Weighing the risks of climate change

Lara Buchak, UC Berkeley

ABSTRACT: This essay argues that when setting climate change policy, we should place more weight on worse possible consequences of a policy, while still placing some weight on better possible consequences. The argument proceeds by elucidating the range of attitudes people can take towards risk, how we must make choices for people when we don't know their risk-attitudes, and the situation we are in with respect to climate policy and the consequences for future people. The result is an alternative to the Precautionary Principle, an alternative that gives similar policy recommendations in many cases but is also sensitive to the costs of precautions.

1. Introduction

How should we take potential bad consequences into account when setting climate policy? One possible answer is to use the Precautionary Principle. In both its strong and weak form, it speaks to situations in which there is a possibility of serious environmental harm, but a lack of scientific certainty about the relevant cause and effect relationships. The weaker version of the principle says that in these situations, lack of certainty is not a reason not to take precautionary measures. The stronger version of the principle, which I will be concerned with here, says that lack of certainty *is* a reason to *take* precautionary measures: if we cannot rule out the possibility of serious harm, we must take serious measures to prevent it.[1]

[1] The precise form of this principle is stated differently by different authors, and there are a number of canonical statements of it in public policy. Gardiner (2006: 35) mentions the Wingspread Statement (1998), the UN

The strong Precautionary Principle has faced three important criticisms. The first criticism is that it does not have adequate justification. The second criticism is that it is imprecise, ill-defined, or vague. The final criticism is that the Precautionary Principle tells us to care only about the worst-case scenario, and only a particular kind of worst-case scenario. It does not take seriously the fact that all policies carry with them some risks of harm.

This essay argues for an alternative to the Precautionary Principle, an alternative that gives similar policy recommendations in many cases but is also sensitive to the costs of precautions. Furthermore, the alternative argued for can withstand all three criticisms. It is justified by a general theory about risk-taking, along with a simple ethical argument about making choices that affect future individuals. It is precise. And it is appropriately sensitive both to the fact that avoiding disaster is important, and to the costs of doing so.

I derive this alternative from a general, formal theory of decision making. Drawing on previous work, I explain how individuals may reasonably adopt a fairly wide range of attitudes towards risk, where an attitude towards risk is understood in terms of how much weight an individual places on worse scenarios versus better scenarios. Second, I argue for an important principle for decisions involving other people: if you don't know someone's risk-preferences, you should presume that he is inclined to place as much weight on worse scenarios as a

---

Framework Convention on Climate Change (1992), the Third North Sea Conference (1990), the Ozone Layer Protocol (1987), the UN Environment Program (1989), the EU's environmental policy (1994), and the US President's Council on Sustainable Development (1996). Other statements include Principle 15 of the Rio Declaration (1992) and the United Nations Framework Convention on Climate Change. See Bodansky (1991) for older examples of policies that presuppose the principle. For interpretive discussions, see Sandin (1999), Sunstein (2005: 15-20), Gardiner (2006), Steele (2006), Gardiner (2010: 13-14), Broome (2012: 118-20), Moellendorf (2014: 62-89), and Steel (2015).

reasonable person could possibly place on them—that he wouldn't take a risk unless all reasonable people would take it. Next, I explain the upshot of this principle for policy choices about climate change, choices that involve the well-being of future people who will deal with the effects of these policies. In particular, I argue that we should adopt very risk-avoidant policies: policies that place significant weight on worse scenarios. In short, we should be willing to curb human activity in a way that is very costly for humans in order to ameliorate the risk of climate disaster, even if this risk is fairly small. Finally, I show that the principles I have argued for support, in many cases, recommendations that are similar to those of the strong Precautionary Principle, but do not fall prey to the worries mentioned above.

2. There are many reasonable attitudes towards risk

Decision theory is the study of rational decision-making. The kind of rationality in question is *means-ends* or *instrumental* rationality. Thus, decision theory answers the question: if you want some consequence ('end'), what should you do (what 'means' should you take) in order to get it? More precisely, since decision theory concerns what to do when you are uncertain about the state of the world—when you are uncertain which actions lead to which consequences—decision theory answers the question: if you value various consequences to various degrees, what should you do in order to get something of higher value rather than lower value? For example, if you know that you would rather have a career as a successful musician than a career as a successful lawyer, but you prefer both careers to being penniless, and you know that whatever career path you embark on may or may not be successful, should you go to music school or law school?

The classical answer to this question is that you should maximize expected utility: where utility values are assigned to the possible consequences of an action and probability values to the possible states of the world if that action is taken, you should pick the act with the highest weighted average of utility, where each utility value is weighted by the probability of the state in which it is realized. (Utility is a measure of how much you value consequences, and probability is a measure of how likely you think various states of the world are to obtain.[2]) For example, where u(x) is the utility of consequence x, let's say that u(successful musician) = 10, u(successful lawyer) = 1, and u(failed musician) = u(failed lawyer) = 0. And where $p(A \boxright S)$ is the probability of state S if you were to do A,[3] let's say that p(law school $\boxright$ successful lawyer) = 0.99, p(law school $\boxright$ failed lawyer) = 0.01, p(music school $\boxright$ successful musician) = 0.1 and p(music school $\boxright$ failed musician) = 0.9. Then the expected utility of law school is (0.99)(1) + (0.01)(0) = 0.99, and the expected utility of music school is (0.1)(10) + (0.9)(0) = 1, so you should go to music school.

However, I have elsewhere argued that this cannot be the whole story.[4] Knowing how much you value various consequences and how likely they are to obtain if you take a given action is not enough to tell you which action you should take. That is because we don't yet know your *strategy* for getting what you care more about. In particular, we don't know how you trade off a high chance of ending up with something pretty good against a lower chance of ending up

---

[2] We can hold that utility is determined in advance of preferences or that it is derived from preferences, and similarly for probability; the difference between these views won't matter here.

[3] This is the formulation used in 'causal' decision theory. The differences between causal decision theory, evidential decision theory, and Savage's (1954/1972) formulation, which is used in n. 8, won't matter here.

[4] Buchak (2013) explains and argues for all of the claims in this section in more detail.

with something much better. You might think a small probability of success as a musician is worth forgoing the safer option, but you might instead want to play it safe and go for a nearly-guaranteed career as a successful lawyer. This is particularly salient when we note that you only have one life to live, a life in which you will either be a successful musician, a successful lawyer, or have a failed career.

