

The moral inefficacy of carbon offsetting*

Tyler M. John, Amanda Askill, & Hayden Wilkinson

Abstract

Many real-world agents recognise that they impose harms by choosing to emit carbon, e.g., by flying. Yet many do so anyway, and then attempt to make things right by offsetting those harms. Such offsetters typically believe that, by offsetting, they change the deontic status of their behaviour, making an otherwise impermissible action permissible. Do they succeed in practice? Some philosophers have argued that they do, since their offsets appear to reverse the adverse effects of their emissions. But we show that they do not. In practice, standard carbon offsetting does not reverse the harms of the original action, nor does it even benefit the same group as was harmed. Standard moral theories hence deny that such offsetting succeeds. Indeed, we show that any moral theory that allows offsetting in this setting faces a dilemma between allowing *any* wrong to be offset, no matter how grievous, and recognising an implausibly sharp discontinuity between offsettable actions and non-offsettable actions. The most plausible response is to accept that carbon offsetting fails to right our climate wrongs.

*We are grateful to Christian Barry, D Black, John Broome, Garrett Cullity, Johann Frick, William MacAskill, Holly Smith, and Alec Walen for helpful comments and discussion, as well as to audiences at the 2017 Ethics of Giving Conference in St. Andrews, the 2017 Rocky Mountain Ethics Congress, the University of Melbourne, and the Global Priorities Institute in Oxford. For financial support, we thank: the Forethought Foundation for Global Priorities Research, which supported John's contribution to this article through a Global Priorities Fellowship; Rutgers University, which supported John's contribution through a Presidential Fellowship and Graduate Mentor Fellowship; and to New York University Graduate School of Arts and Science, which supported Askill's contribution through a Henry M. MacCracken PhD Fellowship; and the Mellon Foundation, which supported Askill's contribution through a Mellon/ACLS Dissertation Completion Fellowship.

1 Introduction

Consider the following, perhaps familiar, decision.

Carbon Offset

Nikita wants to fly from London to Thailand for a holiday. But she is concerned about the resulting greenhouse gas emissions and the harm they will do, particularly to the global poor.¹ So she donates to an organisation that will plant additional trees to capture at least the same quantity of greenhouse gases in the soil.

Nikita may think that, without her donation, it would be morally impermissible to fly. After all, the additional carbon emitted will do great harm in expectation, especially to the global poor—harm of far greater magnitude than the benefit to her in the form of the holiday.² We will assume here and throughout that she would be right, that it is impermissible.³ It also seems *prima facie* plausible that, so long as she *offsets* that act by donating to the tree-planting organisation, her choice to fly is permissible.

But, in cases with a seemingly similar structure, intuition suggests otherwise: that certain harmful acts cannot be rendered permissible by a mere offset. For example, consider the following, more incendiary case:

Arsonist's Indulgence

Marvin gets great pleasure from committing arson. To satisfy a burning desire of his, he torches a house with a family of six still inside. He knows that this action is wrong so, in an attempt to extinguish its wrongfulness, he donates a large sum to an organisation that improves the fire safety of new construction in low-income countries, thereby preventing at least as many similar deaths.⁴

It seems clear that Marvin acts wrongly, whether or not he makes the donation. Yet his actions appear analogous to Nikita's. Both agents perform a harmful action accompanied by a beneficial action, with the harm and the benefit similar to each other in both kind and magnitude. And both perform the beneficial action with the intention of rendering the harmful action permissible.

In practice, many agents attempt to offset their harmful actions. Most commonly, many fossil-fuel-burning consumers and corporations pay non-profit organisations to trap greenhouse gases in the soil. But

¹See, for instance, Morton (2007).

²For instance, the total emissions of her flight will, in expectation, cause 0.0009 additional deaths this century, concentrated among the global poor. This is calculated based on an estimate of 2.26×10^{-4} additional deaths per tonne of CO₂ emitted in 2020 (Bressler, 2021), a round-trip distance from London to Bangkok of 19,063 kilometres, and an estimate of 1 tonne of additional CO₂ emitted per 3,077km travelled per passenger on long-haul flights. Although an expected 0.0009 deaths may not seem like much, it equates to approximately two weeks of life lost, in expectation.

³Readers sceptical of this assumption may substitute "flying to Thailand for a holiday" with any behaviour that causes gratuitous carbon emissions and so is *pro tanto* wrong. Analogous arguments will apply.

⁴A similar example appears in the television serial *Sherlock*, in which a murderous villain states: "There are charities that I support who wouldn't exist without me. If life is a balance sheet, and I think it is, well, I believe I'm in credit!" (Moffatt et al., 2017).

even beyond this context, in environmentally destructive industries, some firms offset actions that result in the loss of biodiversity (Environment Bank, 2012), forests (Zwick, 2014), clean water (Kaye, 2012), and air (Food & Water Watch, 2012). Some meat-eaters even donate to charities that improve the welfare of farmed animals to make up for the harms they cause (Alexander, 2015). A former parody website CheatNet even satirically recommended offsetting cheating on one’s romantic partner by paying someone else not to cheat on their own partner (MacAskill, 2015).⁵

Do such agents succeed, particularly in the context of carbon offsetting? By offsetting the harms they impose, do agents like Nikita make permissible their otherwise impermissible actions, as intuition might suggest? Or does their behaviour remain impermissible, as is suggested by the analogy between their behaviour and Marvin’s in Arsonist’s Indulgence?

Others have claimed that carbon offsetters do succeed (Broome 2012 ch. 5; MacAskill 2015 p. 174; Lawford-Smith 2016). The distinction between Carbon Offset and Arsonist’s Indulgence, they claim, lies in the cases’ different causal structures. When emitting greenhouse gases and then removing some such gases from the air, it seems that an agent can prevent any harm at all from resulting from their prior emission; the wrongmaking features of the action are then removed. And this differs from Arsonist’s Indulgence. There, the agent harms some subjects and then benefits others, failing to prevent that harm to the original victims—behaviour which is typically prohibited by non-consequentialist theories (at least when less harmful actions are available).

But, in practice, this disanalogy does not hold up to scrutiny. In this paper, we argue that carbon offsetting has a causal structure similar to Arsonist’s Indulgence.⁶ Further, we argue that carbon offsetters act impermissibly and, more generally, so do all agents who perform offsetting behaviour with this structure. And this is not based only on the analogy to Arsonist’s Indulgence; we prove a troubling impossibility result, by which we must accept this conclusion or else admit far more implausible implications.

The paper proceeds as follows. In Section 2, we develop a helpful taxonomy of forms of moral offsetting based on their causal structure. In Section 3, we argue that carbon offsetting inevitably falls under a specific such category. In Section 4, we discuss the permissibility of that category of offsetting according to both standard non-consequentialist and consequentialist theories, providing tentative arguments that carbon offsetting must be unsuccessful. We also consider a key objection: that perhaps offsetting behaviours should not be categorised based on the causal structure of the harms they impose and prevent, but instead by how they change the *risky prospects of harm* to victims. By some views—including some forms of *ex ante* contractualism—Carbon Offset will be innocuous, while Arsonist’s Indulgence remains morally problematic. We show that these views give implausible verdicts in other pernicious offsetting cases, so do not provide an adequate account of offsetting. Then, in Section 5, we provide a general argument that carbon offsetting (and indeed *any* offsetting behaviour with the same causal structure as Arsonist’s

⁵A practice similar to offsetting was also sanctioned by the Catholic Church until 1567. By donating to the church or performing good deeds, sinners obtained *indulgences*, which allowed them to avoid divine punishment for their prior sins.

