# Fundamental utilitarianism and intergenerational equity with extinction discounting

**Article:**

**ORIGINAL PAPER**

# Fundamental utilitarianism and intergenerational equity with extinction discounting

Graciela Chichilnisky[1] · Peter J. Hammond[2] · Nicholas Stern[3]

© The Author(s) 2020

## Abstract

Ramsey famously condemned discounting "future enjoyments" as "ethically indefensible". Suppes enunciated an equity criterion which, when social choice is utilitarian, implies giving equal weight to all individuals' utilities. By contrast, Arrow (Contemporary economic issues. International Economic Association Series. Palgrave Macmillan, London, 1999a; Discounting and Intergenerational Effects, Resources for the Future Press, Washington DC, 1999b) accepted, perhaps reluctantly, what he called Koopmans' (Econometrica 28(2):287–309, 1960) "strong argument" implying that no equitable preference ordering exists for a sufficiently unrestricted domain of infinite utility streams. Here we derive an equitable utilitarian objective for a finite population based on a version of the Vickrey–Harsanyi original position, where there is an equal probability of becoming each person. For a potentially infinite population facing an exogenous stochastic process of extinction, an equitable *extinction biased* original position requires equal *conditional* probabilities, given that the individual's generation survives the extinction process. Such a position is well-defined if and only if survival probabilities decline fast enough for the expected total number of individuals who can ever live to be finite. Then, provided that each individual's utility is bounded both above and below, maximizing expected "extinction discounted" total utility—as advocated, *inter alia*, by the *Stern Review* on climate change—provides a coherent

✉ Peter J. Hammond
p.j.hammond@warwick.ac.uk

Graciela Chichilnisky
gc9@columbia.edu

Nicholas Stern
n.stern@lse.ac.uk

[1] Department of Economics, International Affairs Building, Columbia University, 420 W 118 St MC 3308, New York, NY 10027, USA

[2] Department of Economics, and CAGE (Competitive Advantage in the Global Economy), University of Warwick, Coventry CV4 7AL, UK

[3] Department of Economics, and Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

⌂ Springer

and dynamically consistent equitable objective, even when the population size of each generation can be chosen.

# 1 Introduction

## 1.1 The discounting issue: weighting future generations

The question of whether and how to discount the welfare of future generations has played a critical role in discussing the ethical and economic foundations of long-run policy analysis. This is especially true in the case of climate change which, if managed badly enough, could even accelerate the possibility of human extinction. Indeed, how much to discount the welfare of future generations was a particular, and a particularly contentious, issue addressed by Arrow (1999a, b) himself. In those two largely overlapping articles, he attributed to Koopmans (1960) two arguments in favour of discounting the welfare of future generations. The first "strong" argument, set out also by Koopmans et al. (1964) as well as by Diamond (1965) on his own, was that, without giving less weight to utilities far in the future, no complete and transitive preference ordering on the entire space of infinite utility streams could satisfy the usual continuity and weak Pareto conditions.[1] The second "weak" argument was that failure to discount later generations' utilities would imply that earlier generations are condemned to make excessive sacrifices.

Among many subsequent discussions of this discounting issue, two that clearly drew Arrow's attention were Chichilnisky (1996), cited in Arrow (1999a, b), and then the *Stern Review* (Stern 2007), reviewed in Arrow (2007). Indeed, it seems quite likely that carefully and attentively reading the *Stern Review* led KA to have some doubts about his own earlier articles.[2] This helps explain why he appeared so keen to make discounting the subject of one-on-one discussions with one of us (PH) over at least two lunches at Stanford during the approximate period 2014–2015.[3]

## 1.2 Fundamental utilitarianism

The paper we present here follows many others in addressing the discounting issue by applying modern ideas of social choice theory with interpersonal comparisons, some of which Arrow (1977) himself approved in the early years. Inspired by Adam Smith's *Theory of Moral Sentiments*, we consider an "impartial spectator" who uses conse-

---

[1] Peter Diamond kindly prompted us to recall that Koopmans (1960, p. 287) had the "utility function" of an "individual consumer" as his primary concern. When considering optimal growth, Koopmans (1965, 1967) does consider the case when the utility of future generations is discounted.

[2] From now on we frequently refer to Kenneth Arrow and the three authors of this paper using a pair of initials—i.e., KA, GC, PH and NS.

[3] During these discussions PH remained unaware of Arrow (1999a, b), whose existence KA was too modest to mention. But it was clear that KA had been inspired by his late friend Tjalling Koopmans, who had been a mentor during his time at the Cowles Foundation when it was still in Chicago. It is also said that Koopmans had persuaded KA and Gérard Debreu to amalgamate their separate working papers into what became Arrow and Debreu's (1954) classic paper on existence of general competitive equilibrium.

quentialist decision theory (Hammond 1996, 1998) to contemplate how to navigate arbitrary finite decision trees when the consequences belong to a domain of complete lifetime histories of everything that should be relevant to decisions affecting a single individual. In particular, we follow Kolm (1994) but contradict Broome (1993) in asserting that any "cause of preference" which is ethically relevant should be an "object of preference" included in the consequence domain of all ethically relevant lifetime histories.

When contemplating decisions affecting a whole society, we imagine a "social benefactor". This benefactor's concern is with consequence lotteries which emerge from a significant modification of Rawls' (1971) original position, behind a veil of ignorance. Specifically, we avoid extreme risk aversion assumed by Rawls, which was brought out in Harsanyi (1975) and Hammond (1975). Instead we follow both Vickrey (1945) and Harsanyi (1953, 1955) in assuming that a lottery determines, in effect, which individual in society the benefactor will become, and so whose personal consequence is relevant ex post, after this lottery has been resolved. Harsanyi in particular postulates an "ethical observer" who acts as if there were an equal chance of becoming any named individual upon emerging from what we will call the "Vickrey–Harsanyi" original position. Treating all potential individuals equally in this way accords with the concept of equity due to Suppes (1966) and applied to social choice theory by Sen (1970, Chapter 9*), then by Hammond PJ (1976, 1979) and many successors. In this setting, consequentialist rationality implies that the impartial benefactor should maximize the average expected utility over the entire population. With an infinite set of future generations, however, this equal chance lottery is not even well defined. That is one reason why the strong argument for discounting that Arrow attributes to Koopmans seems so persuasive.

## 1.3 Extinction discounting

In this social choice context, in order to escape the iron logic of Koopmans' strong argument, we recognize some relevant physical reality, and the ultimate inevitability of mass extinction. This may occur due to the astrophysics of the sun, whose energy releases will gradually intensify beyond a level consistent with continued life on Earth. Also, though the likelihood of a devastating asteroid impact may ultimately be reduced by human ingenuity, it is unlikely ever to be eliminated entirely. Finally, there is the possibility of a mass extinction event due to a supervolcano on Earth, or to an intense gamma-ray burst from a source close enough to be within our Milky Way galaxy.

Accordingly we postulate that the possibility of such events together determine a background hazard process whose outcome is a risky extinction date beyond which human life will be impossible. We treat this process as exogenous, in the sense that any human action will have a negligible effect on the hazard rates behind it.

Our main result concerns the effect of "extinction discounting", which is when this background process is used to determine what discount factor should apply to each future generation's utility levels. This accords with the ideas of Sidgwick (1907), Ramsey (1931), Mirrlees (1967) and Dasgupta and Heal (1979), as well as Stern (2007,

2008, 2014a, b, 2015).[4] Specifically, we give three conditions that are jointly sufficient for an "extinction discounted" sum of future generations' utilities to give a well-defined Bergson social welfare function whose expected value should be maximized. These sufficient conditions are:

1. following Arrow (1951, 1965, 1971), as well as Blackwell and Girshick (1954) and Hammond (1998), the fundamental utility function that the impartial spectator and benefactor applies to any individual's personal consequences is bounded both above and below;
2. any individual who never exists is assigned a unique specific utility level that is normalized to zero;
3. attention is restricted to a restricted domain of intergenerational consequence streams for which the background extinction process is fast enough, and the rate of population growth slow enough, to ensure that with extinction discounting applied to the size of each generation, the expected total discounted population is finite and uniformly bounded.

Moreover, as discussed in Dasgupta and Heal (1979, ch. 9), the utilitarian objective with extinction discounting that we consider can be derived from an intergenerational variation of the Vickrey–Harsanyi original position.[5] In this variation, rather than each individual having the same probability of being selected, each has the same *conditional* probability given the event that their generation comes into existence early enough to survive the random background extinction process.[6]

## 1.4 Outline

The remainder of the paper begins with some brief recapitulations of key ideas. First, Section 2 recalls the key distinction between discounting the utility as opposed to the consumption of individuals in future generations. It also sets out Arrow's (1999a; 1999b) reasons for claiming that the utilities of future generations should be discounted.

Next, Section 3 discusses how "consequentialized" ethics can be combined with consequentialized decision theory to make a case for an impartial spectator/benefactor to maximize the expected value any fundamental von Neumann–Morgenstern utility function in a unique cardinal equivalence class, defined on an appropriate domain of personal consequences. Moreover, following ideas that Kolm (1972, 1998) ascribes to Tinbergen (1957), this fundamental utility function, including its domain, should be the

---

[4] Here we note the influence of James Mirrlees, whose Ph.D. dissertation on optimal growth (Mirrlees 1963), contrary to some accounts, had KA as an active external examiner rather than supervisor. In Cambridge during the late 1960s co-authors PH and NS learned about exponential discounting from Mirrlees, who was their and Dasgupta's Ph.D. supervisor, while also providing Heal with valued informal advice.

[5] Dasgupta and Heal point out that an intergenerational extension of the Vickrey–Harsanyi original position had already been discussed by Rawls (1971, pp. 287–8), though without any mention of possible extinction. We note that Dasgupta and Heal's analysis considers generations of individuals who live for only one period. It also precludes generations of different sizes, which may also be endogenous because they can be affected by policy choices, including those that increase the endogenous risk of extinction.

[6] See Ord (2020) for an alternative philosophical analysis that, *inter alia*, also supports extinction discounting.

same for all individuals. Finally, to allow for individuals whose potential existence is precluded by extinction, this domain should include a personal consequence associated with non-existence.

Section 4 extends the decision analysis of Sect. 3 to a society of finitely many individuals. To do so, it introduces a generalization of the even chance lottery that arises in the Vickrey–Harsanyi original position. Any generalized lottery reduces a multi-person decision problem to one with a single lottery. In the case of an even chance lottery, this can be regarded as decision-making by an impartial benefactor who abides by the Suppes equity principle.