Following the tradition of rank-dependent utility theory,[5] I've argued that the missing ingredient is your *attitude towards risk*. Specifically, I've proposed that instead of maximizing expected utility, rational individuals maximize *risk-weighted* expected utility: you needn't weight the utility of each consequence by its probability, but instead can weight the utility of each consequence by a factor that depends both on its probability and on how good or bad it is relative to the other consequences of that action. If you are *risk-avoidant*, then worse consequences get proportionally more weight than better ones. For example, instead of weighting the best state by its probability p, you might weight the best state by a smaller value such as p^2, assigning worse states the (larger) remainder of the weight. In this case, the risk-weighted expected utility (REU) of law school is (0.9801)(1) + (0.0199)(0) = 0.9801, and the REU of music school is (0.01)(10) + (0.99)(0) = 0.1, so law school comes out much better. For the risk-avoidant individual, the utility of being a musician would have to be much higher to make going to music school worth the risk: law school would come out better even if the utility of success as a musician were 90, i.e., 9 times as large as it is. To the extent that you are risk-avoidant—that you place more weight on worse scenarios and less on better scenarios—a bad

[5] Early theories in this tradition include anticipated utility (Quiggin 1982), dual theory (Yaari 1987), Choquet expected utility (Schmeidler 1989, Gilboa 1987), and cumulative prospect theory (Kahneman and Tversky 1979, Tversky and Kahneman 1992).

scenario being very bad will make a big difference in what you choose and a good scenario being very good will not. Furthermore, *the more risk-avoidant you are, the more it is worth it to accept a lower utility for better consequences in exchange for a lower probability of worse consequences.*[6] On the other hand, if you are *risk-inclined*, then better consequences get proportionally more weight than worse ones—in this case, music school might come out much better, and it might come out better even if the utility of being a musician were not much greater than that of being a lawyer.

We can represent the difference between EU-maximization and REU-maximization graphically. For clarity, we will use an example of an action with several possible consequences.[7] Consider the action of taking a job at a fledgling start-up. If the start-up is amazingly successful (probability 0.01) then you will be very wealthy, feel a strong sense of personal accomplishment, have a short and pleasant workday, and be able to travel the world and enjoy the finer things in life (utility 7). If it is very successful (probability 0.29) then you will be fairly wealthy, feel accomplished, and have an enjoyable workday (utility 6). If it is moderately successful (probably 0.5), then you will have enough money to pay the bills, but the hours will be long and boring (utility 4). If it fails (probably 0.2), then you will have to leave and get an unpleasant, demanding job where you are merely scraping by (utility 1). The expected utility of this action is represented by the shaded area in each of the graphs in Figure 1.

---

[6] This follows from the fact that the more weight one places on worse consequences, the more one is willing to lower the utility of better consequences in order to raise the utility of worse consequences—even if the utility by which the better consequences are lowered is larger than the utility by which the worse consequences are raised.

[7] Example and graphs taken from Buchak (2017a) and Buchak (2017b). Numbers in both examples in this section were chosen for ease of exposition.

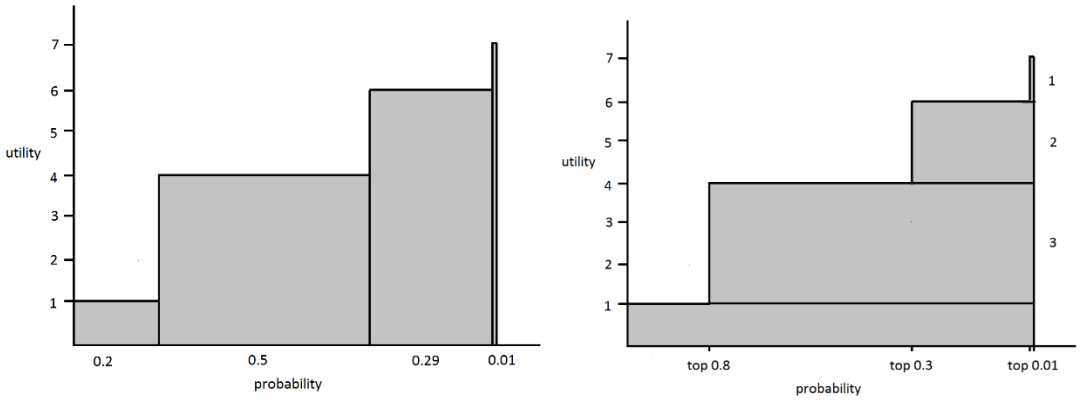

Figure 1: Expected Utility

These two graphs represent the same evaluation, in different ways. The left-hand graph shows that the expected utility of an action is the sum of its possible utility values, each multiplied by the probability of obtaining exactly this value. The right-hand graph shows that, equivalently, the expected utility of an action is the sum of its possible *utility benefits* (the amount which you get if you improve over the next-worst consequence, regardless of what else you get), each multiplied by the probability of obtaining at least this benefit.

Risk-weighted expected utility allows that each individual, instead of weighting each possible utility benefit by its probability, weights each utility benefit by a function of its probability. As benefits obtain in a smaller and smaller portion of scenarios, they matter proportionally less and less to you if you are risk-avoidant; and proportionally more and more to you if you are risk-inclined. The risk-weighted expected utility of taking a job at the start-up, for the risk-avoidant and risk-inclined individual, is represented by the shaded area in the graphs in Figure 2.

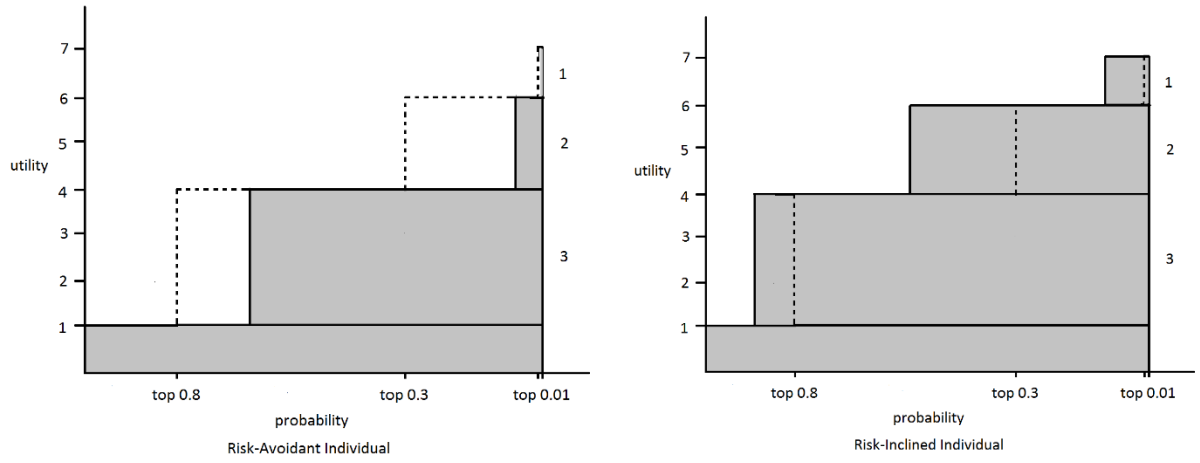

Figure 2: Risk-Weighted Expected Utility

The modification of the classical theory is therefore represented by weighting each utility benefit by a *risk-function* of its probability, rather than its bare probability, where the risk function *r(p)* represents how much weight an individual places on the top *p*-portion of consequences.[8]

