Australasian Journal of Philosophy

ISSN: 0004-8402 (Print) 1471-6828 (Online) Journal homepage: http://www.tandfonline.com/loi/rajp20

The Relatively Infinite Value of the Environment

Paul Bartha & C. Tyler DesRoches

AUSTRALASIAN JOURNAL OF PHILOSOPHY, 2016
http://dx.doi.org/10.1080/00048402.2016.1182196

The Relatively Infinite Value of the Environment

Paul Bartha[a] and C. Tyler DesRoches[b]

[a]University of British Columbia; [b]Arizona State University

ABSTRACT
Some environmental ethicists and economists argue that attributing *infinite value* to the environment is a good way to represent an absolute obligation to protect it. Others argue against modelling the value of the environment in this way: the assignment of infinite value leads to immense technical and philosophical difficulties that undermine the environmentalist project. First, there is a problem of discrimination: saving a large region of habitat is better than saving a small region; yet if both outcomes have infinite value, then decision theory prescribes indifference. Second, there is a problem of swamping probabilities: an act with a small but positive probability of saving an endangered species appears to be on par with an act that has a high probability of achieving this outcome, since both have infinite expected value. Our paper shows that a *relative* (rather than absolute) concept of infinite value can be meaningfully defined, and provides a good model for securing the priority of the natural environment while avoiding the failures noted by sceptics about infinite value. Our claim is not that the relative infinity utility model gets every detail correct, but rather that it provides a rigorous philosophical framework for thinking about decisions affecting the environment.

ARTICLE HISTORY Received 11 December 2015; Revised 3 April 2016

KEYWORDS infinite value; infinite decision theory; environmental ethics; environmental value; cost benefit analysis

1. Introduction

Decisions about the environment are among the most complex that we will ever make. They require trade-offs between conservation goals, economic interests, and the satisfaction of basic human needs. In some cases, we have to make decisions that involve pitting objectives of large but measurable economic value against the conservation of apparently 'priceless' parts of the natural world: lakes and oceans, entire ecosystems, and biodiversity in general. In other cases, our conservation policies may require us to pit these 'priceless' parts of nature against each other. How are we to make such decisions?

The answer depends in part upon how we interpret (and whether we accept) the assertion that a part of nature is priceless. It might be a colourful way of saying that the value is *enormous* or *hard to estimate*. It might represent the view that the value is *incommensurable* with that of ordinary everyday goods (as there is no reasonable way to make comparisons). Most controversially, it might be intended to represent an ascription of *infinite* value. Our paper sets aside the first two interpretations and focuses on cases where 'priceless' can be interpreted in the third way.

Is it ever appropriate to think of parts of the natural environment as having infinite value?[1] In formal terms, is it ever fruitful to model our preferences in a way that treats the environment as infinite in value? We can divide this question into three parts. First, can the concept of infinite value be made meaningful and clear? Second, can we formalize the notion without generating implausible (or even absurd) theoretical and practical consequences? Finally, is there anything to be gained by the assignment of infinite value, as opposed to large finite value, to the environment?

Some authors (notably Broome [2008]) question the very meaning and justification of infinite value. Others, notably Colyvan et al. [2010a, 2010b], countenance versions of decision theory that allow infinite values,[2] but argue persuasively that these are the wrong way to enshrine environmental safeguards. They identify technical and philosophical difficulties with various attempts to formalize the notion, and conclude that infinite values would 'paralyze conservation efforts'. Finally, they maintain that we get everything we want from large finite values, so that there is no need to introduce troublesome infinities. In short, they answer 'no' to questions two and three above.

We argue, however, that one can meaningfully hold preferences of the form, '*A* is infinitely better (or infinitely worse) than *B*.' Far from being absurd, such preferences express deep commitments about the types of trade-offs that we are willing or unwilling to make. We show that these preferences can be represented by assignments of *relatively infinite* value, and we refer to decision-theoretic frameworks that incorporate such values as *infinite-value models*. Relatively infinite values allow us to answer 'yes' to each of our three questions. They can be meaningfully defined, they offer a plausible way to accommodate deontological reasoning, and they generate recommendations distinct from standard (finite) decision theory.

We do not claim that it is always appropriate, in making decisions about the natural environment, to employ infinite-value models. We do, however, aim to disarm criticisms that such models are meaningless or fatally flawed. There are circumstances under which they provide a useful framework for making decisions.

Our paper proceeds as follows. Section 2 reviews two leading strategies for environmental decision-making: *cost-benefit analysis* and *deontological* approaches. Section 3 introduces examples to motivate the idea that parts of nature can be infinitely valuable (in a relative sense), and section 4 shows that relative infinite values can be meaningfully defined. Section 5 adds further examples that prompt us, following Colyvan et al., to reject what we call the *naïve infinite utilities* model. Taken together, the examples of sections 3 and 5 pose challenges for both cost-benefit and deontological approaches. Sections 6 and 7 show that the examples can be handled by two infinite-value models: *lexicographic orderings* and *relative utility theory* (RUT). Although we believe that infinite-utility models are meaningful and useful, we conclude (section 8) by acknowledging some concerns and limitations.

The conclusions reached in this paper may be of interest to at least three groups of scholars. The first is environmental economists, who have generally recognized only

1 By 'parts of the natural environment', following John Stuart Mill [(1874) 2006], we mean those parts of the world that remain relatively (if not wholly) detached from intentional human agency.

2 In particular, Colyvan [2008] develops his *Relative Expectation Theory*. Relative expectation theory is meant to handle some of the puzzles encountered in the St. Petersburg and Pasadena games. Colyvan is also open to other versions of non-standard decision theory, such as that proposed by Easwaran [2008]. As we shall see, however, Colyvan is clearly opposed to modelling the value of the environment as infinite.

the deontological and cost-benefit approaches to environmental decision-making [Howarth 1995; Pearce et al. 2006]. The second group consists of conservation biologists and ecologists currently embroiled in a debate that pits the moral duty to preserve nature against purely anthropocentric motivations to achieve this end [Soulé 1985; Kareiva 2010; Kareiva and Marvier 2012; Soulé 2013; Miller et al. 2014]. In as much as this debate might be construed as involving both deontological and utilitarian reasoning, our paper offers a means for representing considerations of both sorts within a single framework. Finally, relative infinite utility applied to environmental decision-making will be of interest to environmental ethicists and decision theorists who have argued that ascribing infinite value to nature (or anything else) is a mistake [Broome 2004; Colyvan et al. 2010a].

