

Individual responsibility for carbon emissions: Is there anything wrong with overdetermining harm?

Christian Barry (ANU) and Gerhard Øverland (Oslo)

Introduction

Climate change and other harmful large-scale processes challenge our understandings of individual responsibility. People throughout the world suffer harms-severe shortfalls in health, civic status, or standard of living relative to the vital needs of human beings—as a result of physical processes to which many people appear to contribute. Climate change, polluted air and water, and the erosion of grasslands, for example, occur because a great many people emit carbon and pollutants, build excessively, enable their flocks to overgraze, or otherwise stress the environment. If a much smaller number of people engaged in these types of conduct, the harms in question would not occur, or would be substantially lessened. However, the conduct of any particular person (and, in the case of climate change, of even quite large numbers of people) could make no apparent difference to their occurrence. My carbon emissions (and quite possibly the carbon emissions of much larger groups of people dispersed throughout the world) may not make a difference to what happens to anyone. When the conduct of some agent does not make any apparent difference to the occurrence of harm, but this conduct is of a type that brings about harm because many people engage in it, we can call this agent an overdeterminer of that harm, and their conduct overdetermining conduct.¹

What is the moral status of overdeterming harm? Four questions lurk within this broad one. First, are there moral reasons against becoming an overdeterminer of harm? Second, if there are such moral reasons, what is their basis?² Third, do overdeterminers have moral reasons to provide compensation or other forms of reparation to those who suffer overdetermined harms? Fourth, what should overdeterminers of harm do with the

[•] Earlier versions of this chapter were presented at seminars at the Australian National University, University of Oslo, and Centre for Applied Philosophy and Public Ethics. We are very grateful for comments received on those occasions from Geoff Brennan, Garrett Cullity, Keith Dowding, Bob Goodin, Frank Jackson, Seth Lazar, Jonathan Schaffer, Chad Lee-Stronach, RJ Leland, Jeremy Moss, Ingmar Persson, Alex Sandgren, Anne Schwenkenbecher, Kim Sterelny, Daniel Stoljar, and Suzanne Uniacke, and especially to Robert Kirby, Holly Lawford-Smith, RJ Leland and Terry Macdonald for their written comments. Work on this chapter was supported by grants from the Australian Research Council and the Research Council of Norway.

¹ Although individual agents are discussed in this chapter, the analysis can be extended to collective agents as well.

² Following Joseph Raz, let us understand a 'reason' as a consideration in favour of some act, 'which by itself is sufficient to necessitate a certain course of action, provided no other factors defeat it.' (Raz 1990, p. 19). Those who claim that there are overdetermination-based moral reasons against φ -ing are committed at a minimum to the view that this reason will be sufficient to necessitate their refraining from φ -ing, provided no other factors militate in favour of φ -ing. They may differ, of course, about just how weighty overdetermination-based reasons are, relative to other sorts of reasons.

benefits that they have derived from their conduct? While all of the issues raised by these questions are important (and clearly interconnected), the first two—concerning reasons against overdeterming harm and their moral basis—seem most fundamental. If there were moral reasons against overdetermining harm, it would be relatively easy to understand why overdeterminers would owe compensation to those who have been harmed, and why they would be required to share the benefits that their overdetermining conduct produces.

It is not always obvious whether or not the conduct of particular agents overdetermines harms or whether, on the contrary, the conduct of each is necessary for the occurrence of some particular harm. Individual contributions to climate change may be like this. It is certainly *possible* that an individual person might, through their carbon emissions, make some difference to the occurrence of harm to other particular people, in the sense that if they acted differently these people would not have been harmed. If this were generally the case, then the issue of overdetermining harm would not be of critical importance to moral thinking about individual responsibility relating to climate change. But in general it seems very unlikely that individual carbon emissions really make a difference to the occurrence of any particular harm (Sinnott-Armstrong 2005). Some authors avoid the implication that individual carbon emission is a case of overdetermining harm. In a recent article, for example, John Nolt tries to undermine the idea that 'the harm caused by an individual's participation in a greenhouse-gas-intensive economy is negligible' (Nolt 2011, p. 3). He estimates the harmful impact of the 'average American's' carbon emissions by dividing the harm that will be generated by the total amount of carbon emitted by Americans as a group, and then dividing it by the number of Americans. He concludes that by this measure the average American's carbon emissions will be responsible for the deaths or suffering of two future people. To be sure, if each individual American had reason to believe that refraining from emitting carbon excessively might make the difference to whether two future people die prematurely or not, this would give them a very stringent reason not to do so. But while Nolt's accounting exercise may have some uses for moral assessment, his finding is consistent with its being the case that the total amount of harm caused by carbon emissions would be the same whether or not any particular American changed the amount of carbon they emitted. If this is so, then it would be true of each American that their carbon emissions make no difference to the occurrence of harm to particular people-each could vanish from the face of the earth without the reduced emissions meaning that two future people are saved from premature death. And if it is true of each American that they would not make such a difference to particular people, then they would be overdeterminers of harm in our sense. This is not to say that individuals are not overdeterminers of each and every harm engendered by climate change. John Broome, for instance, reports that an average adult in a rich country emits around 800 tonnes of GHGs in her lifetime (Broome 2012, p. 74), and claims that it is very likely that their cumulative emissions will indeed make difference to the occurrence of harm (pp. 74-78). Assuming that Broome is correct, there will nevertheless be many harms that individuals overdetermine through their conduct, additional to whatever harms are counterfactually dependent on their individual carbon emissions. Consequently, the stringency of reasons to refrain from emitting carbon or to compensate for their doing so will depend in some

measure on the moral status of overdetermining harm.

So it is important to examine the issues of moral reasons against overdeterming harm (we'll refer to these as overdetermination-based constraints) and their bases, and that's the task we shall undertake in this chapter. We survey some proposed rationales for these constraints and note some of the criticisms to which they seem vulnerable. We then propose what we take to be a more promising alternative account.

Preliminaries

Doing harm

In exploring the idea of overdetermination-based constraints, we shall restrict our discussion here to cases in which the *type* of conduct in question does harm, rather than merely fails to prevent it or contributes to it in some other way (e.g. by enabling it to occur, or by facilitating its occurrence).³ While the proper way to characterize the distinction between doing harm and merely allowing or enabling it to occur remains controversial, clear-cut instances of doing harm typically have two features, and we will discuss cases in which the agent's conduct possesses them.