With infinitely many potential individuals, the even chance lottery that the impartial benefactor uses in an Vickrey–Harsanyi original position becomes logically incoherent. Section 5 introduces a more structured model with an infinite sequence of generations, each with a finite population, as well as a background extinction process. Then we consider an "extinction biased" original position that equates individuals' *conditional* probabilities of being selected, given the event that they belong to a generation that comes into existence early enough to survive the random extinction process. These equated conditional probabilities exist if and only if the expected total number of individuals in all generations that survive the stochastic extinction process is finite. Then, provided we apply Arrow's boundedness assumption to the fundamental utility function, the corresponding expected utility of an extinction biased benefactor is an absolutely convergent sum. So expected utility maximization for this benefactor is well defined, and is equivalent to applying extinction discounting to each generation's total utility.

Next, Section 6 considers Chichilnisky's (1996) objection that any welfare criterion based on discounted utility must involve a "dictatorship of the present". A mixed criterion is proposed, fulfilling Chichilnisky's definition of sustainable preferences by avoiding not only a dictatorship of the present, but also a dictatorship of the future.

The last main Sect. 7 extends our results to the important case when we the set of individuals belonging to each successive generation is treated as an endogenous variable that is affected by policy choices. The extinction discounted expected utility objective of Sect. 5 can still be applied on a restricted domain with exogenous population bounds on the population of each generation. This objective, however, is equivalent to maximizing the expectation of the extinction discounted total utility of all generations. The objective is well defined, independently of the population bounds, on the restricted domain of generation structures for which the extinction discounted total population exists and is uniformly bounded.

Finally, Sect. 8 offers some concluding discussion.

# 2 Discounting what?

## 2.1 The social rate of discount for future consumption

It is important to distinguish discounting future consumption from discounting future utility. Following what has become standard practice since at least Arrow and Kurz

(1970) as well as Stern (1977), Section 3 of Arrow (1999a) gives the formula $r = \rho + \theta g$ for the social rate of discount to be applied to future consumption, where: [7]

> $\rho$ is the of pure time preference (if any), $\theta$ is the elasticity of marginal utility with respect to income, and $g$ is the of growth of consumption *per capita* …$\rho = 0$ implies equal treatment of present and future.

Thereafter Sect. 4 of Arrow (1999a) starts as follows:

> In [the] formula [$r = \rho + \theta g$], the second term, $\theta g$, is, I think, fairly uncontroversial. If future individuals are going to be better off than we are, then our willingness to sacrifice on their behalf is certainly reduced. It would require a greater rate of return to justify our depriving ourselves of consumption.

> But the presence of pure time preference, denoted by $\rho$, has been very controversial. The English economists, in particular, have tended to be very scornful of pure time preference.

In the ensuing discussion of the term $\theta g$, Arrow (1999a) considers how to value an increment in a good in the future, relative to an increment now. This is the discount factor, which we denote by $\beta$, for that good at that time. In the continuous time model that Arrow uses, the proportionate rate of decrease of the discount factor is given by $-\dot{\beta}/\beta = -\frac{\mathrm{d}}{\mathrm{d}t} \ln \beta$. It is the discount rate for that good at that time; it clearly depends on both the good and the time. In our view, the focus in economic assessments should be on the discount factor, as that is the key shadow price, relative to now, which is needed to find the marginal present value of any change in costs or benefits occurring at any specific time in the future. When we need to evaluate a stream of costs and benefits over time, we can consider the net present value (NPV) of the whole stream, with the costs and benefits at each time $t$ weighted by the discount factor $\beta(t)$. That is, at time 0 one considers NPV $:= \int_0^T \beta(t)\, b(t)\, \mathrm{d}t$, where $b(t)$ denotes net benefit at time $t$, and $T$ denotes the terminal time. When allowing for the inevitable uncertainty surrounding future costs and benefits, one approach is to consider their expected discounted value.

Once we have the right concept of the discount factor that should be applied to future consumption, it becomes immediately clear that this factor will depend on the state of affairs at each relevant time in the future. Unmanaged climate change could make future generations very poor. Then we might place a very high value on extra goods that are available in calamitous circumstances. This could even imply negative discounting, or equivalently, a discount factor greater than one. This possibility also makes it clear that each uncertain future state of the world that could occur is of critical relevance. So, too, is the person or persons who may experience increments in income. Indeed, using the term "*the* discount rate", as if there is just one given rate, clearly constitutes a serious misunderstanding of the basic issues.

So the relevant discount rates in any calculation of expected discounted value are endogenously determined as a result of our planned decisions. Moreover, this endogeneity is potentially severe in the case of climate change. After all, if unmanaged

---

[7] This is sometimes described as the "Ramsey equation". Yet the closest analogy in Ramsey (1928) seems to be Eq. (9) on p. 554, which however involves the elasticity of utility rather than of marginal utility.

climate change causes devastation and deep poverty in the future, that suggests we should weight future consumption higher than present consumption, which implies negative discounting.

## 2.2 Discounting the welfare of future generations

Our concern in this paper, however, is much more with the $\rho$ term of the formula $r = \rho + \theta g$ that Arrow (1999a) gives for the discount rate. This is often called pure time discounting, or pure time preference. It arises when we contemplate policies whose effects, like the climate change induced by greenhouse gas emissions, extend far into the future. This should force us to give some value to the consumption of people who live in the future. Pure time discounting involves, and is even essentially defined as, the relative weight attached to a life in the future compared to a life now, when the two lives are otherwise identical in all respects. That is, the only difference is that one life is in the future, whereas the other is right now.

If the pure time discount rate were 2% per annum, for example, then a life starting 35 years in the future that is otherwise identical to a life that starts now, would have a relative value of $1.02^{-35} \approx 0.5$ compared to a life that starts now. In this sense we are "discounting future lives", which is effectively discriminating by date of birth. It cannot be justified by some notion of the future life being better because it has higher consumption; that would be discounting future consumption, as considered in Sect. 2.1, as opposed to discounting future welfare per se. We emphasize that we are making an ethical comparison between two lives that are identical, except for the dates of birth. It is very difficult to find serious ethical arguments for the kind of discrimination that is involved in giving less weight to future generations' well-being. Indeed, Dasgupta and Heal (1979, p. 262) remind us of a highly relevant passage from Sidgwick (1907, p. 414):

> "How far we are to consider the interests of posterity when they seem to conflict with those of existing human beings? It seems …clear that the time at which a man exists cannot affect the value of his happiness from a universal point of view; and that the interests of posterity must concern a Utilitarian as much as those of his contemporaries, except in so far as the effect of his actions on posterity—and even the existence of human beings to be affected—must necessarily be more uncertain."

In his celebrated paper on optimal saving, Ramsey (1928, p. 261) famously started out by following the spirit of Sidgwick when he refused to discount the welfare of future generations:

> One point should perhaps be emphasised more particularly; it is assumed that we do not discount later enjoyments in comparison with earlier ones, a practice which is ethically indefensible and arises merely from the weakness of the imagination.[8]

---

[8] The latter part of the paper "no longer reckon[s] future utilities and disutilities as equal to present ones, but discount[s] them at a constant rate $\rho$." (Ramsey 1928, p. 553).

Arrow ([1999b](#)) not only quotes this passage from Ramsey, but also adds two later quotes from other English economists. The first is from Pigou ([1932](#), p. 25) stating that pure time preference "implies …our telescopic faculty is defective." The second is Harrod's ([1948](#), p. 40) claim that "[P]ure time preference [is] a polite expression for rapacity and the conquest of reason by passion."[9]

### 2.3 Discounting our own future

When individuals contemplate their own future consumption, a discussion of how to discount it very like that in Sect. 2.1 might easily arise. On the other hand, if the same individuals contemplate their own future standard of living, there may be a closer parallel with Sect. 2.2. For example, an individual saving for retirement might want to discount future income somewhat if there is no longer any need to cover the expense of travelling to work, or of being able to live very close to work.

Such discounting of our own futures is quite different from what concerns us here. We are examining the ethical issue of whether there is any justification for discounting a life simply and only on the grounds that it starts later. It is not clear why the impatience of an individual who may value the future less than the present should be at all relevant to ethics. The fact some people are sometimes impatient in their own decisions does not tell us that there is any moral justification for discriminating between different people just because some are born decades later than others.

At this point it may be worth reminding the reader that market prices and interest rates, or rates of return, are very unlikely to give us ethical evaluations of the kind needed to guide society toward good decisions. Instead, they describe facts concerning the outcomes (equilibrium or otherwise) of the individual choices of many market participants. Indeed, market interest rates typically do not even give ethically appropriate individual marginal valuations, especially given the many interrelated imperfections that seem inevitable in capital markets—as argued, for example, in Hammond ([1992](#)). Looking at market rates is rarely an ethically defensible route to the social evaluations that are necessary here.

Note, however, that just as mortality is one reason for discounting our own futures, so the possibility of human extinction is a reason for discounting social outcomes. This, of course, is the key idea behind extinction discounting, which we take up starting in Sect. 5.

### 2.4 The strong argument

The first reason that KA gave for discounting, impatience, or time perspective was what we will call the "strong argument". This is to be distinguished from what Arrow ([1999b](#)) explicitly describes as the "weak Koopmans argument", which receives brief attention in Sect. 2.6 below, as well as in Sect. 8.3.

---

[9] Among numerous discussions of intergenerational equity, we mention here only Arrow et al. ([1996](#)). For an extensive list of many other works, see our working paper Chichilnisky et al. ([2018](#)).

To quote Arrow (1999b):[10]

> Why then not embrace the idea of zero time perspective? Koopmans in several classic papers (1960, 1964) gave a crushing answer; see also Brown and Lewis (1981) for a more general treatment. The argument seems recondite. Koopmans considers a world which lasts forever. Therefore choice (including ethically-based choice) is based on a preference ordering over infinite-dimensional consumption streams. He argues that if the ordering is continuous and also sensitive (i.e., if one stream is never worse than another and is better at one or more time points, then it must be strictly preferred), it must display impatience.

> A simple restatement of his reasoning can bring out the essential point. I confine myself to the intertemporally separable case. Imagine initially that output consists of a constant stream of completely perishable goods. There can be no investment by definition. Now imagine that an investment opportunity occurs, available only to the first generation. For each unit sacrificed by them, a perpetual stream of $\alpha$ per unit time is generated. If there were no time preference, …we can say that given any investment, short of the entire income, a still greater investment would be preferred.

Thus, Arrow concludes that without impatience the optimal saving rate could become arbitrarily close to 100%. A similar conclusion emerges from the cake-eating example described by Gale (1967, p. 4, Example 2).

Nevertheless, the following passage from Koopmans (1960, pp. 287–288) suggests that he at least intended his results to be applied only in the rather different context of consumer choice:

> This study started out as an attempt to formulate postulates permitting a sharp definition of *impatience*, the short term Irving Fisher has introduced for preference for advanced timing of satisfaction. To avoid complications connected with the advancing age and finite life span of the individual consumer, these postulates were set up for a (continuous) utility function of a consumption program extending over an infinite future period. The surprising result was that only a slight strengthening of the continuity postulate …permits one to conclude from the existence of a utility function satisfying the postulates, that impatience prevails at least in certain areas of the program space.