[8] The general equation for expected utility maximization (in the form represented by the right-hand graph in Figure 1) and the general equation for risk-weighted expected utility maximization are as follows. Where g = {E_1, x_1; …; E_n, x_n} is an ordered gamble that yields consequence x_i in event E_i and $x_1 \leq \cdots \leq x_n$ (here "≤" represents the weak preference relation) we have:

$$EU(g) = \sum_{i=1}^{n}\left[\left(\sum_{j=i}^{n} p(E_j)\right)(u(x_i) - u(x_{i-1}))\right]$$

$$REU(g) = \sum_{i=1}^{n}\left[r\left(\sum_{j=i}^{n} p(E_j)\right)(u(x_i) - u(x_{i-1}))\right]$$

with *u* a utility function of consequences, *p* a probability function of events, and *r* a 'risk function' from [0, 1] to [0, 1], with r(0) = 0, r(1) = 1, and r non-decreasing. Replacing *p(E)* with an objective probability *p* yields anticipated

From this discussion, it should be clear that risk-avoidant individuals are not only concerned with *worst*-case scenarios, but (like expected utility maximizers) with *every* scenario. For the risk-avoidant person, better scenarios make less of a difference to the value of an action, but all scenarios matter to *some* degree. Better scenarios need to be a lot better to offset what happens in worse scenarios, but if they are valuable enough, they can offset the negative value in worse scenarios.

Thus, rational decision-making involves three factors: a utility function, which represents how much you value consequences; a probability function, which represents how likely you think various states are to obtain; and a risk function, which represents how much the top *p*-portion of outcomes matter in your decision making. Importantly, a wide range of risk-attitudes are reasonable.[9] There are two goals in means-ends reasoning: prudence—ensuring that worse scenarios aren't that bad or that likely—and venturesomeness—ensuring that better scenarios are very good. And it is up to each individual to decide how to trade off these two goals: there is no argument to show that it is better to make sure one's life doesn't go too poorly or instead that it is better to ensure a chance that one's life goes extremely well.

---

utility; replacing *r(p(E))* with a non-additive weighting function *w(E)* yields Choquet expected utility. An individual is risk-avoidant if *r* is convex; he is risk-inclined if *r* is concave; and he is risk-neutral (he maximizes expected utility) if *r* is linear. Graphs above use $r(p) = p^2$ for the risk-avoidant individual and $r(p) = p^{1/2}$ for the risk-inclined individual.

[9] Decision theorists sometimes distinguish between *rationality*—an evaluation of formal consistency—and *reasonableness*—an evaluation of substance. Although there are risk-attitudes that are consistent while being unreasonable (just as there are utility and probability assignments that are consistent but unreasonable), my point here is that there is a wide range of risk-attitudes that are both consistent and reasonable.

So, there will be a very wide range of reasonable ways to trade off these goals. In the above choice, it is not unreasonable to pursue music, nor is it unreasonable to pursue law. However, not *every* way of trading off these goals will be reasonable. For example, if you have a 0.98 probability of succeeding as a musician and the above values, it will be unreasonable to pursue law—doing so would amount to being overly cautious and not venturesome enough. Likewise, if you have only a 0.001 probability of succeeding as a musician, it will be unreasonable to pursue music—doing so would amount to being overly venturesome and not prudent enough.

I want to note two facts that follow from the view laid out in this section, facts that will be important in what follows. The first is that even if everyone agrees on the probability and utility of various scenarios, *there can be reasonable disagreement about how much weight to place on these scenarios*. Relatedly, while we might be able to argue about probabilities and utilities with the hopes of converging on what is rational to believe given the evidence and on what is in fact good, *there is little hope of converging on a unique risk-attitude that we each endorse*.

The second important fact concerns when differences in risk-attitudes make a difference to actual choices and when they do not. Differences in risk-attitudes between individuals will be magnified when worse scenarios are *much* worse than other scenarios. Furthermore, differences in risk-attitudes will be magnified when a particular choice is a one-shot gamble rather than a repeated gamble. To see this, note that we should not evaluate each of our choices in isolation—rather, we should evaluate entire choice paths—and note that repeating a gamble minimizes risk. If you and I value money the same, and the only choice each of us faces is whether to take a coin-flip in which we get ten dollars if the coin lands heads and nothing if it lands tails, or

whether instead to take a coin-flip in which we get four dollars however the coin lands, then we might make different choices, depending on how each of us weight worse scenarios. However, if we are each deciding whether to take a *series* of the one kind of coin-flip or a *series* of the other, then unless our risk attitudes diverge very wildly, we will make the same choice. This is because the first series is very nearly certain to yield very close to five dollars times the number of flips, and the second series is certain to yield four dollars times the number of flips. Thus, even if I place much more weight on, say, the worse 80% of consequences than you do, the consequences in the risky but repeated choice (almost certainly) fall within a very small range. To summarize: when a choice involves large differences in consequences or can be thought of as a one-shot gamble, then differences in risk-attitudes will make a large difference to what is chosen.

3. Climate change

We turn now to the particular case of climate change policy. We will assume that the primary charge of those thinking about policy in this domain—whether global policy or a country's individual policy—is to be responsive to the considered, mutually disinterested preferences of the people in its charge. I will assume that 'people in its charge' includes future as well as present people (at least people in the fairly near future, say, 200 years or so). Furthermore, I will assume that the interests of future people must be weighed as highly or nearly as highly as those of present people. This inclusion of future people, and the particular way in which they are included, is no doubt controversial and needs serious philosophical defense[10]—but I hold that the assumptions here map roughly onto how future people should be

[10] See discussion in Broome (2012: 142-43). Note that the assumptions I make here imply that the interests of future people must be taken into account directly, rather than through the preferences of current individuals.

taken into account, ethically speaking. Policies can be as fine-grained as we want, and can include both restrictions and adaptations to possible change.