2. Two Models for Environmental Decisions

We begin with *cost-benefit analysis*, also sometimes referred to as the *ecosystem services approach* [Daily 1997; Costanza et al. 1997]. The term 'cost-benefit analysis' refers to a family of broadly utilitarian frameworks [Sen 2000]. The key assumption common to many such frameworks is that there is a theoretical basis for assigning a determinate, finite, monetary value to any part of nature. To support this assumption, some economists marshal evidence in terms of people's 'willingness to pay' or 'willingness to accept' [Pearce 1998].[3] There is no insistence, however, that only market value should be taken into account. Many parts of the environment included in such analyses are not traded in the marketplace and possess no market value. Other parts have a value determined not solely by market considerations, but also by social and ethical values. Thus, the cost-benefit approach may be characterized more broadly as resting on the assumption that one can assign determinate finite utility to relevant parts of nature.

Given the key assumption, the cost-benefit approach allows us to make complex environmental decisions using the same framework that we use in straightforward economic decisions. We compare the costs and benefits of each option, taking probabilities into account, and apply the principle of maximizing expected monetary value (or expected utility).[4] Environmental economists Bulte and van Kooten characterize the cost-benefit approach as follows: 'Conservation can be promoted by demonstrating that 'saving' species and ecosystems promises higher monetary return than their conversion into other assets' [2000: 114]. The flip side, of course, is that 'every species has to "earn" its place in the sun': extinction is a defensible outcome when conservation is not competitive with the exploitation of a natural resource. Bulte and van Kooten illustrate this, using the example of logging British Columbia's old-growth forests. Cost-benefit analysis recommends a halt to logging at the point where the marginal cost of harvesting additional trees equals the marginal benefit.

Over the last few decades, the cost-benefit approach has been embraced by many ecologists and conservation biologists because it is believed that a direct appeal to social and economic benefits is a winning strategy for protecting the natural environment

[3] Such methods are not uncontroversial (see Diamond and Hausman [1994] and Hausman [2012]). They are not essential on the more liberal characterization of the cost-benefit approach to which we immediately proceed.

[4] Indeed, the cost-benefit approach can be carried out in terms of finite utility, rather than monetary value. The essential assumption here is really *Continuity* of preferences (see section 4). In our discussions below, we represent costs and benefits in terms of expected utility rather than of expected monetary value.

[Daily 1997]. Its advocates claim notable successes, and argue that their approach is conducive to a transparent, objective, and nuanced evaluation process for making environmental decisions [Farber et al. 2002]. Critics point to flaws or unrealistic assumptions in the way that valuations of the environment are obtained, to concerns about whether ethical considerations are properly captured by expressions of 'willingness to pay', and to the worry that it is much easier to quantify the costs of environmental protection than its benefits [Ackerman and Heinzerling 2002; Sagoff 2004]. The successes of the model appear to depend on the whims of the market and ever-shifting tastes. McCauley [2006] warns that cost-benefit analysis provides no protection at all for species that are unloved or deemed to be useless. His approach is entirely different: 'Nature has an intrinsic value that makes it priceless, and that is reason enough to protect it' [ibid.: 28]. McCauley's point seems to be that, even where cost-benefit analysis recommends protection of the environment, it does so for the wrong reasons. We examine this idea further in section 3.

The main alternative to cost-benefit analysis, as illustrated by McCauley's remark, is some type of *deontological framework* for decision-making. The crucial idea is expressed in terms of constraints on human activity.[5] Legislation that puts in place an absolute ban on hunting elephants or developing the Arctic wilderness, for instance, might be motivated by the conviction that such a duty exists—even in the absence of a compelling economic argument. Colyvan and his co-authors offer a succinct characterization of this strategy: 'Such approaches take there to be absolute, non-negotiable duties and obligations, such as the obligation not to wantonly harm other humans' [2010a: 224].

From a philosophical perspective, the tension between cost-benefit and deontological approaches to environmental decision-making is a special case of a more general conflict between decision-theoretic and deontological frameworks. As Colyvan et al. state: 'Problems arise … when we try to reconcile such [deontological] obligations with the maximizing of expected utility in the standard formal decision theoretic framework. Indeed, problems arise … anywhere where trade-offs need to be made' [ibid.]. In this connection, the cost-benefit approach enjoys a significant advantage over deontological approaches: it uses the precise formal language and tools of public policy analysis. As a result, it gets the attention of policy-makers more effectively than traditional, and less easily formalized, ethical arguments that impose limits on what we may do to the natural environment.

It is precisely at this point that infinite utilities enter the picture. The expansion of standard decision theory to include infinite values (*positive* and/or *negative infinity*) looks, at first, like a promising way to incorporate deontological considerations into the decision-theoretic framework. By assigning positive infinite value to some part of nature (for example, an endangered species, such as African elephants), we secure a non-negotiable duty to conserve that good. By assigning negative infinite value to the destruction of the Arctic National Wildlife Refuge, we ensure its protection. Within this expanded decision theory, which we call the *naïve infinite utilities* model, it seems that we can find a home for absolute duties and proscriptions.

[5] Since our purpose in this paper is to explore the representation of such constraints, we set aside debates about their putative justification (such as McCauley's claim that nature has 'intrinsic value'). For a critique of McCauley [2006], see Justus et al. [2009].

In section 5, we discuss objections to this model raised by Colyvan et al. [ibid.], and we agree that they are decisive. Our contention, however, is that these objections can be circumvented with a more refined notion of infinite value. In particular, we show that relative utilities offer a decision-making framework that can accommodate both deontological and utilitarian considerations.