Thus, it seems that Koopmans (and Diamond) started out by considering only consumers who discount their own future selves, as discussed in Sect. 2.3.[11] Nevertheless, it was natural for Arrow to consider the obvious extension to social choice theory, with an infinite series of successive generations. Indeed, this follows the tradition of the

---

[10] The opening question in the quotation is a valuable addition in Arrow (1999b) to the corresponding passage in Arrow (1999a). The second "classic paper" is presumably Koopmans et al. (1964).

[11] Stern (2014b, p. 472; 2015, p. 169) also quotes a recent personal communication in which Peter Diamond, a co-author of Koopmans et al. (1964), had argued to the effect that the results of this line of work, if they would indeed preclude intergenerational equity, should not be applied to the issue of whether to discount the welfare of future generations. Instead Diamond has argued in favour of the kind of "pragmatic" criteria discussed in Sects. 6.7 and 6.8 of Chichilnisky et al. (2018).

later works by Koopmans (1965, 1967), who even devotes part of these surveys to the Ramsey case when the welfare of future generations remains undiscounted. Then the same mathematical analysis which, under some conditions, shows that a consumer cannot treat equally consumption in an infinite number of periods, also rules out intergenerational equity, in the sense of treating all future generations equally.

### 2.5 Two kinds of domain restriction

Note that the strong argument says only that discounting is required if one tries to construct a "sensitive" and continuous complete and transitive preference ordering over an *unrestricted* domain of infinite consumption streams. When one considers a suitably restricted domain of infinite intergenerational consumption streams, then as KA was surely well aware, Ramsey (1928) had already shown that an optimal savings plan could well exist without any discounting. More generally, following Gale (1967), Mirrlees (1967) and others, there is a standard convergence condition for existence of an optimal growth path satisfying the equation $r = \rho + \theta g$ as a first-order condition for the continuous time optimization problem. The convergence condition requires the social rate of discount $r$ to exceed the sum of the rate of growth of population and the rate of technical progress. Indeed, if the elasticity $\theta$ is greater than 1, then the higher is $g$ (the growth rate, or rate of technical progress), the more likely it is that $r$ become large enough so that the relevant welfare integral converges. When it does converge, one avoids the kind of oversaving problem that can arise in the cake-eating example due to Gale (1967).[12]

The main result of our paper, however, considers a second kind of domain restriction. Assuming a bounded utility function and an exogenous background extinction process, we prove that a sensitive and continuous complete and transitive preference ordering does exist on the restricted domain of personal consequence streams for which the expected extinction discounted population is finite. Moreover, the ordering satisfies a suitable version of intergenerational equity.

### 2.6 The weak argument

Arrow (1999a) offered an additional reason for abandoning intergenerational equity:

> I therefore conclude that the strong ethical requirement that all generations be treated alike, itself reasonable, contradicts a very strong intuition that it is not morally acceptable to demand excessively high savings rates of any one generation, or even of every generation.

A very similar argument is adduced in Arrow (1999b), where he adds:

> Not merely is saving arbitrarily close to 100% unacceptable but very high sacrifices are also. I call this the *weak Koopmans argument*.

And in Arrow (2007, p. 4) he writes:

---

12  See Stern (2014b; 2015, p. 166) for some recent discussion.

> Tjalling Koopmans pointed out in effect that the savings rates implied by zero time preference are very much higher than those we observe. (I am myself convinced by this argument.)

This kind of argument, and its relation to intergenerational equity, were discussed by Asheim et al. (2001) and then Asheim and Buchholz (2003). Their main result demonstrates that, given any efficient and non-decreasing consumption allocation, there exists a utility function for which that allocation is a unique maximum of the undiscounted sum of utilities over all future generations. This relies, however, on a key "technological" assumption that any efficient and non-decreasing consumption allocation maximizes the finite present discounted value of consumption with discount factors that decrease over time. This assumption seems to rule out the kind of climate emergency that the world may be facing currently.

Later, the weak argument was discussed in Stern (2007, 2008, 2015)—see especially Stern (2008, p. 16), which cites Mirrlees and Stern (1972). In his comment on the *Stern Report*, however, Arrow (2007, p. 4) himself suggests that, at least in the context of mitigating climate change, the discounting issue may lack practical importance:

> Many have complained about the *Stern Review* adopting a value of zero for $\rho$, the social rate of time preference. However, I find that the case for intervention to keep $CO_2$ levels within bounds (say, aiming to stabilize them at about 550 ppm) is sufficiently strong as to be insensitive to the arguments about $\rho$.[13]

Thus, KA's reasoning included the recognition that, even with substantial pure-time discounting, unmanaged climate change has the potential to cause damage severe enough to justify strong action.[14]

We would agree, while noting that Stern (2015) in particular discusses how many current economic models fail to capture adequately the immense potential damage. We also note the observation from Box 1 in Stern (2008, p. 20) that, for any choice of discount rates, we can construct a stream of damages over time so that the present discounted value of losses from climate change is infinite.

On the other hand we note that, in the years since Arrow wrote in (2007), there has been extraordinary technical progress in developing renewable energy sources, as well as low and even negative carbon technologies. This progress suggests that, as discussed in Sect. 8.3, the "weak argument" in favour of discounting is steadily becoming even weaker.

---

[13] Since KA wrote this, the scientific evidence on the potential damage from unmanaged climate change has grown ever more worrying. The IPCC (2018) report documents the significant risk of potentially very large damages from allowing temperatures to rise by 2 °C instead of 1.5 °C. This supports the argument that the Paris COP21 target agreed in December 2015 of holding temperature increases to "well below 2 °C" should be tightened to 1.5 °C. Achieving this is likely to require net global $CO_2$ emissions being reduced to zero by 2050. In particular, KA's suggestion that $CO_2$ levels should be stabilized at about 550 ppm should now be seen as far too high.

[14] The later joint paper by Arrow et al. (2013) even advocates a declining discount factor, though this may be easier to justify if it is applied to monetary measures of consumer benefit when these are increasing over time.

## 3 An impartial spectator's fundamental utility

### 3.1 A universal domain of personal consequences

Our interest is in applying prescriptive social choice theory to the issue of determining what discount rates one should apply to future generations' welfare. For this application we use an individualistic theory of social consequences that starts with a "universal" *personal consequence domain*, which we take to be a non-empty set $Y$ whose typical member $y$ has many attributes or dimensions.

We postulate that each personal consequence $y \in Y$ is *comprehensive*. That is, it must include everything that should be of concern to an idealized version of the "impartial spectator" who plays such a key role in Adam Smith's *Theory of Moral Sentiments*. In particular, it must include the individual's own preferences and beliefs, insofar as they are deemed relevant. Consequences should also allow for variations in the date and circumstances surrounding an individual's birth, upbringing, and death, including those aspects that also affect parents, partners and family members, especially "demographic consequences" of the kind considered in Hammond (1998). Also included should be any adverse consequences of living in an unequal society, as well as any favourable consequences that are often modelled as altruism. Other relevant dimensions can deal with society's respect for individual and group rights, as discussed in Hammond (1995), as well as "deontological" or agent-relative consequences reflecting how well individuals respect their social obligations.[15]

Of course, when considering intergenerational equity and discounting the future, we also need to allow for intertemporal, even intergenerational consequences. These imply that, in effect, one has overlapping generations. Thus, with one exception, we assume that each consequence $y \in Y$ has attributes which include personal copies of any ethically relevant common or impersonal circumstances that are shared with other persons.[16]

The one exception, which we use repeatedly from Sect. 5 on when discussing extinction, is that we postulate one particular *non-existence* consequence $y_0 \in Y$. This is the unique personal consequence that comes about if and only if the person concerned never exists. It corresponds to the consequence denoted by $d$ in Bommier and Zuber (2008). Note in particular that our framework allows for a variable number of individuals, since those who are excluded and so whose personal consequences are not deemed relevant can be modelled as experiencing the non-existence consequence $y_0$. This device will be especially useful in Sect. 7 when we discuss variable population numbers.

---

[15] This paragraph has been inspired in part by the issues that Patrick Suppes kindly raised when discussing PH's presentation of the paper that became Hammond (1988) to the May 1986 conference on "Distributive Justice and Inequality" at the Wissenschaftskolleg zu Berlin. At the time, the most meaningful part of PH's inadequate response may have been the remark that philosophers like Pat excel at drawing attention to the ethical relevance of consequences that belong to domains much richer than those usually considered by economists. In some ways, our oral discussion foreshadowed the kind of procedure that, following Portmore (2007, 2009), Brown (2011), and Mukerji (2016), philosophers now describe as "consequentialization".

[16] We note that our approach is very different from that of Feng and Ke (2018), whose individuals live for only one period, but care about the future because of altruism.

## 3.2 Fundamental cardinal utility

A standard economist's view is that ethical theory should prescribe normatively appropriate decisions for society. Following Hammond (1996, 1998) in particular, we impose consequentialist rationality and continuity postulates implying that there exists a unique cardinal equivalence class of von Neumann–Morgenstern utility functions $Y \ni y \mapsto u(y) \in \mathbb{R}$ with the property that the consequences of prescribed behaviour in any finite decision tree should be a consequence lottery $\lambda \in \Delta(Y)$ that maximizes lifetime expected utility

$$\mathbb{E}_\lambda u = \sum_{y \in Y} \lambda(y)\, u(y) \tag{1}$$

over the finite set of lifetime consequence lotteries that are feasible in the tree. Here we are using the familiar definition that any two utility functions $y \mapsto u(y)$ and $y \mapsto \tilde{u}(y)$ are *cardinally equivalent* just in case there exist an additive constant $\alpha \in \mathbb{R}$ and a *positive* multiplicative constant $\rho \in \mathbb{R}$ such that

$$\tilde{u}(y) = \alpha + \rho\, u(y) \ \text{ for all } y \in Y \tag{2}$$

Any function $Y \ni y \mapsto u(y)$ in this cardinal equivalence class can be called a *fundamental utility function*. This is because the preferences it represents are the same for all individuals, both potential and actual, and so "fundamental" in the sense considered by Tinbergen (1957) and Kolm (1972, 1994).

## 3.3 General discrete lotteries and bounded utility

Menger (1934) showed how to modify the well-known St. Petersburg paradox so that it applies to any unbounded utility function. This result led Arrow (1951; 1965, pp. 28–44; 1971, ch. 2; 1972) to insist that utility should be bounded. Indeed, suppose that expected utility is to be extended to a continuous function defined not only over simple lotteries whose support is a finite set of possible outcomes, but also over general discrete lotteries whose support is a countably infinite set of possible outcomes, including extinction dates. Then analysis such as that in Hammond (1998, Sect 8), which uses ideas from Blackwell and Girshick (1954), shows that any member of the unique cardinal equivalence class of fundamental utility functions must be bounded. Accordingly, from now on we make the following assumption:

**Assumption 1** There exist both a common lower bound $\underline{u}$ and a common upper bound $\bar{u} > \underline{u}$ such that, for all possible consequences $y \in Y$ that any potential individual may face, one has

$$\underline{u} \leq u(y) \leq \bar{u} \tag{3}$$

### 3.4 A unique normalized fundamental utility function

The extensive literature on population ethics stemming from Sikora and Barry (1978) and from Parfit (1984) commonly designates $u(y_0)$, the utility of non-existence, as the *neutral level* of utility. Following Hammond (1988) and many other works, we invoke what Blackorby et al. (1995, 2005) refer to as the "zero-critical level" principle, which requires the normalization

$$u(y_0) = 0 \tag{4}$$

Obviously, this first normalization is especially convenient when, as in Sect. 7, we need to consider a possibly infinite set of potential persons, of whom a variable finite number come into existence.

