Title	Principled utility discounting under risk
Author(s)	Mintz-Woo, Kian
Publication date	2019-05-02
Original citation	Mintz-Woo, K. (2019) 'Principled Utility Discounting under Risk', Moral Philosophy And Politics, 6 (1), pp. 89-112. doi: 10.1515/mopp-2018-0060
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://doi.org/10.1515/mopp-2018-0060 http://dx.doi.org/10.1515/mopp-2018-0060 Access to the full text of the published version may require a subscription.
Rights	**© 2019 Mintz-Woo, published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 Public License.** https://creativecommons.org/licenses/by/4.0/
Item downloaded from	http://hdl.handle.net/10468/10423

DE GRUYTER Moral Philosophy and Politics 2019; 6(1): 89–112

Kian Mintz-Woo*

Principled Utility Discounting Under Risk

https://doi.org/10.1515/mopp-2018-0060

Abstract: Utility discounting in intertemporal economic modelling has been viewed as problematic, both for descriptive and normative reasons. However, positive utility discount rates can be defended normatively; in particular, it is rational for future utility to be discounted to take into account model-independent outcomes when decision-making under risk. The resultant values will tend to be smaller than descriptive rates under most probability assignments. This also allows us to address some objections that intertemporal considerations will be overdemanding. A principle for utility discount rates is suggested which is rooted in probability discounting. Utility discounting is defended against objections from Parfit (1984) and Broome (2005, 2012). A sample utility discount rate is estimated.

Keywords: climate change, decision-making under risk, discounted utilitarianism, intergenerational justice, overdemandingness, utility discount rate

1 Introduction

When doing cost-benefit analyses for a policy or project, analysts try to consider all the costs and benefits of the project. They then weight them using temporal *discount factors* in order to arrive at present values for those costs and benefits. These discounted effects are summed across different time periods with the goal of determining the *net present value* of the project as a whole; if positive, the benefits outweigh the costs; if negative, the costs outweigh the benefits. There are different types of discount factors and different types of associated discount rates that can be generated by a sequence of discount factors. In this paper, I defend the use of *pure* or *utility discount factors* or, using the usual variable, a positive utility discount rate 'δ' for decision-making under risk.[1]

---

[1] I prefer the term *utility* discounting since it clearly contrasts with *consumption* discounting. Some other terms used in the literature, which are synonymous, are the *pure rate of time preference* and *temporal discounting*. However, these terms are misleading: there is no conceptual

***Corresponding author: Kian Mintz-Woo,** Princeton University, 304 Louis Marx Hall, Princeton, NJ, USA, E-mail: mintzwoo@princeton.edu,
http://orcid.org/0000-0002-9216-9561

There are two ways of thinking about this defence. To those who are comfortable with probability discounting, it is a principled method of assigning value to δ which does not rely upon revealed preference, the usual method. To those who are opposed to probability discounting, or who think that (undiscounted) explicit expected value calculations—ones which comprehensively assign values and probabilities to all outcomes—are the only way to determine net present values, it is a more substantive rebuttal, holding that discounting is a better way of accounting for a special class of outcomes.

The utility discount rate is an important aspect in modelling the value of policies on future generations. The value of this discount rate is a consequential topic for climate policy; the level of discounting helps determine the urgency of actions both to mitigate and to adapt to climate change (cf. Dietz et al. 2007; Godard 2009; Stern 2015; Wahba and Hope 2006; Weisbach and Sunstein 2007). The discussions about such discounting have impact on estimates of the social cost of carbon (e. g. Fleurbaey et al. 2018; Interagency Working Group on Social Cost of Carbon 2013) as well as costs of mitigation and adaptation (Nordhaus 2008; Stern 2007). This is because many dominant integrated assessment models rely on maximizing intertemporal utility, so the discount rate applied to utility has a significant effect on the viability of a project.

There are problems with both having a high positive social utility discount rate and having a zero utility discount rate. For instance, if we accept $\delta = 0$, then we may find that utility streams are non-comparable (Dasgupta 2012; Koopmans 1960, 1965; Ramsey 1928). Furthermore, with no pure discount rate, we take the utility of future people, no matter how distant, to be equally important to current decision-making. This is very problematic; there are almost certainly vastly more future people than present people, so small benefits to future people overwhelm the importance of present benefits. On the one hand, with no pure discounting, large certain losses to current people would be outweighed by trivial certain benefits to future people because of their sheer number. If we could invest large amounts of money in ways that would benefit the far future, we could be morally required to do so. Plausibly, this is overdemanding for current people.[2] On the

---

requirement that utility discounting applies only due to the passage of time; economists allow that time could be used as a proxy for other qualities (this is indeed the position this paper will take). Also, the rate is often written as a ‘preference’ because the initial usage of the term governed individuals who *preferred* experiences to occur at particular points in time; however, ‘preference’ is misleading when it comes to social discounting (as is the issue in this paper), since it is far less obvious that society is a type of entity that can prefer, but these additional terminological complexities need not concern us here.

[2] For instance, I differ from Tarsney (2017, pp. 344ff.), who takes it that overdemandingness cannot support discounting, since it would imply that our demands to the current global poor would

other hand, with a high utility discount rate, the benefits from long-term projects such as climate mitigation end up with negligible present values. This is both morally unacceptable as well as practically nonsensical: even long-term benefits should be able to have significant present values if we have high credence that those benefits will materialise. What would address this challenge are discount factors (and a discount rate) that is principled and low, yet non-zero.

Happily, a principled justification for a positive utility discount rate may, in our present circumstances, generate discount factors which also solve this overdemandingness problem. I argue that such a utility discount rate can be generated by appealing to the subjective probability of a particular set of outcomes, specifically outcomes in which the policy under consideration fails to make a difference. This is a way of both generalizing and theoretically grounding a common economic reason to discount utility, which is the exogenous probability of extinction (e. g. most famously Stern 2007).[3] This is thus what Parfit (1984, p. 482) calls a 'probabilistic discount rate'; it is not a justification of a positive utility discount rate due 'purely to the passage of time', but of a discount rate due to subjective probabilities (Mintz-Woo 2017), which often grow in a predictable manner that correlates with time (in this context, we are considering probabilities that approximately grow or shrink exponentially). The discount factors that I defend can be thought of as the subjective probability that the policies under consideration make their intended or expected difference.

Using this argument, the paper (a) generalizes accounting from a single outcome, extinction, to a broader class of outcomes that I call *model-independent outcomes*; (b) argues that addressing these model-independent outcomes in terms of utility discounting is superior to explicit expected value calculations; and

---

be overdemanding and also subject to spatial discounting (thanks to an anonymous reviewer for pushing this point). I agree with Tarsney that this is not a good basis for introducing spatial discounting; however, I believe an undiscounted far future is so many orders of magnitude more demanding than impartiality with respect to current people it would render all current people slaves to the future, morally required to do everything possible to increase the probability of human survival over any care or consideration for each other. This seems intuitively incredible or—more germane for my purposes—it is at least *plausible on its face* that this is overdemanding. A result of the same magnitude does not follow from impartially considering those currently alive. For instance, if the wealthy members of the world coordinated, the costs of addressing extant extreme poverty could be minor (Singer 2011, pp. 214f); this would not be the case with respect to equally considering centuries or millennia of potential future people.