Let us consider a very simple model, to give us an idea of how to apply the above analysis. We are choosing between policies, where each policy has two main effects: it raises or lowers the probability of drastic climate change, and it has certain non-negligible negative consequences if climate change does not occur. For example, one policy (a 'precautionary' policy) might be to substantially limit carbon emissions and another policy (an 'industrious' policy) might be to allow any level of carbon emissions, where the first policy will carry with it a lower probability of drastic climate change but the second policy will have a much higher utility if drastic climate change does not occur, since it will allow individuals to engage in a wider range of industrious efforts. We will assume that the policies have the same utility if drastic climate change does occur, i.e., that the negative effects of drastic climate change will negate any other effects of the policies. (All of these assumptions could be relaxed, and a more accurate model would consider many different levels of climate change—e.g. 2-degree warming, 3-degree warming, 6-degree warming—and hold that a precautionary policy has lower probabilities of all levels of climate change than the second, but an industrious policy has higher utility at all levels of climate change, perhaps much higher at the least drastic levels and negligibly higher at the most drastic levels. But our crude model will be enough to illustrate how to apply our analysis to climate change policy.)

We will also assume that there are two relevant groups of individuals. One is the group of future individuals, who bear the costs of drastic climate change if it occurs. The other is the group of current individuals, who are not affected very much by drastic climate change (since it would occur in the future), but who still bear the costs of precautionary policies—though these

costs are much lower than the costs to future people of drastic climate change. So, to illustrate, we might think that under an industrious policy, there is a 0.05 probability of a 6-degree increase in temperature,[11] a consequence which has utility -10,000 for future people and -1,000 for present people, and a 0.95 probability of a non-noticeable increase in temperature, the consequence of which, if everyone is free to act without regard to emissions, has a utility of 500 for everyone. And we might think that the policy of limiting emissions reduces the probability of a 6-degree increase to 0.01, but also reduces the utility under a non-noticeable increase to 0 for everyone. Thus, we face the following choice:

	Present People	Future People
Industrious Policy	{-1,000, 0.05; 500, 0.95}	{-10,000, 0.05; 500, 0.95}
Precautionary Policy	{-1,000, 0.01; 0, 0.99}	{-10,000, 0.01; 0, 0.99}

Note that the preferences of the future people concerning their options will depend heavily on their risk-attitudes, since these options have the two features mentioned above: the choice concerns a single, non-repeatable event; and the worse consequences are much worse than the better consequences. For example, some individuals might be willing to accept a slightly increased risk of climate change in exchange for the benefits of the industrious policy, whereas others might not. Standard expected utility maximizers will: it makes a bigger difference that the top 95% of states yield a consequence with utility 500 rather than 0, than that the near bottom 4% of states yield a consequence with utility 0 rather than -10,000. On the other hand, if one is

[11] This is the number suggested by the Fourth Intergovernmental Panel on Climate Change (2007).

fairly risk-avoidant, then the reverse holds.[12] And (I've claimed) all of these attitudes are reasonable. So the question of what to do in cases in which reasonable people could make different choices becomes crucial.

4. Taking Risks for Future People

We first face the question of what the future people prefer. An obvious difficulty is that we do not know their preferences. In light of this difficulty, what risk-preferences should we attribute to them for purposes of decision-making? I have elsewhere argued for a principle that I call the Risk Principle:[13]

> *Risk Principle:* When making a decision for an individual, choose under the assumption that he has the most risk-avoidant attitude within reason unless we know that he has a different risk-attitude, in which case, choose using his risk-attitude.[14]

This principle says to avoid risk, unless we know that the individual would prefer otherwise. (I will talk in terms of *knowledge* that he prefers otherwise, but we could instead substitute *consent*

[12] For example, a future person will prefer adopting the precautionary policy if his risk-preferences are represented by $r(p) = p^{2.5}$, which corresponds to giving 2.5% of the weight to the worst 1% of consequences, giving 12% of the weight to the worst 5% of consequences, and giving 82% of the weight to the worst 50% of consequences (i.e., to caring about the bottom half of consequences nearly five times as much as the top half).

[13] See Buchak (2017b).

[14] I leave open whether one must use his risk-attitude if it is unreasonable, but that case doesn't matter here.

or *hypothetical consent* or *sufficient evidence*—I don't have a view about which of these is the right way to fill out the principle.)

The basic idea is that when we make a decision for another person, we consider what no one could fault us for. Assuming we don't know an individual's risk-attitude, then if no reasonable person would reject an option on the grounds that it is too risky, we are justified in choosing that option. Conversely, if a reasonable person could reject it on these grounds, then we are not justified in choosing it. If we do know the individual's risk-preferences, then no one could fault us for using them to make a decision.

It follows from the Risk Principle that when we use it to take a risk on behalf of another individual, this risk could not reasonably be rejected by him as too risky, whether or not we know his risk- attitudes. More specifically, if we choose according to the Risk Principle, then whatever comes to pass, the individual could not reasonably complain that he would have given more weight to avoiding the possibility in question: either *no one* would reasonably give more weight to this, or he himself would give this exact weight.

What is the lower bound on reasonable risk-avoidance? More work would need to be done here, and I can do no more than speculate. But to give you a rough idea of how I am thinking about this by putting some (very speculative) numbers on it, I think that it is not unreasonable to care about the bottom half of consequences five times as much as the top half,[15] but that is close to the reasonable lower limit. This is way more risk-avoidant than most people are, but it still represents a healthy regard for venturesomeness. Importantly, the strategy known

[15] This coincides roughly with how much weight people have been found to give to the bottom half of individuals relative to the top half of individuals in a particular hypothetical social choice involving medical outcomes. See Nord (1993).

as maximin, which puts all of the weight on the worst outcomes, would not count as reasonable, since it gives no due to venturesomeness.

One might worry that if we make very risk-avoidant choices on behalf of another individual, he could fault us for not being risk-inclined *enough*. There seems to be an odd asymmetry, where an individual who ends up in the bad scenario can fault us for not making the bad scenario better, but an individual who ends up in the good scenario cannot fault us for not making the good scenario (even) better. In response, I do not know why there is an asymmetry here, when there isn't one obviously suggested by decision theory, but the Risk Principle seems to map onto our considered practices, both of decision-making for others and of reasonable complaint. For example, let's say I drive a carpool, and I discover that the seatbelts in the back aren't working. I would need to first get everyone's permission to drive them in this vehicle, but I wouldn't need to first get everyone's permission to be late to pick them up because I was busy fixing the broken seatbelts—even if it turns out that everyone would have preferred riding with the broken seatbelts. Risk-avoidance appears to be the *default*, and the choice *without fault*.