3. Two Examples: A Challenge to the Cost-Benefit Approach

This section presents two examples of reasoning that involve trade-offs between a large but finite good and the preservation of a significant part of the environment. In each case, there is a plausible interpretation of the reasoning that assigns infinite value to a part of nature. Our focus here is on the *subjective* perception of infinite value and how it figures in the agent's reasoning.

We don't claim that the infinite-value interpretation is compulsory, but we do claim that *it is a possibility that should be taken seriously*. Crucially, that is all we need in order to demolish the universality of the cost-benefit approach. Since cost-benefit analysis rejects, *a priori*, the type of reasoning that may be at work in these examples, it cannot serve as a universal framework for environmental decision-making.

Example 1. The Atmosphere. In *The Making of the Atomic Bomb*, Rhodes [1986] describes an incident from July 1942 when Edward Teller, investigating thermonuclear reactions that might result from nuclear fission, expressed the fear that atomic bombs 'might ignite the earth's oceans or its atmosphere and burn up the world'. Robert Oppenheimer took the possibility seriously enough that he raised the issue with Arthur Compton. Compton later wrote ([ibid.: 419; from Compton [1956: 127ff.]):

> Was there really any chance that an atomic bomb would trigger the explosion of the nitrogen in the atmosphere or the hydrogen in the ocean? This would be the ultimate catastrophe. Better to accept the slavery of the Nazis than to run a chance of drawing the final curtain on mankind!
> We agreed that there could only be one answer. Oppenheimer's team must go ahead with their calculations.

In the end, Hans Bethe studied the calculations and, to the satisfaction of Teller and everyone else, 'the impossibility of igniting the atmosphere was assured by science and common sense'.

Compton weighed the destruction of the atmosphere against world domination by the Nazis. He might have meant that the Manhattan Project should proceed, so long as the probability of igniting the atmosphere or oceans was sufficiently low. This would be consistent with assigning a large but finite value to the atmosphere. The tenor of Compton's remarks, however, suggests otherwise. He wanted proof that the catastrophe was *impossible* (or at least that the detonation of an atomic bomb would not increase its chance of occurring). On this interpretation, Compton was asserting that a Nazi triumph might be a vast evil, but that the loss of the atmosphere would be *infinitely* worse, since the atmosphere is indispensable to human survival. Construed in this way, Compton's reasoning departs from conventional cost-benefit analysis, which requires a finite value for the atmosphere.

Example 2. Northern Gateway Pipeline. The Enbridge Northern Gateway Pipeline proposal has recently met with stiff resistance from British Columbians, particularly First Nations people. Under the Enbridge proposal, the pipeline would transport oil from the Athabasca oil sands in northern Alberta to Kitimat on the coast of British Columbia, and from there tankers would load and ship the oil to Asian ports. Both the pipeline and the tankers would pass through ecologically sensitive areas that are of great importance to local populations. The following comments [Swift et al. 2011] are representative of the opposition:

> First Nations have used our ancestral laws to ban Enbridge's pipelines and tankers from our lands, taking up more than half of the proposed pipeline and tanker route from the Rockies, clear across to the Pacific Ocean. Our Nations are the wall this pipeline will not break through. Our lands and waters are not for sale, not at any price. We want no part of Enbridge's project and their offers are worthless to us when compared to the importance of keeping our lands, rivers and the coast free of crude oil spills. What Enbridge is offering is the destruction of our lands to build their project, and the risk of oil spills for decades to come which could hurt everyone's kids and grandkids.
>
> —Chief Larry Nooski, Nadleh Whut'en First Nation, member Nation of the Yinka Dene Alliance, 2011

> Every year my calendar is run by the sea and the land. You can't take away that essence of me. The money from a pipeline is there for a short time. The land is there forever.
>
> —Nancy Nyce, Haisla Nation, Nana ki'la Guardians, 2011

This example differs from the previous one because the preservation of the Kitimat ecosystem, unlike the atmosphere, is not indispensable for continued human existence. Nevertheless, the foregoing remarks express sentiments similar to those of Compton in *Example 1*. Some First Nations people in British Columbia appear to regard the lands and waters as *priceless relative to* (infinitely more valuable than) the potential economic benefits of the proposed pipeline.

4. The Meaning of Infinite Utility

In opposition to the infinite-value interpretation of the preceding examples, we might argue that, for human agents, infinite value is a meaningless concept. Broome offers a direct argument for the case of human life: 'No one's life has infinite value. How could it have? Our human lives are only finite in length, and during them we can experience and achieve only a finite number of things' [2008: 53]. This argument clearly generalizes. As finite beings, we can only experience finite gains and losses. Environmental damage, as in *Example 2*, can never be experienced as infinitely bad.

It is important to distinguish here between two concepts of utility. The first, associated with Jeremy Bentham [1789 (1843)] and John Stuart Mill [1863 (1906)], takes utility as an evaluative measure of subjective experience (happiness, pleasure, etc.). We agree that the idea of *absolutely infinite utility*, construed in this way, is meaningless. The second, in the tradition of von Neumann and Morgenstern [1953], treats utility as a device for representing the structure of an agent's preferences. The concept of infinite utility, construed in this way, is not automatically undermined by Broome's argument, although, as we shall see, other objections can be offered.

Our focus in this paper is on representations of infinite utility, but, as a preliminary task, we show in this section that it can be meaningfully defined and that the objections are far from decisive. To begin with, provided that we allow that there are subjective probabilities, there is a straightforward way to understand what it means for you to prefer outcome B infinitely relative to some other outcome A. It means that you would give up A for any bet that gives you a positive chance, however small, of gaining B. That is almost right, but we need to say what happens if you lose the bet! To that end, we introduce some notation.