[3] For instance, Ng (2016, p. 316) passes from the accurate point 'Since future utility is less certain to be realized, a time or rather an uncertainty discount rate to reflect this is fully justified' to 'there is a small probability of our becoming extinct' without noting that this is not the only way that future utility is less certain to be realized. Thanks to Hayden Wilkinson for pointing me towards this source.

(c) applies this more general utility discounting principle to model-independent outcomes beyond extinction. The purpose of arguing for a general principle is partially because otherwise we may not properly recognize or reflect on non-extinction-based model-independent events and partially because it is of theoretical value to determine general principles underlying particular true claims.

I begin by laying out the different types of discounting in Section 2. I contrast my view with several moral philosophers, and point out that my method of assigning a utility discount rate is broadly acceptable. Then, in Section 3, I present my utility discounting principle, which depends on the probabilities of model-independent outcomes, generalizing the usual appeal to extinction. In Section 4, I defend my principle, and the implied positive utility discount rate, from objections by Broome (2005, 2012); Cowen and Parfit (1992); Parfit (1984), including the argument that this is superior to explicit expected value calculations and that temporal relativity is absurd. In Section 5, I apply this principle by providing a model for how one would compute the utility discount rate using my principle. Finally, I offer some concluding thoughts in Section 6.

2 Groundwork and different types of discounting

In this section, I begin by clarifying the scope of my claim. The scope depends on a familiar, but important, distinction between decision-making under certainty, risk, and uncertainty. My principle applies to decision-making under risk and applies to utility discounting (as opposed to consumption discounting). In short, utility discounting is discounting potential future utility as a function of the time it occurs in order to account for potential utility in present value terms.[4]

Philosophers (Broome 1994; Caney 2008, 2009; Cowen and Parfit 1992; Nolt 2018; Parfit 1984; Sidgwick 1907; Tarsney 2017) and economists (Arrow et al. 1996; Ramsey 1928) have both objected to utility discounting, but it is sometimes unclear what decision-making assumptions their arguments apply under. A helpful distinction is between decisions made under conditions of certainty, risk, and uncertainty. Under certainty, the individual knows the outcome of any of her decisions.[5] Under risk, the individual knows the potential outcomes of any of her decisions, and can assign probabilities to any of those outcomes dependent

---

[4] In this context, the function increases exponentially in time, but in principle the discounting could take different forms.

[5] I use the terms ‘consequence’ and ‘effect’ synonymously, but use ‘outcome’ to indicate the entire set of consequences (or effects) that result from an action. I am using ‘knows’ as shorthand for has credence 1, although technically they may not necessarily always be coextensive.

on particular decisions.[6] Finally, under uncertainty, at least some outcome for at least some decision cannot be assigned a probability.[7] My defence of discounting applies under risk, and perhaps under uncertainty.[8] Since my claim is made primarily under conditions of risk, this means that my claims do not conflict with, for instance, Ramsey (1928), whose framework was explicitly meant for individuals operating under certainty.[9] Neither does it conflict with Sidgwick (1907, p. 414), who held that the uncertainty of one's effects on the future allows for deviations from intertemporal equality.

Ramsey's formula, which relates the social discount rate to the utility discount rate and a growth factor, is

$$r = \delta + \eta \cdot g,$$

where r is the overall *social consumption discount rate*, δ is the *utility discount rate*, η is the *elasticity of marginal utility of consumption*, and g is the *consumption growth rate*.[10] The second term $\eta \cdot g$ tells us that, if we multiply the rate of growth of consumption by the elasticity of marginal utility of consumption,

---

[6] I am thus defining certainty to be the limit case of risk, where all the probabilities are 1; if one prefers distinct categories, that can be done without affecting my arguments. The probabilities of interest are also sometimes called *ex ante* probabilities.

[7] This may be, inter alia, because there are no plausible probability distributions or because there are multiple probability distributions. My discussion applies in either case.

[8] Even if they only apply under risk, this is not necessarily a limitation; Roser (2017) argues that the types of probabilistic information available in the context of climate change are the types that are decision-relevant, meaning that we should act as if we are decision-making under risk, not uncertainty, in climate contexts.

[9] Among the arguments for discounting under conditions of certainty, many are unconvincing (Caney 2008; Parfit 1984). However, Koopmans (1960) and Dasgupta and Heal (1979) have potentially interesting arguments: Koopmans' axioms imply that, to order infinite consumption streams, utility discounting is necessary. Dasgupta and Heal argue that, in the presence of exhaustible resources, utility discounting is strictly preferable to zero time preference. In response, Broome (1992) objects to both of these derivations, taking Koopmans' framework as a reductio of the axiom of continuity and showing that Dasgupta and Heal's approach implies that the utility discount rate, while positive, should approach zero. I can be agnostic about discounting under certainty here, but I think these arguments are at least plausible and should be considered on their merits. Obviously, my principle could be used to generate small utility discount rates which could be applied in the decision-making under certainty of arguments such as those of Koopmans (1960) and Dasgupta and Heal (1979), allowing us to have a principled basis for introducing a discount rate and avoiding their impossibility results, although this would involve using an argument developed under risk and applying it in a different context.

[10] Note that Ramsey's formula is a *conclusion*, in that it is an optimization constraint dependent on many assumptions, such as separability, isoelasticity, identical populations, and many others. If we relax these strong assumptions, the resultant conclusion would be far more complex. One

we get discounting for the diminishing marginal value of consumption between generations.

With these initial clarifications in mind, my claims *do* conflict with the claims of some others, such as Caney (2008, 2009) and Cowen and Parfit (1992); Parfit (1984). My principle applies to the utility discount rate δ. Since many have argued that the utility discount rate should be zero, my principle conflicts with them. Some have also argued that the consumption discount rate r should be zero.[11] Obviously, a fortiori my principle will conflict with them since $r = 0$ practically implies $\delta = 0$ (not vice versa).

For instance, Caney holds that risk is no reason to discount; he holds that the morally appropriate response would include exactly the same level of action 'as would be appropriate if it were known that the malign effects would *definitely* occur' (2009, p. 176, italics added). He concludes that, due to considerations in the context of climate change, 'the upshot is that there should be zero discounting for risk. It makes no difference to the assessment if the probability is one hundred or eighty-five percent, say' (180). These claims conflict with my claim that social discounting is appropriate in the context of, and as a result of, the (risky) probabilities assigned to certain kinds of outcomes.[12]

Furthermore, Cowen and Parfit (1992, pp. 146f.) discusses a 'probabilistic discount rate' of my type but argue that it is wrong to discount for time 'rather than' probability (Broome 1994, takes a similar line). I will respond to this more specifically in Section 4, but my primary points are (a) discounting for time 'rather than' probability is a false choice—one can discount for probability and that generates a positive δ, contra Cowen and Parfit's claim that r should be 0 (and, a fortiori, $\delta = 0$); and (b) if we are using a discount rate instead of explicit expected utility calculations, we can make the comparison of policies simpler and more tractable—this is because we do not need to explicitly determine the value of

---

reason to defend working in this simplified framework is that it helps isolate the purely intergenerational aspects and can be extended separately to account for greater intragenerational realism.