This reduces the class of cardinally equivalent utility functions satisfying (4) to those related by the restricted class of transformations that satisfy (2) for $\alpha = 0$—i.e, transformations of the form

$$\tilde{u}(y) = \rho\, u(y) \text{ for all } y \in Y \tag{5}$$

To single out a unique utility function from the equivalence class of fundamental utility functions related by (5), we impose a second normalization. After ruling out the *trivial case* when $u(y) \equiv 0$ throughout $y$, the particular normalization we adopt is

$$\sup_{y \in Y} |u(y)| = 1 \tag{6}$$

## 4 Utilitarianism for an impartial benefactor

### 4.1 Social consequences as personal consequence profiles

Initially, we consider a fixed finite set $I$ of individuals, of size $n = \#I$. Later, starting in Sect. 5, we will extend our analysis to potentially infinite populations. The latter is of course the case that was of most concern to Arrow (1999a, b) when arguing that future utilities should be discounted.

For now, however, given the fixed set of individuals $I$ of size $n$, we take the *social consequence domain* to be the Cartesian product $Y^I$ of $n$ copies of the personal consequence domain $Y$. So each social consequence $y^I \in Y^I$ is a mapping $I \ni i \mapsto y_i \in Y$ that determines a personal consequence for each of the $n$ individuals in $I$. Equivalently, it is a list

$$y^I = \langle y_i \rangle_{i \in I} \in Y^I \tag{7}$$

## 4.2 Biased and extended original positions

Vickrey (1945, 1960) and Harsanyi (1953, 1955, 1977, 1978, 1979) independently formulated the idea that ethical social decisions would be those that were taken impartially in a version of what Rawls (1971) later described as an "original position", behind a "veil of ignorance" where the decision-maker does not know which person she or he will become eventually. In contrast to Rawls, what we will call the *Vickrey–Harsanyi original position* requires the impartial benefactor to contemplate what to choose if faced with an even chance lottery whose different outcomes are the personal consequences of the various individuals in society—see also the discussion by Mongin (2001) and others of Harsanyi's "impartial observer".

The theory presented here will also accommodate lotteries in the form of *biased original positions* where, upon emerging from behind the probabilistic version of the veil of ignorance, an arbitrary specified probability distribution $\mu \in \Delta(I)$ determines the probability $\mu_i$ of becoming each person $i \in I$. Indeed, we consider *extended original positions* where these biased probabilities of becoming different people can even be chosen.

## 4.3 Expected utility from a biased original position

Thus, we consider decision problems which reduce to choosing from consequences that take the form of a lottery pair $(\lambda, \mu) \in \Delta(Y^I) \times \Delta(I)$ over possible pairs $(y^I, i) \in Y^I \times I$. So, for a suitable utility function $Y^I \times I \ni (y^I, i) \mapsto u^*(y^I, i) \in \mathbb{R}$, expected utility is given by

$$\sum_{y^I \in Y^I} \sum_{i \in I} \lambda(y^I) \, \mu_i \, u^*(y^I, i) \tag{8}$$

The purpose of an original position, however, is to reduce each consequence profile $y^I \in Y^I$ to a single consequence $y_i \in I$ where $i \in I$ is chosen at random. This makes it natural to specify, for each $i \in I$, that $u^*(y^I, i)$ can be reduced to $u(y_i)$.

Consider now, for each joint distribution $\lambda \in \Delta(Y^I)$ and each $i \in I$, the marginal consequence lottery $\text{marg}_i \lambda \in \Delta(Y)$ over personal consequences which, for each $y \in Y$, has probabilities given by

$$\lambda_i(y) := \text{marg}_i \lambda(y) := \lambda(\{y^I \in Y^I \mid y_i = y\}) \tag{9}$$

Then define the compound lottery $\mu \circ \lambda \in \Delta(Y \times I)$ so that, for every $(y, i)$ in the domain $Y \times I$, one has

$$(\mu \circ \lambda)(y, i) := \mu_i \cdot \text{marg}_i \lambda(y) = \mu_i \cdot \lambda_i(y) \tag{10}$$

Along with the assumption that $u^*(y^I, i)$ can be reduced to $u(y_i)$, definitions (9) and (10) evidently imply that the expected value (8) can be written as

$$\mathbb{E}_{\mu \circ \lambda}[u(y)] = \mathbb{E}_\mu \left[ \mathbb{E}_{\lambda_i}[u(y_i)] \right] = \sum_{i \in I} \mu_i \sum_{y_i \in Y} \lambda_i(y_i) \, u(y_i) \tag{11}$$

This is the expectation w.r.t. $\mu$ of different individuals' expected utilities arising from the relevant personal consequence lotteries $\lambda_i$.

### 4.4 An original position with suppes equity

No version of Adam Smith's impartial spectator has the power to choose what individual it should represent upon emerging from behind the veil of ignorance. Thus, it must treat the biased original position $\mu \in \Delta(I)$ as fixed. Then it is reduced to choosing $\lambda \in \Delta(Y^I)$ in order to maximize the function (11) while treating $\mu \in \Delta(I)$ as a fixed vector of probabilistic weights. So the appropriate objective is the weighted utilitarian *Bergson social welfare function* (or BSWF) defined by

$$\Delta(Y^I) \ni \lambda \mapsto W(\lambda; \mu) := \mathbb{E}_{\mu \circ \lambda}[u(y)] = \sum_{i \in I} \mu_i \sum_{y_i \in Y} \lambda_i(y_i) u(y_i) \quad (12)$$

We adapt definition 5 on p. 296 of Suppes (1966) to our context of a weighted utilitarian Bergson social welfare function (12)—see also Sen (1970, Chapter 9*). Specifically, we define a *two-person decision situation* as a pair of lotteries $\nu$, $\rho \in \Delta(Y^I)$ for which there exist two individuals $j, k \in I$ such that for all other individuals $i \in I \setminus \{j, k\}$, the marginal distributions $\nu_i, \rho_i \in \Delta(Y)$ satisfy $\nu_i = \rho_i$. In this case Eq. (12) implies that the welfare difference between $\nu$ and $\rho$ is

$$W(\nu; \mu) - W(\rho; \mu) = \mu_j \sum_{y_j \in Y} [\nu_j(y_j) - \rho_j(y_j)] u(y_j)$$
$$+ \mu_k \sum_{y_k \in Y} [\nu_k(y_k) - \rho_k(y_k)] u(y_k) \quad (13)$$

Now, in this two-person situation, *Suppes equity* insists that interchanging the pairs of marginal lotteries $\nu_j, \nu_k$ and $\rho_j, \rho_k$ of these two individuals should have no effect on the social preference between $\nu$ and $\rho$ in $\Delta(Y^I)$. In other words, we should consider the two new lotteries $\tilde{\nu}, \tilde{\rho} \in \Delta(Y^I)$ whose marginal distributions $\tilde{\nu}_i, \tilde{\rho}_i \in \Delta(Y)$ satisfy $\tilde{\nu}_i = \nu_{\tau^{j,k}(i)}$ and $\tilde{\rho}_i = \rho_{\tau^{j,k}(i)}$ for all $i \in I$, where $I \ni i \mapsto \tau^{j,k}(i) \in I$ is the transposition mapping that interchanges individuals $j$ and $k$ while leaving all other individuals unaffected. Then Suppes equity requires that

$$W(\tilde{\nu}; \mu) \gtreqless W(\tilde{\rho}; \mu) \text{ according as } W(\nu; \mu) \gtreqless W(\rho; \mu) \quad (14)$$

Evidently (13) is consistent with (14) for all pairs $\nu, \rho \in \Delta(Y^I)$ if and only if $\mu_j = \mu_k$. Given that $\mu \in \Delta(I)$ is a probability distribution and $n = \#I$, evidently Suppes equity holds for all pairs of individuals $j, k \in I$ if and only if the weights satisfy $\mu_i = \frac{1}{n}$ for all $i \in I$. So instead of the weighted sum (12), Suppes equity implies that we should have an *unweighted* utilitarian BSWF of the form

$$\Delta(Y^I) \ni \lambda \mapsto W(\lambda) := \frac{1}{n} \sum_{i \in I} \sum_{y_i \in Y} \lambda_i(y_i) u(y_i) \quad (15)$$

This, of course, is precisely the form that Vickrey (1945, 1960) and Harsanyi (1953, 1955, 1977, 1978, 1979) advocated, taking the view that ethical decisions are those that would be made by an impartial benefactor who, by definition, acts as if facing an original position in which there is an equal probability of becoming any of the $n$ individuals $i \in I$.

## 5 Extinction discounting of future generations

### 5.1 Generational structures

Inspired by Arrow (1999a, b, 2007) as well as Dasgupta and Heal (1979, ch. 9), our concern is whether one should discount the utilities of individuals who belong to future generations. To discuss this formally, we need an extended framework which recognizes that different individuals can belong to different generations. So we label each individual $i \in I$ by a pair $(t_i, k_i) \in T \times \mathbb{N}$. Here $T$ denotes the set of possible discrete times or dates, which is taken to be a copy of $\mathbb{N}$, the countably infinite set of natural numbers.

Corresponding to each time $t \in T$, *generation $t$* is defined as the enumerated finite set $G_t$ of all individuals born at date $t$. This set can be identified with the Cartesian product set $\{t\} \times \mathbb{N}_{n_t}$ of individuals whose first label is the date $t$, and whose second label is a number that ranges over the set $\mathbb{N}_{n_t}$ consisting of the first $n_t$ natural numbers. Evidently, $n_t = \#G_t$, and the set of all individuals who are ever born is the countably infinite set

$$I = \cup_{t \in T} (\{t\} \times \mathbb{N}_{n_t}) = \cup_{t \in T} (\{t\} \times G_t) \subset \mathbb{N} \times \mathbb{N} \tag{16}$$

The fact that $I$ is infinite creates the obvious difficulty that there is no way for an ethical benefactor to be impartial by acting as if there were an equal probability of becoming each individual $i \in I$.