The first feature is what might be called *relevant action*. If Sue is linked to John's injuries by relevant action, then there is an answer to the question of *how* she was relevant to his injuries that refers to some act of hers. In a car crash in which Sue's car runs over John, Sue becomes relevant to John's broken leg *by driving into him*. The answer to the how question refers to an action of hers—her driving the car in a particular way and at a particular time. The second feature is that there is a *complete causal process* that links Sue's relevant action to John's injuries.⁴ That is, there is an intact sequence linking the relevant action of Sue with the fracture of John's leg. In a car crash this intact sequence takes the form of a physical process involving the transfer of energy and momentum from Sue to John—a complete energy momentum sequence connects them.⁵

All other things being equal, moral reasons associated with doing harm that possess these features are commonly thought to have several important normative characteristics. First and foremost, there are stringent constraints against engaging in such conduct. They are stringent in the sense that prospective doers of harm cannot easily justify their conduct by appealing to the costs to themselves of refraining from doing harm, nor by appealing to the overall good that their conduct will bring about. And they are stringent in the sense that they demand much of agents who have ignored these constraints, but are now in a position to mitigate or alleviate the harm that they have done.

³ The distinctions between doing, allowing, and enabling harm are discussed in detail in Barry and Øverland (2012).

⁴ Ned Hall (2002; 2004) refers to causal connections of this kind as exhibiting 'locality'.

⁵ This is not the only form that such processes can take. If John opens a dam and the water floods the town below, he does not *transfer* energy to the dam or the water, but rather *releases* stored energy that was being held at bay by the dam. For a short useful discussion of the different ways the idea of a complete process might be conceived, see Schaffer (2003, p. 32).

Sue has a stringent moral reason not to drive into and maim an innocent person, even if it is the only way she can avoid losing her own hand, or protect her child from suffering a broken leg. On the other hand, the fact that Sue would lose a hand or that her child would suffer a broken leg were she instead to intervene in a traffic incident to help an innocent person escape significant injury is ordinarily thought to provide justification (or at least a very good excuse) for her failure to help. In addition, the potential victims have claims against prospective doers of harm that they not harm them. Moreover, these claims are *enforceable*—potential victims (or third parties acting on their behalf) can enforce these claims through the proportional use of force. For instance, Sue may be prevented from killing a pedestrian with her car even if this involves injuring her significantly. Even when it is, all things considered, permissible or even obligatory to do harm to innocent non-threatening people (bystanders), compensation is typically owed to those who are harmed. This is because their stringent claims against having harm done to them have been infringed, however justifiably.⁶

All we have done so far is describe what we take to be widely accepted views about the normative significance of doing harm, and the way this category differs from allowing harm. Whether moral reasons associated with doing harm really have such significances remains a matter of philosophical debate. We won't join in this controversy here. We shall simply assume that reasons based on doing harm possess the characteristics that commonsense morality accord to them, and plumb the significance of this assumption for the issue of overdetermining harm. Do particular overdeterminers harm others? The answer to this question depends very much on the conception of harm that is adopted. Some conceptions of harm that make counterfactual dependence essential for the attribution of harm will clearly not treat individual overdeterminers as harming others.⁷ Conceptions of harm that make the attribution of harm depend on whether an individual is involved in the production of bad states of others, on the other hand, may count individual overdeterminers as harming.⁸ Since the questions we have raised concerning the moral status of overdetermining harm are of interest whether or not individual overdeterminers are taken to harm others, we shall bypass discussion of the appropriate account of harm.

What is overdetermining harm?

To help fix ideas and make our discussion more concrete, let us introduce a few variations of a simple imaginary case that we hope can shed light on the more complex cases—such as individual carbon emissions—with which this volume is concerned.

A person—Robinson—is living relatively well on his small island. However, the water in the lake surrounding his island starts to rise. At some point the island is swallowed up as a result of the rising water levels, and Robinson drowns. 51 people have engaged in a type of conduct—shovelling excess waste from their gardens into the lake—that has led to the rising water levels. Tom is one of the 51: he digs in his garden to extract some valuables, and shovels the excess waste into the lake. Tom's conduct

⁶ For discussion, see Thomson (1977).

⁷ For discussion see Hanser (2008), Thomson (2011), and Bradley (2012).

⁸ For discussion see Shiffrin (2010) and Harman (2009).

alone would not have resulted in Robinson's death. Unfortunately, 50 other people also shovel their waste into the lake. As a result, water levels rise and Robinson is drowned.

Consider three versions of this story (only one of which (it turns out) is an instance of overdetermining harm, as we have defined this notion).

Robinson I: Tom's disposing of his waste is necessary for the flooding of the island and the drowning of Robinson. Although 50 others shovel waste into the lake, it is nevertheless the case that if Tom were to have abstained, Robinson would not have drowned.

Robinson II: Tom's disposing of his waste makes no apparent difference to the drowning of Robinson. 50 others shovel waste into the lake and that number, or less than that number, is enough to drown Robinson. If Tom were to have abstained, the island would still have been flooded, and Robinson would still have drowned.

Robinson III: Tom's disposing of his waste makes no apparent difference to the drowning of Robinson. He shovels his waste into the lake after many others have disposed of their waste, and after this has already led to the flooding of the island and the drowning of Robinson.

We'll assume throughout this chapter (unless otherwise specified) that the situations are transparent to Tom and others in each of these cases. In Robinson I, Tom is *not* an overdeterminer with respect to Robinson's death. Why? Because Robinson's death would in this case be counterfactually dependent on Tom's conduct. Had Tom refrained from shovelling his waste into the lake, Robinson would not have drowned. In this scenario Robinson's vulnerability is obviously a very good reason for Tom not to shovel his waste into the lake.⁹

In Robinson III, Robinson's death is not counterfactually dependent on Tom's conduct. Moreover, it is clear that there is no intact causal sequence that links Tom's conduct with Robinson's death. Robinson has already drowned when Tom shovels his waste into the lake. Consequently, Tom is not an overdeterminer of Robinson's death in this case. In the absence of some independent and negative consequence of his disposing of his waste in this way (perhaps it shows disrespect of some sort), it is hard to see why Tom has any reason to refrain from acting as he does in Robinson III. It does not seem that there is any claim of Robinson's that would be infringed by Tom's conduct.¹⁰

In Robinson II, by contrast, Tom is an overdeterminer of Robinson's death. Robinson's death is not counterfactually dependent on Tom's conduct (he would have died whether or not Tom disposed his waste), but it is nevertheless the case that if Tom and enough others shovel their waste into the lake, Robinson will drown. Further, there may be an intact sequence linking Tom's conduct and Robinson's death in Robinson II (but, since it is overdetermined, there may not). If there are overdetermination-based

⁹ This is precisely the sort of reason Broome stresses when condemning carbon emissions without counterveiling offsets (see Broome 2012, esp. pp.50-55, 96).