Another objection is that instead of attributing to an individual very risk-avoidant preferences—on the grounds that he could not reasonably complain about the consequences of a decision made on this basis—we should attribute to him a typical or average risk function, perhaps typical or average of all people or of his reference class (those with the same gender, nationality, age, or social class, for example). After all, these are ways of having evidence about someone else's risk-preferences, and so they are ways of answering the individual's hypothetical complaint ("what justified you in taking this risk on my behalf?" "because that's the risk I (reasonably) thought you'd want to take!").

I am not sure whether this way of attributing risk- attitudes is justified. But whether it is or not, the Risk Principle can account for it. To the extent that information about an individual's class amounts to knowledge about his preferences (or provides hypothetical consent), the Risk Principle says we are justified in choosing on its basis. And notice that if we do think that such information serves as a basis for choosing for another person, we will also think it provides good evidence about his preferences. If we think that it does not, then that will be because we think it does not provide much evidence at all, or does not provide evidence that can be used to form justified beliefs. Either way, the background presumption is that we cannot choose a more-than-minimally-risky gamble for another person unless we have some reason to think that he would take that gamble himself: in the absence of a strong reason to do something else, we default to risk-avoidance.

Compare this to a case in which you do not know someone else's preferences about (non-risky) consequences, as when you are choosing ice cream for a friend and you don't know whether he likes chocolate or vanilla. In this case, you are justified in choosing arbitrarily, or on the slimmest of bases: you know a slight majority of men prefer chocolate, or you vaguely remember the friend selecting vanilla once. That you are justified in selecting on these bases is precisely because, in contrast to risk-preferences, there really is no default for preferences about consequences, and therefore no justification needed to move away from it: you have no prima facie reason to choose one or the other, so a weak reason of any kind will do the trick. (It may be objected that there is prima facie reason to choose one or the other, on the grounds that (say) chocolate is more offensive to people that dislike it than vanilla is, so one should choose vanilla—but then this is precisely because chocolate is the risky choice, and one ought to be risk-avoidant when choosing for another.)

Again, the default is risk-avoidance: take only the risks that no one could reasonably reject. To the extent that you can be justified in knowing someone else's risk-attitude, you can move away from the default in making choices for someone else. This means that you are more justified in making risky choices for your acquaintances than for a stranger; and it may mean that you are justified in making risky choices for someone to the extent that you have reliable information about risk- attitudes within his reference class. But if someone is distant from you or you do not have reliable information about him, then you must default to risk-avoidance: you must weight worse scenarios as highly as one can reasonably weight them.

To summarize, the default risk attitude is the one that gives maximal weight to worse possibilities, within the bounds of reason: for each *p*, weight the bottom *p*-portion of outcomes as much as a person could reasonably weight them.[16] And we may move away from the default risk attitude, in making a decision that affects a single individual, just in case we know or can be justified in believing that he would choose to move away from the default.

To return to the case of group decision-making, there is a short step from the Risk Principle to attributing a maximally risk-avoidant attitude to those group members whose risk-attitude are unknown, so that we can make policy decisions on the basis of these attributed preferences plus the known preferences of the remaining group members. (In our case, this will amount to assuming that the future people prefer the precautionary policy.) The remaining objection to consider is this: it is a feature of group decision-making that a smaller group is sometimes authorized to decide on behalf of a larger group (as when elected officials make

[16] Formally, if R = {r(p) | r is a reasonable risk function} is the set of reasonable risk-attitudes, then the most risk-avoidant of the reasonable risk-attitudes is $r'(p) = min_{r \in R} r(p)$. (We will tentatively assume that *r'* itself has the properties of a risk function, and is convex.)

decisions for their constituents or a sub-committee is formed to resolve an issue), or taken to be representative of the larger group (as when a randomly-selected group is polled for their views concerning some issue). In these cases, one does not need to directly know the preferences of the larger group. So why, in our case, do we not simply take the risk-attitudes of the sub-group whose risk-attitudes are known, to be the appropriate ones on which to base our decision? (If present people have risk-attitudes that imply they would prefer the industrious policy were they in the situation of the future people, then why don't we assume the future people would also prefer this?)

In response, note that when a smaller group is authorized to make decisions on behalf of the whole, the larger group (or an agreed-upon representative thereof) must agree to this. In the case of future people, we don't have this authorization. Alternatively, when a smaller group is taken to be representative of the larger group, then it must be representative in the right way: it must statistically mirror the relevant features of the larger group, or we must have reason to think it does. But, as I've argued above, since future people are distant from us, we don't have good reason to think that their risk-preferences are similar to ours. Nor can we reliably speculate about their risk-preferences on the basis of the risk-preferences of present individuals.

Therefore, when thinking about climate change policy, we should make the decision as if future individuals are maximally risk-avoidant. We should presume that they are willing to pay significant costs to reduce the risk of worse consequences, even to reduce this risk by a fairly small amount. So, in our example, we must attribute to future individuals a strong preference for the precautionary policy.

What does this mean about the policy decision, given that the group affected includes both present and future people? That depends on how we should aggregate the preferences of our entire group.

One view is that we cannot take any risks that anyone in our group would not consent to. This follows from one generalization of the Risk Principle to group decision-making: in order to move away from the default, everyone involved must prefer to do so. If this is the correct view, then the conclusion is that we must make maximally risk-avoidant policy choices.

I'm not going to claim that this is the correct view, however. There is something to be said for it, but there is also something to be said for the thought that it is too risk-avoidant, since it does not take the preferences of the less-risk-avoidant into account at all. Luckily, we do not need such a strong view to reach the conclusion that our policy choices must be very risk-avoidant. For our choice has two features. First, it is the future people who bear the costs of the risks—who have lower utility in worse scenarios. Second, there are a large number of future people—at least as many as there are present people. And when one subgroup is larger than another subgroup *and* is more affected by the decision than the other, the preferences of the former can't be weighted *less* heavily than those of the latter. Thus, the group choice can be no riskier than an even balance between the preferences of the present people and those of the future people.