### 5.2 A hazard process of extinction and survival

Following Dasgupta and Heal (1979, ch. 9), our argument for extinction discounting relies on treating extinction as a stochastic process. We do, however, focus on unavoidable background extinction hazards, as opposed to new hazards that might arise because of foolish policy. So, for each time $t \in T$, let $E_{>t}$ and $E_{\leq t}$ respectively denote the events that the background process does not result in extinction until after time $t$, and the complementary event that extinction will occur at or before time $t$. Let $\sigma_t \in [0, 1]$ denote the *survival probability* of event $E_{>t}$, and $\eta_t := 1 - \sigma_t$ as the probability of event $E_{\leq t}$.

The *hazard rate* associated with the extinction process determines the conditional probability that extinction occurs at time $t$ exactly given the event $E_{>t-1}$ that it had not already occurred by time $t - 1$. Thus, it is the function given by

$$T \ni t \mapsto h_t := \frac{\eta_t - \eta_{t-1}}{\sigma_{t-1}} = \frac{\sigma_{t-1} - \sigma_t}{\sigma_{t-1}} = 1 - \frac{\sigma_t}{\sigma_{t-1}} \in [0, 1] \qquad (17)$$

We assume that neither immediate extinction nor continued existence can ever be predicted with probability 1, implying that $0 < h_t < 1$ for all $t \in T$. Evidently (17) then implies that the functions $t \mapsto \sigma_t$ and $t \mapsto \eta_t$ are respectively strictly decreasing and strictly increasing in $t$. We also assume that, in the limit as $t \to \infty$, the extinction and survival probabilities satisfy

$$\eta_t \to 1 \quad \text{and} \quad \sigma_t \to 0 \qquad (18)$$

### 5.3 Extinction adjusted intergenerational equity

In Sect. 4.2 we represented an original position by a probability distribution $\mu \in \Delta(I)$, where $I$ was finite. Now we consider the case introduced in Sect. 5.1 when $I$ is countably infinite. Instead of $\mu \in \Delta(I)$, we require it to belong to $\Delta^*(I)$, defined as the set of *discrete distributions* over $I$. That is, the distribution $\mu$ takes the form of a mapping $I \ni i \mapsto \mu_i \in [0, 1]$ for which the countably infinite sum $\sum_{i \in I} \mu_i$ equals 1.

In this new setting, given any original position $\mu \in \Delta^*(I)$, any date $t \in T$, and any individual $i \in \{t\} \times G_t$, define the *extinction adjusted probability* $\nu_i := \mu_i/\sigma_t$ as the *conditional* probability of selecting $i$ given the event $E_{>t}$. Previously, when $I$ was finite with $n = \#I$, intergenerational equity required that $\mu_i = \bar{\mu} = 1/n$ for all $i \in I$. In this new setting when $I$ is infinite but there is a background extinction process, we say that the original position $\mu \in \Delta^*(I)$ satisfies *extinction adjusted intergenerational equity* just in case the extinction adjusted probabilities satisfy $\nu_i = \bar{\nu}$ for all $i \in I$, implying that $\mu_i = \bar{\nu} \sigma_t$ for all $i \in \{t\} \times G_t$. For these to be probabilities in $\Delta^*(I)$, evidently they must satisfy

$$1 = \sum_{i \in I} \mu_i = \sum_{t \in T} \sum_{i \in \{t\} \times G_t} \bar{\nu} \sigma_t = \bar{\nu} \sum_{t \in T} \sigma_t n_t \qquad (19)$$

Given that, for each $t \in T$, the population of generation $t$ is $n_t = \#G_t$, let us now define the *extinction discounted total population* as the infinite sum

$$n(G^T) := \sum_{t \in T} \sigma_t n_t \qquad (20)$$

of non-negative terms, provided this sum converges to a finite number. But (20) converges if and only if there exists $\bar{\nu} > 0$ that solves Eq. (19). So the following result follows immediately:

**Proposition 1** *There exists a unique original position $\mu \in \Delta^*(I)$ satisfying extinction adjusted intergenerational equity if and only if the discounted expected total population given by (20) is finite.*

### 5.4 Utilitarianism with an infinite population

Consider any individual $i$ who belongs to the generation $G_t$ that comes into existence starting at the specific date $t \in T$. In this new setting with a background extinction process, we reinterpret the marginal personal consequence lottery $\lambda_i \in \Delta(Y)$ that faces any individual $i \in I$ as the *conditional* lottery that $i$ would experience given the event $E_{>t}$ that extinction has not occurred by date $t$.

With this interpretation, the weighted utilitarian BSWF can be defined by (12) even in our setting with an infinite set $I$ provided that one has a well-defined infinite sum

$$\sum\nolimits_{i \in I} \mu_i \sum\nolimits_{y_i \in Y} \lambda_i(y_i)\, u(y_i) = \sum\nolimits_{i \in I} \mu_i\, \mathbb{E}_{\lambda_i} u \qquad (21)$$

By definition, any original position $\mu \in \Delta^*(I)$ satisfies $\mu_i \geq 0$ for all $I \in I$ and $\sum_{i \in I} \mu_i = 1$. This makes the following result entirely evident:

**Proposition 2** *Under the bounded utility assumption 1 in Sect. 3.3, the utilitarian sum (21) converges absolutely for all original positions $\mu \in \Delta^*(I)$ and for all profiles $\langle \lambda_i \rangle_{i \in I}$ of personal consequence lotteries $\lambda_i \in \Delta(Y)$.*

### 5.5 Extinction discounted utilitarianism

Say that the objective (21) is *extinction discounted utilitarian* just in case the original position $\mu \in \Delta^*(I)$ satisfies extinction adjusted intergenerational equity. Putting Propositions 1 and 2 together leads to our main result:

**Theorem 1** *Given the fixed generation structure $G^T$, under the normalization imposed in Sect. 3.4, there is a uniquely specified extinction discounted utilitarian welfare objective represented by the expected value of*

$$Y^I \ni y^I \mapsto W(y^I) = \frac{1}{n(G^T)} \sum\nolimits_{t \in T} \sigma_t \sum\nolimits_{i \in G_t} u(y_i) \qquad (22)$$

*if and only if $n(G^T)$ defined by (20) is finite.*

## 6 Sustainable preferences

### 6.1 Sustainability

In Brundtland (1987), sustainable development was famously defined as:

> …development that meets the needs of the present without compromising the ability of future generations to meet their own needs.

Decades before, Hicks (1946, p. 174) had a similar idea when he defined an individual's "income" as

> …the maximum amount of money which the individual can spend this week, and still expect to be able to spend the same amount in real terms in each ensuing week.

In this spirit, and following Solow (1991, 2012), sustainability might be defined as giving each generation access to an opportunity set that allows it to be no worse off than it would have been with the opportunity set that was available to any of its predecessors. This suggests trying to maximize the initial generation's welfare level subject to *monotone sustainability*—i.e., requiring successive generations' welfare levels to be non-decreasing over time.[17]

In 1993 KA invited GC to present work related to the concept of sustainable development to a Stanford workshop on "The Reconsideration of Values", in which PH also participated. This led to the articles Chichilnisky (1996, 1997, 2009) which went beyond the extensively discussed criterion of "sustainable development" and pioneered the new concept of "sustainable preferences" or "sustainable welfare criterion".

In the framework of generation structures set out in Sect. 5.1, it is natural to consider sustainable preferences over lotteries with infinite-dimensional personal consequence profiles $y^I \in Y^I$ as outcomes. Moreover, the relevant welfare criterion should be the expected value of some uniformly bounded von Neumann–Morgenstern social utility function

$$Y^I \supseteq \mathcal{D} \ni y^I \mapsto V(y^I) \in \mathbb{R} \tag{23}$$

defined on a possibly restricted domain $\mathcal{D}$ of personal consequence streams $y^I$ that belong to the countably infinite Cartesian product set $Y^I$. Moreover, the function $y^I \mapsto V(y^I)$ should satisfy the Pareto criterion.

## 6.2 Dictatorship of the present or the future

Following Chichilnisky (1996, pp. 240–241), a "dictatorship of the present" is a welfare criterion which, after some generation that depends on the choices at hand, is insensitive to the welfare of all succeeding generations. In other words, a dictatorship of the present occurs if a strict preference for one personal consequence stream $y^I$ over an alternative stream $\tilde{y}^I$ can never be overturned by any changes in these two streams that affect only sufficiently distant generations.

By contrast, a "dictatorship of the future" is insensitive to the welfare of the present, disregarding the welfare of all generations that precede some generation. In other words, a dictatorship of the future occurs if a strict preference for one personal consequence stream $y^I$ over an alternative stream $\tilde{y}^I$ can never be overturned by any changes in these two streams that affect only generations sufficiently close to the present.

Using alternative terminology suggested by Heal (1998, p. 69), a criterion displaying dictatorship of the present is *insensitive to the long-run future*; whereas one displaying dictatorship of the future is *insensitive to the present*. Sustainability allows

---

[17] For a fuller discussion of sustainability in this sense, see *inter alia* Hammond (1993) as well as Pezzey (1997, p. 451) and Asheim (2007). See also Llavador et al. (2015).

neither form of temporal dictatorship. Somewhat surprisingly, it is relatively easy to find welfare criteria that are sustainable in this sense.

## 6.3 Formal definitions

The following formal definitions adapt those of Chichilnisky ([1996]) to the current setting where: (i) there is a fixed generational structure $G^T$, as defined in Sect. [5.1]; (ii) the social objective is the expected value of a von Neumann–Morgenstern social utility function as in ([23]).

**Definition 1** Given a generational structure $G^T$, a welfare criterion $V : \mathcal{D} \to \mathbb{R}$ on the domain $\mathcal{D} \subseteq Y^I$ is a *dictatorship of the present* if for all pairs $y^I, \tilde{y}^I \in \mathcal{D}$ with $V(y^I) > V(\tilde{y}^I)$, there exists a date $s = s(y^I, \tilde{y}^I) \in T$ such that if any pair $z^I, \tilde{z}^I \in Y^I$ satisfies

$$i \in \cup_{t=1}^{s} G_t \implies z_i = y_i \quad \text{and} \quad \tilde{z}_i = \tilde{y}_i \tag{24}$$

then $z^I, \tilde{z}^I \in \mathcal{D}$ and $V(z^I) > V(\tilde{z}^I)$.

**Definition 2** Given a generational structure $G^T$, a welfare criterion $V : \mathcal{D} \to \mathbb{R}$ on the domain $\mathcal{D} \subseteq Y^I$ is a *dictatorship of the future* if for all pairs $y^I, \tilde{y}^I \in \mathcal{D}$ with $V(y^I) > V(\tilde{y}^I)$, there exists a date $s = s(y^I, \tilde{y}^I) \in \mathbb{N}$ such that if the pair $z^I, \tilde{z}^I \in Y^I$ satisfies

$$i \in \cup_{t=s}^{\infty} G_t \implies z_i = y_i \quad \text{and} \quad \tilde{z}_i = \tilde{y}_i \tag{25}$$

then $z^I, \tilde{z}^I \in \mathcal{D}$ and $V(z^I) > V(\tilde{z}^I)$.