¹⁰ Insofar as compensation is owed to Robinson's relatives, it would seem to be owed by the others who shovel waste prior to the water reaching the stage where it causes his death, but not by Tom.

constraints, then all else being equal it would be wrong of Tom to dispose of his waste in Robinson II.

Existing approaches to overdetermining harm

Scepticism

It might be argued that there is no constraint against Tom's disposing of his waste in Robinson II. After all, one might reasonably believe that the reason why there are constraints against certain types of conduct is that by engaging in them one tends to make other people worse off, or at least imposes a risk on others that they might be made worse off. In situations of overdetermination, however, this is not the case. Robinson will not be made worse off if Tom shovels his waste into the lake on this particular occasion. Recall that we are assuming that the fact that the conduct will make Tom an overdeterminer is transparent to him and others. The sceptic claims that while it might be true that the agent's particular action would ordinarily have been wrong because it would have done harm to the victim, he does nothing wrong by performing that action in this particular case, because he makes no difference to what happens to the victim. There is some evidence that the law, at least, frees overdeterminers of harm from responsibility in some cases.¹¹ This line seems consonant with the general view of morality endorsed by Frank Jackson (1997):

the morality of an action depends on the difference it makes; it depends, that is, on the relationship between what would be the case were the act performed and what would be the case were the act not performed.

This view has been sympathetically explored (though not endorsed) in application to the case of climate change by Walter Sinnott-Armstrong (2005).¹²

The sceptic claims that there are no overdetermination-based constraints. He believes that all else being equal Tom has no reason not to dispose of his waste in Robinson II. When Tom shovels his waste into the lake, he obtains benefits without worsening anyone's situation, and this gives him a good reason to do it. Correspondingly, Robinson has no claim against Tom that he refrain from the overdetermining conduct. Of course, many people seem to think that there *is* something wrong about Tom's conduct in Robinson II. They find it problematic that Tom can φ when φ -ing is a type of action that alone or in conjunction with the actions of others could bring about severe harm, even though on this particular occasion the harm will occur whether or not Tom φ s, because many others will φ . But the skeptic may be able to provide plausible explanations for this intuition without attributing any moral significance to conduct that overdetermines harm. For example, they can argue (in a Lockean vein) that because we strongly associate φ -ing with certain harms, this leaves a kind of psychological trace that we cannot easily free ourselves from. This inclines us to condemn an agent's φ -ing even when her φ -ing is not necessary for the harm to occur.

¹¹ These are the asymmetrical concurrent cause cases, the symmetrical concurrent cause cases where one sufficient cause is a natural event, and the damage limitation rule in cases of preemptive causation. For discussion, see Moore (2003, pp. 1264–66).

¹² See also the discussion in Garrett Cullity's chapter in this volume.

Moreover, intuitions about overdetermining harm cases could stem from the belief that we rarely can be certain whether our conduct will make a difference to what happens to others. And when we are uncertain whether harm counterfactually depends on our φ -ing, this obviously gives us a reason not to φ . Perhaps many people condemn agents like Tom in cases like Robinson II because they feel that no one could ever really be so sure that their shovelling of waste would make no difference to the outcome. Without a positive justification for constraints against overdetermining harm, this sort of 'debunking' explanation of people's intuitions about them would be hard to resist. This does not mean that the skeptic's position comes without cost. In particular, it seems plausible that, where it suddenly to become possible for Robinson or a third party to intervene to protect him in a way that would impose significant cost on each overdeterminer, it would be permissible to do this. How much cost? Certainly much more cost than could be imposed on an innocent bystander who was in no way related to Robinson's plight. In general, how much cost we can impose on the agent without violating his rights seems to depend in large measure on how much cost this agent would be required to bear in the first place. For example, it seems impermissible for one agent to compel another to do X unless that agent is morally required to do X. So the skeptic will either need to claim that it is in fact impermissible to impose more cost on the overdeterminers than on an innocent bystander, or explain how it overdeterminers can at once be liable to bear more cost to avert the prospective harm to Robinson without having a duty to take on cost to avoid harming him.

Still, skepticism about overdetermining harm is a powerful position, and one that we shall not attempt to disprove here. The appeal of scepticism could be blunted considerably if a plausible positive account of these constraints can be provided, however, and that is what we shall try to do.

So let's explore some existing justifications for overdetermination-based constraints to see whether they hold promise.¹³ Our aim in exploring them is not to provide knockdown arguments against them. Rather, we indicate the sort of objections to which they are vulnerable. We think these objections are serious (even if they could ultimately be met) and warrant exploration of alternative justifications for constraints against overdetermining harm that are not vulnerable to them.

Absolutism

It might be argued that since φ ing is the sort of action that typically has bad consequences, it is simply wrong to φ , even if in certain circumstances φ ing makes no one worse off. Following Jonathan Bennett, let us distinguish between *absolutism* and *relativism* about types of conduct.¹⁴ Absolutism about φ ing is the thesis that it could *never*

¹³ Our list includes what we take to be the most promising existing justifications for the constraints in question, and is not exhaustive (see Sinnott-Armstrong 2005 for some other arguments that we do not consider).

¹⁴ As Bennett (1995, pp. 165–71) notes, when contrasting absolutism and relativism we ought not to describe the kind of behaviour whose moral status is in question in terms of its *overall* consequences. There is substantive dispute between absolutism and relativism about the wrongness of engaging in some conduct when that conduct is described as 'killing an innocent person when there is no good to be achieved by so

be right to φ . Relativism about φ ing is the thesis that it can be permissible to φ in some possible circumstances, but that it may not be permissible in others.

It is controversial whether absolutism about any type of conduct is plausible (Anscombe 1961; Bennett 1995). However, even if one could make a case for absolutism about some types of conduct—shooting or torturing innocent non-threatening people, for example—it is extremely unlikely that a case can be made for absolutism about many of the types of conduct that are involved in overdetermination cases. If no one were made worse off or rendered unduly vulnerable as a result of our emitting high levels of carbon or shovelling waste into lakes, how could it possibly be wrong for us to do these things?