[11] Sometimes, such as with Kelleher (2012); Parfit (1984), there is imprecision about whether philosophers argue that $\delta = 0$ or that $r = 0$, but I think it more charitable to read them all as arguing for the former.

[12] At some points, Caney (2008, 2009) makes it clear that he is objecting only to discounting human rights, but, since the authors he is objecting to are not discussing human rights but the evaluation of (the moral worth of) personal experiences and satisfied preferences, there is a potential worry that there is no substantive engagement. However, if we treat the relevant consequences as *being* rights-violations (or threats of rights-violations), then there is engagement. It is also worth noting that there are independent reasons to object to his claims about probability discounting as well (Mintz-Woo 2017).

certain kinds of outcomes which are selected to be equally likely regardless of the policies chosen.

However, even if one were to differ on the scope of my disagreement with these authors, another way to read the paper is as of interest to many potential advocates of net present value.[13] Suppose that one was interested in determining a net present value for a project in a morally respectable way. In particular, suppose one wanted to determine how to compare costs and benefits over time in a way which did not depend on our revealed preferences with respect to time. This might be because we think that our revealed preferences do not generate moral reasons to socially prefer when harms and benefits occur. Since the following principle appeals to our *beliefs* and *credences* regarding non-moral matters, it is not morally objectionable in the way that the dominant revealed time preference approach may be taken to be. Furthermore, although Caney and Parfit do not, many would endorse probability-based discounting, so this would be theoretically amenable. On these grounds, one might be interested in a principled method of assigning a value to δ which is less morally problematic than the dominant approach to net present value. This is what I offer.

3 A principle for utility discounting

In this section, I suggest a simple principle which allows one to determine appropriate values for δ. Assignments to this parameter are determined by the subjective probabilities for specific outcomes which are exogenous to the model under consideration.[14] These outcomes are not *unknown* (nor are their probabilities indeterminate—we are decision-making under risk and those probabilities will be needed to determine the utility discount rate); however, I will argue that they are best accounted for in modelling through a utility discount rate, given that explicitly modelling them requires assigning values to outcomes which are difficult to evaluate. In other words, when we are making decisions with reference to a particular model which will leave some outcomes out, those outcomes can be incorporated by discounting, and this discounting can be done in a principled manner. The level of discounting is determined by the assigned probabilities (which we can assume are given as we are deciding under risk).

---

[13] Thanks to an anonymous reviewer for suggesting this framing.

[14] As Kelleher (2017b) points out, it may be ‘highly misleading’ to take $\delta > 0$ to be due simply to the fact that some cost or benefit is in the future. Beyond being justified because there is uncertainty about the future, it may also be justified by considering the demandingness of $\delta = 0$ for current people. This is a way of systematizing his intuition.

When we are making policy decisions, just as when we are making personal decisions, we rely upon models which simplify the world. In personal life, when making decisions, we consider some potential consequences—usually those that appear most salient. We lack the time or intellectual capacity to consider the full range of outcomes. In policy settings, we also are limited. In economic climate contexts, this is exemplified by Integrated Assessment Models (IAMs), which combine a simplified physical model with economic equations representing economic and social inputs and outputs resulting from these physical parameters.

It is in the context of such simplified models that my principle is appropriate.[15] Instead of explicitly introducing these special outcomes, it is permissible to treat the model as deterministic (i. e. with decision-making under certainty), but then to add this discount rate in order to account for the special outcomes.[16] The purpose of discounting, on my principle, is to incorporate, in a simple manner, certain outcomes which are not in the model. The following conditions are meant to be necessary and jointly sufficient. The principle is:

Principle 3.1 (Discounting Under Risk). *For a given model which is used to make decisions under risk, it is permissible to discount for any outcome following an effect ('positive' or 'negative') by its subjective probability such that the effect*

1. *is exogenous to the model; and,*
2. *the outcome would have the same value subsequent to the effect regardless of the initial decision (in other words, if the effect occurred, it would make the decision-maker indifferent between the decisions (policies) under consideration with respect to the subsequent outcomes, i. e. those after the effect).*

Exogenous elements are those that do not arise within the given model, i. e. those which are determined by forces which are external to the model. One can think of exogenous effects as those whose probability is independent of any other part of the model. Endogenous elements are those which are driven by forces internal to the model. The first condition is needed because any effect which is endogenous to the model is accounted for; its probability and value is incorporated. The discounting is a way of introducing outcomes which would otherwise be excluded

---

[15] By a model, I mean a set of equations and parameters, some of which are given values exogenously and some endogenously. By definition, a model is simpler than reality and, in this context, it is important (a) that some outcomes have not already been explicitly modelled and (b) that not all potentially endogenous aspects are taken to be endogenous. Why I bring attention to these points should become clearer through the following discussion.

[16] Mathematically, treating the model as deterministic and applying a discount rate is equivalent to explicitly treating the model as probabilistic, as long as the probability of the relevant outcomes obey some particular assumptions. This has been known since at least Yaari (1965).

from the model. The second condition implies that the value of an outcome is independent of any other part of the model, subsequent to the effect occurring. Together, the probability and value of these events are independent of the rest of the model. This means that their expected value (the product of their probability and value) is independent of the model. For instance, suppose our effect is a global extinction event at some particular time and our outcome is all of the future *after* the extinction. It could satisfy the first condition if, for instance, the extinction's probability is independent of other factors in the model. It would satisfy the second condition if the outcome of extinction would have the same value regardless of which decision were adopted.[17] Since extinction would have the same value no matter which climate policies we adopt now, this would usually be satisfied as well. To account for this, we discount for the probability that the effect occurs at that particular time (the outcome *prior* to the effect should be undiscounted). Call outcomes satisfying both conditions *model-independent outcomes* (with all other outcomes *model-dependent outcomes*).

Therefore, my claims only apply in the context of such (by definition, incomplete) models. In the real world, the conditions I lay out in my principle are never met. Exogeneity, for instance, requires that the model be incomplete. In the real world, large-scale policies have all kinds of intended and unintended side-effects and interactions which are not captured in models. So my principled form of discounting requires that the models are not complete. It also requires that the kind of effects I am interested in are not already captured in the models, or else there could be double-counting. However, in extant IAMs, these conditions are met so this is not a problem.

A short proof that adopting this type of discounting would be implied by a explicit expected value calculation (summing over different outcomes each weighted by their (expected) probability) is in order.

Implication of Principle from Expected Value. Without loss of generality, we can adopt the expected value V for n effects with each effect $i \leq n$ being valued at u_i with subjective probability p_i. Technically, V, u_i and p_i are all indexed to some time $t \in T$, but the subscripts are suppressed for readability, and we begin by just considering one arbitrary timeslice. For convenience, we sort the model-independent effects first; for instance as the first j effects ($j = 0$ is a possible trivial limit case).

