Climate Change and Decision Theory

Andrea S. Asker[1] & H. Orri Stefánsson[2]

Abstract. Many people are worried about the harmful effects of climate change but nevertheless enjoy some activities that contribute to the emission of greenhouse gas (driving, flying, eating meat, etc.), the main cause of climate change. How should such people make choices between engaging in and refraining from enjoyable greenhouse-gas-emitting activities? In this chapter we look at the answer provided by decision theory. Some scholars think that the right answer is given by *interactive* decision theory, or game theory; and moreover think that since private climate decisions are instances of the *prisoner's dilemma*, one rationally should engage in these activities provided that one enjoys them. Others think that the right answer is given by *expected utility theory*, the best-known version of individual decision theory under risk and uncertainty. In this chapter we review these different answers, with a special focus on the latter answer and the debate it has generated.

Keywords: Catastrophic risk; Climate change; Climate thresholds; Decision theory; Expected utility; Expected harm; Extreme uncertainty; Imperceptible harm; Risk

Index terms: Causal decision theory; Climate catastrophes; Climate change; Climate harm; Climate thresholds; Decision theory; Decision-maker; Decision problem; Evidential decision theory; Expected utility; Expected Harm; Extreme Uncertainty; Imperceptible harm; Joyguzzling; Prisoner's dilemma; Rationality; Risk; Tragedy of the commons; Thresholds

1. <u>Introduction</u>

As a reader of this handbook, you are likely to be worried about climate change, and would presumably be willing to accept some cost to lessen the harm of climate change. But, like most other people, you are also likely to enjoy some activity that is associated with greenhouse gas emission, the main cause of anthropogenic climate change. If so, then this chapter might help you make rational decisions about whether to engage in such activities. It might be worth emphasising from the start that our intention is not to tell you what to do when faced with the option to engage in an enjoyable greenhouse-gas-emitting activity.

[1] Stockholm University and Institute for Futures Studies.
[2] Stockholm University, Swedish Collegium for Advanced Studies, and Institute for Futures Studies.

(By "greenhouse-gas-emitting activity" we mean an activity that requires the emission of greenhouse gas; examples include eating meat, flying, and driving a fossil fuel powered vehicle.) Instead, our aim is to discuss the tools with which you can evaluating such options.

The outcomes of actions that are associated with greenhouse gas emission are highly uncertain (see, e.g., Broome 2012, Bradley and Steele 2015, Pindyck 2020). That is, while we know that, on an aggregate level, greenhouse gas emissions cause climate change and associated harmful events (such as floods, droughts, famines, etc.), we can never know beforehand what effect (if any) an individual act of greenhouse gas emission has on the climate. As decision theory is a tool (in fact, *the* tool, some might say) for choosing when you don't know what outcomes your options will bring, decision theory should be able to help you decide whether to engage in a greenhouse-gas-emitting activity. But what answer does decision theory provide? And, more generally, how should a climate-change-conscious agent reason about an enjoyable greenhouse-gas-emitting activity, according to decision theory? These are the main questions that this chapter will address.

We start by considering the idea that *private climate decisions*—as we shall call decisions by individuals about whether to engage in greenhouse-gas-emitting activities—are examples of the *prisoner's dilemma*, which, if correct, means that orthodox decision theory recommends engaging in enjoyable greenhouse-gas-emitting activities. We provide some arguments against thinking that the prisoner's dilemma is the correct way to model such decisions, and suggest instead that they should be modelled using expected utility theory. Moreover, we argue that expected utility theory often recommends that climate-change-conscious agents should *not* engage in enjoyable greenhouse-gas-emitting activities. We also discuss why some have been sceptical of this recommendation, and why others have been sceptical of the applicability of expected utility theory to private climate decisions.

2. Climate change and decision theory

2.1 Prisoner's dilemma

In this subsection we first describe prisoner's dilemma in its standard formulation, before connecting it to climate decisions. Imagine that Abel and Beatrice were caught by the police while entering a bank they intended to rob. Without confession, the prosecutor can only get them sentenced for unlawfully carrying weapons, which lands them each one year in prison, in comparisons to ten years for attempted robbery. The prosecutor promises each of Abel and Beatrice that if one of them confesses but the other does not, then she will not press charges against the confessor. If, however, both confess, then the prosecutor will charge both for attempted robbery, but they can expect to get a couple of years knocked off the ten-year sentence in return for having collaborated with the prosecutor.

Assuming that the only thing that matters to Abel and Beatrice, when deciding between confessing and not, is the length of their own prison sentence, we can formulate each criminal's decision problem as in Table 1. The leftmost column represents the actions available to the criminal whose decision problem is being modelled, the top row represents the possible actions that the criminal's comrade could take. The four cells specifying length in prison are the different outcomes, for the criminal whose problem is being modelled, from the corresponding combination of actions by the criminal and their comrade.

	My comrade confesses	My comrade doesn't confess
I confess	I get 8 years in prison	I get 0 years in prison
I don't confess	I get 10 years in prison	I get 1 year in prison

Table 1: Abel and Beatrice's prisoner dilemma

Given what we have assumed about Abel and Beatrice's preferences, that is, that in this decision problem they only care about the length of their own prison sentence, it should be evident that it is the interest of each to confess. Whatever the other person does, each does best by confessing—which, in the language of game theory, means that both confessing is the only *Nash equilibrium* of the game (named after Nash 1950). For instance, if Abel confesses, then by also confessing Beatrice can ensure 8 rather than 10 years in prison; whereas if Abel doesn't confess, then by confessing Beatrice can walk free rather than sitting one year in prison. The same is true for Abel; whatever Beatrice does, Abel does better by confessing.

We can state the prisoner's dilemma more generally and formally. Let *P* and *Q* be two action types, e.g., confessing or not confessing. We assume that there are two "players" who each can choose between *P* and *Q*; one player can choose between the instances of *P* and *Q* in the first column, the other player can choose between the instances of *P* and *Q* in the top row. We moreover now assume that each player's preferences can be represented by *ordinal utilities*, that is, numbers where a higher number represents a more strongly desired outcome.[3] Finally, (x,y) denotes that the player choosing from the first column gets outcome represented by utility x while the player choosing from the top row gets outcome represented by utility y. Then two players find themselves in a prisoner's dilemma just in case their decision problem has the structure of Table 2.