The universalization requirement

One might appeal to Kant-inspired arguments as a basis for overdetermination-based moral reasons. That is, in deliberating about whether or not she can φ , the agent must not think in terms of whether or not her qing would make a morally relevant difference in this instance, but whether it would make a morally relevant difference if everyone qed, or (to take another version of the universalization test) if everyone who was *inclined to* φ did so.¹⁵ Clearly, the conduct that universalization tests would rule out as impermissible would seem to depend very much on how the conduct is described. If Tom's conduct in Robinson II is described as 'shovelling waste into the lake', then perhaps this test will require that the agent not engage in it, since morally unacceptable consequences (e.g. Robinson's death) would result were everyone to do it, or even if everyone who is inclined to do so does it. If the conduct in Robinson II is instead described in terms of 'shovelling waste into the river when doing so makes no difference to anyone', then it is likely that the conduct would not be ruled out by the universalization test. As Parfit puts it, 'if my contribution would make no difference, I can rationally will that everyone else does what I do. I can rationally will that no one contributes when he knows that his contribution would make no difference.' (Parfit 1983, p. 87)

One could insist that in formulating the universalization test we should describe the conduct in a way that does not include the incorporation of any contextual features based on what others will do or are likely to do. Allowing the incorporation of such features seems to invite agents to concoct 'gimmicky' maxims that effectively *circumvent* the universalization tests, rather than apply them sincerely (Pogge 1998). The main drawback to this proposal is that, if we exclude such contextual information, the universalization test is very likely to yield results about how agents ought to behave that in many cases are difficult to accept. It seems very important that we be able to consider the likelihood that other people will behave in certain ways when we are deliberating about what we are permitted to do, rather than assuming away such information. To borrow an example of Jonathan Glover's, I may not be able to rationally will that everyone be heavily armed, but it does not follow from this that I may not heavily arm myself when I know that others around me are heavily armed (Glover 1977, p. 131).

doing'. The conduct must be described in such a way as not to invoke an overall moral assessment—it is not controversial that there is an absolute moral prohibition against 'impermissibly killing someone'. ¹⁵ See Pogge (1998) for discussion of this version of the test.

Overdetermination as contribution: necessary elements of a sufficient set

A very different proposal for justifying overdetermination-based constraints is to say that overdeterminers contribute to overdetermined harms, not by becoming a difference maker with respect to their occurrence, but by becoming necessary elements in a set of actual antecedent conditions that is sufficient for bringing those harms about—a 'NESS' condition for their occurrence.¹⁶ An agent's conduct is a NESS for some harm only if it was a necessary element of a set of antecedent actual conditions that was sufficient for its occurrence. Richard Wright (1985) explains this notion as follows:

[A] particular condition was a cause of (condition contributing to) a specific consequence if and only if it was a Necessary Element of a Set of antecedent actual conditions that was Sufficient for the occurrence of the consequence.¹⁷

According to some theorists, NESS conditions provide the best available analysis of the concept of causation.¹⁸ Wright, for example, claims that the NESS idea 'not only resolves but also clarifies and illuminates the causal issues in the problematic causation cases that have plagued tort scholars for centuries', but that it 'is the essence of the concept of causation' (Wright 1985, pp. 1802, 1805).

In Robinson II, Tom's conduct is indeed a necessary element of some actual set of conditions that was sufficient for its occurrence, even if disposing of the waste of 40 would suffice to flood the island. When Tom throws waste into the lake in this case, his conduct becomes a necessary element in *an* actual set of conditions comprising it and, for instance, the conduct of others $(1, 2, 3 \dots 38 \text{ and } 39)$ who shovel waste in the lake. Together these 40 people would flood the island. And this, of course, is true of each of the 51 waste throwers. It is true of each that their waste disposal was a necessary element of *a* set of actual antecedent conditions—namely, any set excluding the waste disposal of 11 others, but including that of 39 others, which was sufficient to bring about the outcome. A plurality of sets of actual antecedent conditions was sufficient for Robinson's death. It is not the case that the conduct of any one of the 51 waste throwers was necessary for the sufficiency of all of these sets, but the conduct of each was necessary for the sufficiency of some of them.

On this view, there are overdetermination-based constraints against Tom's throwing waste in Robinson II, because by doing so he would become a contributor (NESS) to the outcome. Insofar as we accept that the reasons against becoming a contributor to harm are ordinarily quite stringent, then there is nothing mysterious about there being such reasons against Tom's shovelling his waste into the lake in Robinson II.

While this approach might rescue the idea that there are overdetermination-based constraints, in our view it does so by invoking an extremely implausible notion of

¹⁶ This view of contribution is developed by Hart and Honoré (1985). They acknowledge that it builds on J. S. Mill's notion of a jointly sufficient set of conditions, as well as on Mackie's (1980) idea, in the context of causal generalizations, of an INUS condition (an insufficient but necessary part of an unnecessary but sufficient condition).

¹⁷ Capitals added to clarify the acronym.

¹⁸ Hart and Honoré do not make any such claim.

contributing to harm. A well-known example from the literature on causation (cf. Lewis 2000) helps illustrate this. Bill and Ben each throw a rock towards a window. The window shatters. They are convicted of vandalism, and are required to compensate the shopkeeper. Alice has it all on tape. A closer look reveals that it was only Bill's rock that hit the window. Ben's rock followed its trajectory soon thereafter, passing through the shattered glass, but without hitting it—it did not change the timing or the manner of the occurrence of the shattering of the window. To say that Ben contributed to the shattering of the window when neither he nor his rock came into contact with the glass (and where no energy was transferred from the approaching rock to the glass) is implausible. Ben was accordingly not a contributor to the breaking of the window, and any understanding of contribution that maintains that he did contribute to this outcome is implausible.

Some legal theorists nevertheless seem to favour this NESS account of contribution in allocating costs for accidents and punishments for crime (Stapleton 2008). It is not difficult to understand why, given that many legal systems currently make causal contribution to harm a condition for liability to bear cost to address it (American Law Institute 1965). One might be concerned that overdeterminers will be let off the legal hook unless one adopts a NESS account of causal contribution, or something akin to it. The correct response to this, however, is not to pretend that an agent contributes to something by saying that he is a NESS of it. Ben does not contribute to (although he is a NESS of) the breaking of the window in the case imagined above. If we believe that overdeterminers should be made liable for the costs of their conduct, we should instead conclude that there are good grounds to change the requirements for liability. There might be good reasons to make both Bill and Ben bear the cost of repairing the window, and good reasons perhaps to punish them. Making actual contribution to harm a requirement for duties to provide compensation to the victims of harm could be questionable, morally speaking (Waldron 1995).