To remind the reader: if we assume risk-neutrality, then both present and future people will prefer the industrious policy; but as we increase risk-avoidance, future people will prefer the precautionary policy, and by a larger and larger margin the more we increase risk-avoidance. If we assume maximal risk-avoidance for future people, then they will very strongly prefer the precautionary policy. And so both the numbers of and stakes for future people mean that the

extent to which they prefer the precautionary policy outweighs the extent to which present people prefer the industrious policy, if indeed present people do prefer it.

Thus, even though it would be reasonable to eschew the policy that limits emissions—in the sense that some individuals might reasonably eschew it—we are required to adopt it, since a reasonable but very risk-avoidant future person would adopt it. We have to be particularly cautious with respect to risks faced by future people, because we don't know whether these people would be willing to accept them.

The general lesson that can be drawn from this section is this:

> *Future Risk-Avoidance Principle:* If we are making a decision whose largest effects concern a large group of future individuals, then we should make a very risk-avoidant choice: a choice which weights the worse consequences proportionally much more heavily than the better consequences.

5. The Precautionary Principle and Future Risk Avoidance

Recall the (strong) Precautionary Principle, from above:

> *Precautionary Principle:* If we cannot rule out the possibility of serious environmental harm, then we should take precautions to prevent this harm.

Although the antecedent of the principle is usually stated in terms of 'uncertain' cause and effect relationships, I take this formulation to capture the upshot of this uncertainty presumed important: that we cannot be sufficiently certain that environmental harm won't occur.

The argument in the present essay suggests that the Precautionary Principle can be vindicated, to a limited extent, since it falls under the scope of the Future Risk-Avoidance Principle. It follows from the Future Risk-Avoidance Principle that as long as we hold that the probability of some serious negative consequence is non-negligible when precautions are not taken and that taking precautions makes some difference to this probability (easily outweighing loss of best case utility), we are required to take precautions. Furthermore, on the analysis here, we are required to take precautions precisely because we are in the situation described by the Precautionary Principle: the potential negative consequence is a serious harm and its probability is above some small threshold. In addition, the analysis here gives a specific reason *why* the situation described by the Precautionary Principle has this upshot. The negative consequence would affect individuals who we cannot assume would consent to the risk even given large costs to avoiding it—and therefore must be assumed to prefer to avoid it (even if we ourselves might take it).

Thus, the Future Risk-Avoidance Principle can recover some of the judgments that many apparently find intuitive. In addition, I claim, the Risk Principle avoids some of the criticisms that have plagued the Precautionary Principle.

I will discuss three major sets of criticisms of the Precautionary Principle. First, the Precautionary Principle may be claimed to be unjustified, in a number of ways. It may be criticized as *lacking an adequate philosophical justification*, for example that there is no more general principle that it follows from. It also may be criticized as *possessing a justification that is misguided*. In this vein, for example, Cass Sunstein claims that the Precautionary Principle arises from irrational decision-making tendencies surrounding small probabilities, specifically the tendency to focus on the magnitude of harms and not pay enough attention to the probability

of these harms.[17] As evidence, he notes that people are not sensitive to risks of harms in a way that is proportional to the probabilities of these harms: they do not weight consequences linearly in probability. He holds that the Precautionary Principle is simply a codification of our irrational attitudes. Finally, the Precautionary Principle may be criticized as being *incompatible* with desiderata for decision-making that we do take to be justified. Martin Peterson presents a result to show that the Precautionary Principle is jointly incompatible with several widely accepted principles of decision-making.[18]

The Future Risk-Avoidance Principle avoids the criticism that it lacks positive justification. It is grounded in a general, well-justified theory of decision-making and of risk-taking for other individuals and for groups. It is grounded in what we owe to future people when we are uncertain about what will happen: while we are allowed to a certain extent to weigh their interests against ours, when we think about what their interests are, we must presume that their risk-attitudes lie on a particular edge of reasonableness. We must presume that they care a lot more about worse scenarios than better scenarios, and therefore that there need only be a small probability of a very bad consequence in order for them to prefer to expend a lot of resources to avoid that consequence. In addition, the introduction of risk-attitudes can explain why ordinary decision-making tendencies need not be irrational, and so principles based on them need not be misguided.[19] Risk-avoidant individuals—and therefore the policy recommendations arising from

---

[17] Sunstein (2005: 64-88).

[18] Peterson (2006, 2007).

[19] I don't claim that all of the documented tendencies of ordinary people are rational, just that there is reason to care a lot about worse scenarios, even if they are unlikely. For example, I agree that people tend not to be good at estimating probabilities, and that the tendency to make different choices in response to different descriptions of a

the principles in this paper—are sensitive to probabilities, but not in a linear fashion. Finally, if we make a very weak assumption, namely that reasonableness requires that moving probability from a bad consequence to a good consequence always make the resulting gamble strictly preferred to the original gamble, then the Future Risk-Avoidance Principle is compatible—and, indeed, entails—the principles Peterson mentions.[20]

The second set of criticisms of the Precautionary Principle concern its imprecision. It seems to make *unclear* recommendations.[21] The circumstances in which the principle is supposed to apply—when there is 'serious harm' whose possibility 'we cannot rule out', for example[22]—are *not adequately defined*. And even if they are adequately defined,[23] then the principle faces a common problem for principles based on threshold concepts, namely that small differences in input can lead to large differences in output. In this case, small differences in circumstances can lead to very different policies: if a harm is close to 'serious' but not quite over the threshold then the principle will not tell us to take precautions, but if it is just over the threshold then the principle will tell us to take serious precautions.

---

problem is irrational. The probabilities that we use when we employ the Risk Principle must themselves be reasonable, and we must use a privileged or correct description of the outcome of a policy. See Buchak (2013: 74-81) for a discussion of the relationship between REU-maximization and documented behavioral deviations from EU-maximization.

[20] The technical form of the assumption is that r is increasing, rather than merely non-decreasing. The satisfaction of the conditions in Peterson (2006, 2007) follows from REU-maximization with a strictly increasing risk function; see Buchak (2013: chapter 3).

[21] See Bodansky (1991: 5), Posner (2004: 140), Sunstein (2005: 26), and Gardiner (2006: 36).

[22] Or when we are 'uncertain about the relevant cause and effect relationships'.

[23] See Sandin (1999), Resnik (2003), and Steel (2015: 27-8) for examples of precisifications of the principle.

The Future Risk-Avoidance Principle is precise, and makes clear recommendations.[24] Again, it is simply an application of ordinary decision-making principles. There is no discontinuity between how our analysis deals with climate change and how it deals with less serious cases, because all cases of decision-making are subsumed under the same general rules.