---

[17] It is conceptually important to note that I am *not* claiming the extinction value is '0', just that its value is *independent* of any current choices of (climate) policy—once extinction occurs, the value of the *rest* of the future is unaffected by climate policy.

$$\begin{aligned} V &= \sum_{i=1}^{n}(p_i \cdot u_i) \\ &= \sum_{i=1}^{j}(p_i \cdot u_i) + \sum_{i=j+1}^{n}(p_i \cdot u_i) \end{aligned}$$

Consider the values V_1 and V_2 which are the timeslice values of two arbitrary policies under consideration. Using the equations above, we can write them as follows:

$$\begin{aligned} V_1 &= \sum_{i=1}^{j}(p_i \cdot u_i) + \sum_{i=j+1}^{n}(p_i \cdot u_i) \\ V_2 &= \sum_{i=1}^{j}(p_i \cdot u_i) + \sum_{i=j+1}^{m}(p_i \cdot u_i) \end{aligned}$$

In the trivial case $j = 0$, the sum of probabilities $p_1 + \cdots + p_j = 0$, so the Principle permits discount factors of 1 (meaning $\delta = 0$). In other words, it is permissible not to discount. This is trivially the case.

In the case where $j > 0$, we justify, for each $i \leq j$, that p_i and u_i in V_1 are identical to p_i and u_i in V_2 using the conditions of Principle 3.1. By the first condition, model-independent effects are exogenous to the model; their probabilities are independent of the policy selected, i. e. for all $i \leq j$, p_i in V_1 is identical to p_i in V_2. By the second condition, model-independent effects have the same value regardless of the policy chosen, i. e. for all $i \leq j$, u_i in V_1 is identical to u_i in V_2. (Note that this does not depend on whether the effect is 'positive' or 'negative', i. e. whether $u_i \geq 0$ or $u_i \leq 0$.)

Thus, for each $i \leq j$, the associated $p_i \cdot u_i$ will be equal between V_1 and V_2. We are permitted to discount for these model-independent effects because they make no difference between the values of the policies under consideration. Intuitively, if we were to determine which policy was *better* between V_1 and V_2, we could ignore the values of $p_i \cdot u_i$ for $i \leq j$ because they are identical; we compare policies based on the states of world where they actually differ.

We have been considering the value of an arbitrary timeslice (for the value of policies V_1 and V_2). For a given $i \leq j$, the associated model-independent effect may have lower probability in a subsequent year, but we would expect the cumulative sum over time of the p_i to increase (recall that by definition these probabilities are independent of the choice between V_1 and V_2).

In a simple case where p_i simply grows exponentially with rate δ_i (i. e. $p_{i+1} = p_i(1 + \delta_i)$), we can calculate that rate from a given probability p_i and the timeslice t it comes from with the formula

$$\delta_i = \frac{1}{\sqrt[t]{p_i}} - 1.$$

This justifies, for each $i \leq j$, using a discount rate of δ_i associated with the p_i. It also justifies

$$\delta = \sum_{i=1}^{j} \delta_i$$

as an overall pure discount rate. □

There are two equivalent ways of intuitively understanding discounting according to this principle. First, we can think of removing these model-independent outcomes because they are independent of the model and intuitively ‘cancel each other out’ when comparing adopting the policy to not adopting it. The aggregated probability for the retained model-dependent outcomes are then less than certain, so we should discount them because they are less probable. Second, we can think of assuming that the model-dependent outcomes are certain, but that they might be ‘prevented’ by the model-independent outcomes. This prevention places a limit on how we value that model-dependent outcome, since we are not certain that the outcomes will depend on chosen policies in the way the model suggests. Then we should discount the model-dependent outcomes in line with the probability of the prevention (i. e. in accordance with the model-independent outcomes).

Since models already incorporate most relevant outcomes—in particular, they should include the model-dependent outcomes—these do not need to be introduced through changing the value of δ. Furthermore, very few such outcomes can appropriately be introduced by changing the value of δ. By definition, those outcomes which are endogenous to the model are accounted for in the model. For those model-dependent outcomes which are exogenous, changing δ would not be a way of accounting for their impact, since changing δ affects both the value of adopting or not adopting the policies symmetrically, whereas such exogenous model-dependent outcomes would, by definition, change the value of the policies asymmetrically. So, for instance, it might be exogenous to a model whether there is some type of climate-independent plague event, which usually kills or sickens its victims. This would be exogenous, but not model-independent, because the disvalue of the event depends on the previous choices made, for instance because these choices have implications for the size of the population. With a smaller population, this exogenous event may have less negative effects since there are fewer people to infect. That means that this exogenous event would still be model-dependent.

How does this principle depend on my previous arguments that it is rational to discount for subjective probability? If one accepted a decision-making view without probability discounting, then the (often small) probabilities that such exogenous outcomes would interfere with the expected outcomes would be ignored.

One might worry that this does not yield $\delta > 0$, i. e. a positive utility discount rate, because there is nothing about *time* in this principle. That is accurate; this is a case of discounting for something which correlates with time, not with time itself. The thing that correlates with time is the cumulative probability that one of these model-independent outcomes has occurred.[18] Of course, this probability must increase with respect to time. This is because the probability in question is cumulative, in short it is the probability of the event occurring *or* having occurred, so any subsequent probability is at least as great as the earlier time. However, it usually will be increasing at a decreasing rate. If it were increasing at a constant rate, or at an increasing rate, it will—given sufficient time steps—have probability greater than 1. Of course, these points do not imply that it take the form of a standard discount rate, but it would lead to a curve with a similar structure (i. e. positive first derivative and negative second derivative, asymptotically approaching 1. Furthermore, it is logically possible that there are no model-independent outcomes (equivalently, that we assign all model-independent outcomes zero probability), but it is difficult to imagine this being the case in any actual instance of long-term decision-making.[19]

One could question whether this would actually generate a meaningful discount rate.[20] Assigning defensible probabilities to model-independent events in the far future is extremely difficult and the probabilities of model-independent events in the near future might be thought to be negligibly small. This could yield twin worries: in the near future, this principle gives us (effectively) no pure discounting, while in the far future, the principle gives us no guidance (or perhaps also no discounting).

---

[18] I assume the relevant probability type is subjective, but the account could be adapted to account for objective probability assignments.

[19] It is also logically possible that some outcomes—ones I call model-dependent outcomes, since they satisfy the first condition but not the second—will *increase* the value of particular outcomes. For instance, it might be that some type of asteroid hitting the earth would double the sensitivity of mitigation outcomes. I do not think discounting (or 'reverse discounting') such possibilities is the best way to account for such model-dependent outcomes, but one would get equivalent analyses regardless of whether one accounted for these by reverse discounting or by making them endogenous or explicit within the model. I think this is an unimportant issue practically, however, since I am unaware of the mechanism by which such an exogenous event would work.

[20] I thank an anonymous reviewer for suggesting this objection.