⁶An analogous point has been made independently by Stefánsson (2019), which came to our attention while preparing the final version of this article.

Indulgence) must fail. We do so via a troubling impossibility result, which implies that either all such offsetting attempts must fail or else we must face some even more implausible implication. Given this result, we do not think that carbon offsetting, as it is currently practiced, is morally defensible.⁷

2 A taxonomy of offsetting

Take an agent who has performed some action that, without any further intervention, will do harm (in the presence of permissible, less harmful alternatives). As we define it, for that agent to *offset* that action is for them to perform some separate, beneficial action—beneficial to at least the same degree as the harmful action was harmful—with the intention of rendering the harmful action permissible.⁸ Defined this way, there are several distinct categories of actions that count as offsetting. And which of these categories an act falls under will affect whether it succeeds in affecting the permissibility of the harmful action.

The first such category is offsetting that prevents the harms of the original action. Suppose for instance that Hiroki drops a brick over Gemma’s foot, but then catches the brick before it strikes her. This latter action—the offset—counts as *Harm Prevention*, as he reduces the harm of his initial action to zero.⁹

Harm Prevention: An action performed in conjunction with another, otherwise impermissible, action that prevents that other action from resulting in any harm.

Other offsets do not prevent the original harm, and instead attempt to make up for that harm with a separate benefit. Among such offsets are those that provide benefits to the same individual who is harmed. For instance, suppose Hiroki has dropped the brick on top of Gemma’s foot, it has struck her, and her cuboid has fractured. With the intention of making up for it, Hiroki pays Gemma’s hospital bills, helps her with tasks around the house, and perhaps gives her additional cash compensation. These acts count as *Compensatory Offsetting*¹⁰ as, no matter how much compensation Hiroki provides Gemma, he has still harmed her and wronged her.

Compensatory Offsetting: An action performed in conjunction with another, otherwise

⁷Note that there are several related issues that we will not address. The first is that offsetting raises various practical and political worries, including: that common offsetting strategies are ineffectual in providing benefits (Beder, 2014); and that they give public legitimacy to (or ‘greenwash’) harmful practices and industries. We set aside such worries aside here; we are concerned with the first-order normative question of whether offsetters act impermissibly. The second issue is: the closely-related normative question of whether we ought to encourage offsetting by others. We remain agnostic on this, as it is not directly relevant to our discussion. And the third issue: whether offsetting succeeds in contexts other than carbon emissions; our discussion will have implications for other contexts but, due to the relative ubiquity of carbon offsetting, we will focus on carbon emissions here.

⁸Alternatively, offsetting might be defined as any such action performed with *or without* that intention. Or an offset may be defined as a mere *feature* of an action, rather than a separate action itself—the behaviour being offset and the offset itself may be combined into a single action. We suspect that such modified definitions fit less neatly with our ordinary language notion of offsetting. It also seems somewhat less plausible that these less demanding forms of offsetting succeed—if the more demanding form of offsetting that we consider fails in rendering the initial action permissible, then these surely fail too.

⁹We can assume that Hiroki is certain that he will catch the brick and that the act does not constitute a wrongful expression.

¹⁰Paying Gemma’s hospital bills may also count as a partial form of Harm Prevention if Gemma would otherwise not have healed as quickly or completely.

impermissible, action that benefits an individual who has been or will still be harmed as a result of the harmful action.

Other offsets provide a separate benefit but to different individuals altogether. For instance, Hiroki might instead make a donation to Steel Toes International, a (fictional) charity that helps protect the feet of bricklayers from on-the-job accidents. Hiroki does not thereby prevent his initial harm nor compensate Gemma. Instead, like Marvin in Arsonist’s Indulgence, he attempts to make right his action by preventing others from coming to similar harm; he engages in *Non-Compensatory, On-Target offsetting*.

Non-Compensatory, On-Target Offsetting: An action performed in conjunction with another, otherwise impermissible, action that benefits people who were not and will not necessarily be harmed as a result of that other action and that benefits them by averting similar misfortune.

But Hiroki might instead make that donation to UNICEF, a charity that provides humanitarian and developmental assistance to children and mothers in developing countries. Hiroki thereby does not compensate Gemma, he does not prevent his initial harm, nor does he prevent the same type of harms from befalling others. He instead prevents quite different harms to quite different people; he engages in *Non-Compensatory, Off-target Offsetting*.

Non-Compensatory, Off-Target Offsetting: An action performed in conjunction with another, otherwise impermissible, action that benefits individuals who have not and will not necessarily be harmed as a result of that other action and that benefits them in a manner different from averting misfortune similar to that which that action imposes.

3 Under which category does carbon offsetting fall?

Several philosophers supportive of carbon offsetting practices have suggested that they are instances of Harm Prevention. For instance, Broome (2012, p. 85) claims that: “If you successfully offset all your emissions, you do no harm by emissions. You therefore do no injustice by them.” Similarly, MacAskill (2015, p. 174) argues that “...carbon offsetting prevents anyone from ever being harmed by your emissions; it’s the ‘equivalent’ of never...” emitting greenhouse gases in the first place. And Lawford-Smith (2016) claims that the obligation to offset carbon arises from the fact that “[w]hen we offset our GHG emissions, we neutralize the harms they might otherwise do.” According to these philosophers, to emit and then offset is the same as not emitting in the first place, both in terms of permissibility and in terms of the harm to each of your potential victims.¹¹ We suspect that the typical non-philosopher who offsets their emissions believes much the same.

¹¹Partial exceptions to this pattern include Cullity (2019) and Barry and Cullity (2022). Both deny that emitting and then offsetting is (always) morally equivalent to not emitting at all (in particular, see Cullity, 2019, p. 26). But both appear to endorse the descriptive claim: that the outcome of emitting and offsetting is the same as not emitting at all.

If the Earth's climate and weather systems worked in a particular way, they would be correct. If all of the harms done by releasing greenhouse gases were determined only by their total mass in the atmosphere at some later date, then there would be no harm done by increasing that total mass and then reducing it by the same amount.