	P	*Q*
P	(2,2)	(4,1)
Q	(1,4)	(3,3)

Table 2: An abstract prisoner's dilemma

What is fascinating about the prisoner's dilemma, is that if both players are instrumentally rational—that is, if they choose the option that they believe will best serve their ends—then they will both choose *P*, and thus get the outcome (2,2) which is *strictly Pareto dominated* by the outcome (3,3), meaning that they both strictly prefer the latter to the former. So, by each acting rationally, they get an outcome that is, from each player's

[3] "Ordinal" utilities *only* represent how the outcomes are ordered by the agent's desires. So, for instance, the magnitude of the difference between ordinal utilities is not significant. In contrast, "cardinal" utilities, which will play a role later, represent how differences between outcomes compare, in addition to how the outcomes themselves compare, according to some agent's desires.

perspective, strictly worse than an outcome that they could have obtained by "cooperating". To see that they will, if rational, each choose *P*, note that if one player chooses *Q*, then the other player can get 4 by choosing *P* rather than 3 by choosing *Q*; but if one player chooses *P*, then the other player can get 2 by choosing *P* rather than 1 by choosing *Q*. So, whatever the "opponent" does, one does better by choosing *P*.

Intuitively, it might seem that private climate decisions have important features in common with the prisoner's dilemma (see, e.g., Gardiner 2001, 2011, and Johnson 2003). For concreteness, suppose that the enjoyable greenhouse-gas-emitting activity under consideration is to go for a Sunday drive in a gas-guzzling SUV (Sinnott-Armstrong 2005). Now, the occurrence of harmful climate change—the "climate crisis", as we shall from now on call it—does not depend on whether you go for the drive or not; it will (let us assume) be avoided if and only if the majority of *other agents* (including corporations) change their behaviour. (Harmful climate change is of course already happening, but to simplify, let's use the term "climate crisis" for some particularly harmful climate change that could, at least in theory, still be averted.) Therefore, Table 3 might seem to be a way of formulating the climate decision in question.

	Others don't change their ways	Others change their ways
Go for Sunday drive	Climate crisis occurs but I get to enjoy Sunday drive	Climate crisis is averted and I get to enjoy Sunday drive
Don't go for Sunday drive	Climate crisis occurs and I don't get to enjoy Sunday drive	Climate crisis is averted but I don't get to enjoy Sunday drive

Table 3: Climate decision as prisoner's dilemma

If this is how to formulate a private climate decision, for a climate-change-conscious person who enjoys Sunday drives, then it would seem that there is only one rational choice: to go for the drive! The same holds for other enjoyable greenhouse-gas-emitting activities.

But if that is true, then it is rational for people—even those who are concerned by climate change—to continue pursuing their greenhouse gas emitting activities, as long as they enjoy them. The climate crisis is going to occur anyway, whatever any individual does, so she might just as well enjoy these activities while she can.

Partly for the above reason,[4] climate change is often said to be an example of Hardin's (1968) *tragedy of the commons*, that is, an outcome that obtains as a result of actions by a collection of agents, where each agent acted rationally—in light of their preferences—but still the outcome that results is worse—from each agent's perspective—than an outcome that they could in theory have brought about by acting together.

Moreover, and for similar reasons, a number of climate ethicists have argued that individuals have no moral obligation to refrain from private activities (driving, flying, etc.) that cause greenhouse gas emission (see, e.g., Johnson 2003, Sinnott-Armstrong 2005, Kingston and Sinnott-Armstrong 2018). To put it simply, these "individual denialists" (as Broome 2019 calls them) contend that each such private act of greenhouse gas emission does no harm—in particular, the occurrence of the climate crisis is independent of the act—and therefore we do not have a moral obligation to not engage in it. As Sinnott-Armstrong (2005: 297) puts it:

> Global warming will still occur even if I do not drive just for fun. Moreover, even if I do drive a gas-guzzler just for fun for a long time, global warming will not occur unless lots of other people also expel greenhouse gases. So my individual act is neither necessary nor sufficient for global warming.

A weakness in individual denialism is that it overlooks the fact that although the climate crisis will occur independently of any particular private climate action, such an action could affect *how harmful* the effects of climate change will be. Moreover, even if—as some hold—any single private climate action will *most likely* make no difference, it

[4] Why only *partly* for this reason? Because we have only considered the decisions of (private) individuals. But climate change is of course mainly the result of the decisions of *group agents*, such as governments and large corporations. Thus, the tragedy of climate change may only partly be due to private climate decisions being instances of the prisoner's dilemma.

could carry a great *expected harm* (and benefit). In sum, individual denialists seem to commit two related fallacies. First, they wrongly view climate change and the harm it will cause in binary terms rather in terms of degrees. Second, they fail to consider the fact that a decision can carry great expected harm even if it most likely causes no harm. In section 2.3 we consider some estimates of the expected harm done by private climate decisions.

If the observations from the last paragraph are correct, then it would be incorrect to think of private climate decisions as prisoners' dilemmas; more generally, it is then inappropriate to model private climate decisions as a game where the uncertainty pertains *only* to the actions of others. Instead, we should evaluate private climate decisions using *expected utility theory*, which requires that we formulate them in the framework of individual decision-making under uncertainty[5] (where the uncertainty concerns the true "state of the world"). That is what we do in the next subsection. In subsequent subsections we consider some objections to how we apply expected utility theory to private climate decisions, and some more general worries about the applicability of expected utility theory to such decisions.

2.2. Expected Utility Theory

We start by describing expected utility theory as an abstract theory, before explaining how it can be applied in private climate decisions. Let **O** be a finite set of *n* outcomes, denoted o_1, o_2, etc, up to o_n, while **A** is a finite set of alternatives, or options, that is, whatever the agent of interest can choose between, with individual alternatives denoted a_1, a_2, etc. u is a cardinal[6] utility function on the set of outcomes, which numerically represents how desirable the agent in question considers the outcomes, while p is a (subjective) probability function, which numerically represents the agent's degrees of belief (i.e., her uncertainty),

[5] The term "expected utility theory" is sometimes reserved for what economists like to call decision-making *under risk*, that is, when the options carry with them given (or "objective") probabilities for their potential outcomes. The term "subjective expected utility theory" is then used for what economists like to call decision-making *under uncertainty*, that is, when the decision-maker has to consult her own degrees of belief (or subjective probabilities) for the potential outcomes (see, e.g., Steele and Stefánsson 2015, section 3). Since we assume that when it comes to climate decisions, all probabilities are subjective—and no options come with given probabilities—we ignore this distinction.

[6] Recall footnote 1.

and where $p(o_j \mid\mid a_i)$ denotes the degree with which the agent believes that outcome o_j obtains *under the supposition that* she chooses alternative a_i. For now, we leave open how to interpret this supposition, but it is an issue to which we return below.