Let us consider these in turn. In terms of the near future, Rees (2003) reminds us that past experience is no guide. At no time in the past has our species had the technological prowess to both destroy large numbers of people and to sculpt our habitats according to our whims without knowledge of the consequences. This may turn out for better or for worse. We may be able to create synthetic viruses which lead to widespread pandemics or bacteria that consume CO2 out of the atmosphere. This technological capability appears to be—if anything—accelerating. This has led some theorists, the transhumanists, to propose that the technological capability will reach a level of acceleration which is self-perpetuating. Whether or not this will occur, we are far from the endpoint of our technological capabilities. And they certainly could have major effects on our forms of life. Whether utopias or dystopias, it seems plausible to me that, even in the near term, there are non-negligible probabilities of model-independent outcomes (cf. 5 for my estimates of these probabilities). I welcome responses in terms of better estimates, but I do not think them as negligible as this questioner might.

In terms of the further future, it is worth trying to structure the situation. One way that this could be spelled out is by following the timeline mentioned in Dasgupta (2008, p. 164); he did not specify a particular period of risk, but he claimed that periods *beyond* 100–200 years into the future are *not* periods of risk. So we could think of the medium term (e. g. 50–100 years into the future) as reasonably being a period for social decision-making under risk, with uncertainty beyond that. (Of course, the argument does not depend on these particular ranges of dates.) If this is acceptable, it is also worth pointing out that many models—including the climate IAMs under consideration in this paper—run 'only' to 2100 so this worry may be more of a theoretical concern than a practical one.

In this case, where there is a period of risk, followed by a period of uncertainty, I tentatively want to suggest my argument may be applicable to the latter period of decision-making under uncertainty. In particular, if we follow the methodology suggested for calculating a utility discount rate in Section 5, that can generate a value of δ applicable in the period of risk. The curve generated might then be simply extended out (i. e. retaining the same value of δ) beyond into the subsequent period of uncertainty.

I am not certain whether this is justifiable, however, since the basis of the value in the period of risk is the particular probability assignment given in that period. After we no longer have a probability assignment, it is more difficult to know whether retaining the same value of δ is sensible. It would amount to assuming that the discount factors would continue to grow in a similar manner as during the period of risk even though we are ignorant of the probabilities. However, one could respond that—as any particular year is an arbitrary cut-off date between risk and uncertainty—we should expect risk to ‘gradually’ shift to uncertainty, so this continuity can be justified during such a transition period. For these reasons, I would suggest that my principle could also apply to a period of uncertainty following a period of risk, even if, perhaps, that period of uncertainty is not the entire future.

If this is reasonable, the discount rate in the near future may reasonably be non-negligible (albeit much smaller than extant utility discount rates) and these discount rates might apply broadly into the future. Regardless, the worry about the far future can partially be mitigated by the practical consideration that many of these models do not address the far future.

Now that I have introduced discounting under risk with my principle, I address several important objections.

4 Objections

Many objections can, and have been, raised against arguments of the preceding type. In this section, I address three. In the process, I argue that this way of addressing model-independent outcomes is superior to explicit expected value calculations because it does not require explicitly determining the value of model-independent outcome.

The first objection is raised by Parfit (1984). He claims that it is morally abhorrent to discount future people (or their utility) *simply* in virtue of the fact that they are temporally distant. Would it be morally permissible to discount people who are *spatially* distant? Modulo the details of what occurs in a life, it is worth as much as any other no matter where or when it occurs (Caney 2014).

Of course this last point is right. But the conclusion (that we should not have a positive utility discount rate) does not follow from the premise (that future people’s utility is worth the same as present people) (also cf. Heath 2017; Purves 2016).

The claim is not that, for example, we are saying that some member of a future generation is actually worth one-fifth of a current person. We are saying that, given the live possibilities that this future will be unaffected by our policies—for instance, because the future generation she is a part of does not exist or they live such a radically different existence that she is not Earth-bound—we can discount her utility.[21] This is the familiar thought from *expected utility*: when being paid

---

[21] More precisely, we are discounting the utility associated with any effects *after* that model-independent effect.

£10 for heads and £0 for tails on a coin, we do not expect you will get £5 for something between heads and tails—that is not one of the possible outcomes! No, you will get either £10 or £0, but *on average* you should expect to get £5 so you should (or could) treat it as worth £5 in present expected terms.

This leads to a second objection. Parfit (1984, p. 482) could grant this point and still object to my principle because 'it mis-states our moral view. It makes us claim, not that more remote bad consequences are less likely, but that they are less important'. But this is a very odd objection; assigning a value to δ does not tell us anything about 'our' moral view at all; it only tells us how to weight various costs and benefits. It is worth emphasizing that the intention of a cost-benefit analysis is to find and aggregate the *present* value of *future* costs and benefits. Counting a future cost less in present terms is consistent with holding that its future (or 'actual' or 'moral') value is undiscounted. The question is what the present value of future utility is; discounters agree that the future value of future utility is undiscounted.

An objector could push the point in a slightly different manner. Broome (2012) argues that the modelling of uncertainty should be in terms of different scenarios which have different probabilities attached to them, i. e. weighted expected value. Once we consider the options, we can evaluate them by the weighted sum of the outcomes that may result. Suppose I am right and discounting for these events would give us the same assessment. One could argue that there is still a good reason to avoid having $\delta > 0$. What someone who is advocating my principle should be doing, on this argument, in order to be as transparent as possible, is laying out all the different potential outcomes (including these exogenous outcomes) along with their probabilities and then doing a traditional expected value calculation. There are two reasons for this: theoretically, it is more transparent what is going on (that the discounting is occurring for probabilities, and not for something correlated with things like time); and, practically, it is more flexible. Instead of just considering these special model-independent outcomes and using this ad-hoc principle of introducing them, one would be doing the same for each potential outcome (namely, identifying its value, assigning it a probability, and multiplying the two).

These are, I believe, plausible reasons for thinking we should not use a positive δ according to my principle, although I do not think they as weighty as one might initially take them to be. Note that the objection is not that our analyses will give different results; this is instead a question of what methodology to adopt when doing cost-benefit analyses. I suggest that we can address the transparency worry by having the modellers be fully aware that time is acting as a proxy for changes in probabilities of model-independent outcomes. Given that discount rates are familiar concepts for most modellers, and discount rates often are proxies for other things (e. g. diminishing marginal value of money), this would not be difficult to account for.

A more important point is that, although Broome is right that it would be ideal to have such a complete calculation, it would be less feasible and tractable. This is because there is a benefit to using a positive δ which I suspect many in this area have not recognized. This benefit may make the suggested method superior, given that the results would converge. The benefit is that, in adopting my principle, we do not need to assign values to these model-independent outcomes in order to evaluate the policies. All we need to know is that the model-independent outcomes have *equal* value, regardless of the policy chosen. In short, we do not need to know the value *itself*! So, for instance, the value of human extinction, on some views should be 0 (in pure terms like utility or welfare) whereas, on other views it might be undefined or incomparable, due to variable population size. When using my discounting principle, this complexity is avoided, yielding a simpler working system. Instead of explicitly laying out the expected value calculation and, thereby, requiring probabilities and values for every outcome, by discounting we avoid evaluating the model-independent outcomes. When working with models, simplicity is an important virtue. A modeller may well agree that, ideally, all outcomes would be individually labelled, assigned probabilities, and evaluated. But, again, we are finite operators and require models at least partially because we cannot address all aspects of the world explicitly. Having a discount rate to account for model-independent outcomes makes the models far more tractable, since it is unnecessary to determine the actual values of these model-independent outcomes. Evaluating an outcome like ‘extinction of humanity’ is difficult at best.