A consequentialist solution?

Shelly Kagan has recently tried to address our question regarding overdeterminationbased constraints from a consequentialist perspective. At first glance, the idea of consequentialist support for overdetermination-based constraints does not appear promising. After all, consequentialism enjoins agents to focus on the differences that they can make to outcomes that are morally significant; and the cases we have been discussing are precisely those where it seems that no one person makes a difference to harmful outcomes by qoing. However, Kagan argues, on the contrary, that in the case that concerns us the expected disvalue of engaging in overdetermining conduct will outweigh its expected value, and is thus easily handled with ordinary consequentialist machinery. Kagan distinguishes between two basic types of overdetermination cases, which he refers to as 'triggering cases' and 'imperceptible difference cases' (Kagan 2011, pp. 117-118). Our Robinson example is clearly a triggering case, so we will focus on Kagan's treatment of this type.

In triggering cases, he claims, the action of each person involved most likely makes 'no difference at all' (Kagan 2011, p. 118). However, there is *some* act that triggers the morally relevant outcome. He illustrates this idea through the example of individual

contributions to the killing of commercially raised chickens. He proposes that poultry sales typically function by way of thresholds.¹⁹

presumably it works something like this: there are, perhaps, 25 chickens in a given crate of chickens. So the butcher looks to see if 25 chickens have been sold, so as to order 25 more ... When 25 have been sold this triggers the call to the chicken farm, and 25 more chickens are killed, and another 25 eggs are hatched and raised in torture. (Kagan 2011, p. 122).

In this case, he claims, it is true that a great many chicken purchases make no difference to the number of chickens tortured and killed. This is because it makes no difference 'whether 7, 13 or 23 chickens have been sold' (Kagan 2011, p. 122). So those who buy the 7th, 13th or 23rd chicken will make no difference to the number of chickens tortured and slaughtered. The person who buys the 25th chicken, however, will make a difference to that number: their purchase will lead to a call to the chicken farm. On Kagan's view then, what could make it wrong to shovel waste into the lake in our Robinson case is that when each person acts there is some probability that *their* shovelling will be what triggers—makes a difference to—the morally relevant outcome (in this case Robinson's death).

This argument shows, we think, that the consequentialist can often give reasons why people should refrain from acts that they *believe* may overdetermine harm. We don't think, however, that it shows that the consequentialist has a reason why people should refrain from such acts when—as in our Robinson case— it is *known* that they will overdetermine harm. In our Robinson case, it is not true of the agent that there is some chance that their conduct will make a difference to the morally relevant outcome. In fact, they know that it won't. Kagan himself admits that in variations of his cases where the number of people buying chickens is not a multiple of 25, *all* can truly say that their purchase made no difference to the amount of chicken suffering. (Kagan 2011, p. 128). So our question is whether in cases like this there is any reason for a particular person to refrain from chicken purchasing. And here it is hard to see how the consequentialist can respond. They can only say, as Kagan himself seems to admit, that we often have no idea whether or not we are necessary elements of a triggering cohort that makes a difference to the occurrence of harm or members of a cohort that overdetermines it (such that no member is a necessary element of the set that triggers it.)

It seems that Kagan also overstates the force of his response even in cases where agents lack knowledge about whether they will be overdeterminers of harm. For example, there may be many cases where the chance of making a morally relevant difference by φ ing is extremely small, whereas the expected value of φ ing is quite significant. So it is not true that appeals to expected utility will 'guarantee' that it is wrong to perform the overdetermining act (Kagan 2011, p. 120; as pointed out by Nefsky 2012, at p. 369). How likely is it that cases of overdetermination will be such that the expected disvalue of possibly being an agent that triggers harmful consequence by φ ing will give

¹⁹ This same argument, made through the same example, was made previously by Peter Singer (Singer 1980 at p.335). Curiously, Kagan does not cite Singer's earlier treatment of this issue (in the very same journal.)

you reason to refrain from it, given what else is at stake? That seems to depend largely on the case that is being considered.

In cases involving markets, such as those discussed by Kagan, there may be good reason to believe that individuals will have such reason. This is because, as he points out, there may be market incentives for poultry shops to want to ensure that the number of chickens they order is more or less the same as the number of purchases required before a new order is triggered. In this case, the chance that individuals will have made a difference to the overall number of chickens ordered may be (relatively) high. But this will not be true of all cases. With respect to environmental harms and climate change, for instance, there will be no mechanism of this sort. Here it may often be the case that overall contributions wildly overshoot the number needed to bring about the morally relevant outcome. And in this case the chance that any will have made a difference to the potential harm caused by it and the likelihood that the particular φ ing will bring it about, it may vary a great deal from case to case, and consequently be more or less likely to outweigh the expected value of φ ing.

An alternative approach

Levels of description

Does overdetermination even exist? Some simply deny the existence of overdetermination in a world like ours, where there are certain lawful physical processes.²⁰ To be sure, there are many situations in which some set of people *appears* to consist of overdeterminers with respect to some injury, but perhaps this is only because we lack the necessary detailed microanalysis of the situation and how it has come about. Accordingly, the question 'what is wrong with being an overdeterminer of harm?' would not even get off the ground.

Harms and other outcomes can be described generally—a window is broken, Robinson dies—or they can be described more specifically—the window breaks and Robinson dies at particular times and in particular ways. One can clearly be an overdeterminer of harms that are described at a very general, rough-grained level. However, it is very often the case that a person is not an overdeterminer with respect to the outcome when it is described in a more fine-grained fashion. Whether a person is an overdeterminer with respect to some outcome, then, seems to depend on the level of detail at which that outcome is described. It may well be that in a world such as ours no outcomes are overdetermined when they are described at a sufficiently fine-grained level of detail.