- *Weak preference*. $B \succcurlyeq A$ (also $A \preccurlyeq B$) means B is at least as good as A.
- *Strict preference*. $B \succ A$ (also $A \prec B$) means that B is strictly preferred to A.
- *Indifference*. $B \approx A$ means that the agent is indifferent between B and A.
- *Gambles*. $[\lambda B, (1 - \lambda)Z]$ is the gamble that gives the agent chance λ of winning B and chance (1-λ) of winning Z, where $0 \leq \lambda \leq 1$.
(A fair coin toss between B and Z, for instance, is represented as $[{}^1\!/_2\, B, {}^1\!/_2\, Z]$.)

We define relative infinite utility as a three-place relation, in terms of a *base-point* that specifies the losing outcome (the worst of the three under consideration).

Relative Infinite Utility.
Let A, B, and Z be any three outcomes, where $B \succ A \succ Z$. An agent *values B infinitely relative to A and base-point Z* if
$[\lambda B, (1\text{-}\lambda)Z] \succ A$ for any $\lambda > 0$.
(The agent is willing to trade A for *any* gamble that offers a positive chance of B, when the 'losing outcome' or base-point is Z.)

This definition establishes, at minimum, that relative infinite utility is a meaningful concept. It employs no machinery apart from ordinary real-valued probabilities and preferences among ordinary gambles.

There does not seem to be any *prima facie* reason why a rational agent could not have such preferences. Nevertheless, it is well known that standard decision theory cannot accommodate them, by virtue of the *Continuity Axiom*, which states that for any three outcomes Z, A, and B such that B is preferred to A and A is preferred to Z, the agent must be indifferent between A and *some gamble* between B and Z.

Continuity.
Whenever $B \succcurlyeq A \succcurlyeq Z$, there exists some $0 \leq \lambda \leq 1$ such that $[\lambda B, (1-\lambda)Z] \approx A$.

The following picture may be helpful. Gambles between Z and B are represented as points along the interval from Z to B. The probability λ of winning B is represented as a proportion of the total interval. Given an outcome A that is intermediate between Z and B, an agent whose preferences satisfy *Continuity* will always be able to find some point (gamble) in this interval that is equivalent to A (such that the agent is indifferent between A and the gamble).

Clearly, it is impossible for an agent whose preferences satisfy *Continuity* to prefer one outcome infinitely relative to another. Suppose that B is strictly preferred to A and A is strictly preferred to Z, as in *Figure 1*. By *Continuity*, there is

Figure 1: Continuity

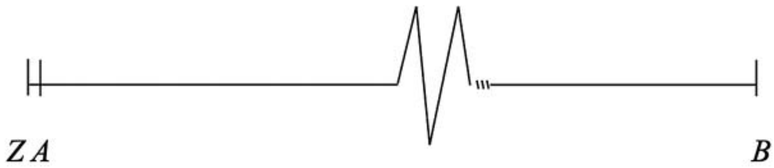

Figure 2: Violation of Continuity

a value λ strictly between 0 and 1 such that the agent is indifferent between A and $[\lambda B, (1-\lambda)Z]$. The agent therefore prefers A to gambles between B and Z where the probability of B is positive but less than λ. Consequently, she does not value B infinitely relative to A and Z.

The case where the agent *does* prefer B infinitely relative to A and Z is illustrated in Figure 2. The zigzag line indicates a discontinuity in the agent's preferences: outcomes on the right side are preferred infinitely to outcomes on the left side (with base-point Z).

Continuity is a basic assumption of standard decision theory. Its plausibility rests on the observation that its failure 'prevents any sort of trade-off or balancing compensation' [Fishburn 1974: 1444]. Feldman, for example, offers a succinct argument that we don't value our own lives as relatively infinite with respect to ordinary goods: 'If no one were to accept a positive increment in the probability of death in exchange for a modest money gain, the world as we know it would grind to a halt' [1994: 6]. In a similar vein, Colyvan et al. [2010a: 226] write this:

> People do not act as though they attribute infinite value to anything. If they did, they would sacrifice any finite amount of money, goods, or well-being for even a miniscule chance at achieving what they putatively infinitely value and moreover, would not care about the probabilities in question … [But] people do care about the probabilities and this only makes sense if the values in question are finite.

However, Luce and Raiffa [1957], Dreier [1996], Hájek [1997] and many others note that violations of *Continuity* are not necessarily irrational.[6] Some empirical studies also appear to support attributions of infinite value [Spash and Hanley 1995].

For present purposes, it suffices to say that the matter is unsettled, and that there is a well-established body of work on preferences that violate *Continuity*.[7] Our view is that people do sometimes offer deontological arguments that suggest discontinuous preferences, and that their arguments should not be ruled out *a priori*. The next few sections discuss three approaches for representing discontinuous preferences: naïve infinite utilities, lexicographic orderings[8], and the *relative utilities* approach.[9]

---

6 Dreier's example is the following: you might prefer B, *avoiding financial loss*, to A, *financial loss*, to C, *committing a sin*, and yet prefer A to any gamble with probability p of B and $(1-p)$ of C, as long as $(1-p) > 0$. In Dreier's view, this is not obviously irrational. This case bears some analogy to those discussed in section 3.

7 Skala [1975] provides an extended treatment of non-Archimedean utility.

8 Fishburn [1974] reviews theoretical results that specify the axioms for representation of preferences in terms of lexicographically ordered utilities.

9 Yet another approach is to use hyperreal or surreal numbers. For an application of these ideas to Pascal's Wager, see Hájek [2003].

5. The Naïve Infinite Utilities Model

The naïve infinite utilities model brings deontological intuitions into decision theory in a very simple way. The essential idea is to allow the utility function that represents an agent's preferences to take the values $+\infty$ and $-\infty$, in addition to finite real values. Expected utilities of acts are then calculated more or less as usual.[10] To see how this provides protection for the environment, consider the examples from section 3.