Note that the only difference between these two definitions concerns the conditions ([24]) and ([25]). These determine whether the strict preference between two personal consequence streams remains unchanged after alterations in the consequence streams only for all generations that originate: either (i), in the case of Definition [1], after date $s$, so sufficiently far into the future; or (ii), in the case of Definition [2], before date $s$, so sufficiently close to the present;

The last definition of this Section also adapts that of Chichilnisky ([1996]) to the current setting where, with fixed generation structure $G^T$, the social objective is the expected value of a von Neumann–Morgenstern social welfare function defined on a restricted domain $\mathcal{D} \subset Y^I$.

**Definition 3** Given a restricted domain $\mathcal{D} \subset Y^I$ of admissible personal consequence streams, a *sustainable* von Neumann–Morgenstern social utility function is a mapping $\mathcal{D} \ni y^I \mapsto V(y^I) \in \mathbb{R}$ which satisfies the Pareto criterion, and is neither a dictatorship of the present, nor a dictatorship of the future.

## 6.4 Example of sustainable preferences

The following example offers a two-dimensional parametric class of sustainable preferences.

**Example 1** Given the fixed generation structure $G^T$, consider any fixed stream of discount factors $\beta^T \in \mathbb{R}_{++}^T$ for which total discounted population $\sum_{t \in T} \beta_t \#G_t$ is finite. Suppose the boundedness assumption 1 of Sect. 3.3 is satisfied. Then, for any pair of parameters $\alpha \in [0, 1]$ and $\omega \in (0, 1)$, consider the von Neumann–Morgenstern welfare function

$$Y^I \ni y^I \mapsto \Psi(y^I; \beta^T, \alpha, \omega) := (1 - \omega) \sum_{t \in T} \beta_t \sum_{i \in G_t(y^I)} u(y_i)$$
$$+ \omega \left[ \alpha \liminf_{i \to \infty} u(y_i) + (1 - \alpha) \limsup_{i \to \infty} u(y_i) \right] \tag{26}$$

This puts positive weight $1 - \omega$ on the first term, which is a dictatorship of the present, and positive weight $\omega$ on the second term, which is a dictatorship of the future. The strict convex combination of these two defines sustainable preferences.

For a possible interpretation of (26), suppose we relax condition (18) so that, for some $\omega \in (0, 1)$, it becomes $\eta_t \to 1 - \omega$ and $\sigma_t \to \omega$. Suppose too that for each $t \in T$ we interpret $\beta_t$ as the conditional probability of generation $G_t$ coming into existence given that extinction does occur at some finite time. Then (26) becomes expected utility when:

1. with probability $1 - \omega$, extinction does occur in finite time, with survival probabilities specified by $\beta_t$;
2. with probability $\omega$, extinction never occurs, and the utility of the resulting infinite consequence stream is given by the second term of (26).

Note too that the second term of (26) is an asymptotic form of the "Hurwicz criterion" for decisions under uncertainty that was discussed, *inter alia*, in Arrow and Hurwicz (1972). Here the parameter $\alpha$ can be regarded as a "coefficient of pessimism".

## 6.5 Other sustainable preferences

Going beyond the class of preferences set out in Example 1, Chichilnisky (1996) characterizes sustainable preferences for the important special case when the preference ordering over sure consequence streams is represented by a utilitarian social welfare function defined on the Banach space $\ell^\infty$ of all bounded utility streams in $\mathbb{R}^\infty$ that is not only strictly increasing and continuous, but is also linear. For this important special case, Theorem 2 of Chichilnisky (1996) offers a complete characterization of all sustainable welfare criteria using finitely additive measures that represent linear functionals in the dual of $\ell^\infty$. See Chichilnisky (1996, 1997, 2009), Heal (1998), and Lauwers (2017) for further discussion.

Among other work that discusses sustainable preferences, especially in the context of renewable and exhaustible resources, we mention Figuières and Tidball (2012). This builds on work by Chichilnisky et al. (1995) that considers "the green golden rule".

# 7 Endogenous population

## 7.1 A domain with endogenous generation structures

The framework of our earlier working paper Chichilnisky et al. (2018) allowed the generation structure $G^T$ to be chosen. In this extended framework, since the set of individuals $I$ is variable, the concept of an original position $\mu \in \Delta^*(I)$ needs to be changed.

First, we replace the fixed generation structure $G^T = \langle G_t \rangle_{t \in T}$ of Sect. 5.1, with $G_t = \mathbb{N}_{n_t}$ for each $t \in T$ and $I = \cup_{t \in T}(\{t\} \times G_t)$, by:

1. an extended generation structure $M^T = \langle M_t \rangle_{t \in T}$ of finite *maximal population sets* for each generation, with $M_t = \mathbb{N}_{m_t}$ for each $t \in T$;
2. an overall maximal set $M := \cup_{t \in T}(\{t\} \times M_t)$.

Given the structure $M^T$, we limit attention to a restricted domain $\mathcal{D}(M^T)$ of sequences $(G^T, y^I)$ that satisfy:

1. $G_t \subseteq M_t$ for all $t \in T$;
2. $y_i \in Y \setminus \{y_0\}$ for all $i \in I = \cup_{t \in T} G_t$.

Thus, the personal consequence of each $i \in G_t$ excludes non-existence.

## 7.2 An extended original position

We can now extend this domain $\mathcal{D}(M^T)$ to a new domain $\hat{\mathcal{D}}(M^T)$ that includes any personal consequence profile $y^M \in Y^M$ for the fixed maximal set of potential persons. To do this, simply recognize that all individuals $i \in \cup_{t \in T}(M_t \setminus G_t)$ do not exist by defining

$$y_i = y_0 \text{ for all } i \in \cup_{t \in T}(M_t \setminus G_t) \tag{27}$$

This brings us back to our previous framework with the fixed generation structure $M^T$ replacing $G^T$.

If an original position $\nu \in \Delta^*(M)$ exists that satisfies the appropriate intergenerational equity condition given extinction, there must exist a constant $\kappa > 0$ such that $\nu_i = \kappa \, \sigma_t$ for all $t \in T$ and $i \in M_t$. As in the derivation of (19) in Sect. 5.3, a necessary and sufficient condition for such an original position to exist is that the extinction discounted sum of the maximum population number in each generation given by

$$m(M^T) := \sum_{t \in T} \sigma_t \, \#M_t = \sum_{t \in T} \sigma_t \, m_t \tag{28}$$

is finite. When it is finite, the obvious counterpart of (22) is the NM welfare function

$$Y^M \ni y^M \mapsto W^M(M^T, y^M) := \frac{1}{m(M^T)} \sum_{t \in T} \sigma_t \sum_{i \in M_t} u(y_i) \tag{29}$$

Because (27) implies that $u(y_i) = 0$ for all $i \in \cup_{t \in T}[\{t\} \times (M_t \setminus G_t)]$, then provided that $m(M^T)$ defined by (28) is finite, the extinction discounted welfare sum (29) can obviously be reduced to

$$W(G^T, y^I) := \frac{1}{m(M^T)} \sum_{t \in T} \sigma_t \sum_{i \in G_t} u(y_i) \tag{30}$$

That is, we exclude the zero utilities of all the individuals whose existence is precluded by (27). The result is a welfare function defined on the restricted domain $\mathcal{D}(M^T)$ of pairs $(G^T, y^I)$ satisfying restrictions 1 and 2 above.

## 7.3 Total expected utility

The only role that the extended generation structure $M^T$ and, when it is well-defined, the associated constant $m(M^T)$ play in (30) is in allowing $W(G^T, y^I)$ to be interpreted as expected welfare for an impartial benefactor facing an original position with extinction discounted intergenerational equity. This interpretation relies only on the maximal generational structure $M^T$ and implied extinction discounted total population $m(M^T)$ being large enough to include all relevant generational structures $G^T$ that could result from policy choices. This allows us to treat $M^T$ and $m(M^T)$ as exogenous provided we restrict attention to the domain $\hat{\mathcal{D}}(M^T)$—or equivalently, after removing non-existent individuals, to the domain $\mathcal{D}(M^T)$.

Once we do treat $M^T$ and $m(M^T)$ as exogenous, and so fixed constants, maximizing the expectation of (30) is obviously equivalent to maximizing extinction discounting *total* utility, defined by

$$W^{\text{total}}(G^T, y^I) := \sum_{t \in T} \sigma_t \sum_{i \in G_t} u(y_i) \tag{31}$$

This definition makes the extended generation structure $M^T$ and, when it is well-defined, the associated constant $m(M^T)$ entirely irrelevant, except insofar as one is able to construct them throughout the domain of the function $\mathcal{D} \ni (G^T, y^I) \mapsto W^{\text{total}}(G^T, y^I)$.

## 7.4 The domain for total expected utility

We now construct the domain $\mathcal{D}$ over which, for some large enough set of potential individuals, the total utilitarian objective (31) can be interpreted as rescaled expected utility of an impartial benefactor in a suitable original position with intergenerational equity and extinction discounting. To do so, first define for each non-negative real $r \in \mathbb{R}_+$:

1. the set

$$\mathcal{G}_r := \{G^T \mid \sum_{t \in T} \sigma_t \, \#G_t \leq r\}$$

of generation structures for which the extinction discounted total population does not exceed $r$;

2. the associated domain

$$\mathcal{D}_r := \{(G^T, y^I) \in \mathcal{G}_r \times y^I \mid \forall i \in \cup_{t \in T} G_t : y_i \neq y_0\}$$

Then the domain $\mathcal{D}$ we are seeking is evidently the infinite union $\mathcal{D} := \cup_{r \in \mathbb{R}_+} \mathcal{D}_r$. This, of course, is the domain of associated pairs $(G^T, y^I)$ over which the extinction discounted total population $\sum_{t \in T} \sigma_t \#G_t$ is not only finite, but uniformly bounded. Indeed, because $W^{\text{total}}(G^T, y^I)$ is uniformly bounded over this domain $\mathcal{D}$, its expected value represents the appropriate preference ordering over the space $\Delta^*(\mathcal{D})$ of discrete lotteries over this domain.

Let $\mathcal{G}$ denote the domain of all generation structures $G^T$ for which $\sum_{t \in T} \sigma_t \#G_t$ is finite. We remark that then (31) remains valid as a definition of a BSWF even over the extended domain

$$\hat{\mathcal{D}} := (G^T, y^I) \in \mathcal{G} \times y^I \mid \forall i \in \cup_{t \in T} G_t : y_i \neq y_0\}$$

Thus, the extinction discounted total population $\sum_{t \in T} \sigma_t \#G_t$ need not be uniformly bounded if one does not require there to be one extinction discounted original position that works throughout the domain $\hat{\mathcal{D}}$ of pairs $(G^T, y^I)$.