Consider again the case of the two rock throwers. Bill and Ben each throw rocks. The window shatters. They are convicted of vandalism, and required to compensate the shopkeeper. Alice has it all on tape. A close look reveals how the two rocks together break the window. Suppose that on the basis of this footage we can safely infer that the

²⁰ See Bunzl (1979). For arguments against this view, see Schaffer (2003), who claims that overdetermination is 'everywhere'.

breaking and shattering of the window would have been slightly different had only one of the two rocks hit the window. At the rough-grained level of description, the outcome may be said to be the same as what would have occurred had either (but not both) of them thrown the rock—the window is shattered. The outcomes are different, however, when described in a finer-grained way—the manner in which the window shattered would have been different if one or the other rock alone had hit the window. There are consequently no overdeterminers with regard to the outcome described in this finegrained way—the precise manner in which the window was shattered.

Consider cases that are commonly described as involving pre-emptive overdetermination.²¹ In such cases, there are two events, each of which could be sufficient to cause some harm. However, one process pre-empts the ability of the other to bring about the harm. For instance, Al and Jonathan both start fires to burn down Martin's house. Al's fire reaches the house first and burns it down. While both Al and Jonathan are overdeterminers relative to the rough-grained description of the outcome (burning down the house), neither Al nor Jonathan is an overdeterminer relative to the fine-grained description of the outcome: Al's fire burns down the house, and Jonathan's doesn't burn down anything.

Of course, there are many other agents who may affect the timing and manner of occurrence of outcomes—including individuals who try to prevent these outcomes. Susanna's phone call to distract Al from his mission may *delay* the burning down of Martin's house, and change other details about its occurrence. However, it is not true of Susanna that there is an intact sequence linking her with the burning down of Martin's house. She *enabled* the occurrence of the burning of the house in the way that it actually happened, but she clearly didn't burn it down.

In Robinson II, is Tom an overdeterminer with respect to the outcome when it is described in this fine-grained way? Suppose that 40 loads of waste would suffice to flood the island. Tom is among the 51 who throw waste into the lake. We need to know whether they are doing it at exactly the same time. If not, and they are at the same distance from the island, a detailed analysis is likely to reveal that it was the waste shovelled into the lake by the first 40 only that caused the water to rise above the island, and that the last 11 were not relevant to the killing of Robinson. Hence, the 40 who first threw their waste into the lake together killed Robinson by flooding the island, while the 11 others did not contribute to his death. In this case, neither Tom nor the others would be overdeterminers of Robinson's death when it is described in a fine-grained manner. Some of them would have contributed to it, others would not have.

By contrast, if the 51 people threw the waste into the lake at *exactly* the same time, and they were at *exactly* the same distance from the lake, a detailed analysis of the situation is likely to show that *each* of them contributed to Robinson's death—there would be intact causal sequences linking each person's waste disposal with this drowning—some of the waste of each would be part of the process involving the raising of the water levels. They flooded the island with 40/51 = roughly 78 per cent of their waste. The rest of the waste of each was redundant. Hence, all 51 people contributed to

²¹ These are sometimes referred to as Frankfurt-type cases, after Harry Frankfurt, who is credited with introducing them into the literature (Frankfurt 1969).

the flooding of the island, and there were no overdeterminers of the outcome described in a sufficiently fine-grained way.

A proposal: probability of being a member of the actual set

A fine-grained analysis seems to us to provide the correct diagnosis of who contributed to Robinson's death. It fails, however, to explain the moral reasons that apply to Tom when he is considering whether he can throw waste into the lake. Learning whether or not Tom actually did or did not contribute to Robinson's death as described in this finegrained way doesn't really explain why it is wrong for him to dispose of his waste in the first place.

The proposal here is that overdetermination-based constraints are based on the possibility that some agent will become a necessary element of *the* set of actual conditions that *actually* brings about the overdetermined outcome. The significance of the overdetermination-based constraint is therefore in an important way a function of the probability that by φ ing an agent will be member of *this* set. If the agent φ s, then an overall moral appraisal of his conduct would also depend on the costs to him of refraining from φ ing and on his knowledge of the situation that he is in, as well as the benefits his conduct may produce for others. The higher the costs to him, the less transparent the situation, and the larger the benefits his conduct will produce for others, the less culpable he will be.

This account of overdetermination-based constraints can be summed up as follows:

The significance of overdetermination-based constraints (against φ ing) is a function of how bad the overdetermined outcome is and the probability that by φ ing this agent will be an element of the set of actual antecedent conditions that brings this outcome about.

Assuming that the badness of the outcome is held constant, then the more unlikely it is that he will be among the actual set by his φ ing, the less significant his constraint against φ ing will be. If there is no risk that by φ ing the agent will be among the actual set of conditions that in fact does the harm, then there are no overdetermination-based constraints against her φ ing.²²

Consider how our proposed account applies to variations of Robinson II. Assuming that the necessary number of waste throwers is 40, then the significance of Tom's overdetermination-based constraints against throwing waste into the lake is reduced when the number of other people throwing waste into the lake increases. This is because when the number of waste throwers increases, the probability that Tom will be a member of the actual set decreases. Holding the necessary number of waste throwers constant at 40 his probability will be 40/40, 40/41, 40/42, and so on.

When holding the number of people throwing waste constant at 51, the significance of Tom's overdetermination-based constraint against throwing waste into the lake will increase as the number of waste throwers required to flood the island

²² The conduct may, of course, be wrong for some other reason.

increases. This is because the larger the number of waste throwers necessary to bring about Robinson's drowning, the greater the likelihood that Tom will be a member of this set. For instance, if only one person is required, then the probability that he is a member of the actual set is 1/51, if two people are required, then the chance that he is among them is 2/51, and so forth, until the required number is 51, in which case he will certainly be a member of this set.

Is the number of contributors relevant?

When only one person needs to φ to bring about an outcome, and there are 51 people who φ , then the probability that any particular individual will be member of the actual set is much lower than if a larger number of people were required to bring it about. According to our proposal, this means that the overdetermination-based constraint against throwing waste will be correspondingly weaker on each than if a larger number of people were required. But perhaps there is an additional factor that we have failed to consider that might militate against this conclusion: the absolute number of people whose conduct will in fact bring about the outcome.²³ If Tom should happen to be a member of the actual set that brought about the flooding when only one person is required to flood the island, he would be the lone contributor to Robinson's death. This might seem to make a difference. Perhaps the size of the actual set is independently relevant for determining the significance of an agent's constraint against becoming an overdeterminer.