With the above notation in hand, we can define the *expected utility* (EU) of any alternative, a_i, as follows:

$$EU(a_i) = \sum_{j=1}^{n} u(o_j)p(o_j \mid\mid a_i)$$

It is worth noting that everything we say about expected utility in what follows also holds for normative generalisations of the theory, such as *risk-weighted* expected utility theory (Buchak 2013; for discussion, see Bradley and Stefánsson 2019). In other words, we shall remain quite liberal about what counts as "expected utility".

Now, according to expected utility *theory*, any rational agent has preferences that satisfy a number of constraints, which mathematically entail that the agent *maximises expected utility* (for some results of this kind, see, e.g., Ramsey 1926, von Neumann and Morgenstern 1944, Savage 1954, Bolker 1967, Bradley 2017). These results—the mathematical entailments—are called "representation theorems". Put slightly more precisely, any version of expected utility theory holds that a rational agent has preferences between the alternatives in **A** such that, for any two alternatives a_i and a_j in **A**, the agent prefers a_i to a_j just in case the expected utility (according to the agent) of the former is greater than the expected utility (according to the agent) of the latter. We shall get back to the potential practical importance of representation theorems in section 2.5.

So, what does expected utility theory imply for the question of whether to engage in a greenhouse-gas-emitting activity, such as going for a Sunday drive in a gas guzzling SUV? That of course depends on one's desires (or values), as represented by the utility function, u, and one's degrees of belief or uncertainty, as represented by the probability function, p. As to the former, we have assumed that the agent in question enjoys the activity but is worried about the harmful effects of climate change. That is, she prefers the effects of climate change to be less rather than more harmful, other things being equal; but she also enjoys the greenhouse-gas-emitting activity and so prefers to engage in it, other things being equal. The question we shall soon address is how weight these two preferences against each other when they conflict.

What about the probability function? Let us assume that the agent's uncertainty, as represented by p, is in line with the currently available evidence. So, she knows that climate change is happening, and that it is largely caused by the atmospheric concentration of greenhouse gas, and that further greenhouse gas emission will make climate change even worse. But one might wonder: does that hold for *any* amount of further greenhouse gas emission? Some think not: the greenhouse gas emission associated with a *single* Sunday drive, say, is so insignificant that it won't make climate change any more harmful (see, e.g., Sinnott-Armstrong 2005, Budolfson 2017, Cripps, 2013). Others disagree (see, e.g., Hiller 2011, Broome 2012, 2019, Morgan-Knapp and Goodman 2015, Lawford-Smith 2016). We shall get back to this disagreement in a moment. But first, we note that the practical implication of this disagreement partly depends on which version of expected utility theory one adopts.

Recall that we defined the expected utility of alternative a_i as follows: $EU(a_i) = \sum_{j=1}^{n} u(o_j)p(o_j \,||\, a_i)$. But we left open the interpretation of $p(o_j \,||\, a_i)$, saying only that it denotes the degree with which the agent believes that outcome o_j obtains *under the supposition that* she chooses alternative a_i. This supposition can be interpreted *evidentially*, as a traditional conditional probability (as in the *evidential* decision theory of Jeffrey 1965), or it can be interpreted *counterfactually* (or *subjunctively*), as the probability that o_j *would* come about if one *were to* choose a_i (as in, e.g., the *causal* decision theory of Joyce 1999). The difference between these two interpretations can have interesting implications for what expected utility theory implies for private climate decisions. To see this, suppose that a person reasons as follows:

> If I decide to go for a Sunday drive in my gas guzzling SUV, then that is evidence that others will do so too; not because my going causes them to go, but because if I conclude that—despite the evils of greenhouse gas emission—I will allow myself to

enjoy a Sunday drive, then other people like me will probably come to that same conclusion.

A person who reasons like this will judge the decision to go for a Sunday drive differently depending on whether she uses causal decision theory or evidential decision theory. Take the rather vaguely specified outcome that atmospheric greenhouse gas reaches concentration level X, which we assume to be particularly harmful. If a person uses causal decision theory when reasoning about a Sunday drive, then the outcome where atmospheric greenhouse gas concentration reaches this level X will be weighted by the probability that *her* Sunday drive *causes* this outcome. In contrast, if she uses evidential decision theory, then the outcome in question will be weighted by the conditional probability that this concentration will be reached *given that* she goes on a Sunday drive. So, due to the person's belief about the evidential but not causal relationship between her decision and what others decide, the possibility of reaching concentration X will count more heavily against the Sunday drive if she uses evidential decision theory than if she uses causal decision theory.

To establish more formally the above difference between causal and evidential decision theory, consider first the implication of interpreting $p(o_j \,||\, a_i)$ evidentially, that is, as the conditional probability of outcome o_j obtaining given that alternative a_i is chosen. Let o_j be the outcome that atmospheric greenhouse gas reaches concentration level X and a_j the alternative to go for a Sunday drive on some particular evening. Then, since the person takes her decision to go for a Sunday drive to be evidence that others will do so too—and since she knows that driving increases the atmospheric concentration of greenhouse gas—she will find that $p(o_j \,||\, a_i)$ is considerably higher than $p(o_j)$, and, more importantly, considerably higher than $p(o_j \,||\, \text{not-}a_i)$, the conditional probability of greenhouse gas reaching level X given that she does *not* go for that drive. Less formally, she finds that the probability that atmospheric concentration of greenhouse gas reaches particularly harmful levels is considerably higher if she goes for a Sunday drive than if she does not. If she applies evidential decision theory and thus uses evidential suppositions when choosing what to do, this may suffice to deter her from going on that Sunday drive.

A person who applies causal decision theory will reason quite differently. When considering the probability that o_j *would* come about if one *were to* choose alternative a_i, one should (to simplify somewhat) keep fixed everything not causally affected by the choice of a_i. And recall that we are assuming that the decision-maker takes her decision to go for a Sunday drive to be evidence that others will do so without causing them to do so. Thus, since o_j is the outcome that atmospheric greenhouse gas reaches level X and a_j is the alternative to go for a Sunday drive on some particular evening, interpreting '||' counterfactually means that $p(o_j \mid\mid a_i)$ may not be (significantly) higher than $p(o_j \mid\mid \text{not-}a_i)$. Since this is the probability that figures in causal decision theory, the possibility of reaching harmful levels of atmospheric concentration of greenhouse gas may count less heavily against going on that Sunday drive according to causal decision theory than according to evidential decision theory.