Example 1 (continued). Represent the utilities of the relevant outcomes as follows, where N and S are finite positive numbers:

$$u(\text{Atmosphere destroyed}) = -\infty$$
$$u(\text{Nazi victory}) = -N$$
$$u(\text{Allied victory}) = S$$

Finally, suppose that $Pr(\text{Atmosphere destroyed} / \textit{Build bomb}) = p > 0$: the creation of the atomic bomb has a small positive probability p of destroying the atmosphere. Given these assumptions (and assuming that the allies win if they have the bomb), the expected utility of building the bomb is $-\infty$:

$$\begin{aligned} \text{EU}(\textit{Build bomb}) &= p \cdot u(\text{Atmosphere destroyed}) + (1-p) \cdot u(\text{Allied victory}) \\ &= p(-\infty) + (1-p)S \\ &= -\infty. \end{aligned}$$

The expected utility of *not* building the bomb is finite, with value between S and $-$N, depending on the probability that the allies win without the bomb. On the naïve infinite-utilities model, the bomb should not be built if there is *any* positive chance that it will burn up the atmosphere.

By contrast, on a cost-benefit model, we assign a *very large* negative utility to the destruction of the atmosphere. A decision to build the bomb is justified (and indeed, rationally required) if the probability of catastrophe is sufficiently small.

Example 2 (continued). Suppose that $u(\text{Spill}) = -\infty$: an oil spill from the pipeline or from a tanker is infinitely bad. Suppose that $Pr(\text{Spill}/\textit{Build pipeline}) = p > 0$. Then

$$\text{EU}(\textit{Build pipeline}) = p(-\infty) + (1-p)(\textit{finite}) = -\infty.$$

The expected utility of going ahead with the pipeline is $-\infty$. The Northern Gateway project should be rejected if there is *any* positive chance of a major oil spill. Once again, the reasoning here is qualitatively different from the cost-benefit approach, where the correct decision will vary depending upon the probability and the utility values.

Unfortunately, these apparent successes for naïve infinite utilities are eclipsed by enormous difficulties that arise in other scenarios. Consider a modified version of the *Northern Gateway* example: we are trying to choose between one shipping option that has very low associated risk of an oil spill, and a second option that has high associated risk. The first option should be preferable; yet on the naïve infinite-utilities approach both options have negatively infinite expected utility! The model provides no guidance: there is no basis for choosing among acts with equal expected utility.

[10] For arithmetic involving $+\infty$ and $-\infty$, see Rudin [1976]. Note that some acts will have undefined expected utility, since $\infty + -\infty$ is undefined.

Colyvan and his co-authors identify a host of such difficulties for the naïve infinite-utilities model. The key problem relates to the very definition of a utility scale. If u is a function that assigns a numerical value to each outcome, then u must satisfy the following two properties in order to count as a representation of an agent's preferences:

(U1) $u(x) > u(y)$ if and only if the agent strictly prefers x to y; and
(U2) $u(x) = u(y)$ if and only if the agent is indifferent between x and y.

In particular, u must discriminate between two outcomes if and only if the agent discriminates between them.

It is worth noting that, even though (U1) and (U2) are standard in decision theory, Easwaran [2014] has recently argued that, in some cases where u represents expected utility, we should weaken the conditions.[11] Decision theory needs only the idea that strict inequality corresponds to strict preference: there can be cases where two acts have equal expected utility, but we nevertheless strictly prefer one to the other. Easwaran's amendment is motivated by cases involving partitions where some possible state has a probability that is zero or undefined. Since no such cases occur in this paper, we retain (U1) and (U2).

The following examples, based on the discussion in Colyvan et al. [2010a], describe cases of intuitively straightforward decisions that become problematic when we admit infinite utilities. In both cases, conditions (U1) and (U2) are violated: the utility function fails to discriminate properly.

Example 3. Non-discrimination: the mangrove forest.
Colyvan et al. write [ibid.: 225]:

> Infinite value is insufficiently discriminative of salient outcomes. For example, if some habitat is assigned infinite value (e.g., mangrove forests), attributing meaningful values to larger or smaller regions of that habitat is problematic because, according to standard accounts of infinite value, all infinitely valued items are equal… A region of a specific habitat may be highly valuable but, all else being equal, saving more of the habitat is surely more valuable. … Assigning infinite value precludes such finer discriminations.

Let $W \equiv$ *The whole forest is saved* and $H \equiv$ *Half the forest is saved*. The simplest version of the problem is that W is obviously preferable to H, but $u(W) = u(H) = \infty$. The utility function does not discriminate.

There is a closely related probabilistic version of the problem. Suppose that we have a choice between *Save half the forest* and *Save the entire forest*. Suppose that in each case we have the same probability $p > 0$ of success, and the outcome if we fail is identical (the forest is lost). Finally, assume that there are no other relevant considerations. Since W is strictly preferable to H, standard preference axioms imply that a rational agent must prefer *Save the entire forest* to *Save half the forest*.[12] But, on the naïve infinite utilities model, since both W and H have value $+\infty$, both acts have infinite expected utility. Again, the utility function fails to discriminate properly.

[11] Our thanks to one of the referees for pointing this out.
[12] This is a direct consequence of the *Independence Axiom*: if an agent prefers x over y, then the agent should also prefer a bet that yields x with probability p over a bet that yields y with probability p, provided that the other outcome is identical in both bets. For discussion, see Resnik [1987]. (Resnik calls this the *Better-Prizes Condition*.)

Example 4. Probability swamping: The endangered species problem. Colyvan et al. write [ibid.: 255]:

> If an outcome (e.g., protection of threatened habitat) is assigned infinite value, the expected value of actions that have the slightest chance of producing that outcome is infinite.… For example, if persistence of an endangered species is considered infinitely valuable, actions with *any* nonzero chance of ensuring its survival will have infinite expected value. In particular, actions with high and actions with low probabilities of species survival would have identical expected value. This would imply indifference about actively protecting endangered species and passively doing nothing, and this is patently the wrong result.