If this were true, then the size of the actual set necessary to do harm would be relevant for two separate assessments that together would determine the significance of an agent's overdetermination-based constraint: (1) the probability that any given person will be a member of the actual set is relevant for gauging the significance of the overdetermination-based constraint against φ ing (the smaller the set, the less likely they are in it, and the less wrong their φ ing) and; (2) the amount of harm that each person in that set is responsible for is relevant for the general wrongness of being a member of the actual set (the smaller the set, the more harm each member of the set is responsible for).

These two assessments could pull in different directions. The probability that an agent will become a member of the actual set by φ ing could be low. Correspondingly, the overdetermination-based constraint against φ ing is weak (on the first dimension). But if the agent nevertheless ends up as a member of the actual set, his role in bringing it about is substantial—he might even be the only one in the set—in which case his overdetermination-based constraint against φ ing would be stringent (on the second dimension.)

However, we are not convinced that the size of the set matters. Why should it matter that you are one of many when you are a necessary element in the actual set? Consider Robinson II. If it takes 51 people to kill Robinson, and that is well known by all, it is not clear that any one of them is any less to blame for his death than if one of them were to drown him alone. After all, each of them is in a position to avoid killing Robinson by refraining from pushing waste into the lake. That one person's contribution depends on other factors to be effective doesn't seem to alter the extent of her

²³ Thanks to Geoff Brennan for pointing out this possibility.

responsibility when she knows that these factors obtain, regardless of how many additional necessary factors there are. Looked at from another angle, each of these contributors would, it would seem, be liable to defensive force (should it become possible to defend Robinson) that would be comparable to that of one person whose conduct alone would be sufficient to drown him. As noted above, this seems a good indicator of the cost they would be obliged to take on to avoid contributing to harming Robinson.

Nor is it obvious that the size of the actual set is of independent relevance, even when the agents involved do not know whether the other necessary factors obtain. When Tom considers whether or not to shovel waste into the lake, his uncertainty about whether or not enough others have done so that his conduct would combine with theirs to bring about Robinson's death is obviously relevant to his decision. But it is not clear whether the *numbers* of factors that would have to be present for the outcome to be brought about is relevant in itself. When the conduct of many people is necessary for bringing about an outcome, then it will often be more difficult for any particular person to ascertain that they will become a contributor to this outcome if they φ . The true significance of numbers could therefore simply be that the more factors that are necessary to bring about some outcome, the more likely an agent is to be uncertain about whether his conduct will be necessary for it to occur.

Stringency

Our focus so far has been on proposing a way of understanding overdeterminationbased constraints that undermines the appeal of scepticism about reasons against overdetermining harm. But what of the view that Tom's being an overdeterminer (rather than a necessary condition) of Robinson's drowning makes no difference to the moral reasons that apply to Tom?²⁴ That is, the constraints that apply to him are just as stringent as those that would apply were his conduct instead to be a necessary condition for the harm.

This position is implausible. The fact that the bad outcome will happen whether or not Tom φ s clearly should make a difference to the weight of reasons against his φ ing. This is supported by considered judgements in a range of cases. It seems to make a significant moral difference whether one is exposing someone to 1/51 risk of death or is instead an overdeterminer with a 1/51 risk of being among the actual set that kills a person. In both cases there is some possibility that you will kill him. But only in the former case is there a possibility that you will make a difference to whether or not he is killed. There seems also to be significant moral difference between Tom's pulling the trigger of a Russian roulette gun with 51 chambers and his being one of 51 people shovelling waste into a lake (or one of 51 shooting at a person's head, for that matter), when it takes only one to kill the victim.

The fact that the bad outcome is going to happen whether or not one φ s seems to function as a factor that reduces the stringency of moral reasons against φ ing. Moreover, the degree to which it reduces the significance of the overdetermination-based constraint depends crucially on the motives of the agent who is φ ing, and the interests

²⁴ This seems to be the position defended in Goldman (1999) and Tuck (2009, pp. 32-60).

that are served when he φ s. As originally described, Tom shovels waste into the lake in Robinson II simply to get rid of it, because he can reap benefits from doing so, and for no other reason. However, we could modify the case to allow him other reasons for throwing the waste into the lake. For instance, he might need to get rid of it to save another person from drowning. In that case, it seems that the reasons against his throwing waste into the lake are outweighed by his positive reasons to save that person. And that seems to be the case even when the probability that he will be a member of the set that actually does harm to Robinson is very high. By contrast, it does not seem that he would be permitted to φ in order to save another person if his φ ing would drown Robinson when it is *not* the case that Robinson would have drowned anyway—we cannot do harm to one innocent person to save one other innocent person from a comparable harm to which we would not be a contributor.

Earlier in this essay we noted that it was widely accepted that doers of harm cannot easily justify their conduct by appealing to the costs to them of refraining from doing harm, or to comparable harms that their conduct will prevent for others if they do harm. Our claim here is that it is much easier for overdeterminers to appeal to costs of these kinds to justify their conduct. Overdetermination-based constraints are simply not as significant as constraints against doing harm, when your conduct is necessary for harm being done. That the bad outcome is going to happen increases the plausibility of the appeal to cost or to prospective benefits for others.

Could Tom invoke just *any* reason to protect another person from harm to override his overdetermination-based constraint against throwing waste when the island will be flooded whether or not he does so? If he could, this constraint would be extremely weak—hardly worth its name. It may seem plausible that these constraints are very weak when the probability that Tom will be a member of the actual set is very low. On the other hand, when the probability that he will be among the actual set is high, the fact that his conduct would bring minor benefits to him or others does not seem sufficient to override this constraint. If Alice needs to tidy up her garden, it seems wrong of Tom to help her by throwing the waste into the lake if the situation is as described in Robinson II. This is so even if the probability of his being among the actual set that causes the flooding is quite low—if 15 rather than 40 would suffice to bring it about. Although Robinson is going to drown anyway, Tom shouldn't throw waste into the lake unless an interest of considerable moral significance is served by his doing so. And it seems that if he throws waste into the lake without this sort of justification, he would become liable to defensive force by Robinson or others acting on his behalf.

What if Tom were to throw waste into the lake in Robinson II in order to benefit himself? It seems pretty obvious that he could do so if it were necessary to save his own life or avoid a very significant loss. But what if he instead stands to gain only significant monetary benefits? Suppose he is not throwing the waste simply to get rid of it, but because he must get dispose of it in order to extract some precious stones on his property. Suppose also that the only way to get rid of the waste is to throw it into the lake.