Some philosophers find it obvious that causal decision theory is the correct normative decision theory. One should, on this view, not choose the course of action that generates the "best news", i.e., one shouldn't focus on what is *evidence* for the best outcome; instead, one should choose the course of action that in expectation *causes* the best outcome (see, e.g., Joyce 1999). Others disagree and point out that a decision-maker who chooses the actions that are evidentially (even if not causally) related to the better outcomes will do better in the long-run (e.g., Ahmed 2014). Yet others have a more conciliatory approach and suggest that there may be no fact of the matter as to which of these theories is the correct one (see, e.g., Briggs 2010).

This is not the place to enter into the debate between causal and evidential decision theory. However, in what follows, we shall consider in more detail what the causal interpretation of expected utility theory recommends when making private climate decisions. The reason we assume causal decision theory in what follows is that a recent and ongoing debate about the ethics of private climate decisions can, as we shall see, be fruitfully interpreted as a debate over what causal decision theory recommends, when a climate-change-conscious decision-maker reasons about an enjoyable greenhouse-gas-emitting activity. In the following subsections we review that debate.

2.3 How you could make a *large* difference

Some ethicists have recently argued—largely in response to individual denialists (recall section 2.1)—that due to the chaotic nature of the climate, it is possible that small changes in the atmospheric concentration of greenhouse gas have surprisingly large effects.[7] The idea is that because of the atmosphere's extreme instability and the so-called "butterfly effect", some small intervention like a Sunday drive can cause a large disturbance to the weather (Broome, 2019, pp. 112–113; Cullity, 2019, p. 24). For instance, a Sunday drive in a gas guzzling vehicle could cause the passing of some *metrological threshold* (Morgan-Knapp and Goodman 2015) which could trigger some extreme and potentially very harmful weather events.[8] Moreover, some (e.g., Broome 2019) hold that in the long-run, the time and place of some harmful events associated with climate change—such as storms, floods, and disease outbreaks—will almost certainly be altered by a Sunday drive, thus affecting who experiences harm, in addition to increasing the severity and frequency of the climate-change induced harm.

The notion of thresholds can also be used to illustrate how individual greenhouse-gas-emitting activities can trigger greater changes in the *total amount* of emissions. For instance, due to *market thresholds*, a seemingly insignificant consumption decision—such as to not purchase a flight ticket to Thailand—might trigger a significant change in a large corporation's behaviour—such as an airline deciding to stop flying to Thailand—which could have a substantial impact on total greenhouse gas emission (cf. Singer 1980 on meat consumption; see also Kagan 2011 and Nefsky 2011).

So, even if a decision about going for a Sunday drive will not affect whether the climate crisis *occurs*—and even if it is unlikely to have *any* effect on the climate (an

---

[7] Morgan-Knapp and Goodman (2015), Lawford-Smith (2016), and Broome (2019) discuss the relevance of this for climate ethics; for the original source, see Lorenz (1963, 1969)

[8] Is worth noting that this talk of 'thresholds' may be a little misleading, since it may seem to suggest that a particular weather event is guaranteed to occur if and only if the threshold is passed, whereas in fact the relationship between greenhouse gas emission and the weather events in question is more stochastic.

assumption that some question; see, e.g., Broome 2019)—it could potentially cause extreme weather events that kill many people. In other words, even if the probability that a Sunday drive causes some climate related harm is small, that probability is not zero. Moreover, the stakes are, by the above argument, very high: the individual action could cause an extremely harmful event. Thus, while going for a Sunday drive is, by assumption, pleasant, it also carries with it *expected* harm. (The "expected harm" that an activity carries is the sum we get when we, first, multiply each magnitude of harm that the activity *could* result in by the probability that it results in that magnitude of harm, and, second, add up these probability weighted magnitudes of harm.)

Kingston and Sinnott-Armstrong (2018) have argued, in response to the claim that going for a Sunday drive carries with it expected harm, that such a drive could at most change the *timing* of when some meteorological threshold is passed or when some harmful weather event occurs. The reason is that humanity is constantly causing greenhouse gas to be emitted, and thus even if you refrain from going for a Sunday drive, the greenhouse gas concentration that your drive *would* have resulted in will instead be reached "only *a fraction of a second later*" (177, emphasis in original). Moreover, they contend: "Even if we grant, for the sake of argument, that a change in timing of a given concentration threshold did correspond to a change in the timing of the relevant effects, it does not matter whether those impacts happen now or a second later." (ibid)

The above argument is mistaken. Emitted greenhouse gas stays in the atmosphere between 300 and 1,000 years. Consequently—and whether or not a peak concentration will be reached (cf. Nefsky 2021)—the small difference in timing does *over time* correspond to a potentially very large increase in expected harm. Perhaps the easiest way to see this is by considering Figure 1. Suppose that the two lines represent the predicted atmospheric concentration of greenhouse gas over one hundred years; the orange line is the concentration with the Sunday drive in question, the blue line the concentration without the Sunday drive. Now, at each point, the difference in concentration—as represented by the distance between the two lines—is small; correspondingly, at each time, the increased risk of harmful weather events (storms, floods, droughts, etc.) may be small. But when evaluating the increased risk of climate harms due to the Sunday drive, what matters is not the distance between the two lines at any point in time, but rather the *total area* between the two lines (over at least 300 years). We know that, at the current and close by concentrations of greenhouse gas in the atmosphere, an increased concentration is correlated with an increase in the number of harmful weather events. Therefore, the total area between the two lines—extended for some additional hundreds of years—corresponds to a potentially very large increase in the risk of harmful weather events. Therefore, even if we think that "it does not matter whether [*harmful events*] happen now or a second later", it does not follow that it does not matter whether *the increased greenhouse gas concentration* happens now or a second later.

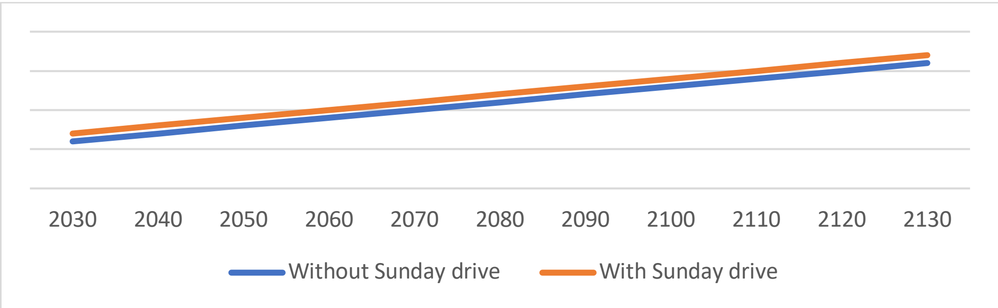

Figure 1: Difference in greenhouse gas concentration over time

Nefsky (2021) seems to overlook the importance (in particular, in the long run) of this probabilistic correlation between increased greenhouse gas concentration and harmful weather events when she says, in response to a similar point made by Broome (2019), that:

> The chance of making a difference via the butterfly effect [i.e., due to the chaotic nature of the climate] is no more of a chance of making a difference for the worse than it is a chance of making a difference for the better. So, simply delaying a threshold-crossing by an imperceptible fraction of a second is no more likely to make the harmful outcomes more severe than it is to make them less severe.