The final objection also comes from Broome (2005, 2012). His objection to temporal discounting is that it introduces temporal relativity, and that such relativity is absurd. Since discounting is only with respect to future persons, he claims that a model that includes discounts can (correctly) value a future individual less than an individual after that individual would (correctly) value them. Broome’s example is people who die in war. Individuals prior to Caesar’s wars (call them Early Individuals) would judge the utility lost in the World Wars as less bad than Caesar’s wars since the discount rate would lead to heavily discounting the harms from the World Wars due to their large temporal distance. But to us (Later Individuals), this seems absurd. Surely the World Wars are many times more harmful than Caesar’s wars. So the relativism of discounting is problematic.

But the difficulties compound, Broome continues. The Early Individuals could be *aware* that the World Wars would be so much larger and that they (were they in our shoes) would view the Early Individuals’ judgment as absurd. So the relativism is genuine; there is no factual information that differs—just temporal placement.

But how do we explain our Later Individual judgments about the World Wars' (much) greater disutility when compared to Caesar's wars? Well, recall that we discount *because of risk*. Being the type of creatures that we are, we do not have uncertainty about how our decisions affect the past (i. e. they will not). So it is irrational for us to discount the utility of events we are certain of; they are not subject to risk.[22]

When Broome claims there is no factual difference between the Early Individuals and the Later Individuals, he is begging the question. For the Later Individuals, the past *is* a fact, whereas for Early Individuals the future is only conjecture, which they can assign probabilities to as best they can. So the relativism is not surprising or problematic at all: it is united in the epistemic capabilities of the different agents. The Later Individuals count past lives equally, since they are unaffected by the policies of Later Individuals; the Early Individuals discount in line with their risk over how and whether their actions will affect those yet to come. If we forget our bounded epistemic status, we are bound to make irrational choices.

5 A sample estimate of a utility discount rate for climate policies

Having defended a positive social utility discount rate for certain model-independent outcomes, it is worth considering what the implications would be concretely. I do so by providing a sample calculation to indicate how would could provide values and to point out what types of information would be relevant to such a calculation. Note that the considerations in this section are entirely independent of the preceding discussions; one can accept my principle while rejecting the (tentative!) values I propose here. Conversely, one can accept the values I suggest here while disagreeing with my utility discounting principle. What I do in this section is provide a sample estimation of a utility discount rate given by my Principle 3.1. I am estimating in the context of climate change mitigation and adaptation policies with IAMs.[23]

---

[22] If we were beings that knew only the future, and did not remember any of the past, this would be reversed. And a being with supreme knowledge—i. e. one for whom nothing was uncertain—any discounting would be irrational, since their 'model' would be complete: nothing would be exogenous to it. Of course, this being would also be decision-making under certainty.

[23] This needs to be specified since, as in Principle 3.1, which classes of events are model-independent depends on both the models and the types of policies under consideration. The

In doing so, it is worth keeping in mind that projections into the future are difficult at best, and that we should keep in mind the wise exhortations of Dasgupta (2008, p. 164): ‘we shouldn’t believe any model that explicitly models risk when the horizon extends 100–200 years into the future. We simply don’t know what the probabilities are’. However, in this context, we are assuming that—at least for the medium term (e. g. 50–100 years into the future)—it is reasonable to treat our decision-making as decision-making under risk (also cf. Roser 2017). This is often implicit in models and discussions about long-term decision-making but it is an important assumption. In this manner, it is helpful as a heuristic to consider what discount rates would imply for the probability of a class of outcomes by 2100. Of course, any suggested probabilities can and should be updated as more information becomes accessible. However, an important difficulty in doing so is that a probabilistic prediction is difficult to falsify (Armstrong et al. 2014).

Furthermore, the relevant i classes of outcomes must be disjoint and satisfy the Principle 3.1: they have to be exogenous to the model and be such that, were they to obtain, their value would be unchanged by the policies of climate mitigation and adaptation under consideration. I call the values of δ for each of the $i \leq j$ relevant classes of events δ_i.

We are assuming that the only utility relevant to the evaluation is human, so the obvious *first class* of events is human extinction from exogenous—non-climate—sources. Stern (2007) famously utility discounts for this class, but the usual assumption is that this is discounted because this would lead to ‘zero’ utility after extinction.[24] But, as mentioned above, there is nothing morally special about having zero units of utility result (just as there would be nothing morally unique about having four units result). The actual reason that such outcomes are salient, as Principle 3.1 explicates, is because *climate policies would not have an effect on an extinct species*. Stern’s discount rate was $\delta = 0.001$ which implied probability of human extinction by 2100 of approximately 10%. It turns out that 10% is not an unusual figure for those concerned about such risks (Bostrom 2013; Sandberg and Bostrom 2008), and estimates range as high as even odds of survival by the end of the century (Rees 2003). Out of deference to Stern, and to make this estimate more easily comparable, let us adopt $\delta_1 = 0.001$ for this class of events and, for the purposes of imprecise decision-making, a range of $0.0005 < \delta_1 < 0.0042$, corresponding to a range of $\approx$ 5% to $\approx$ 50% by the end of 2100.

---

values suggested for δ_2 and δ_3, in particular, would be very different if different policies and/or models were under evaluation.

[24] This should be suspect for several reasons, such as that zeroing any utility or utility scale is subject to stipulation and arbitrariness.

Even if humanity survives, there is some probability that humanity is able to leave Earth and settle on other planets, generating a *second class* of outcomes. This is obviously not explicitly modelled in any IAMs, so it is exogenous in the relevant sense. Furthermore, if humanity were to leave Earth and settle on other planets, then the mitigation and adaptation policies under consideration in climate contexts would not affect the value of the outcome of leaving Earth—at least under anthropocentric assumptions. Thus, this class would also satisfy Principle 3.1.