As noted initially, the issue of what to do with the *benefits* of overdetermining conduct is a distinct matter that we cannot address in detail in this chapter. However, it

seems to us that Tom should not proceed in this case and claim all the benefits for himself in Robinson II. Even when the probability that an agent is member of the actual set that brings about some harm by φ ing is very low, he seems required to share the benefits with those harmed when this is possible, or with relevant third parties (e.g. relatives) when it is not. When the overdeterminer φ s, he violates a constraint and correspondingly infringes the claim of the person that suffers the overdetermined harm. That he infringes a claim explains why it is impermissible for him to keep all the benefits for himself. If Tom ends up as an element in the actual set that brings about the outcome (described in a fine-grained way), then he seems to have no claim on the benefits he gains at all, and ought to compensate the victim as well.

Compensation

It is not obvious why anyone should provide compensation for victims of some harmful activity unless their conduct was in some way actually implicated in the harms that they have suffered. Compensation is due, in thef first instance, from those who ended up among the set that actually brought about the bad outcome, described in a fine-grained way. Sometimes, of course, we will be unable to determine who actually brought about the outcome described in such a way. In these cases, we propose that all of those who engaged in the activity ought to share equally the cost of compensation, all else being equal.²⁵ How is it that we can require these people to compensate when we cannot tell for sure that they actually contributed to the outcome? We are entitled to do that, because these agents have all engaged in the wrongful action of taking a risk of becoming an element in the actual set that brought about the outcome. It is therefore fair that they compensate the victim, since the overdeterminers are owed no benefit of the doubt.²⁶

References

American Law Institute (1965). Restatement (Second) of Torts. §§431–433.

Anscombe, G. E. M. (1961). War and Murder. In W. Stein (ed.), *Nuclear Weapons:* A Catholic Response (pp. 43–62). London: Burns & Oates.

Barry, C. and Øverland, G. (2012) The Feasible Alternatives Thesis. *Politics, Philosophy and Economics*, 11:1, 97-119

Bennett, J. (1995). The Act Itself. New York: Oxford University Press.

Bunzl, M. (1979). Causal Overdetermination. Journal of Philosophy 76(3), 134-150.

²⁵ Things would not be equal if the probabilities that different agents are elements in the actual set differ, or if some of these agents are more culpable than others. In this case those with higher probabilities of being in the actual set or who are more culpable would owe correspondingly more.

²⁶ We do not mean to deny that those who are not members of the set who actually bring about the harm may also be liable to provide compensation, but we are not endorsing this view either.

Frankfurt, H. (1969). Alternate Possibilities and Moral Responsibility. *Journal of Philosophy* 66(4), 829–39.

Glover, J. (1977). Causing Death and Saving Lives. Harmondsworth: Penguin.

Goldman, A. I. (1999). Why Citizens Should Vote: A Causal Responsibility Approach. *Social Philosophy and Policy* 2, 201–217.

Hall, N. (2002). Non-locality on the Cheap? A New Problem for Counterfactual Analyses of Causation, *Noûs* 36(2), 276–294.

Hall, N. (2004). Two Concepts of Causation. In J. Collins, N. Hall and L. A. Paul (eds), *Causation and Counterfactuals* (pp. 225–276). Cambridge, MA: MIT Press.

Harman, Elizabeth. (2009). Harming as Causing Harm. In M.A. Roberts and D. Wasserman (eds.) *Harming Future Persons*, (pp.137-154) Dordrecht: Springer.

Hart, H. L. A. and Honoré, T. (1985). *Causation in the Law*, 2nd edition. Oxford: Clarendon Press.

Jackson, F. (1997). Which Effects? In J. Dancy (ed.) Reading Parfit (pp. 42–53). Oxford: Blackwell.

Kagan, S. (2011). Do I Make a Difference? *Philosophy & Public Affairs* 39, 105–141.

Klockseim, J. (2012). A Defense of the Counterfactual Comparative Account of Harm," *American Philosophical Quarterly*, 48 (4), 285–300.

Lewis, D. (2000). Causation as Influence, Journal of Philosophy 97(4), 182–197.

Mackie, J. L. (1980). *Cement of the Universe*, 2nd edition. Oxford: Oxford University Press.

Moore, M. (2003). For What Must We Pay? Causation and Counterfactual Baselines, *San Diego Law Review* 40, 1181–1271,

Nefsky, J. (2012). Consequentialism and the Problem of Collective Harm: A Reply to Kagan. *Philosophy & Public Affairs*. 39, 364-395

Nolt, J. (2011). How Harmful Are the Average American's Greenhouse Gas Emissions?, *Ethics, Policy & Environment* 14(1), 3–10; DOI: 10.1080/21550085.2011.561584.

Parfit, D. (1983). Reasons and Persons. Oxford: Clarendon Press.

Pogge, T. (1998). The Categorical Imperative, in P. Guyer, ed., *Kant's "Groundwork of the Metaphysics of Morals": Critical Essays.* Lanham, MD: Rowman and Littlefield, 189-213.

Raz, J. (1990). Practical Reason and Norms, 2nd edition. Oxford: Clarendon Press.

Schaffer, J. (2003). Overdetermining Causes, Philosophical Studies 114, 23-45.

Shiffrin, S. (1999) Wrongful Life, Procreative Responsibility, and the

Significance of Harm. Legal Theory 5, 117-148

Singer, P. (1980) Utilitarianism and Vegetarianism, Philosophy & Public Affairs 9, 325-337.

Sinnott-Armstrong, W. (2005). 'It's Not My Fault': Global Warming and Individual Moral Obligations. In W. Sinnott-Armstrong and R. B. Howarth (eds) *Perspectives on Climate Change: Science, Economics, Politics, Ethics.* Amsterdam: Elsevier.

Stapleton, J. (2008). Choosing What We Mean by 'Causation' in Law. *Missouri Law Review* 73, 433–480.

Thomson, J. (1977). Some Ruminations on Rights, Arizona Law Review 19, 45-60.

Thomson, J. (2011). More on the Metaphysics of Harm. *Philosophy and Phenomenological Research* 82, 436–458.

Tuck, R. (2009) Free Riding. Cambridge, MA: Harvard University Press.

Waldron, J. (1995). Moments of Carelessness and Massive Loss. In D. G. Owen (ed.), *The Philosophical Foundations of Tort Law* (pp. 387–40). Oxford: Clarendon Press.

Wright, R. (1985). Causation in Tort Law, California Law Review 73, 1788-1813.