# 8 Concluding discussion

## 8.1 Revisiting intergenerational equity

In Sects. 3 and 4 we reviewed the ethical arguments we favour for choosing policies whose consequences maximize the expected utility of an impartial benefactor. In particular, this benefactor should act as if placed in a Vickrey–Harsanyi original position with an equal probability of becoming any individual, as the Suppes equity criterion requires. When confronted, however, with long-term consequences such as those connected to climate change, then as argued in Hammond (1973), since any finite horizon will eventually be reached, consistent planning requires us to allow an indefinite future. With a potential infinite set of individuals as well as infinite time, we are confronted with what Arrow (1999a, b) called the "strong argument" for discounting the welfare of future generations. This is that, given other standard assumptions, avoiding discounting would produce some form of logical contradiction.

To circumvent this argument, we first dispose of a third source of possible divergence in expected utility sums—unbounded utility. Specifically, we follow Arrow (1951, 1965, 1971, 1972) himself and assume that the common fundamental utility function of each potential individual is bounded both above and below. Then we circumvent the problem of infinitely many time periods by first recognizing an exogenous background extinction process ensuring that with probability 1 all life will cease in finite time. Then we modify intergenerational equity by considering an extinction adjusted original position in which individuals in different generations get chosen

with equal *conditional* probabilities given that their generation survives the extinction process. We show that this adjusted original position exists if and only if, given the fixed background extinction process, the expected discounted population is uniformly bounded by some finite number. When this condition is satisfied, the assumption of bounded utility ensures that the impartial benefactor's expected utility sum converges absolutely.

Our great friend KA had espoused both strong and weak arguments in favour of discounting the welfare of future generations not only in writings such as Arrow (1999a, b, 2007), but also in later oral discussions with PH and others. Yet his willingness to engage in these discussions suggests that, in the end, he may have been starting to experience some doubt about whether these two arguments, which he attributed to Koopmans, really had settled the discounting issue. Perhaps he was moving rapidly toward our own current view that the purely technical difficulties captured by the strong argument should not be allowed to overwhelm pragmatic ethical arguments which recognize the existence of a background extinction process, and allow extinction discounting.

## 8.2 Beyond welfarism

As in most of public economics, including work by KA himself, the approach set out in this paper is utilitarian. It is *a fortiori* welfarist as well as consequentialist, though with an all-encompassing domain of personal consequences, as explained in Sect. 3.1. Nevertheless, it is just one way of looking at the ethics of public and private decision-making. Indeed, several other approaches have commanded the attention of philosophers when they undertake ethical arguments. Among these others are: (i) contractarianism; (ii) Kantian ethics; (iii) Aristotelian ethics; (iv) arguments based on rights or liberty. This is not the place for a detailed discussion, some elements of which are provided in Stern (2014a; 2015, Chapter 6), along with references to further discussion of these different ethical perspectives.

Here we simply note that these four perspectives that go beyond welfarism would all seem to exclude discounting the welfare of individuals who belong to future generations. It follows that all urge strong action on climate change. Specifically, a contractarian would likely regard a reasonable or acceptable social contract as excluding decisions such as those resulting from an unacceptable dictatorship of the present which rides roughshod over future lives by refusing to manage climate change. A Kantian categorical imperative is to behave as you would have others behave; that would likely involve respecting others' livelihoods, even if they live decades later. Similarly, an Aristotelian notion of virtue would likely prohibit seriously damaging others' lives in pursuit of narrow self-interest. Finally, any approach based on rights would surely include respecting the rights of future generations.

We conclude that ethical arguments against discrimination by date of birth also apply in ethical frameworks that transcend utilitarian consequentialism. "Pure-time discounting" is not merely a narrow technical concern for nerdy economists. Avoiding this kind of discrimination is fundamental to most ethical doctrines.

### 8.3 The "weak argument": a more optimistic scenario?

Much remains to be done in reforming the world's economic system in order to avoid the serious risk of catastrophic climate change due to excessive greenhouse gas emissions. Other serious risks include excessive acidification by dissolved $CO_2$, not only of oceans, but also of reserves of fresh water. Another important concern is the loss of biodiversity.

Nevertheless, there are two reasons why the sacrifices required of either the current or future generations may turn out ultimately to be considerably less than had been feared back in the late 1990s. This was the period when negotiations were conducted which led to the Kyoto Protocol being approved by 160 nations in 1997, and becoming part of international law in 2005 after enough nations had ratified it.[18]

First, as is becoming widely recognized, the last ten or fifteen years have seen quite extraordinary technical progress, especially in using wind and solar power instead of steam to generate electricity. Along with technologies such as those that allow vehicles to use electric power to varying degrees, this is part of a general process whereby clean or zero-carbon technologies have become cheaper than high-carbon technologies in many sectors and geographical areas. In large measure this transformation has come about as a result of combining changes in social priorities with a process that KA did so much to illuminate in Arrow (1962)—namely the dynamics of learning by doing. Indeed, we may have within reach a new model of sustainable growth which is much more attractive and more inclusive than those experienced hitherto. It could include, for example, cities where the inhabitants and commuting workers can move and breathe, There can also be far higher resource productivity, as well as ecosystems which are robust and fruitful. Achieving such goals will require radical change that relies on major innovation and large investments. Those investments, however, would likely yield very high social returns.

Second, it may be appropriate to draw attention to the much less well known yet highly promising technologies for $CO_2$ removal by direct air capture that one of us (GC) has been helping to pioneer.[19] Moreover, there is some prospect in the next few years of complementary technical changes enhancing the use of $CO_2$ as a valuable industrial input. Indeed, a carbon X-prize was recently announced, intended to "challenge the world to reimagine what we can do with CO2 emissions by incentivizing and accelerating the development of technologies that convert CO2 into valuable products."[20]

Until relatively recently, it had seemed that the world would find it difficult to reduce the atmospheric concentration of greenhouse gases to safer levels without significant sacrifices of economic progress. The prospect of such sacrifices had become of great concern to KA who, during the last two decades of his life, addressed this kind of issue in several co-authored publications such as Arrow et al. (2003, 2004, 2012, 2013, 2014). Yet thanks in large part to recent technological developments, the world,

---

[18] Chichilnisky and Sheeran (2009) offer one account of these negotiations, particularly the final stages that took place in Kyoto.

[19] GC is the co-inventor and co-patentee of Carbon Negative Technology™, as described by Eisenberger et al. (2009), Chichilnisky (2011), Chichilnisky and Eisenberger (2011), as well as Choi et al. (2011).

[20] See http://www.carbon.xprize.org.

including we think KA himself during his last years, has been thinking more and more about policies that can manage change rather than discussing what "sacrifices" might be worthwhile. Indeed, there is the real prospect that directly capturing $CO_2$ from the air could soon replace existing sources of this widely used industrial gas and earn sufficient profit to make the sacrifice disappear entirely. In any case, this emphasis on how to manage change becomes even more urgent once we begin to consider other issues, such as those concerning human health and the robustness of ecosystems; here too KA has had so much to say. His legacy is truly extraordinary.

# References

Arrow KJ (1951) Alternative approaches to the theory of choice in risk-taking situations. Econometrica 19(4):404–437

Arrow KJ (1962) The economic implications of learning by doing. Rev Econ Stud 29(3):155–173

Arrow KJ (1963) Social choice and individual values, 2nd edn. Yale University Press, New Haven

Arrow KJ (1965) Aspects of the theory of risk-bearing. North- Holland, Amsterdam

Arrow KJ (1971) Essays in the theory of risk-bearing. Markham; North-Holland, Chicago; Amsterdam

Arrow KJ (1972) Exposition of the theory of choice under uncertainty. In: McGuire CB, Radner R (eds) Decision and organization. North-Holland, Amsterdam, pp 19–55 ch. 2

Arrow KJ (1977) Extended sympathy and the possibility of social choice. Am Econ Rev Pap Proc 67(1):219–225

Arrow KJ (1999a) Inter-generational equity and the rate of discount in long-term social investment. In: Sertel MR (ed) Contemporary economic issues. International Economic Association Series, Palgrave Macmillan, London, pp 89–102

Arrow KJ (1999b) Discounting, morality, and gaming. In: Portney PR, Weyant JP (eds) Discounting and intergenerational effects. Resources for the Future Press, Washington DC, pp 13–21

Arrow KJ (2007) Global climate change: a challenge to policy. The economists' voice. Published Online: 2007-06-30 https://doi.org/10.2202/1553-3832.1270

Arrow KJ, Cline WR, Mäler K-G, Munasinghe M, Stiglitz JE (1996) Intertemporal equity, discounting, and economic efficiency. In: Bruce JP, Lee H, Haites EF (eds) Climate change 1995: economic and social dimensions of climate change, contribution of Working Group III to the second assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 1–32

Arrow KJ, Cropper ML, Gollier C, Groom B, Heal GM, Newell RG, Nordhaus WD, Pindyck RS, Pizer WA, Portney PR, Sterner T, Tol RSJ, Weitzman ML (2013) Determining benefits and costs for future generations. Science 341:349–350

Arrow KJ, Cropper ML, Gollier C, Groom B, Heal GM, Newell RG, Nordhaus WD, Pindyck RS, Pizer WA, Portney PR, Sterner T, Tol RSJ, Weitzman ML (2014) Should governments use a declining discount rate in Project analysis? Rev Environ Econ Policy 8(2):145–163

Arrow KJ, Dasgupta P, Goulder L, Daily G, Ehrlich P, Heal G, Levin S, Mäler K-G, Schneider S, Starrett D, Walker B (2004) Are we consuming too much? J Econ Perspect 18(3):147–172

Arrow KJ, Dasgupta P, Goulder LH, Mumford KJ, Oleson K (2012) Sustainability and the measurement of wealth. Environ Dev Econ 17(3):317–353

Arrow KJ, Dasgupta P, Mäler K-G (2003) The genuine savings criterion and the value of population. Econ Theor 21(2):217–225

Arrow KJ, Debreu G (1954) Existence of an equilibrium for a competitive economy. Econometrica 22(3):265–290

Arrow KJ, Hurwicz L (1972) An optimality criterion for decision-making under ignorance. In: Carter DF, Ford F (eds) Uncertainty and expectations in economics. Oxford University Press, Oxford

Arrow KJ, Kurz M (1970) Public investment, the rate of return and optimal fiscal policy. Johns Hopkins University Press, Baltimore

Arrow KJ, Lind RC (1970) Uncertainty and the evaluation of public investment decisions. Am Econ Rev 60(3):364–378

Asheim GB (2007) "Economic analysis of sustainability" Chapter 1 of justifying, characterizing and indicating sustainability. Springer, Dordrecht

Asheim GB, Buchholz W (2003) The malleability of undiscounted utilitarianism as a criterion of intergenerational justice. Economica 70:405–422

Asheim GB, Buchholz W, Tungodden B (2001) Justifying sustainability. J Environ Econ Manag 41:252–268