If what climate scientists say about the probabilistic relationship between atmospheric greenhouse gas concentration and harmful weather events is correct—which we see no reason to doubt—then Nefsky's response fails. The butterfly effect means that we cannot be *sure* that a Sunday drive won't prevent some harmful weather event; for all we know, it might prevent such an event, or it might cause such an event, or it might make no difference whatsoever. However, the aforementioned probabilistic relationship should make us *believe* that, at least over time, the increased concentration that a Sunday drive results in is more likely to cause such harmful events than to prevent them (Morgan-Knapp and Goodman 2015 and Lawford-Smith 2016 make a similar point).[9] In fact, some think that the Sunday drive is, over the very long run, almost certain to *both* cause some such events and prevent other such events (see, e.g., Broome 2019), but that it will on balance very likely cause more rather than fewer such events. In sum, the Sunday drive increases the expected number of harmful weather events.

Further criticism has been directed at this threshold argument based on expected harm, for instance, by Budolfson (2019) and Kingston and Armstrong (2018) (and to some extent by Nefsky 2011, 2021). Responses can be found in, for instance, Broome (2019) and Hedden (2020). We shall not however discuss this threshold argument further, but instead turn our attention to another argument, based not on thresholds but on *imperceptible* harm.

2.4 How you probably (also) make an *imperceptible* difference

Some might think that, even if private climate decisions—such as going for a Sunday drive—may cause some harm, the harm *to each person* is *imperceptible*. Moreover, it might seem that such imperceptible—or hardly perceptible—harms are morally insignificant. For instance, according to some estimates, the *lifetime* greenhouse gas emission of an "average American" reduces life expectancy by six months *in total*, that is, when the effects on all of

[9] Does this reasoning assume evidential rather than causal decision theory? We think not. Recall that we are assuming that the agent we are considering does not believe that her private climate choices causally effect the choices of others. But we assumed that her climate beliefs are based on the currently best evidence. Therefore, she believes that, holding fixed everything that is not causally related to her choice (such as the choices of others), her emission is more likely to cause than prevent harmful weather events.

humanity is aggregated (see, e.g., Broome 2012). According to other estimates, based on *the social cost of carbon*—which is an attempt at monetizing the harm done by greenhouse gas emission—the harm caused by the emission of a "a typical academic who travels" (Broome 2019: 111) is about $40,000, when all harmful effects are added up. In light of these figures, it may be tempting to draw the conclusion that the expected harm of your lifetime emission on any one individual is so small as to be imperceptible, or not measurable by our current instruments. It may be even more tempting to draw the conclusion that the expected harm of a single Sunday drive on any one individual is imperceptible.

The above argument clearly does not hold for *all* harms that your lifetime emissions or your Sunday drive may cause. It may however hold for the *ex ante* or *expected* harm to each person. For instance, when you go for a Sunday drive, the expected harm that the corresponding increase in greenhouse gas concentration imposes on each person may be sufficiently small to be deemed "imperceptible". And it is probably not measurable with any precision or certainty. However, it is possible, as we have seen, that the *actual* (or *ex post*) harm is very perceptible and measurable. If the climate is as chaotic as climate scientists think, then seemingly trivial acts such as going for a Sunday drive could trigger some extreme weather event that ends up killing someone. Thus, even if the expected harm to each person may be tiny, some of the possible actual (i.e, *ex post*) harm is far from tiny. So, what we have here is, on the one hand, a very small (possibly imperceptible or not measurable) increase in *expected* harm to each person, which is very (subjectively) *likely* given our current evidence, and, on the other hand, a large (and certainly perceptible) possible actual harm to some persons, which may be very (subjectively) *unlikely*.

But there are also examples of possible *ex post* harms that may be imperceptible. For instance, the increased greenhouse gas emission associated with your Sunday drive may warm the planet (or some region of it) by some amount that is not perceptible to anyone, which could result in the water supply in some region falling by some imperceptible amount, and so on. And this, in turn, might result in some imperceptible increase in pain or discomfort. So, in this case we have a *continuous ex post* harm: Whenever we emit, there may be some imperceptible increase in actual harm (see, e.g., Parfit 1984 and Broome 2019).

To finally come back to the expected utility answer to the question of how to evaluate private climate decisions: both of these tiny harms—that is, both the tiny increase in expected harm to each person and the imperceptible increase in *ex post* harm—should, we contend, figure in the expected utility calculation of a climate-change-conscious agent when she reasons about an enjoyable greenhouse-emitting activity. But some think that this is a mistake—there are no imperceptible harms (see, e.g., Kagan 2011). Discussing this matter in detail would get us too far off track, so let us simply assume, for now, that there are imperceptible harms. One might still wonder to what extent imperceptible harms—and, more generally, tiny harms—could be morally significant (see, e.g., Sandberg 2011).

An argument from Parfit (1984) is often used to argue that even imperceptible harms are morally significant. In a simplified version, Parfit's argument was this: Suppose that we can give a person a number of electric shocks, each of which is imperceptible to the person, but one thousand of which is excruciatingly painful. If one thinks that imperceptible harms are morally insignificant, then it is not worse to give the person one shock than to give the person zero shocks, and not worse to give the person two shocks than one shock, and so on, and not worse to give the person one thousand shocks than nine hundred ninety-nine shocks. Assuming that the "not worse than" relation is transitive—which means that if A is not worse than B which is not worse than C then A is not worse than C—it follows that it is not worse to give the person one thousand shocks than to give them no shock. But then we have clearly reached a false conclusion from a valid argument, and so we should reject the argument's weakest premises. Parfit suggests we reject the premise that an imperceptible harm is not morally significant.

