that their adherents take them to have, including consequentialist theories, contractalist theories, Kantian theories, universalization theories, enlightened self-interest theories, and many other normative theories. It also means that morality and other forms of normativity cannot be relied upon to solve collective action problems even in a world of normatively flawless agents. One practical upshot is that even when a disaster will ensue if everyone acts in a particular way or on a particular principle, that doesn’t settle the question of whether individuals are permitted to act in that way or on that principle.

In sum, I provide answers to the most pressing questions about what we should do about difficult collective action problems, especially when non-cooperation would lead to catastrophe.

Advisors: Gideon Rosen, Peter Singer

Full Text: www.budolfson.com/papers/BudolfsonDissertation.pdf