But this is not how the Earth's climate and weather systems work. As Broome (2019) points out, our atmosphere forms a highly chaotic system, such that "...a very small change at one time and place can escalate to cause very large differences in the weather all over the world just a few weeks later." Broome argues:

Given this instability, we should expect global weather in, say, fifty years time to be entirely different if you go joyriding on Sunday from what it would have been if you stayed at home. Your increasing emissions will not cause gradual changes punctuated by occasional discrete events such as a typhoon or a child's death from cholera... Instead they will cause typhoons to form at quite different times and places, and they will lead to a completely different distribution of cholera outbreaks. Your Sunday drive will cause a completely different group of people to be exposed to cholera and other risks of death. Some who would have died will survive because of your drive, and others who would have survived will die. (Broome, 2019, p. 113)

As Edward Lorenz (1972) famously speculated, the chaotic nature of the Earth's atmosphere suggests that an additional flap of a butterfly's wings in Brazil can change the path of a tornado in Texas a week later. As it turns out, a butterfly's wing flap cannot (as it carries too little energy to cut through the air's viscosity). But removing multiple kilograms of carbon dioxide at one location and adding the same at another location could, according to more recent atmospheric physics.¹² As Lorenz puts it, "Small errors in the coarser structure of the weather...tend to double in about three days. ... Small errors in the finer structure—e.g., the positions of individual clouds—tend to grow much more rapidly, doubling in hours or less. ... Errors in the finer structure, having attained appreciable size, tend to induce errors in the coarser structure." Suppose you make *even a small change* to the atmospheric system—equivalent to a small 'error'—such as moving multiple kilograms of air from one location to another. The effects of doing so will, in almost all cases, grow and grow, cascading into larger and larger changes to weather patterns. Almost always, this results in weather events (including cyclones, heat waves, and monsoons) decades in the future occurring at times and places very different to those at which they'd occur otherwise. So, make such a change and you cause some people to lose homes, fall ill, and even die when they otherwise wouldn't, and some to avoid suffering such calamities when they otherwise would. And this is exactly what we do when we emit in one location and offset by preventing emissions elsewhere.

This empirical fact deals a serious blow to the Broome-MacAskill-Lawford-Smith account of carbon offsetting. Emitting and then offsetting does not leave the world as it would otherwise be, and thereby

¹²For a contemporary, technical discussion of this result, of Lorenz' claims below, and of the implications we draw from them, see Palmer et al. (2014). As Palmer and colleagues show (in greater technicality than can be adequately summarised here), these claims do appear to hold up under our current understanding of atmospheric physics. See also Tribbia and Baumhelfner (2004); Morss et al. (2009).

prevent all harm. Instead, emitting greenhouse gases during one’s flight over the Atlantic Ocean in May results in one sequence of cascading changes to future weather events, while trapping greenhouse gases by reforesting the Congo Basin in June will result in another such sequence, which does not cancel out the first. Overall, countless people will still be left worse off due to changes in the incidence, timing, and location of future hurricanes, tornadoes, monsoons, and droughts.¹³

Has one therefore done those people harm? It is not entirely clear. Although one’s actions do cause people to be worse off, the mechanism by which this occurs is indirect, unpredictable, and inscrutable. And it is certainly counterintuitive to think that merely by a minor carbon-emitting act one could be responsible for, say, all of the harm done by a tornado or cyclone. But if we deny that one has harmed those people, we face an even more counterintuitive implication: that one cannot do *any* harm through one’s emissions. After all, the harms of carbon emissions are in general indirect, inscrutable and unpredictable in a similar way. But it would be absurd to say that one therefore cannot do harm by emitting.¹⁴

So, yes, standard practices of carbon offsetting benefit some and harm others, leaving the original victims of one’s emissions uncompensated. Such offsetting does not count as Harm Prevention but as Non-Compensatory, On-Target Offsetting.¹⁵ Notably, this is the same category as that of Marvin’s behaviour in Arsonist’s Indulgence. But, like Marvin, do carbon offsetters engaging in Non-Compensatory, On-Target Offsetting still act wrongly? We address this in the next section.

4 What do normative theories have to say?

Can Non-Compensatory, On-Target Offsetting succeed in rendering the prior agent’s prior action permissible? Can any other forms of offsetting? The answer will depend, in part, on other normative commitments.

4.1 Standard non-consequentialist theories

For some of the forms of offsetting described above, it is fairly clear what standard non-consequentialist theories will say. Harm Prevention, for instance, may indeed succeed—if the harm is completely eliminated, a theory that takes act sequences as its objects of evaluation may well permit an act sequence involving an otherwise wrongful action and Harm Prevention. And Compensatory Offsetting is already widely discussed

¹³Others may be made better off too (cf. Mogensen and MacAskill, 2021). See also Barry and Cullity (2022, p. 362).

¹⁴To deny that one can do harm through indirect, unpredictable, and inscrutable mechanisms leads to further absurdity. For instance, consider the following case:

Mystery Machine

A zany inventor shows you her latest creation: a large, elaborate machine, with a single button. It is an assembly of countless distinct mechanisms that interact in some indiscernibly complex manner. She tells you that, if you press the button, the machine will perform its convoluted operation and, once complete, will thereby cause a randomly-selected person in a nearby town to die a painful death.

Yes, pressing the button would cause someone’s death only via an indirect, inscrutable, and unpredictable mechanism. But it seems absurd to say that you would not thereby do harm to the victim.

¹⁵Some carbon offsetting approaches would still count as Harm Prevention, e.g., adding a filter to your factory’s smokestack to sequester any of the carbon that would otherwise be emitted. But such practices are rarely available to individual offsetters.

in the criminal justice literature, where it is often taken to be governed by the logic of conditional obligation: making a victim whole does not erase one's wrong, but it might be an additional wrong to *fail* to make a victim whole (Gardner, 2011; Raz, 2004).

But it is not so clear what such theories have to say about Non-Compensatory Offsetting. After all, such offsetting behaviour has rarely been discussed in philosophical literature.¹⁶

Victor Tadros (2011) has argued to the effect that Non-Compensatory Offsetting can play the normative role of Compensatory Offsetting only in very particular cases: if *A* has wronged *U* and *B* has wronged *V*, and *A* can more efficiently compensate *V* than *U* and *B* can more efficiently compensate *U* than *V*, then *A* and *B* can be required to 'swap' duties, such that *A* is required to compensate *V* and *B* is required to compensate *U*. In such highly specific cases, argues Tadros, Non-Compensatory Offsetting, too, can be governed by conditional obligation, compensating for wrongs done. Ordinarily, however, we regard Non-Compensatory Offsetting as inadequate to play the normative role of compensating for wrongful action. If I have harmed you, I owe it to you to make it up to you, and cannot make up for that wrong by benefiting someone else entirely. But this does not make the initial harmful action permissible.

Christian Barry and Garrett Cullity (2022, p. 363) also object to actions resembling Non-Compensatory, On-Target Offsetting in their case of *Two Rivers*.

Two Rivers There are two local rivers in which our townsfolk engage in [leisure boating, which discharges toxins that exposes those living downstream to risks of death], with a different downstream town on each. I go boating in one river, adding my toxic discharge to it. I run my extractor in the other [thereby reducing the toxicity of that river by the same amount as my boating increases it in the other river].

Barry and Cullity note that the (risks of) harms to each downstream town do change here. The harms to the first population require a justification and, by standard non-consequentialist theories, comparable benefits to a *separate* population do not provide sufficient justification. So, even if those harms and benefits resulted from a single action, that action would not be permissible. And if split them into distinct actions, the agent surely still acts impermissibly.

There is a further non-consequentialist argument against the moral efficacy of Non-Compensatory On-Target Offsetting. And that is that it may stand or fall with *Off-Target* offsetting and, intuitively, *Off-Target* Offsetting does not succeed.