Assessing the probability of this second class of events is even more speculative than the first. We should be wary of unbridled optimism. Futurists in the 1960s thought that humanity would be able to travel away from Earth very easily by the beginning of the current millennium (incidentally, Stanley Kubrick's *2001: A Space Odyssey* was motivated by this optimism). We can also remember H. G. Wells's triumphal speech of 1902 to the Royal Institution:

> Worlds may freeze and suns may perish, but there stirs something within us that can never die. . . a day will come, one day in the unending succession of days, when beings—beings who are now latent in our thoughts, and hidden in our loins—shall stand upon this earth as one stands upon a footstool, and shall laugh and reach out their hands amidst the stars.[25]

So what do experts say today?[26] On the optimistic side, we have astronomers like Chris Impey (*Beyond: Our Future in Space*) who predict frontier colonies beyond Earth in three decades with 2115 bringing human beings who have never been on Earth. His optimism is outstripped by journalists like Stephen L. Petranek (*How We'll Live on Mars*), who thinks people will get to Mars within a couple of decades. On the more pessimistic side, we have science historian Erik M. Conway of the Jet Propulsion Laboratory (*Exploration and Engineering: The Jet Propulsion Laboratory and the Quest for Mars*), who thinks that human desires to get to Mars have been instrumental in damaging and undermining the efforts of Martian scientists and that the technical challenges are daunting or insurmountable. My subjective probability assessment is that it is certainly not impossible that by 2100 humanity will have left Earth, but I am far more pessimistic than the optimists cited above, due to the scale of the engineering and technical challenges, along with the recent lack of interest in funding extraplanetary research. I would suggest the probability to be an order of magnitude smaller than the probability of extinction (i. e. $\approx$ 1% in 2100 as opposed to $\approx$ 10% for the class of extinction outcomes); this implies an associated $\delta_2 = 0.0001$.

---

[25] Of course, it may still be that he was right, and that his day was more than a century hence, but it should still give us pause.

[26] A helpful popular treatment was written by Kolbert (2015).

The first class contains outcomes where humanity goes extinct (rendering adaptation and mitigation unimportant after the extinction event); the second class contains outcomes where humanity is not extinct but not on Earth (again making adaptation and mitigation on Earth unimportant). Primarily for the sake of completeness, I introduce another class, but I think a reasonable probability assignment will render it close to negligible. In such a *third class* of outcomes, humanity survives, stays on Earth, but is somehow unaffected by climate mitigation and adaptation policies. It might be that some Earth-bound habitation could mitigate the effects of climate change, either below or above ground (cf. 'dome world' in Gardiner 2011, p. 43).[27] Some futurists propose radical human augmentation or downloading our consciousness onto computers (etc.) in such a manner that we are not affected by environmental or physical limits. These futurists believe that this would require some beneficial Artificial Intelligence or singularity to be reached, but I think such outcomes should be assigned negligible probability. In terms of such AI, it is worth noting that surveys of people in relevant fields believe that such an event could occur quite soon (i. e. 50% probability of human level AI ranged from 2035 to 2050; 90% probability of human level AI from 2065 to 2150). However, Armstrong et al. (2014, p. 326) tell us that 'There is a strong tendency to predict the development of AI within 15-25 years from when the prediction is made... Experts, non-experts and failed predictions all exhibit this same pattern'.[28] For these reasons, I do not think we should take these estimates seriously. So, on the one hand, this third class of events has to be small because it only includes those which are unaffected by chosen policies. However, on the other hand, I believe that there is reason to think that a dome world scenario should be given at least as large a probability to settling away from Earth, because there would be fewer engineering challenges, both in terms of mundane things like vegetation and water but also for more difficult factors like gravity. So I am assigning it the same level of magnitude as the second class of events (i. e. $\approx$ 1% in 2100) for $\delta_3 \approx 0.0001$.

This implies a value $\delta = \sum_{i=1}^{j} \delta_i = \delta_1 + \delta_2 + \delta_3 \approxeq 0.0012$ (or 0.12%). Since many probability assignments—for these policies and models—to the second and third class of events will be relatively small compared to the first, δ will often be fairly close to δ_1, making the usual exogenous extinction probability at least a good heuristic. However, δ will always be strictly greater than δ_1 for IAMs, since previous estimates for δ, to my knowledge, have never incorporated any $i > 1$

---

[27] Although this outcome would have to be specified very carefully, because climate effects would still have impacts, e. g. on the upkeep of the dome or on groundwater levels underground.

[28] As some involved are aware, there are also many other biases that could explain the near timing of such estimates: http://aiimpacts.org/short-prediction-publication-biases/.

classes. Furthermore, it may be that more plausible probability assignments in the future will have $\delta_2 \gg \delta_1$ or $\delta_3 \gg \delta_1$, in which case $\delta_1 \approx \delta$ would be a *very poor* heuristic, e. g. if certain technological capabilities advance very rapidly. Beyond the theoretical advantages, this is an important practical reason to be principled about determinations of the value of δ.

6 Conclusion

It is worth returning to the point that modelling the costs of the future involves significant uncertainty, and should not be the only (perhaps not even the most) relevant input into social decision-making (Kelleher 2017a). However, it is of great value to put such modelling on a strong philosophical and ethical foundation, especially given the outsize influence it has had over policy and the foundations of utilitarianism (Nordhaus 2007a,b). The purpose of this paper was to make such parameter assignments more principled and to lay out clearly the presumptions needed to defend a positive parameter assignment to δ. For individual decision-making, as it is not of such social importance, it is plausibly not part of the moral realm whether individuals discount the value of events in their own future—which psychology suggests that they do (Frederick et al. 2002) and where I agree with some economists who think that this is morally defensible or perhaps innocuous (Arrow et al. 1996). However, social decision-making plausibly requires a higher standard of rationality and more perspicuous reasoning.

The generation of a particular value of δ in Section 5 is meant to provoke discussion, since it is worth thinking explicitly about what level of magnitude we should assign to such complex and difficult to foresee events. I am optimistic that if we consider these probabilities more explicitly, we might converge to reasonable values (which may even be negligible). This is necessary because, for beings like us, little—if any!—of the future is certain and we should contemplate the risks of the future with our eyes open.

Acknowledgment: Appreciation for all of the people who have read and commented on drafts of this including: Mark Budolfson, Marc Davidson, Steve Gardiner, Ben Henry, Paul Kelleher, Ewan Kingston, Simo Kyllönen, Aaron Malthias, Lukas Meyer, Darrel Moellendorf, John O'Neill, Duncan Purves, Matthew Rendall, Dominic Roser, Karl Steininger, Kevin Tobia, Sara Uckelman, and three very helpful anonymous reviewers.

Versions of this paper also benefited from participants at the 89th Pacific Division Meeting of the American Philosophical Association (and my commentator Hun Chung), Societas Ethica's Conference on Climate Change, the University of

Graz' Climate Ethics and Climate Economics Symposium, the European Consortium for Political Research's 42nd Joint Sessions, and most of all at the University of Münster Demandingness in Practice workshop, whose organizers also guest edited this special issue.

Funding: This work was funded by the Austrian Science Foundation (FWF) under research grant W 1256-G15 (Doctoral Programme Climate Change—Uncertainties, Thresholds and Coping Strategies).

References

Armstrong, S., Sotala, K. and Ó hÉigeartaigh, S. (2014). 'The Errors, Insights and Lessons of Famous AI Predictions — and what they Mean for the Future', *Journal of Experimental & Theoretical Artificial Intelligence* 26 (3): 317–342.

Arrow, K.J., Cline, W., Mäper, K., Squitieri, R. and Stiglitz, J. (1996). 'Intertemporal Equity, Discounting, and Economic Efficiency', in J. Bruce, H. Lee, and E. Haites (eds.). *Climate Change 1995—Economic and Social Dimensions of Climate Change* (Cambridge: Cambridge University Press).