Let $S \equiv$ *Species is saved*. Suppose that $u(S) = \infty$ and $u(\sim S) = k$ (finite). Let P represent the option of passively doing nothing, while I represents active intervention to protect the endangered species. For definiteness, assume that $Pr(S / P) = 0.01$ and $Pr(S / I) = 0.9$. Then

$$\begin{aligned} \mathrm{EU}(P) &= (0.01)(\infty) + 0.99k &&= \infty, \text{ and} \\ \mathrm{EU}(I) &= (0.9)(\infty) + 0.1k &&= \infty. \end{aligned}$$

Plainly, act I is preferable to act P; in fact, standard decision theory implies that a rational agent must prefer act I.[13] Yet the counter-intuitive result of the calculation is that both acts have equal infinite expected utility. Once again, our utility function fails to discriminate properly and, as a result, it does not even qualify as a utility scale.[14]

Colyvan et al. provide additional examples of 'counterintuitive and counterproductive' results that seem to follow from the use of the values $+\infty$ and $-\infty$, or formally equivalent approaches. For instance: 'actions with a minute chance of preserving two endangered species are always better than actions that guarantee saving one species' [ibid.: 227]. This example will be discussed later (sections 5 and 7).

We can describe the problem in two different ways. If we treat the expected utility calculations as *prescribing* a rational course of action, the problem is that, in both examples, we have a counterintuitive recommendation to be indifferent between two acts that are plainly *not* equally good. If we treat the utility calculations as *representing* the agent's well-defined preferences, the problem is that, in both examples, we have a failure of representation. Either way, we have a disaster: the utility function fails to satisfy the discrimination requirements (U1) and (U2).

Taking stock, it is clear that the cost-benefit approach fares much better than the naïve utilities approach with Examples 3 and 4. We can assign higher finite utility to preserving the whole mangrove forest than to preserving half the forest. Similarly, if we assign a high finite utility to saving the endangered species, then we obtain higher expected utility for active intervention than for passively doing nothing.

With Examples 1 and 2, of course, the cost-benefit approach fails to duplicate deontological reasoning. Nevertheless, in a different paper, Colyvan et al. [2010b] outline a way in which the cost-benefit model might approximately represent deontological reasoning by using *large finite utilities*. The essential idea is the following: '[The] absolute difference between the utilities of outcomes associated with an obligation (prohibition)

[13] This is also a consequence of the *Independence Axiom*.

[14] The problem generalizes. Echoing a point made by Hájek [2003] in a discussion of *Pascal's Wager*, just about *any* act that you might perform is consistent with some positive probability that the species will survive. If every act open to you has equal infinite expected utility, then decision theory doesn't rule out any option.

and a permissible act should be much greater than the absolute difference between any permissible acts' [ibid.: 515]. This idea is then extended to more complex cases where we have a ranking of obligations or prohibitions: the absolute difference between the utilities of an outcome associated with an obligation (prohibition) of higher rank and one of lower rank should be much greater than the absolute difference between two outcomes at the lower rank. *Example 1* is modelled by assigning a very large negative utility to Nazi victory, and a vastly larger negative utility to the outcome in which the atmosphere is destroyed. *Example 2* can be handled in a similar way. Call this the *CCS model*.

The CCS model generates the right *conclusions*, but it does not reproduce Compton's reasoning or the reasoning of Nooski or Nyce. It leads to reasoning that is qualitatively different from the no-compromise approach associated with infinite values. However, as we have seen, the naïve infinite utilities model cannot claim Examples 1 and 2 as successes either, and it faces devastating objections in Examples 3 and 4. In the absence of any better way to represent infinite utilities, the CCS model is the best formal tool for approximately representing deontological reasoning within decision theory.

There is an alternative response. Some people in the deontological camp urge us to step back from the use of formal models altogether. Here is a representative opinion [Ackerman and Heinzerling 2002: 1576]:

> The real debate is not between rival cost-benefit analyses. Rather, it is between environmental advocates who frame the issue as a matter of rights and ethics, and others who see it as an acceptable area for economic calculation. That debate is inescapable, and it logically comes before the details of evaluating costs and benefits.

Ackerman and Heinzerling offer a set of principles and strategies. In our view, however, it is just as troubling for these hierarchical theories to operate with a framework that rules out utilitarian reasoning *a priori* as it is for the cost-benefit model to rule out deontological reasoning *a priori*. We should aim for a decision theory that accommodates both types of reasoning.

The challenge can now be formulated clearly. We want to define a utility function that accomplishes two things:

(i) *Representation of infinity*. It allows us to assign infinite value to certain outcomes or parts of nature.
(ii) *Discrimination*. It qualifies as a utility scale by satisfying the discrimination requirements (U1) and (U2).

In the next two sections, we show how this can be done.

6. Lexicographic Utilities

As a first example of a lexicographic ordering, consider the ordering $<^L$ on $\mathbb{R}^2$, the set of ordered pairs of real numbers:

$$(x_1, x_2) <^L (y_1, y_2) \text{ if and only if } x_1 < y_1 \text{ or } (x_1 = y_1 \text{ and } x_2 < y_2).$$

The relation between any two vectors in $\mathbb{R}^2$ is determined by the ordering of their first components except in the case of a tie, where it is determined by the ordering of their second components. This idea can be extended to define a standard lexicographic ordering on any finite-dimensional vector space. To determine the order relation between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$, compare the first components x_1 and y_1, then the second components x_2 and y_2, and so forth for all n components. Then $\boldsymbol{x} <^{L} \boldsymbol{y}$ if and only if $x_k < y_k$ for the first component k where they differ.[15]

We shall make use of an important consequence of a representation theorem due to Hausner [1954]: if an agent's preferences over gambles based on a finite set A of $(n+1)$ pure outcomes satisfy the standard axioms of decision theory apart from *Continuity*, then those preferences can be represented by an n-dimensional lexicographic ordering. This means that there are linear real-valued functions $u_1, \ldots, u_n$ of gambles based on the outcomes in A, such that, for any gambles x and y,

$$x \succ y \quad \text{if and only if} \quad (u_1(x), \ldots, u_n(x)) >^{L} (u_1(y), \ldots, u_n(y)).$$

In effect, the utility of each outcome is an n-dimensional vector, and the preference ordering on outcomes corresponds to the standard lexicographic ordering on $\mathbb{R}^n$.