Intuitively, On- and Off-Target Non-Compensatory Offsetting may seem morally very different. Recall the other practical examples offsetting raised earlier: of offsetting cheating, consumption of animal products, and environmental harms. In these cases, an offsetter cheats on their romantic partner and, in exchange, prevents someone else from doing the same; an offsetter causes the suffering of nonhuman animals and, in exchange, prevents the suffering of other animals; an offsetter pollutes an ecosystem and,

¹⁶See the recent Barry and Cullity (2022) for a notable recent exception.

in exchange, cleans a different ecosystem. Each of these offsetting actions is *On-Target*. Indeed, we know of no action popularly described as offsetting that is *Off-Target*, preventing harms very different from those harms which they are intended to make up for. This suggests that we intuitively regard On-Target offsets as more morally appropriate than Off-Target offsets.¹⁷

What could explain why On-Target Non-Compensatory offsets are morally preferable to Off-Target Non-Compensatory offsets? The only difference between Non-Compensatory, On-Target Offsetting and Non-Compensatory, Off-Target Offsetting is that the former prevents harms that are *similar to* those caused. But it hardly seems that *mere similarity* to harms that one has committed or will commit could in itself give one greater moral reason to prevent such harms.

One explanation for the moral difference between On-Target Offsetting to Off-Target Offsetting is that we can compensate groups for harms that occur to other members of that group. But this seems implausible.¹⁸ First, when attempting On-Target offsetting of carbon emissions, it is not clear that we are even compensating members of a coherent group. Entirely different people around the world seem to share little morally-relevant resemblance simply because they are victims of climate change. Second, even if there is a coherent group of victims, it is highly unclear that we can justly compensate some people for harms committed against others *merely* because of group membership. And, if not, then On-Target offsetting cannot succeed if Off-Target offsetting fails.

But we think this final argument is not ultimately successful. There are disanalogies between On-Target and Off-Target offsetting. The primary disanalogy is based on our *obligations to signal*.¹⁹ Individuals responsible for harms of particular kinds to particular groups incur obligations to signal their moral commitment to avoiding and averting such harms. These obligations arise from two places. First, signalling that such harms are bad and to-be-avoided has positive expected welfare effects insofar as it encourages others to see these harms as bad and to-be-avoided, rather than embracing an attitude of complacency towards these harms. Second, and derivatively, such signalling can serve as a first step in collectivizing to act against the unjust social practices in which one has participated (Lawford-Smith, 2015). On such a view, mere similarity does not matter as such; it only matters instrumentally. Preventing harms similar to harms we cause correlates with our duties to signal our commitment to opposing and eradicating the harms and wrongs in which we are complicit.

Taking stock, on standard forms of non-consequentialism, On-Target Non-Compensatory fails to nullify one's wrongdoing. Indeed, Harm Prevention is the only form of offsetting that can succeed in this manner.

¹⁷When faced with a hypothetical case where the offsetting is off-target, as in the fictional example noted in Footnote 4, intuition suggests that such actions are morally abhorrent.

¹⁸*Cf.* Independent discussion of this problem in Chan and Crummett (2019).

¹⁹A second disanalogy is based on disrespect. Unlike Off-Target Offsetting, On-Target Offsetting may help one to adopt respectful attitudes towards the victims of one's harmful actions. And whether one holds or adopts such attitudes can plausibly make the difference to an act's permissibility (see Liao, 2012; Sverdlik, 2011)—change one's attitude and one's action may well become less seriously wrong.

4.2 Ex ante non-consequentialist views

You might object that, in the above, we have mischaracterised both Harm Prevention and what non-consequentialist views have to say about it. You might think that the actual harms our actions cause are not morally relevant *per se*; that the morally relevant consideration is instead the *risks* of harm we impose on others (see, e.g., Frick, 2015). If so, you may wish to distinguish some of the cases we have called Non-Compensatory, On-Target Offsetting by labelling some as *Ex Ante Harm Prevention*.

Ex Ante Harm Prevention: An action performed in conjunction with another, otherwise impermissible, action that prevents that action from making worse the *ex ante* prospects faced by any person.

On some moral theories—such as *ex ante* contractualism—it is not a wrong-making feature of an action for it to impose harms, so long as that actions does not impose additional *risk* of harm for any subject. By such views, offsetting via *Ex Ante* Harm Prevention does succeed. And this would include most carbon offsetting practices. Since agents like us are radically uncertain of how our emissions and offsets will affect weather systems, we cannot consider the actual outcome of each action but must instead consider the gambles they produce. And as Broome (2019) points out, from that radically uncertain vantage point, the risks of harm we impose on the potential victims of climate change and adverse weather events scale with the total mass of greenhouse gases in the atmosphere. By emitting a kilogram of carbon dioxide, we increase everyone’s probability of a dire outcome; by removing a kilogram elsewhere, we return those probabilities to their previous levels. So, carbon offsetting will count as *Ex Ante* Harm Prevention. And so, according to moral theories that consider only risks of harm rather than harm *per se*, it seems that carbon offsetting may succeed.

While *ex ante* contractualism and related theories offer an apparently satisfying explanation of many of our intuitions about offsetting, they encounter serious problems. In certain offsetting cases, they deliver verdicts that will seem abhorrent to anti-aggregationist intuitions. If we want our moral theory to give a plausible precissification of commonsense non-consequentialist morality, these theories are not the answer. To see why, consider one such case:

Mercenary’s *Ex Ante* Offset

Carolina works as a mercenary and has been deployed by a repressive government to kill civilians. She spots a crowd of civilians huddling together alongside a concrete retaining wall, attempting to evade a hailstorm of bullets from her colleagues. She decides to kill several of the civilians with a trick shot—she fires several rounds almost directly upwards, carefully calculating the angle so that they will return to Earth (in approximately 30 seconds) and strike several civilians. But she is then struck by the realisation that killing civilians for money is a morally appalling thing to do. So she tosses her bulletproof vest into the crowd. She predicts

that someone will pick it up, put it on, and have greater protection from her and her colleagues' bullets, but she has no idea which person it will be. We can suppose that, from her perspective, this later action reduces each person's risk of injury and of death by an amount equal to the increase due to her incoming bullet.

By firing those rounds, Carolina has raised the probability of death for each civilian in the crowd. And by then tossing the bulletproof vest, she has reduced that probability for each (and, we can suppose, reduced it by at least as much as she increased it). So her later action counts as *Ex Ante* Harm Prevention, and is structurally identical to most carbon offsetting practices. But it seems clear that Carolina still acts wrongly. Likewise, she would act wrongly even if the crowd had ten billion people in it and the bullets took a causally complex path toward their victims (much like our carbon emissions do). So it cannot be that any old action can be offset through *Ex Ante* Harm Prevention—it cannot be that reducing each potential victim's prospects of harm to previous levels is enough to erase the wrongness of the harmful action. And so *ex ante* contractualism and other theories that imply differently do not give us an adequate account of offsetting. They violate the spirit of non-consequentialism: they allow trade-offs between harms and benefits to different people, in a way that permits actions like Carolina's actions in Mercenary's *Ex Ante* Offset.²⁰

Note, further, that the *ex ante* contractualist cannot defend offsetting simply by identifying an account of *ex ante* compensation that forbids actions with the basic structure of Mercenary's *Ex Ante* Offset. Because Carolina's actions have the same causal structure as offsetting practices, any principled non-consequentialist account of wrongdoing that forbids Carolina's actions will also forbid more quotidian practices of offsetting. If *ex ante* contractualism and related theories can plausibly permit offsetting while forbidding Carolina's actions, it is not clear how.