The final criticism is that the Precautionary Principle is *myopic*.[25] It tells us to care only about serious harm, and perhaps only a certain kind of harm—environmental harm, rather than other types of harm. It does not take seriously the fact that all policies, including precautionary policies, carry with them some risks of harm. Furthermore, these policies have costs in non-disastrous scenarios—sometimes large, known costs—and because the Precautionary Principle tells us to think only about how to prevent serious harms, the Precautionary Principle is insensitive to the size of these costs.

The Future Risk-Avoidance Principle avoids this criticism, because (1) every relevant harm and benefit goes into determining the utility value of a consequence, and (2) risk-avoidance does not council us to ignore scenarios in which there is not serious harm, but simply to weight them proportionally less. Although scenarios without serious harm, or with harm that is less serious, are weighted less by virtue of the fact that they are relatively better, if costs in these scenarios are high enough then we should not adopt precautionary policies. So, the analysis here allows us both to acknowledge that the costs of precautionary measures can possibly outweigh

[24] At least, if we can say roughly what constitutes the most risk-avoidant attitude within reason, which, as mentioned, is a question for further study. Note that while 'largest risks' and 'large group' are imprecise notions, this is because the Future Risk-Avoidant Principle is a gloss on the conclusion of the more precise principles from the previous sections, not because it represents the limits of precision.

[25] Sunstein (2005: 26-34), Gardiner (2006: 53-54).

their benefits *and* that we ought to take them across a very wide range of costs even the probability of serious harm is low without them.

It will be helpful to contrast the Future Risk-Avoidance Principle with two other proposed rules for climate change policy.[26] The first is Maximin, which says to choose the action whose worst-case scenario has the highest utility.[27] Unlike the Future Risk-Avoidance Principle, Maximin is *only* senstitive to what goes on in the worst-case scenario. Those who advocate using Maximin advocate using it only in certain circumstances, but problems arise for applying it to the case of climate change no matter what these circumstances are taken to be. If these circumstances include that the utility differences between non-disastrous outcomes are small or unimportant, then Maximin will simply be silent when there are serious harms and benefits in these other scenarios. But since there are serious differences between the non-disastrous consequences of most policy decisions, including policy decisions about climate change, it will not be applicable in most policy decisions.[28] If we instead hold that we should use Maximin even when there are serious differences between consequences in non-worst scenarios,

[26] Since all of principles in this section speak of the effects of climate change full stop, rather than the effects on some particular person, we must assume that there is a single evaluation of or utility assignment to each consequence, perhaps arrived at through aggregating the well-being of each individual.

[27] See, e.g., Hansson (1997), Gardiner (2006, 2010), McKinnon (2009), and Moellendorf (2014); much of the discussion in these papers cites Rawls (1974, 1999). Maximin can be thought of either as an *interpretation* of the Precautionary Principle, or as an *alternative* to it. Maximin is only a plausible rule under uncertainty, rather than risk, but the rule is still important to contrast with the Future Risk-Avoidance Principle, particularly since one needs to introduce some threshold probability for states to include in the decision problem in order to apply it.

[28] See also Sunstein (2005: 112).

then the rule would tell us to simply ignore these differences, so it would fall prey to the myopia criticism: it would tell us to ignore something that we should not ignore.[29]

The second alternative to the Precautionary Principle is Expected Utility Maximization.[30] Again, expected utility maximization in the individual case says: to figure out the expected utility of an action, weight the utility of each consequence by the probability of that consequence if that action is taken; then pick the action with the highest expected utility. One then adds: not only should this principle be used for individual decision-making, it should be used for group decision-making of the relevant sort. This principle is obviously sensitive to all of the costs and benefits of an action. However, compared to the Future Risk-Avoidance Principle, it is too sensitive to what goes on in better scenarios, especially the very best scenarios. (That its recommendations are *very different* from those of the Precautionary Principle tell us that it does not track intuitions very closely, although of course its proponents could argue that that's no mark against it.) If I am right about the correct theory of rational decision-making, then expected

[29] See Steele (2006: 27); see also my criticisms of Rawls's difference principle in Buchak (2017b). Another principle for decision-making (also plausible only under conditions of uncertainty) is Minimax Regret. Minimax Regret says to choose the action that has the smallest maximum regret, where the 'regret' of an action A in each state is defined as the value difference between the highest-valued action in that state and the value of A in that state. See Chisholm and Clarke (1993), Hansson (1997), and Steel (2015) for discussion of this principle in this context of climate ethics. A criticism similar to the above can be leveled against Minimax Regret: if the rule only applies when there is serious regret in only one scenario, then it will not be applicable to most decisions; and if it applies more generally, it tells us to completely ignore all possible regrets except the *most* serious regret, which we should not ignore.

[30] EU-maximization has many defenders. For an extended defense in the particular case of climate ethics, see Broome (2012).

utility maximization is actually under-motivated in the individual decision-making case: there is no reason to weight consequences linearly in probability rather than by some other weighting. And if this is right, then the idea that we need to use expected utility maximization in group decision-making has not yet been given a justification.

Thus, there are important differences between the Future Risk-Avoidance approach and other approaches in the literature on climate change policy. Unlike the Maximin and Precautionary approaches, my approach recognizes that there might be significant costs of precautionary policies and that these could in principle be high enough to counterbalance the risks of serious climate change under industrious policies. However, unlike the Expected Utility approach, my approach recognizes that these costs, because they are relatively less bad, need to be much, much higher in order to counterbalance the risks of industrious policies. The Future Risk-Avoidance Principle is thus a well-justified principle that represents a middle option between the existing approaches.

6. Conclusion

In this essay, I have argued that climate change policy must be very risk-avoidant: we must place a great deal of weight on the worst possible consequences. We thus must be willing to incur high costs in order to make these consequences less likely. I've argued that this holds because we must presume that individuals whose risk- attitudes we don't know would only take risks all reasonable people would take. And so we must presume this of future people, who make up a large portion of our group, and are significantly affected by the choice.

I've offered an example to show how this conclusion might bear on climate change policy. Our actual policy recommendations will depend heavily on the numbers involved, which

must be worked out by both scientists and ethicists, but the general lesson is that we ought to be willing to sacrifice a lot to ameliorate the risks of climate change—but not without limit. This conclusion contrasts with two major principles in the decision theory literature: Expected Utility Maximization, which holds that we ought to sacrifice only in proportion to the probability of the risks; and Maximin, which holds that we ought to sacrifice without limit. It also contrasts with the Precautionary Principle. While my approach argues for some of the same practical conclusions as the Precautionary Principle, it avoids some of the problems plaguing that principle, and is part of a more general theory of rational and moral decision-making.