Blackorby C, Bossert W, Donaldson DJ (1995) Intertemporal population ethics: critical-level utilitarian principles. Econometrica 63(6):1303–1320

Blackorby C, Bossert W, Donaldson DJ (2005) Population issues in social choice theory, welfare economics, and ethics. Cambridge University Press, Cambridge

Blackorby C, Bossert W, Donaldson D, Fleurbaey M (1998) Critical levels and the (Reverse) repugnant conclusion. J Econ (Zeitschrift für Nationalökonomie) 67(1):1–15

Blackwell D, Girshick MA (1954) Theory of games and statistical decisions. Wiley, Hoboken

Bommier A, Zuber S (2008) Can preferences for catastrophe avoidance reconcile social discounting with intergenerational equity? Soc Choice Welf 31:415–434

Broome J (1993) A cause of preference is not an object of preference. Soc Choice Welf 10(1):57–68

Brown C (2011) Consequentialize this. Ethics 121(4):749–771

Brown DJ, Lewis LM (1981) Myopic economic agents. Econometrica 49(2):359–368

Brundtland GH (1987) Report of the World Commission on environment and development: our common future. Oxford University Press, for the United Nations, Oxford

Chichilnisky G (1996) An axiomatic approach to sustainable development. Soc Choice Welf 13(2):231–257

Chichilnisky G (1997) What is sustainable development? Land Econ 73(4):467–491

Chichilnisky G (2009) Avoiding extinction: equal treatment of the present and the future. Econ Open-Access Open-Assess E-J 3(2009–32):1–25. https://doi.org/10.5018/economics-ejournal.ja.2009-32

Chichilnisky G (2011) Green capitalism: negative carbon and the green power fund. Int J Green Econ Indersci Enterp Ltd 5(4):321–333

Chichilnisky G, Eisenberger P (2011) Carbon negative power plants. CryoGas International, May, p. 36

Chichilnisky G, Hammond P, Stern NH (2018) Should We Discount the Welfare of Future Generations? Ramsey and Suppes versus Koopmans and Arrow. Working paper, http://wrap.warwick.ac.uk/107726

Chichilnisky G, Heal G, Beltratti A (1995) The green golden rule. Econ Lett 49(2):175–179

Chichilnisky G, Sheeran KA (2009) Saving Kyoto: an insider's guide to what it is, how it works and what it means for the future. New Holland Publishers Ltd, Cape Town

Choi S, Drese JH, Eisenberger PM, Jones CW (2011) Application of amine-tethered solid sorbents for direct CO2 capture from the ambient air. Environ Sci Technol 45(6):2420–7

Dasgupta PS, Heal GM (1979) Economic theory and exhaustible resources. Cambridge University Press, Cambridge

Diamond PA (1965) The evaluation of infinite utility streams. Econometrica 33:170–177

Eisenberger PM, Cohen RW, Chichilnisky G et al (2009) Global warming and carbon-negative technology: prospects for a lower-cost route to a lower-risk atmosphere. Energy Environ 20:973–984

Feng T, Ke S (2018) Social discounting and intergenerational Pareto. Econometrica 86(5):1537–1567

Figuières C, Tidball M (2012) Sustainable exploitation of a natural resource: a satisfying use of Chichilnisky's criterion. Econ Theor 49(2):243–265

Gale D (1967) On optimal development in a multi-sector economy. Rev Econ Stud 34(1):1–18

Hammond PJ (1973) Consistent Planning and Intertemporal Welfare Economics. Ph.D. dissertation, University of Cambridge

Hammond PJ (1975) A note on extreme inequality aversion. J Econ Theory 11:465–467

Hammond PJ (1976) Equity, Arrow's conditions, and Rawls' difference principle. Econometrica 44(4):793–804

Hammond PJ (1979) Equity in two-person situations: some consequences. Econometrica 47(5):1127–1135

Hammond PJ (1988) Consequentialist demographic norms and parenting rights. Soc Choice Welf 5:127–145

Hammond PJ (1992) On the impossibility of perfect capital markets. In: Dasgupta P, Gale D, Hart O, Maskin E (eds) Economic analysis of markets and games: essays in Honor of Frank Hahn. Press, Cambridge, M.I.T, pp 527–560

Hammond PJ (1993) Is there anything new in the concept of sustainable development? In: Campiglio L, Pineschi L, Siniscalco D, Treves T (eds) The environment after Rio: international law and economics. Graham & Trotman, 1994, London, pp 185–194 ch. 13

Hammond PJ (1995) Social choice of individual and group rights. In: Barnett WA, Moulin H, Salles M, Schofield N (eds) Social choice, welfare, and ethics. Cambridge University Press, Cambridge, pp 55–77

Hammond PJ (1996) Consequentialism, structural rationality and game theory. In: Arrow KJ, Colombatto E, Perlman M, Schmidt C (eds) The rational foundations of economic behaviour (IEA Conference volume No. 114). Macmillan, London, pp 25–42 ch. 2

Hammond PJ (1998) Objective expected utility: a consequentialist perspective. In: Barberà S, Hammond PJ, Seidl C (eds) Handbook of utility theory, Vol. 1: principles. Kluwer Academic Publishers, Dordrecht, pp 145–211 ch. 5

Harrod RF (1948) Towards a dynamic economics. Macmillan, London

Harsanyi JC (1953) Cardinal utility in welfare economics and in the theory of risk-taking. J Polit Econ 61:434–435

Harsanyi JC (1955) Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. J Polit Econ 63(4):309–321

Harsanyi JC (1975) Can the maximin principle serve as a basis for morality? a critique of John Rawls's theory. Am Polit Sci Rev 69:594–606

Harsanyi JC (1977) Morality and the theory of rational behavior. Soc Res 44:623–656

Harsanyi JC (1978) Bayesian decision theory and utilitarian ethics. Am Econ Rev (Papers and Proceedings) 68:223–228

Harsanyi JC (1979) Bayesian decision theory, rule utilitarianism, and Arrow's impossibility theorem. Theor Decis 11:289–317

Heal GM (1998) Valuing the future: economic theory and sustainability. Columbia University Press, New York

Hicks JR (1946) Value and capital: an inquiry into some fundamental principles of economic theory. Clarendon Press, Oxford

IPCC (Intergovernmental Panel on Climate Change) (2018) Global Warming of 1.5°C: An IPCC special report, www.ipcc.ch/sr15

Kolm S-C (1972, 1998) Justice et équité (Paris: CNRS); translated as Justice and Equity. MIT Press, Cambridge

Kolm S-C (1994) The meaning of 'fundamental preferences'. Soc Choice Welf 11(3):193–198

Koopmans TC (1960) Stationary ordinal utility and impatience. Econometrica 28(2):287–309

Koopmans TC (1965) On the concept of optimal economic growth. Acad Sci Scr Varia 28(1):225–287 reprinted as Cowles Foundation Paper 238

Koopmans TC (1967) Objectives, constraints, and outcomes in optimal growth models. Econometrica 35(1):1–15

Koopmans TC, Diamond PA, Williamson RE (1964) Stationary utility and time perspective. Econometrica 32(1–2):82–100

Lauwers L (2017) The axiomatic approach to the ranking of infinite utility streams. In: Chichilnisky G, Rezai A (eds) The economics of the global environment: catastrophic risks in theory and policy. Springer International, New York, pp 231–255

Llavador H, Roemer JE, Silvestre J (2015) Sustainability for a warming planet. Harvard University Press, Cambridge

Menger K (1934) Das Unsicherheitsmoment in der Wertlehre, Betrachtungen im Anschluss an das sogenannte Petersburger Spiel. Zeitschrift für Nationalökonomie 5: 459–485; translated under the title The Role of Uncertainty in Economics. In: Shubik M (ed.) (1967) Essays in Mathematical Economics in Honor of Oskar Morgenstern. Princeton University Press, Princeton, ch. 16, pp 211–231

Mirrlees JA (1963) Optimum Accumulation Under Uncertainty. Ph.D. dissertation, University of Cambridge

Mirrlees JA (1967) Optimum growth when technology is changing. Rev Econ Stud 34(1):95–124

Mirrlees JA, Stern NH (1972) Fairly good plans. J Econ Theory 4(2):268–288

Mongin P (2001) The impartial observer theorem of social ethics. Econ Philos 17:147–179

Mukerji N (2016) The case against consequentialism reconsidered. Springer, New York

Ord T (2020) The precipice: existential risk and the future of humanity. Bloomsbury Press, London

Parfit D (1984) Reasons and persons. Oxford University Press, Oxford

Pezzey J (1997) Sustainability constraints versus 'optimality' versus intertemporal concern, and axioms versus data. Land Econ 73:448–466

Pigou AC (1932) The economics of welfare, 4th edn. Macmillan, London

Portmore DW (2007) Consequentializing moral theories. Pac Philos Q 88:39–73

Portmore DW (2009) Consequentializing. Philos Compass 4:329–347

Ramsey FP (1928) A mathematical theory of saving. Econ J 38(4):543–559

Ramsey FP (1931) Epilogue. In: Braithwaite RB, Ramsey FP (eds) The foundations of mathematics and other logical essays. Routledge and Kegan Paul, London, pp 287–292

Rawls J (1971) A theory of justice. Belknap Press, Cambridge

Sen AK (1970) 2017 Collective choice and social welfare, Expanded edn. Penguin Books, London

Sidgwick H (1907) The methods of ethics. Macmillan, London

Sikora RI, Barry B (1978) (eds.) Obligations to future generations. Temple University Press, Philadelphia

Solow RM (1991) Sustainability: an economist's perspective. The Eighteenth J. Seward Johnson Lecture, Marine Policy Center, Woods Hole Oceanographic Institution

Solow R (2012) A few comments on 'Sustainability and the Measurement of Wealth'. Environ Dev Econ 17(3):354–355

Stern NH (1977) Welfare weights and the elasticity of the marginal valuation of income. In: M. Artis and R. Nobay (eds.) Proceedings of the AUTE edinburgh meeting of April 1976. Blackwell, Basil, ch. 8

Stern NH (2007) The economics of climate change: the Stern review. Cambridge University Press, Cambridge

Stern NH (2008) The economics of climate change. Am Econ Rev 98(2):1–37

Stern NH (2014a) Ethics, equity and the economics of climate change. Paper 1: science and philosophy. Econ Philos 30(3):397–444

Stern NH (2014b) Ethics, equity and the economics of climate change. Paper 2: economics and politics. Econ Philos 30(3):445–501

Stern NH (2015) Why are we waiting? the logic, urgency, and promise of tackling climate change. MIT Press, Cambridge

Suppes PC (1966) Some formal models of grading principles. Synthese 6:284–306

Tinbergen J (1957) Welfare economics and income distribution. Am Econ Rev Pap Proc 47(2):490–503

Vickrey WS (1945) Measuring marginal utility by reactions to risk. Econometrica 13(4):319–333

Vickrey WS (1960) Utility, strategy and social decision rules. Quart J Econ 74(4):507–535