It is easy to represent relatively infinite utilities using a lexicographic ordering. The first dimension of utility is infinitely more valuable than the second, the second is infinitely more valuable than the third, and so forth. Thus, if $u_1(x) = 1$ and $u_1(y) = 0$, then any gamble with positive probability p of x is preferred to y.[16] In this case, x has infinite utility relative to y.

Returning to the examples from sections 3 and 5, we find that lexicographic orderings allow us to represent the deontological reasoning while avoiding the objections raised by Colyvan et al. to infinite utility.

Example 1: Let $B \equiv$ *Allied victory*, $A \equiv$ *Nazi victory* and Z $\equiv$ *The atmosphere is destroyed*. Consider Compton's assertion that he would prefer A to any chance of Z:

$$A \succ [pB, (1-p)Z] \text{ for any } p < 1.$$

We can represent this situation with two-dimensional lexicographic utilities:

$$u(B) = (1, S)$$
$$u(A) = (1, -N)$$
$$u(Z) = (0, -M).$$

Intuitively, the first dimension represents the viability of the atmosphere (1 if viable, 0 if not) while the second represents the social desirability of the outcome ($S > -N > -M$). Assuming that each component of the utility function is linear,

$$\begin{aligned} u([pB, (1-p)Z]) &= (p, pS + (1-p)(-M)) \\ &<^{L} u(A) \end{aligned}$$

whenever $p < 1$. So, Z is infinitely worse than A. Hence, if $Pr(Z \,/\, Build\ bomb) = (1 - p) > 0$, then

$$\mathrm{EU}(\textit{Build bomb}) = u([pB, (1-p)Z]) <^{L} u(A),$$

so the bomb should not be built.[17]

[15] Lexicographic orderings can be defined in a much more general way (see Fishburn [1974]), but here we need only the finite-dimensional case.

[16] This is a consequence of the linearity of each u_i.

[17] Bayesian epistemologists often embrace the epistemic norm of *Regularity*: a rational agent should assign non-zero credence to any proposition that is not a logical or mathematical impossibility. The application of this norm does not undermine our analysis of *Example 1*: if $Pr(Z \,/\, Build\ bomb) > Pr(Z) > 0$, a similar calculation still shows that the bomb should not be built. The same point applies to the analysis of this example with relative utilities, in section 7.

We omit the analysis of *Example* 2, which is parallel to the analysis of *Example* 1.

Example 3: Let $W \equiv$ *The whole forest is saved*, $H \equiv$ *Half the forest is saved*, and $D \equiv$ *The forest is destroyed*. We can represent the situation using two-dimensional lexicographic utilities:

$u(W) = (1, a)$
$u(H) = (^1\!/_2, b)$
$u(D) = (0, c).$

Think of the first dimension as representing the portion of the forest that is saved, while the second dimension represents other social and economic considerations that are (on this model) infinitely less important. The lexicographic representation clearly reflects the preference ordering: $W \succ H \succ D$. Furthermore, the probabilistic version of the example amounts to a choice between two gambles: $G_W = [pW, (1-p)D]$, chance p of saving the whole forest, and $G_H = [pH, (1-p)D]$, chance p of saving half the forest. Since $u(G_W) = (p, pa + (1-p)c)$ and $u(G_H) = (^1\!/_2 p, pb + (1-p)c)$, it is clear that $u(G_W) >^L u(G_H)$. The problem of non-discrimination, raised by Colyvan et al. [2010a], does not arise.

Example 4: Let $S \equiv$ *The species is saved*, let P represent the *passive act* of doing nothing, and let I represent *active intervention* to save the species. As earlier, $Pr(S / P) = 0.01$ while $Pr(S / I) = 0.9$. Once again, we can represent the relatively infinite value of S using two-dimensional lexicographic utilities:

$u(S) = (1, a)$
$u(\sim S) = (0, b)$

The first dimension represents the survival of the species (1 if it survives, 0 if not) and the second represents other considerations. In this case, $\mathrm{EU}(P) = (0.01, 0.01a + 0.99b)$ while $\mathrm{EU}(I) = (0.9, 0.9a + 0.1b)$, and we have correctly represented the case: active intervention is preferable to doing nothing. Lexicographic utilities reflect the relatively infinite value of saving the species, and they correctly discriminate between the two relevant acts.

It appears that lexicographic utilities are well-suited to meet the challenge formulated at the end of section 5. They answer the main objections raised against infinite value by Colyvan et al. [ibid.]. They enable us to model deontological reasoning precisely. Furthermore, lexicographic models provide a generalization of standard decision theory: Hausner's representation theorem guarantees that, in any situation where *Continuity* is violated (but the other standard preference axioms are not), we can find a representation of the agent's preferences in terms of lexicographic utilities. It is tempting to conclude that the lexicographic approach fully solves the problem of modelling parts of nature with relatively infinite value.

Against this conclusion, there is an important philosophical difficulty with the lexicographic approach. In specifying a lexicographic representation, we have to fix a definite number of dimensions, n, for the utility function. The immediate difficulty is that of how to justify the choice of n: any fixed number of dimensions can appear arbitrary and rigid. To illustrate the problem, consider the following example, based on the discussion of lexicographic (or 'lexical') orderings in [ibid.].

Example 5. Saving multiple species. Let $B \equiv$ *Two species are saved*, $A \equiv$ *One species is saved*, and $Z \equiv$ *No species is saved*. Certainly, $B \succ A \succ Z$, but is each outcome in the series infinitely better than the next? Colyvan et al. [ibid.: 227] allege that we are no better off modelling this situation lexicographically than using naïve infinite utilities:

> A lexical ordering might hold that actions with a minute chance of preserving two endangered species are always better than actions that guarantee saving one species. Such results are counterintuitive and counterproductive. … Sensible nontrivial trade-offs between outcomes at different lexical levels must be possible, but this is precluded by both the explicitly infinitary and lexical approaches to environmental value.