Others however reject the premise that the "not worse than" relation is transitive (see, for instance, Temkin 1987, 1996, and Rachels 1998; see also Spiekermann 2014). And actually, one might reasonably think that this relation is transitive without claiming that the *better than* relation is not transitive. For note that one way in which B could be no worse than A is for A and B to be *incomparable* or *incommensurable*, that is, realising different values that cannot be put on a single scale. But incomparability (and incommensurability) is arguably not transitive. For instance, a career as a lawyer with a

$150,000 yearly salary might be incommensurable with—or, as some (e.g., Chang 2002) would say, *on a par with*—a career as a banker with a $200,000 yearly salary; similarly, a career as a banker with a $200,000 yearly salary may be incommensurable with a career as a lawyer with a $140,000 yearly salary; but clearly, a career as a lawyer with a $150,000 yearly salary is *not* incommensurable with a career as a lawyer with a $140,000 yearly salary. So, the assumption that the "not worse than" relation is transitive can plausibly be questioned, which means that Parfit's argument for the moral significance of imperceptible harms contained a questionable premise.

But, the reader might wonder, is the above really relevant to climate decisions? Perhaps not. Let's consider the small (perhaps imperceptible) harms that a private greenhouse-gas-emitting activity, such as going for a Sunday drive, imposes on each individual. If your lifetime greenhouse gas emission reduces total life expectancy by at least six months—i.e., if Broome (2012, 2019) is right—then your (or at least some of your) individual acts of greenhouse gas emission reduce life expectancy by some smaller amount. And, when divided amongst billions of people, this smaller amount becomes *very* small indeed. Perhaps we want to call it "imperceptible"; at the very least, it will probably not be measurable (by our current instruments). However, when applied to this harm—imperceptible or not—the "not worse than" relation has the structure of the "not shorter than" relation between lives (Broome 2019: 125). And, of course, the "not shorter than" relation is transitive—and, some (sufficiently large) shortenings of life expectancy are unquestionably harmful. So, Parfit's argument for why imperceptible harms are morally significant holds for at least this potentially imperceptible harm of private greenhouse-gas-emitting activities.

The same can be said of many other such harms. For instance, if your emission causes an imperceptible decrease in the water supply, then when the "not worse than" relation is applied to the harm that this causes, it has the structure of the "no less than" relation as applied to water supplies. That relation leaves no room for incommensurability and is of course transitive. The same moreover holds for the increased probability of death, say, that private greenhouse-gas-emitting activities carry. Whether an increased *probability* of death in itself counts as a *harm*, rather than just an increased probability of harm, is contentious, so let's simply assume that you (typically) *wrong* someone when you increase the probability of their death by (at least) some sufficient amount. Now, your Sunday drive may result in a tiny, perhaps imperceptible (as in, not measurable by our current instruments), increase in the probability that some person dies due to, say, an extreme weather event. Such increases in probability of death may be highly *imprecise*. But as long as each act of emission increases the probability (imprecise or not), Parfit's argument holds; the transitivity of the "no more probable than" relation would show that even an imperceptible increase in the risk of death can wrong the person.

But perhaps some incomparability remains. For instance, although it would, we assume, be better if everyone refrained from Sunday drives, it might be that the imperceptible increase in temperature caused by some particular Sunday drive makes life more comfortable for some people or animals but less comfortable and/or riskier for others (Hedden 2020: 549). In that case, the two options—going on a Sunday drive and refraining from doing so—might not be equally good even though neither is better than the other. How to make decisions when two options are incommensurable, or incomparable, or on a par, is a highly debated question. Different proposals have been suggested (see, e.g., Hare 2010, Bales et. al 2014, Shoenfield 2014), but it seems fair to say that consensus is lacking. Discussing these different proposals in detail would take us too far from the topic of this paper. But, as Hedden (2020) points out, some of them would recommend going for the Sunday drive, in the case just imagined, while others would recommend against it.

2.5 But can decision theory handle *extreme* uncertainty?

A final family of objections to the expected utility approach to private climate decisions is that the uncertainty involved in such decisions is simply *too extreme* for the approach to be of any use. There are (at least) two somewhat different objections within this family. One worry is that some possible but very unlikely outcomes are too extreme for the expected utility approach to give any reasonable practical guidance; the other worry is that we lack sufficient knowledge of the relevant probabilities and utilities to be able to calculate meaningful expected utilities.

The first of the above worries could be attributed to Martin Weitzman, who in a number of papers proposed his (in)famous *dismal theorem* (first stated in Weitzman 2009), which he interpreted as undermining the use of standard cost benefit analysis (CBA) when evaluating climate policies. For the present purposes, the importance of Weitzman's theorem is the claim that since climate catastrophes could be *infinitely bad*, that is, since there is no bound on the possible bad of a climate catastrophe, such catastrophes either cannot be taken into account in traditional CBA or else they completely dominate the analysis. After all, the result of multiplying an infinite negative utility by any arbitrarily small probability[10] is still infinitely negative; hence, the possibility of such a catastrophe, no matter how unlikely, will dominate the analysis unless completely ignored. The same of course holds, more generally, for *any* expected utility analysis. An option that could result in an outcome that has infinitively negative utility has an infinitely negative expected utility.

Now, there is, of course, some chance of a climate catastrophe no matter what we do; in particular, whether you do or don't go for that Sunday drive, there is some chance of climate catastrophe. However, expected utility theory, properly applied, asks, in this case, (amongst other things) about the *difference* in the chance of a climate catastrophe depending on which of these two options is chosen. But then the infinitely negative utility of a climate catastrophe might seem to suggest that those who apply expected utility theory to reason about enjoyable but greenhouse-gas-emitting activities will *always* come to the conclusion that they should not engage in such activities, since any increase in the atmospheric concentration of greenhouse gas arguably increases the probability of a climate catastrophe by *some* tiny amount.

But that seems counterintuitive. For instance, imagine a lonely pensioneer who has finally gotten a sought-after grandchild. Visiting the grandchild gives her immense happiness and not being able to see the child gives her great sorrow. The only way she has

[10] Note that probabilities are real-valued. Hence, an arbitrarily small probability cannot be infinitesimally small.

of visiting the grandchild is by driving her hybrid car, which will, say, get her back and forth on electricity except for the very last couple of miles. These couple of miles cause some greenhouse gas emission. If the person uses expected utility theory to reason about what to do, and moreover understands that any increase in the concentration of greenhouse gas increases the probability of a climate catastrophe, which she assigns an infinitely negative utility, then she will find that she should never visit the grandchild. But that would seem too extreme.

Some might respond that what went wrong in the above analysis was that we focused on *individual* decisions, instead of something like a person's *lifestyle*, or lifetime emission. If the person in question lives a climate friendly lifestyle in general, then she can justifiably go visit the grandchild, even if that particular action has negative expected utility (Nefsky 2021). Others might respond that while *traditional* expected utility gives the wrong answer in this case, we can instead apply generalisations of expected utility that make use of hyperreal probabilities (see, e.g., Herzberg 2011) or that handle infinites by some other means (see, e.g., Bartha & DesRoches 2017). For instance, since an infinitely negative utility multiplied by an *infinitesimal* probability need not be infinitely negative, some generalisations of expected utility theory—e.g., those involving hyperreal probabilities—might recommend driving to meet the grandchild (assuming that the increase in the risk of catastrophe is infinitesimal).