Of course, there are a few questions that need to be answered in order to fully apply my analysis to climate change policy-making. The first is how a group of the relevant sort ought to proceed when probabilities are imprecise or when the relevant individuals do not agree about the probabilities. This does not mean we are totally in the dark about what to do in these cases. On the approach here, we must care about relatively small probabilities of bad scenarios, and if all models or individuals agree that there is at least some threshold probability of a bad scenario, all models or individuals may recommend precautionary policies. Furthermore, unless someone is prepared to argue that we should only use the most optimistic predictions, as long as the threshold probability is suggested by many of the models or individuals, we will have to adopt precautionary policies

The second question is how to take different risk-attitudes into account if the situation is not so neat: if, for example, a policy has both potentially really wonderful consequences as well as potentially really terrible consequences, where the former accrue to present people and the latter to future people, and the present people are known to be relatively risk-inclined. This will take us deeper into questions of distributive justice, and go beyond the simple scenario presented

here. Still, this essay is a starting point in presenting a view about how to attribute risk-attitudes to the group whose preferences are unknown. The third questions is what precisely characterizes the set of reasonable risk-preferences.

The argument presented here reveals that how we ought to respond to climate change depends in large part on how we ought to attribute risk-attitudes to future people, and on how we ought to make group decisions when different individuals face different risks and have different preferences. Even if one is not convinced by all of the steps in my argument, one can use its general form to frame the debate. For example, one could argue that we should attribute risk-attitudes to those with unknown preferences differently than the Risk Principle suggests. Or, one could accept that risk-weighted expected utility maximization is the correct theory of rational decision-making—that there are a variety of attitudes one could adopt for individual decision-making—but argue that there is a particular risk-attitude we should adopt for group decision-making, independent of the attitudes of the individuals in the group; for example, that a group should always use the most risk-avoidant preferences within reason or that a group should always maximize expected utility. Regardless of whether one agrees with my conclusions, it should be clear that the literature on climate change has not sufficiently addressed the fact that different people have different risk-attitudes and therefore might reasonably prefer different policies, and it is my hope that this essay goes some way towards bringing this fact into the debate.

Bibliography

Bodansky, Daniel (1991), 'Law: Scientific Uncertainty and the Precautionary Principle', *Environment* 33(7): 4-44.

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

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

Buchak (2017a), 'Precis of *Risk and Rationality*', *Philosophical Studies* 174(9): 2363–2368.

Buchak (2017b), 'Taking Risks Behind the Veil of Ignorance', *Ethics* 127(3): 610-644.

Chisholm, Anthony H. and Harry R. Clarke (1993), 'Natural Resource Management and the Precautionary Principle', *Fair Principles for Sustainable Development: Essays on Environmental Policy and Developing Countries*, ed. Edward Dommen (Aldershot, England: Edward Elgar Publishing Ltd.).

Gardiner, Stephen M. (2006), 'A Core Precautionary Principle', *Journal of Political Philosophy* 14(1): 33-60.

Gardiner, Stephen M. (2010), 'Ethics and Global Climate Change', *Climate Ethics: Essential Readings*, eds. Stephen M. Gardiner, Simon Caney, Dale Jamieson, Henry Shue (New York, NY: Oxford University Press), pp. 1-35.

Gilboa, Itzhak (1987), 'Expected Utility with Purely Subjective Non-Additive Probabilities', *Journal of Mathematical Economics* 16: 65-88.

Hansson, Sven Ove (1997), 'The Limits of Precaution', *Foundations of Science* 2: 293-306.

Intergovernmental Panel on Climate Change, Working Group 1, *Fourth Assessment Report* (Cambridge: Cambridge University Press, 2007).

Kahneman, Daniel and Amos Tversky (1979), 'Prospect Theory: An Analysis of Decision under Risk', *Econometric*a 47: 263-291

McKinnon, Catriona (2009), 'Runaway Climate Change: A Justice-Based Case for Precautions', *Journal of Social Philosophy* 40(2): 187-203.

Moellendorf, Darrel (2014), *The Moral Challenge of Dangerous Climate Change: Values, Poverty, and Policy* (Cambridge, UK: Cambridge University Press).

Nord, Eric (1993), 'The trade-off between severity of illness and treatment effect in cost-value analysis of health care', *Health Policy* 24: 227-38.

Peterson, Martin (2006), 'The Precautionary Principle is Incoherent', *Risk Analysis* 26(3): 595-601.

Peterson, Martin (2007), 'Should the precautionary principle guide our actions or our beliefs?', *Journal of Medical Ethics* 33: 5-10.

Posner, Richard A. (2004), *Risk and Response* (New York, NY: Oxford University Press).

Quiggin, John (1982), 'A Theory of Anticipated Utility', *Journal of Economic Behavior and Organization* 3 (1982): 323-343

Rawls, John (1974), 'Some Reasons for the Maximin Criteria', *The American Economic Review* 64(2): 141-146

Rawls, John (1999), *A Theory of Justice* (Cambridge: Belknap). Original edition 1971

Resnik, David B. (2003), 'Is the precautionary principle unscientific?', *Studies in History and Philosophy of Biological and Biomedical Sciences* 34: 329-44.

Sandin, Per (1999), 'Dimensions of the Precautionary Principle', *Human and Ecological Risk Assessment* 5(5): 889-907.

Savage, Leonard (1972), *The Foundations of Statistics* (Dover). First published in 1954 (John Wiley and Sons, Inc.).

Schmeidler, David (1989), 'Subjective Probability and Expected Utility without Additivity', *Econometrica* 57(3): 571-587.

Steel, Daniel (2015), *Philosophy and the Precautionary Principle: Science, Evidence, and Environmental Policy* (Cambridge, UK: Cambridge University Press).

Steele, Katie (2006), 'The Precautionary Principle: A New Approach to Public Decision-Making?', *Law, Probability and Risk* 5: 19-31.

Sunstein, Cass (2005), *Laws of Fear: Beyond the Precautionary Principle* (Cambridge, UK: Cambridge University Press).

Tversky, Amos and Daniel Kahneman (1992), 'Advances in Prospect Theory: Cumulative Representation of Uncertainty', *Journal of Risk and Uncertainty* 5: 297-323.

Yaari, Menahem E. (1987), 'The Dual Theory of Choice under Risk', *Econometrica* 55(1): 95-115