4.3 Consequentialist theories

Standard forms of act-consequentialism deliver verdicts similar to non-consequentialist theories here, or else face serious problems.²¹

Take any act-consequentialist theory that is *maximising*—that permits an action only if it brings about the best available outcome (or *prospect* over outcomes). For an action to be otherwise wrong, and so a candidate to be offset, it must be one that brings about a suboptimal outcome (prospect). Suppose that the agent follows up that suboptimal action with another action that ensures the outcome (prospect) is just as good as if they had performed the optimal one in the first place—they engage in Harm Prevention. Consequentialists may agree that this makes the earlier action permissible. After all, it ends up resulting in an optimal outcome (prospect) after all.

²⁰See further criticism of the *ex ante* view of harm in Horton (2017), Mogensen and MacAskill (2021), and Lerner (n.d.).

²¹Much of what follows will hold for *rule*-consequentialism too, at least for sophisticated forms of rule-consequentialism. Meanwhile, unsophisticated forms of rule-consequentialism—under which rules simply prohibit or recommend coarse-grained types of acts—will say straightforwardly that any typically harmful act is prohibited, whether or not it is followed by an offsetting action.

But maximising act-consequentialist theories imply that any other form of offsetting fails, at least in most circumstances. When engaging in Non-Compensatory Offsetting (or, indeed, any offsetting that isn't Harm Prevention), in all but rare circumstances, the agent has the option to perform the beneficial offsetting action *whether or not* they perform the otherwise impermissible one.²² In both Carbon Offset and Arsonist's Indulgence, the agent can donate the same sum to charity whether or not they have acted wrongly. (The same goes for cases of Harm Prevention where the harm-preventing action is costly.) And maximising consequentialism will require that they do so whether or not they have acted wrongly (or else use their money for something even more impactful). By performing some suboptimal action and then offsetting that wrong, the agent brings about an outcome worse than if had performed the *optimal* action followed by the offsetting behaviour (or an even more impactful behaviour). Their harmful action still brings about a worse outcome than would otherwise obtain, so they still act wrongly.

What about act-consequentialist theories that aren't maximising? Consider act-consequentialist theories that are *satisficing*: that permit actions only if they bring about a *sufficiently good* outcome (or *prospect* over outcomes) among those available. Such theories will agree with the above about Harm Prevention—if the agent brings about what will otherwise be a suboptimal outcome (or lottery), but in the knowledge that they will later perform a complete Harm Prevention and restore an optimal outcome (lottery), then their earlier action is optimal, and so must be sufficiently good. But what do they say about the other types: Compensatory and Non-Compensatory offsetting? When an agent acts to make the world worse than it would otherwise be (sufficiently worse that doing so would normally be impermissible), but knows that they will later make it better by a similar amount, it may seem that the outcome of their earlier action may still be sufficiently good. But this depends on how the theory specifies 'sufficiently good'; this could be either by some *universal threshold* of value that applies for all decisions, or by some *relative threshold* according to the individual decision. The first, a universal threshold, is implausible. To see why, consider a world in which all available outcomes are guaranteed to be *far, far* better than that universal threshold; an agent could act to make the world far worse, in some appalling manner (as Marvin does in Arsonist's Indulgence), and this would still be permissible so long as the value of the world still lay above the threshold. But this is absurd; Marvin clearly acts wrongly. So we are left with the second option: a relative threshold of goodness, proportional to how good the outcomes in that particular decision are. And, by this form of satisficing consequentialism, all types of offsetting other than Harm Prevention will typically fail. As before, the agent can later perform the beneficial offsetting action with or without the earlier action. And performing the harmful action is just as much worse than not performing it *regardless* of whether the agent then performs the beneficial offsetting action or not. If the harming without offsetting is wrong, then so too is the harming *with* offsetting. Either way, the harmful action is *just as much worse* relative to their option set. So, any plausible form of satisficing act-consequentialism will say that the agent still acts wrongly if they perform the harmful action, regardless of whether they offset.

²²This does not hold if the agent only obtains the means to offset by performing the harmful action, such as emitting carbon by taking a business trip but earning a great deal of money from taking that trip. But these cases will be uncommon in practice. And we are sceptical that they count as genuine instances of offsetting, rather than cases of simply choosing the lesser evil.

Note that, for both of the above forms of act-consequentialism, we reasoned as though the agent could choose to later perform the offsetting action whether or not they performed the earlier harmful action, at least in cases of Compensatory and Non-Compensatory Offsetting. But this is not a valid way to reason under one class of act-consequentialist views: *actualist* theories.²³ Actualism is the view that an agent ought to Φ iff what will happen if they Φ is better than what will happen if they don't Φ , conditional on the actions that they will *actually* perform in future (or given their uncertainty of which future actions they will perform).²⁴ For example, I may know that I have a miserly character and will not donate to charity later unless I act badly now and make myself feel guilty. From my perspective while choosing *now* whether to act badly, I cannot independently choose to donate to charity—that is not under my control, except through the ways I manipulate my future self. Acting badly now must be considered to also bring the (prospect of) future benefit conferred by my later, guilt-motivated actions. So the initial harmful action may well be rendered permissible—on net, it is at least as good as the alternative, since it causes my future self to make the donation necessary to make the outcome just as good as it would otherwise have been. And this may be so regardless of whether I donate in a way that compensates the victims, prevents similar harms to others, or benefits in completely different ways. So actualist consequentialism allows any of the four categories of offsetting to succeed. But they can only succeed where the agent is (nearly) certain that they will not otherwise perform the later beneficial action, and they cannot simply perform it now. This is quite a rare circumstance for carbon offsetters, we would suggest. And so, even under actualist act-consequentialism, many instances of carbon offsetting will be unsuccessful—after all, in most practical cases, potential offsetters could make that donation with trivial ease even without booking a flight.

5 A general argument against Non-Compensatory Offsetting

So far, we have seen that standard carbon offsetting is not Harm Prevention (as we defined it) but instead Non-Compensatory, On-Target Offsetting—much like Marvin's actions in Arsonist's Indulgence. And we have seen that standard moral theories agree that agents who engage in Non-Compensatory, On-Target Offsetting still act impermissibly (in most circumstances). But you might still object: perhaps some *other* moral theory says otherwise, drawing further distinctions within the category of Non-Compensatory, On-Target Offsetting (as *ex ante* contractualism does), and judging certain instances of it as permissible; and perhaps that theory is the correct one. For any readers who object in this way, whether they are *ex ante* contractualists or not, we offer here a more general argument against carbon offsetting.²⁵

We show that, if any instance of Non-Compensatory, On-Target Offsetting is to succeed, we must deny one of several highly plausible principles. That is, if what we call *Minimal Offsetting* holds, then the principles we list below cannot all hold.