But actually, it is unclear whether either of these responses is needed to avoid the extreme conclusion in the above example. For it is unclear why we should assume that a climate catastrophe could be infinitely bad. As Broome (2013: S30) puts it: "Weitzman assumes we can put no bound on its badness. But that is clearly false. We are a finite species living on a finite planet. There has to be a finite limit on the badness of anything that can happen here."

Take, for instance, an extreme *existential* catastrophe; suppose that climate change makes life on earth impossible for both humans and animals. This would be enormously bad. Most obviously, it would mean that an enormous number of people and animals who would otherwise exist and lead good lives never come into existence and so don't get to lead good lives. The total (expected) welfare loss would thus be gigantic. But it would not be infinite. We can safely assume that no individual animal or person has the capacity to enjoy infinite welfare. Moreover, even without climate change, the lifespan of our planet is finite. Therefore, the total welfare loss from this existential catastrophe would be finite. Consider next an even worse climate catastrophe: imagine that all people and animals who would lead good lives without the catastrophe still come to exist but lead lives with negative lifetime welfare as a result of the catastrophe—meaning that they would have been better off never existing. Since we can safely assume that no individual animal or person has the capacity to suffer an infinite harm—and since the lifespan of our planet is finite—the welfare loss would be finite even from this catastrophe. In sum, since the bad (i.e., the negative utility) of a climate catastrophe is finite, the possibility of a climate catastrophe does not undermine the use of (some version of) expected utility theory for making climate decisions.

The extreme lack of knowledge involved when making climate decisions raises a somewhat different type of worry, which we only briefly consider. Recall (from section 2.2) that the values that we are meant to plug into the expected utility formula are precise numbers. So, to use it when reasoning about a Sunday drive, one needs to have precise estimates for the probabilities of all the outcomes that the drive could result in. In addition, one needs precise utilities for all these possible outcomes. But, some say, that is too demanding—surely, we don't have precise utilities and probabilities for all the outcomes in which our climate decisions might result (see, e.g., Cullity 2015). Even worse, there might be relevant outcomes that we are not even aware of—and we might moreover *suspect* this to be the case when making our climate decision—in which case it would seem even less plausible that we could assign all the relevant outcomes precise utilities and probabilities (for discussion of this latter issue, see, e.g., Bradley 2017, and Steele and Stefánsson 2021).

In response, some point out that a number of decision theoretic representation theorems (recall from section 2.2) show that if a person's preferences satisfy certain constraints, which are interpreted as requirements of rationality, then it turns out that her graded beliefs can be quantified with precise probabilities and her desires by precise utilities. One worry about this response is that even if these results show that *ideally rational* people's desires and beliefs can be quantified with precise numbers, the same is not true of most ordinary people, who of course fail to satisfy some of these purported rationality constraints on preference. Another worry is that the representation theorems don't establish that rational people have *access* to these precise utilities and probabilities, nor that they use them to evaluate alternatives; these theorems only establish that if a person's preferences satisfy the rationality constraints, then these preferences can be *represented as* maximising expected utility relative to some pair of utility and probability function. But that does not mean that the person in question has access to these utilities and probabilities.

So, it is questionable that any person—perfectly rational or not—has access to all the precise utilities and probabilities that would be required to calculate the expected utility of, say, going for a Sunday drive. But that does not mean that the expected utility approach to private climate decisions is hopeless. In fact, some think that some sort of expected utility theory can always be used. Broome (2012: 129), for instance, says: "The lack of firm probabilities is not a reason to give up expected [utility] theory. You might despair and adopt some other way of coping with uncertainty … That would be a mistake. Stick with expected [utility] theory, since it is very well founded, and do your best with probabilities and [utilities]."

How, precisely, to "do your best with" probabilities and utilities in such cases is, of course, a very difficult question (see, e.g., Bradley and Steele 2015 and Bradley 2017). But the general point is that even if a person does not have access to precise probabilities and utilities, she could use the framework of expected utility theory to try to come to *some* estimate of the relative choiceworthiness of her options. Simply framing the problem in terms of expected utility could be of some help, for instance by reminding her not to ignore very unlikely events. Moreover, in the easy case where one option is *sufficiently* better than another, then a lack of *precise* utilities and probabilities need not be a hinder to see that one option has a higher expected utility than another. But even when the evaluation is not so easy, one can use expected utility theory to guide and discipline one's reasoning. Expected utility theory, in some form or another, is the ideal for which any reasoner should strive, when making decisions under uncertainty. Even if one cannot precisely quantify one's uncertainty and values (or desires), one can, in practice, get far by considering the expected utilities of one's options given different plausible quantifications of one's uncertainty and values.

3 Summary

We have now illustrated how decision theory can be applied to private decisions about options that require greenhouse gas emission, be it going for a drive in a gas guzzling SUV, flying to Thailand for a holiday, or eating meat. Some want to frame private climate decisions in terms of interactive decision theory, or game theory, and moreover think that these decisions are instances of the prisoner's dilemma. Therefore, they believe that decision theory recommends any enjoyable greenhouse-gas-emitting activity.

In contrast, we have argued that the prisoner's dilemma framing ignores the fact that even if an individual private climate decision is neither sufficient nor necessary for climate change, such a decision can make the effects of climate change either better and worse. In particular, even if a single private climate decision is unlikely to make a large difference, it can carry a great expectation of harm or benefit. Therefore, the correct framing of such decisions is the one provided by expected utility theory.

Figuring out what expected utility theory recommends, for some particular enjoyable but greenhouse-gas-emitting activity, is no simple matter, as we have seen. But expected utility theory is the ideal for which any reasoner should strive, even when she is lacking when it comes to some of the parameters needed for the theory to provide an unambiguous answer. In particular, the expected utility framework can and should be used to guide and discipline our reasoning, for instance, when evaluating an enjoyable but greenhouse-gas-emitting activity.[11]

[11] We are grateful to Richard Bradley for helpful comments. Stefánsson's work is supported by Riksbankens Jubileumsfond (Pro Futura Scientia XIII).

REFERENCES

Ahmed, A. (2014). *Evidence, Decision and Causality*. Cambridge University Press.

Bales, A., Cohen, D. & Handfield, T. (2014). Decision theory for agents with incomplete preferences. *Australasian Journal of Philosophy*, 92(3), 453-70.