Bostrom, N. (2013). 'Existential Risk Prevention as Global Priority', *Global Policy* 4 (1): 15–31.

Broome, J. (1992). *Counting the Cost of Global Warming* (Cambridge: White Horse Press).

Broome, J. (1994). 'Discounting the Future', *Philosophy & Public Affairs* 23 (2): 128–156.

Broome, J. (2005). 'Should we Value Population?', *The Journal of Political Philosophy* 13 (4): 399–413.

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

Caney, S. (2008). 'Human Rights, Climate Change, and Discounting', *Environmental Politics* 17 (4): 536–555.

Caney, S. (2009). 'Climate Change and the Future: Discounting for Time, Wealth, and Risk', *Journal of Social Philosophy* 40 (2): 163–186.

Caney, S. (2014). 'Climate Change, Intergenerational Equity and the Social Discount Rate', *Politics, Philosophy & Economics* 13 (4): 320–342.

Cowen, T. and Parfit, D. (1992). 'Against the Social Discount Rate', in P. Laslett and J.S. Fishkin (eds.). *Justice Between Age Groups and Generations* (New Haven: Yale University Press), pp. 144–161.

Dasgupta, P. (2008). 'Discounting Climate Change', *Journal of Risk and Uncertainty* 37: 141–169.

Dasgupta, P. (2012). 'Time and the Generations', in R.W. Hahn and A. Ulph (eds.). *Climate Change and Common Sense: Essays in Honour of Thomas Schelling*, chapter 7 (Oxford: Oxford University Press), pp. 101–130.

Dasgupta, P. and Heal, G. (1979). *Cambridge Economic Handbooks: Economic Theory and Exhaustible Resources* (Cambridge: Cambridge University Press).

Dietz, S., Hope, C. and Patmore, N. (2007). 'Some Economics of 'Dangerous' Climate Change: Reflections on the Stern Review', *Global Environmental Change* 17: 311–325.

Fleurbaey, M., Ferranna, M., Budolfson, M.B., Dennig, F., Mintz-Woo, K., Socolow, R., Spears, D. and Zuber, S. (2018). 'The Social Cost of Carbon: Valuing Inequality, Risk, and Population for Climate Policy', *The Monist* 102 (1): 84–109.

Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002). 'Time Discounting and Time Preference: A Critical Review', *Journal of Economic Literature* 40 (2): 351–401.

Gardiner, S. (2011). *A Perfect Moral Storm: The Ethical Tragedy of Climate Change* (Oxford: Oxford University Press).

Godard, O. (2009). 'Time Discounting and Long-run Issues: The Controversy Raised by the Stern Review of the Economics of limate Change', *OPEC Energy Review* 33 (1): 1–22.

Heath, J. (2017). 'Climate Ethics: Justifying a Positive Social Time Preference', *Journal of Moral Philosophy* 14 (4): 435–462.

Interagency Working Group on Social Cost of Carbon (2013). 'Technical Update. The Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866', *United States Government*, pp. 1–21. Available at https://obamawhitehouse.archives.gov/sites/default/files/omb/inforeg/social_cost_of_carbon_for_ria_2013_update.pdf (accessed on May 10, 2018).

Kelleher, J.P. (2012). 'Energy Policy and the Social Discount Rate', *Ethics, Policy & Environment* 15 (1): 45–50.

Kelleher, J.P. (2017a). 'Descriptive Versus Prescriptive Discounting in Climate Change Policy Analysis', *Georgetown Journal of Law & Public Policy* 15: 957–977.

Kelleher, J.P. (2017b). 'Pure Time Preference in Intertemporal Welfare Economics', *Economics and Philosophy* 33 (3): 441–473.

Kolbert, E. (2015). 'Project Exodus', *The New Yorker* 91 (15): 76–79.

Koopmans, T. (1960). 'Stationary Ordinal Utility and Impatience', *Econometrica* 28 (2): 287–309.

Koopmans, T. (1965). 'On the Concept of Optimal Economic Growth', in Pontifical Academy of Sciences (ed.). *The Econometric Approach to Development Planning*, chapter 4 (Amsterdam: North Holland), pp. 255–87.

Mintz-Woo, K. (2017). 'A New Defence of Probability Discounting', in A. Walsh, S. Hormio, and D. Purves (eds.). *The Ethical Underpinnings of Climate Economics*, chapter 6 (Oxford: Routledge), pp. 87–102.

Ng, Y.-K. (2016). 'The Importance of Global Extinction in Climate Change Policy', *Global Policy* 7 (3): 315–322.

Nolt, J. (2018). 'Cumulative Harm as a Function of Carbon Emissions', in C. Murphy, P. Gardoni and R. McKim (eds.). *Climate Change and Its Impacts: Risks and Inequalities*, chapter 3 (New York: Springer), pp. 39–52.

Nordhaus, W. (2007a). 'A Review of the "Stern Review on the Economics of Climate Change"', *Journal of Economic Literature* 45 (3): 686–702.

Nordhaus, W. (2007b). 'Critical Assumptions in the Stern Review on Climate Change', *Science* 317 (5835): 201–202.

Nordhaus, W. (2008). *A Question of Balance* (New Haven: Yale University Press).

Parfit, D. (1984). *Reasons and Persons* (Oxford: Oxford University Press).

Purves, D. (2016). 'The Case for Discounting the Future', *Ethics, Policy & Environment* 19 (2): 213–230.

Ramsey, F.P. (1928). 'A Mathematical Theory of Saving', *The Economic Journal* 38 (152): 543–559.

Rees, M. (2003). *Our Final Hour* (New York: Basic Books).

Roser, D. (2017). 'The Irrelevance of the Risk-Uncertainty Distinction', *Science and Engineering Ethics* 23 (5): 1387–1407.

Sandberg, A. and Bostrom, N. (2008). 'Global Catastrophic Risks Survey', *Technical Report 2008/1*, pp. 1–5. Available at http://www.global-catastrophic-risks.com/docs/2008-1.pdf (accessed on May 10, 2018).

Sidgwick, H. (1981/1907). *The Methods of Ethics* (Indianapolis: Hackett).

Singer, P. (2011). *Practical Ethics: Third Edition* (Cambridge: Cambridge University Press).

Stern, N. (2007). *The Economics of Climate Change* (Cambridge: Cambridge University Press).

Stern, N. (2015). *Why Are We Waiting?* (Cambridge, Mass: MIT Press).

Tarsney, C. (2017). 'Does a Discount Rate Measure the Costs of Climate Change?', *Economics and Philosophy* 33 (3): 337–365.

Wahba, M. and Hope, C. (2006). 'The Marginal Impact of Carbon Dioxide Under Two Scenarios of Future Emissions', *Energy Policy* 34:3305–3316.

Weisbach, D. and Sunstein, C. (2007). 'Introduction', *The University of Chicago Law Review* 74 (1): 1–3.

Yaari, M. (1965). 'Uncertain Lifetime, Life Insurance, and the Theory of the Consumer, *The Review of Economic Studies* 32 (2): 137–150.