²³This is noted independently in Timmerman (2019).

²⁴See Cohen and Timmerman (2016a,b).

²⁵The argument also extends to offsetting behaviours in other contexts too, insofar as they are Non-Compensatory.

Minimal Offsetting: For at least one action (or set of actions) b that would otherwise be impermissible, there is at least one (Non-Compensatory, On-Target) offsetting action a such that b is made permissible if followed by a .

Of the other conditions that together generate the impossibility, consider next the principle of *No Extreme Offsets*. This principle states that not every otherwise impermissible action can be offset. Some actions, such as Marvin’s behaviour in Arsonist’s Indulgence, are just too seriously wrong to ever be rendered permissible.

No Extreme Offsets: For at least *some* impermissible action (or set of actions) b , there is no offsetting action or actions a such that b is made permissible if followed by a .²⁶

The third condition is that, if otherwise impermissible action can be offset, then so too any action that is only slightly more seriously wrong. There is no arbitrary, and arbitrarily precise, cut-off between what can be offset and what cannot.

Non-Arbitrariness: For any otherwise impermissible actions (or sets of actions) b and $b-$ such that $b-$ is more seriously wrong than b by a negligible amount, if there is an offsetting action a such that b is made permissible if followed by a , then there is some action $a+$ such that $b-$ is made permissible if followed by $a+$.

The remaining two horns are: that the relation of “At Least As Seriously Wrong As” is transitive; and that the set of all possible impermissible acts forms a smooth spectrum in terms of how seriously wrong they each are.

Transitivity: If A , B , and C are impermissible, and A is at least as seriously wrong as B and B is at least as seriously wrong as C , then A is at least as seriously wrong as C .

Deontic Density: If act a is more seriously wrong than act c by a non-negligible amount, then there is a possible act b such that a is strictly more seriously wrong than b and b is strictly more seriously wrong than c .²⁷

Note that those last three conditions appeal to *degrees* of seriousness of wrongness (or, you might say, degrees of impermissibility), rather than just the binary notions of permissibility and impermissibility. We think that it is very plausible that impermissible actions comes in degrees, and that this is supported both by commonsense moral reasoning and by compelling arguments made by others (e.g., ?Gustafsson, 2016). But those who believe that impermissibility is a binary notion need not fear, since the argument

²⁶By “no action” we mean *no metaphysically possible action*.

²⁷A stronger formulation would only require that a is more seriously wrong than c , not that it be more so by a non-negligible amount. But we need not assume such a strong density principle for our proof. We need only that the range of possible wrongness is sufficiently dense that there are no unintuitively large gaps.

can proceed without the assumption that it comes in degrees.²⁸

No moral theory can satisfy all five of the above conditions. The proof of this is brief and non-essential:

Proof: Assume Minimal Offsetting. Then for at least one otherwise impermissible action b_x , there is at least one offsetting action a_x such that b_x is made permissible if followed by a_x . Assume Deontic Density and Transitivity. Then there is an ordered set of actions $\langle b_1, b_2, \dots, b_x, \dots, b_n \rangle$ in which each action is impermissible if performed alone, each is more seriously wrong than the action before it by a very small amount, and b_n is arbitrarily seriously wrong. By Non-Arbitrariness, each action b_i following b_x can be offset with some a_i . This includes b_n and, since b_n can be any arbitrarily seriously wrong action, any action can then be offset. This violates No Extreme Offsets, so we have a contradiction.²⁹ \square

Thus, any comprehensive moral theory must reject at least one of the above five principles. But we take each of those principles to be highly plausible. To reject any of them is to incur a serious intuitive cost. So how should we respond to the above result? Several possible responses are available.

The first and, we think, most plausible is to reject Minimal Offsetting: to accept that no action can be Non-Compensatorily offset. This includes carbon offsetting. We suspect that this response will be counterintuitive for many. But perhaps this is not as implausible as it seems. While people frequently offset certain actions—such as emitting carbon—this is the exception rather than the rule. We often feel uncomfortable with the idea of offsetting other minor wrongdoings. For example, it seems impermissible for me to punch a colleague in the face even if I then donate a large sum to an anti-bullying charity. In fact, carbon offsetting is a case about which we might be quite biased or otherwise morally disengaged, given that many of the victims and beneficiaries are anonymous poor people and/or members of future generations.

Moreover, abandoning Minimal Offsetting need not stop us from judging offsetting practices as permissible, or indeed required, for wrongdoers. The passenger who flies across the world for a holiday may then be obligated to donate a sum to an emission mitigation charity, but they do not thereby succeed in offsetting—their initial act of flying remains impermissible. We suspect that this explains our intuitions in favour of Minimal Offsetting—we mistake the judgement that we ought to perform an offsetting action for the judgement that, if we perform that action, we have done no wrong at all. Still, if we accept this explanation and reject Minimal Offsetting, we bring on revisionary implications, morally indicting all real-world offsetters. But no one ever said morality was supposed to be easy.

In the remainder of this section, we discuss each of the other possible responses and argue that they

²⁸All we need are the trivial assumptions that (i) any permissible action is strictly less seriously wrong than any impermissible action, and (ii) there is some measurable normative notion of ‘worseness’ that captures the fact that, if not offset, Marvin’s behaviour is ‘worse’ than Nikita’s. Worseness may be determined by how bad the action’s outcome is (for consequentialists especially), how great a harm the action imposes, how blameworthy the agent is for performing the action, or what have you. We take it as overwhelmingly plausible that Non-Arbitrariness, Transitivity, and Deontic Density hold if defined in terms of some such notion, in which case the argument still holds.

²⁹This proof will ring familiar to those acquainted with the literature on lexicality and continuum arguments, from which we draw inspiration (Rachels, 2005; Temkin, 2014).

are each more implausible than abandoning Minimal Offsetting—than accepting that carbon offsetting, and all other Non-Compensatory Offsetting, fails.

The second response would be to reject No Extreme Offsets. To do so is straightforwardly very implausible—it renders permissible the imposition of enormous massive harms, even murder or genocide or Marvin’s behaviour in Arsonist’s Indulgence, so long as they are accompanied by large enough offsets. For a moral theory to say this is an unambiguous *reductio ad absurdum*.

A third response is to reject Non-Arbitrariness. This would mean that there exist two actions, differing only slightly in how seriously wrong they are, such that the one can be offset and the other cannot. For example, perhaps breaking a colleague’s arm is offsettable with a donation of \$10,000 to an organisation advocating for workplace safety, but even millions of dollars donated to this charity would not offset breaking the same person’s arm ever so slightly more egregiously! This result could be rendered less implausible if we could find a principled reason to draw a sharp cut-off between wrongful actions we could offset and those we could not. But, by assumption, the difference in wrongness between these actions would be arbitrarily small. As such, it seems that the prospects of plausibly abandoning Non-Arbitrariness are quite dim.