Bartha, P. & DesRoches, C. T. (2017). The Relatively Infinite Value of the Environment. *Australasian Journal of Philosophy*, 95(2), 328-353.

Bolker, E. D. (1967). A simultaneous axiomatization of utility and subjective probability. *Philosophy of Science*, 34(4), 333-340.

Briggs, R. (2010). Decision-theoretic paradoxes as voting paradoxes. *Philosophical Review*, 119(1), 1-30.

Broome, J. (2012). *Climate Matters: Ethics in a Warming World.* New York: W. W. Norton & Company.

------ (2013). A Small Chance of Disaster. *European Review,* 21(S1), S27-S31.

------ (2019). Against Denialism. *The Monist*, 102(1), 110–129.

Bradley, R. (2017). *Decision Theory with a Human Face*. Cambridge: Cambridge University Press.

Bradley, R. & Steele, K. (2015). Making Climate Decisions. *Philosophy Compass*, 10(11), 799-810.

Buchak, L. (2013). *Risk and Rationality*. Oxford University Press.

Budolfson, M. (2019). The inefficacy objection to consequentialism and the problem with the expected consequences response. *Philosophical Studies*, 176(7), 1711–1724.

Chang, R. (2002). The possibility of parity. *Ethics*, 112(4), 659-688.

Cripps, E. (2013). *Climate Change and the Moral Agent: Individual Duties in an Interdependent World*. Oxford: Oxford University Press.

Cullity, G. (2015). Acts, Omissions, Emissions. In Climate Change and Justice, ed. J. Moss. Cambridge: Cambridge University Press, pp. 148–164.

------ (2019). Climate Harms. *The Monist*, 102(1), 22-41

Gardiner, S. (2001). The Real Tragedy of the Commons. *Philosophy and Public Affairs*, 30(4), 387 – 416.

------ (2011). *A Perfect Moral Storm: The Ethical Tragedy of Climate Change*. New York: Oxford University Press.

Hare, C. (2010). Take the sugar. *Analysis*, 70(2), 237-247.

Hardin, G. (1968). The Tragedy of the Commons. *Science* 162 (3859):1243-1248.

Hedden, B. (2020). Consequentialism and Collective Action. *Ethics,* 130(4), 530–554.

Herzberg, F. (2011). Hyperreal Expected Utilities and Pascal's Wager. *Logique Et Analyse*, 54(213), 69-108.

Hiller, A. (2011). Climate Change and Individual Responsibility. *The Monist*, 94(3), 349–368.

Jeffrey, R. C. (1965). *The Logic of Decision*. University of Chicago Press.

Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. Cambridge University Press.

Johnson, B. L. (2003). Ethical Obligations in a Tragedy of the Commons. *Environmental Values*, 12(3), 271-287.

Kagan, S. (2011). Do I Make a Difference? *Philosophy and Public Affairs*, 39(2), 105–141.

Kingston, E., and Sinnott-Armstrong, W. (2018) Kingston. What's Wrong with Joyguzzling? *Ethical Theory and Moral Practice*, 21(1), 169–186.

Lawford-Smith, H. (2016). Difference-Making and Individuals' Climate-Related Obligations. In Hayward, C. & Roser, D. (Eds.), *Climate Justice in a Non-Ideal World* (pp. 64–82). Oxford: Oxford University Press.

Lorenz, E. (1963). Deterministic Nonperiodic Flow*. Journal of the Atmospheric Sciences*, 20, 130–141.

------ (1969). The Predictability of a Flow Which Possesses Many Scales of Motion. Tellus 3, 290–307.

Morgan-Knapp, C. & Goodman, C. (2015). Consequentialism, Climate Harm and Individual Obligations. *Ethical Theory and Moral Practice*, 18(1), 177–190.

Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences,* 36(1), 48-49.

Nefsky, J. (2011). Consequentialism and the Problem of Collective Harm: A Reply to Kagan. *Philosophy and Public Affairs*, 39(4), 364-395.

------ (2021). Climate Change and Individual Obligations: A Dilemma for the Expected Utility Approach, and The Need for an Imperfect View. In Budolfson, M., Plunkett, D., & McPherson, T. (Eds.), *Philosophy and Climate Change* (pp. 201-2021). Oxford: Oxford University Press.

Nozick, R. (1969). Newcomb's problem and two principles of choice. In Rescher, N. (Ed.), *Essays in Honor of Carl G. Hempel* (pp. 114 – 146). D. Reidel.

Parfit, D. (1987). *Reasons and Persons*. Oxford: Oxford University Press.

Pindyck, R. S. (2020). What we Know and Don't Know about Climate Change, and Implications for Policy. National Bureau of Economic Research, Working Paper 27304, http://www.nber.org/papers/w27304.

Rachels, S. (1998). Counterexamples to the Transitivity of Better Than. *Australasian Journal of Philosophy*, 76(1), 71–83.

Ramsey, F. (1926). Truth and probability. In Antony Eagle (Ed.), *Philosophy of Probability: Contemporary Readings* (pp. 52-94). Routledge.

Sandberg, J. (2011). "My Emissions Make No Difference": Climate Change and the Argument from Inconsequentialism. *Environmental Ethics*, 33(3), 229–248.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley Publications in Statistics.

Schoenfield, M. (2014). Decision making in the face of parity. *Philosophical Perspectives*, 28(1), 263-277.

Sinnott-Armstrong, W. (2005). It's Not My Fault: Global Warming and Individual Moral Obligations. In Sinnott-Armstrong, W & Howarth, R. (Eds.), *Perspectives on Climate Change* (pp. 221–253). Amsterdam: Elsevier.

Singer, P. (1980). Utilitarianism and vegetarianism. *Philosophy and Public Affairs*, 9(4), 325–337.

Spiekermann, K. (2014). Small Impacts and Imperceptible Effects: Causing Harm with Others. *Midwest Studies in Philosophy*, 38(1), 75-90.

Steele, K. & Stefánsson, H. O. (2015). Decision Theory. In *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/decision-theory/.

------ (2021). *Beyond Uncertainty Reasoning with Unknown Possibilities*, Cambridge: Cambridge University Press.

Stefánsson, H. O. & Bradley, R. (2019). What Is Risk Aversion? *British Journal for the Philosophy of Science*, 70(1), 77-102.

Temkin, L. (1987). Intransitivity and the Mere Addition Paradox. *Philosophy and Public* Affairs, 16(2), 138–187.

------ (1996). A Continuum Argument for Intransitivity. *Philosophy and Public Affairs* 25(3), 175–210.

von Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Weitzman, M. L. (2009). On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics*, 91(1), 1–19.