A fourth response is to reject the Transitivity of At-Least-As-Seriously-Wrong-As. One might be inspired to do so by arguments concerning the transitivity of betterness among outcomes (see Temkin, 2014), or the transitivity of ‘permissible-when-compared-pairwise’ among actions (see Kamm, 1985). But, while non-transitivity can be motivated in those settings by appeal to particular cases, it does not seem it can be for the relation of ‘At-Least-As-Seriously-Wrong-As’ over actions (relative to the agent’s option set in the decision where that action is chosen). How seriously wrong an action is seems an easy-to-measure quantity. Intuitively, any two impermissible actions must be comparable in terms of which is more seriously wrong; and if an action *a* (within the particular option set it is chosen from) is more seriously wrong than action *b* (within its option set), and *b* more seriously wrong than *c* (within its option set), then surely *a* is more seriously wrong than *c*. The degree of seriousness of a wrong does not seem like a quantity that will be susceptible to problem cases, and so rejecting its transitivity is not a promising approach.

A fifth and final response to the impossibility result is to reject Deontic Density. This amounts to the claim that there are very large value gaps in the space of all actions that could ever possibly be performed. There are some common moral theories that seem to give up on this kind of density. These *Clump Theories* say that the wrongness of actions is clumpy, such that certain act properties make these acts qualitatively more seriously wrong than acts without these properties. For instance, the property of being first-degree manslaughter might set an act apart from all other acts of mere *second*-degree manslaughter (and other acts), making the act of first-degree manslaughter much more seriously wrong than those other acts. Clump Theorists have a good explanation of why some actions can be Non-Compensatorily offset and not others: the most seriously wrong actions are much more seriously wrong than all other actions, and this explains why they can’t be offset. All other actions are not nearly as bad, and their lesser wrongness explains why they can be offset.

But Clump Theories face significant problems of their own. First, for most properties that separate such possible clumps, we can often imagine actions lacking the property that would place them in the more seriously wrong clump yet which seem at least as bad the least seriously wrong action from that clump. For instance, take the property of being first-degree manslaughter distinguishes one clump of actions from the supposedly less seriously wrong actions that are not. (Note that first-degree manslaughter requires that the agent intended to injure the victim severely, while second-degree manslaughter involves only recklessness and awareness of the possible effects.) We can imagine cases of first-degree manslaughter in which the agent has only the weakest of intentions, and perhaps is subject to numerous mitigating factors. Meanwhile, we can imagine cases of second-degree manslaughter in which the agent is extraordinarily reckless and vicious in myriad other ways. It is at least dubious that none of the latter cases are more seriously wrong than any of the former cases.

The second, far more troubling, problem for Clump Theories is one of risk. For any two actions b and $b-$ that lie in distinct clumps (perhaps the least of the more seriously wrong clump, and the greatest of the less seriously wrong clump), we can consider a risky action which offers some probability of having all of the relevant features of act b and the remaining probability of having the relevant features of $b-$. It seems overwhelmingly plausible that this risky action must be less seriously wrong than the worse one, and more seriously wrong than the better one. Indeed, all plausible non-consequentialist theories (in our view) will agree with this. All such views discount the seriousness of wrongness of actions with wrongmaking properties under conditions of risk. For example, firing a gun into a crowd has a probability less than 1 of being murder, and such theories will say that doing so is less seriously wrong than committing murder with probability 1. If this did not hold, actions that have *any* probability of ending up as seriously prohibited actions would also be seriously wrong, no matter how small the probability and no matter the benefits of those actions. But, if we deny such absolutism, then Clump Theorists are still committed to Deontic Density.³⁰ The sequence of actions $\langle b_1, b_2, \dots, b_n \rangle$ invoked in the proof above may identify b_1 as causing someone a papercut, b_n as murdering them, and each b_i in between as an action which may be harmless but has probability proportional to i of being murder. So long as the seriousness of wrongness of murder is discounted by a continuous function of i , Deontic Density will then hold. And if that discounting function is *discontinuous*, then the theory says that there is some probability $0 < p < 1$ such that an action that has probability p of causing a murder is *far* less seriously wrong than an action with probability $p + \epsilon$ of causing a murder, where ϵ is arbitrarily small. This is an implausible result for a theory to have.

As well, Deontic Density seems highly plausible given non-absolutism. By non-absolutism, any otherwise wrongful action can be made permissible if performing it is necessary to save some number n people from a sufficiently horrible fate (but not by saving merely $n - 1$ people from that fate). Plausibly, the number n will vary based on how seriously wrong that action otherwise is. But if we were to commit some such wrongful action and thereby save just $n - 1$ people, or $n - 2$ people, or so on, the action would be progressively more seriously wrong. There would be no such number $n - k$ (above zero) at which there

³⁰One might worry that this argument fails because it equivocates between subjective and objective wronging. But this is not so. If, as Johann Frick argues, objective reasons are evidence-relative or if objective chance is even *metaphysically possible*, then this argument applies to objective reasons as well as subjective reasons.

is a sudden, rapid increase in how seriously wrong that action is. Pretheoretically, it seems much more plausible that the function of how serious the wrong is of killing one to save some number of people is continuous. This should generalise for all non-absolutist constraints on action. Perhaps absolutists such as Kant could reject Deontic Density without too many further theoretical costs, but such absolutists are committed to (what we consider) other morally abhorrent verdicts.

It therefore seems that this impossibility result raises a genuine paradox. Each principle is highly plausible, but together they entail a contradiction. We suggest that the most plausible response is to abandon Minimal Offsetting, which means that all Non-Compensatory Offsetting fails. As such, this result provides a compelling argument that carbon offsetters still act impermissibly. In effect, their actions are analogous to Marvin’s in Arsonist’s Indulgence.

6 Conclusion

Many real-world consumers attempt to offset their wrongful emissions of greenhouse gases. But can they succeed in making their actions permissible? We argue that they cannot. *Contra* other philosophers, we have argued that carbon offsetting is a form of Non-Compensatory Offsetting, not Harm Prevention. So, offsetting one’s emissions does not mean that no one is harmed by those emissions. It follows that carbon offsetters act wrongly according to standard non-consequentialist theories and, in most circumstances, standard consequentialist theories too. Those non-standard non-consequentialist theories that say otherwise happen to deliver implausible verdicts in other offsetting cases. And we have shown that any theory that *does* say that Non-Compensatory On-Target Offsetting can be permissible—non-consequentialist or consequentialist, *ex ante* or *ex post*—must abandon some other highly plausible desideratum for moral theories. If not, they will vindicate carbon offsetting if and only if they also vindicate Marvin’s behaviour in Arsonist’s Indulgence. And vindicating the offsetting of such acts of arson and murder is absurd. So, carbon offsetting cannot succeed.

This is not to say that we ought *not* engage in offsetting behaviour if we have already emitted, nor that we have no moral reason to, nor that donating to mitigate emissions is not a worthwhile activity. But doing so will not erase our climate wrongs; we still ought not emit in the first place. ³¹

References

- ALEXANDER, S., 2015. Vegetarianism for meat-eaters. <https://slatestarcodex.com/2015/09/23/vegetarianism-for-meat-eaters/>. (accessed: 27.06.2021). (cited on page 2)
- BARRY, C. AND CULLITY, G., 2022. Offsetting and risk imposition. *Ethics*, 132, 2 (2022), pp. 352–81. (cited on pages 4, 6, and 7)

³¹This implication may carry over to other forms of offsetting, such as offsetting pollution or the consumption of animal products. But this will depend on exactly which form of offsetting is being attempted. In general, Harm Prevention will likely be morally acceptable, while most other forms cannot ever make things right.

- BEDER, S., 2014. Carbon offsets can do more environmental harm than good. <https://theconversation.com/carbon-offsets-can-do-more-environmental-harm-than-good-26593>. (accessed: 27.06.2021). (cited on page 3)
- BRESSLER, R. D., 2021. The mortality cost of carbon. *Nature Communications*, 12 (2021). (cited on page 1)
- BROOME, J., 2012. *Climate Matters*. W.W. Norton Company, London. (cited on pages 2 and 4)
- BROOME, J., 2019. Against denialism. *The Monist*, 102, 1 (2019), p. 110–29. (cited on pages 5 and 9)
- CHAN, R. AND CRUMMETT, D., 2019. Moral indulgences: When offsetting is wrong. *Oxford Studies in Philosophy of Religion*, 9 (2019). (cited on page 8)
- COHEN, Y. AND TIMMERMAN, T., 2016a. Actualism and possibilism. *The Philosophers' Magazine*, 72 (2016), pp. 107–8. (cited on page 12)
- COHEN, Y. AND TIMMERMAN, T., 2016b. Actualism has control issues. *Journal of Ethics and Social Philosophy*, 10, 3 (2016), pp. 1–18. (cited on page 12)
- CULLITY, G., 2019. Climate harms. *The Monist*, 102, 1 (2019), pp. 22–41. (cited on page 4)
- ENVIRONMENT BANK, 2012. Biodiversity offsetting: A general guide. <https://static1.squarespace.com/static/595ca91bebbd1a1d0aaab285/t/5a316bb5c830250f4f8ea6e5/1513188282549/Biodiversity+Offsetting+-+A+general+guide.pdf>. (accessed: 27.06.2021). (cited on page 2)
- FOOD & WATER WATCH, 2012. Pollution trading. https://web.archive.org/web/20170509174527/https://staging.foodandwaterwatch.org/sites/default/files/pollution_trading_ib_dec_2012.pdf. (accessed: 27.06.2021). (cited on page 2)
- FRICK, J., 2015. Contractualism and social risk. *Philosophy and Public Affairs*, 43, 3 (2015), pp. 175–223. (cited on page 9)
- GARDNER, J., 2011. What is tort law for? part 1. the place of corrective justice. *Law in Philosophy*, 30, 1 (2011), pp. 1–50. (cited on page 7)
- GUSTAFSSON, J., 2016. Consequentialism with wrongness depending on the difficulty of doing better. *Thought*, 5, 2 (2016), pp. 108–18. (cited on page 13)
- HORTON, J., 2017. Aggregation, complaints, and risk. *Philosophy and Public Affairs*, 45, 1 (2017), pp. 54–81. (cited on page 10)
- KAMM, F., 1985. Supererogation and obligation. *Journal of Philosophy*, 82, 3 (1985), pp. 118–38. (cited on page 15)
- KAYE, L., 2012. Water offsets could become a tool for water stewardship. <https://www.theguardian.com/sustainable-business/water-offsets-new-tool-stewardship>. (accessed: 27.06.2021). (cited on page 2)
- LAWFORD-SMITH, H., 2015. Unethical consumption and obligations to signal. *Ethics and International Affairs*, 29, 3 (2015), pp. 315–30. (cited on page 8)
- LAWFORD-SMITH, H., 2016. Offsetting class privilege. *Journal of Practical Ethics*, 4, 1 (2016), pp. 23–51. (cited on pages 2 and 4)
- LERNER, A., n.d. The irrelevance of identifiability. Unpublished manuscript. (cited on page 10)
- LIAO, M., 2012. Intentions and moral permissibility: The case of acting permissible with bad intentions. *Law and Philosophy*, 31 (2012), pp. 703–24. (cited on page 8)
- LORENZ, E., 1972. Does the flap of a butterfly's wings in Brazil set off a tornado in Texas? , transcript of a lecture given to the 139th meeting of the American Association for the Advancement of Science, in Washington, DC (1972). (cited on page 5)

- MACASKILL, W., 2015. *Doing Good Better*. Random House, London. (cited on pages 2 and 4)
- MOFFATT, S.; GATISS, M.; AND HURRAN, N., 2017. The lying detective [television series episode], in S. Vertue (producer), *Sherlock*, Cardiff, UK: Hartswood Films. (cited on page 1)
- MOGENSEN, A. AND MACASKILL, W., 2021. The paralysis argument. *Philosophers' Imprint*, 21, 15 (2021). (cited on pages 6 and 10)
- MORSS, R. E.; SNYDER, C.; AND ROTUNNO, R., 2009. Spectra, spatial scales, and predictability in a quasigeostrophic model. *Journal of the Atmospheric Sciences*, 66, 10 (2009), p. 3115–30. (cited on page 5)
- MORTON, J. F., 2007. The impact of climate change on smallholder and subsistence agriculture. *Proceedings of the National Academy of Sciences*, 104, 50 (2007), pp. 19680–5. (cited on page 1)
- PALMER, T.; DÖRING, A.; AND SEREGIN, G., 2014. The real butterfly effect. *Nonlinearity*, 27, 9 (2014), R123. (cited on page 5)
- RACHELS, S., 2005. Counterexamples to the transitivity of better than. *Australasian Journal of Philosophy*, 76 (2005), pp. 71–83. (cited on page 14)
- RAZ, J., 2004. Personal practical conflicts. In *Practical Conflicts: New Philosophical Essays* (Eds. P. BAUMANN AND M. BETZLER), 172–96. Cambridge University Press, Cambridge. (cited on page 7)
- STEFÁNSSON, H. O., 2019. Should i offset or should i do more good? *Ethics, Policy and Environment*, 25, 3 (2019), p. 225–41. (cited on page 2)
- SVERDLIK, S., 2011. *Motive and Rightness*. Oxford University Press, Oxford. (cited on page 8)
- TADROS, V., 2011. *The Ends of Harm*. Oxford University Press, Oxford. (cited on page 7)
- TEMKIN, L., 2014. *Rethinking the Good*. Oxford University Press, Oxford. (cited on pages 14 and 15)
- TIMMERMAN, T., 2019. Effective altruism's underspecification problem. In *Effective Altruism: Philosophical Issues* (Eds. H. GREAVES AND T. PUMMER). Oxford University Press, Oxford. (cited on page 12)
- TRIBBIA, J. J. AND BAUMHEFNER, D. P., 2004. Scale interactions and atmospheric predictability: An updated perspective. *Monthly Weather Review*, 132, 3 (2004), p. 703–13. (cited on page 5)
- ZWICK, S., 2014. Brazilian landowners can now offset past deforestation with future protection. *Ecosystem Marketplace*, May 7 (2014). (cited on page 2)