In order to evaluate this argument, let us represent the situation with three-dimensional lexicographic utilities:

$$u(B) = (1, 0, b)$$
$$u(A) = (0, 1, a)$$
$$u(Z) = (0, 0, z)$$

Intuitively, the first dimension represents the preservation of two species, the second represents the preservation of one species, and the third dimension rates outcomes in terms of their associated economic cost (here, $b < a < z$ since the less we do, the lower the cost). Under this representation, an act with a tiny chance of achieving B is indeed ranked ahead of an act that guarantees A. If $p>0$, then the gamble $[pB, (1-p)X]$ has higher utility than A, no matter what X is.

However, we could provide a different lexicographic representation of the situation. We could add an initial component that represents *preservation of at least one species*: 1 for 'yes' and 0 for 'no'. By making this the first component of a four-dimensional utility vector, we give highest priority to saving at least one species. The representation becomes this:

$$u(B) = (1, 1, 0, b)$$
$$u(A) = (1, 0, 1, a)$$
$$u(Z) = (0, 0, 0, z)$$

Under this new representation, an act with a tiny probability of achieving B is not necessarily ranked ahead of A. It depends on what happens if the act fails! In particular, the gamble [pB, $(1-p)$Z] has lower utility than A. Hence, the argument of Colyvan et al. [ibid.] is invalid.

This technical 'solution' to the objection raised by Colyvan et al., however, drives home the problem of arbitrariness in lexicographic representations. Imagine that, after we have settled on four-dimensional utilities, somebody asks us how we should represent our preference for $C \equiv$ *Three species are saved*. It seems that we now have to make our utilities five-dimensional. We can foresee that we might need to add more dimensions later, to take additional outcomes into account. Although a lexicographic representation can always be found, any particular choice appears arbitrary and unstable under the introduction of new outcomes.

This problem has no analogue in standard decision theory: with *Continuity* in place, preferences are represented with a scalar utility function. When a new outcome is introduced, we don't have to change the old utility values. We simply slot the new outcome into place.

7. Relative Utilities

7.1 Definition of Relative Utility

Relative utility theory (RUT) and the lexicographic approach have much in common.[18] Both are generalizations of standard decision theory: they represent preference orderings that satisfy all of the standard axioms apart from *Continuity*. Where preferences satisfy *Continuity*, RUT and the lexicographic approach both agree with standard decision theory. Furthermore, like the lexicographic approach, RUT has a representation theorem. If an agent's preferences satisfy the standard axioms apart from *Continuity*, then there is a *relative* utility function that represents those preferences such that rational choices maximize *relative* expected utility.

What are relative utilities? Let $\preccurlyeq$ be a weak preference ordering among outcomes. Whenever $Z \preccurlyeq A$ and $Z \preccurlyeq B$, there will be a unique number $U(A, B; Z)$, with $0 \leq U(A, B; Z) \leq \infty$, that we call the utility of A *relative to* B with *base-point* Z. $U(A, B; Z)$ is a three-place function, and its values are non-negative extended real numbers. In order to define this function, we rely upon the following proposition [Fishburn 1974], which assumes that the agent's preferences satisfy all standard axioms apart from *Continuity*:

Proposition. If $Z \preccurlyeq A \preccurlyeq B$, then there is a unique number λ, $0 \leq \lambda \leq 1$, such that the agent prefers A to any gamble $[pB, (1-p)Z]$ when $p < \lambda$ and prefers $[pB, (1-p)Z]$ to A if $p > \lambda$.

Think of gambles between Z and B as points along the interval between 0 and 1. The *Proposition* delivers a unique value λ ($0 \leq \lambda \leq 1$) such that points to the left of λ are inferior to A and points to the right of λ are superior to A. The relative utility, λ, tells us where A is located in this interval. With reference to *Figures 1* and *2*, the relative utility $U(A, B; Z)$ may be pictured as the ratio of the intervals $Z-A$ and $Z-B$.[19] Relative utilities are thus like ratios of utility differences, and relative utility theory (RUT) works directly with these generalized utility ratios. The only change from section 4 is that, since *Continuity* is not assumed, the agent may or may not be indifferent between A and $[\lambda B, (1\text{-}\lambda)Z]$. The following definition shows that this idea is easily extended to relative utilities greater than 1.

Definition of relative utility. Suppose that a preference ordering $\succcurlyeq$ satisfies the standard axioms with the possible exception of *Continuity*, and that $B \succcurlyeq Z$ and $A \succcurlyeq Z$.

(1) If $B \succcurlyeq A \succcurlyeq Z$, then
$U(A, B; Z) = \lambda$ iff $A \succcurlyeq [pB, (1-p)Z]$ whenever $p < \lambda$ and $[pB, (1-p)Z] \succcurlyeq A$ whenever $p > \lambda$.

(2) If $A \succcurlyeq B \succcurlyeq Z$, then
$U(A, B; Z) = 1 / U(B, A; Z)$ (where 0 and ∞ are taken as reciprocals).

Three special cases. The definition includes three important special cases: relative utilities of ∞, 0, or 1. We give a picture for each of these cases, once again representing discontinuity by the zigzag: outcomes on the right are infinitely preferred to those on the left.

[18]The relative utilities framework is developed in Bartha [2007, forthcoming]. The term *relative utility theory* (RUT) was coined by Colyvan and Hájek [forthcoming].

[19]Indeed, if a standard one-dimensional utility u can be defined, then $U(A, B; Z) = [u(A) - u(Z)] / [u(B) - u(Z)]$.

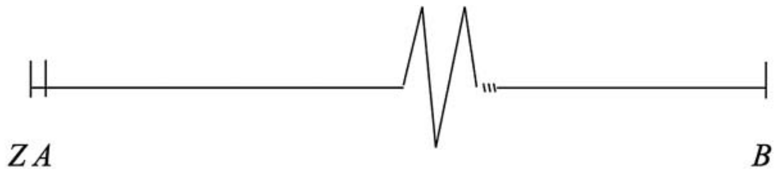

Figure 3: Infinite relative utility

Case 1: *Infinite relative utility* (of B relative to A and Z).

$$U(B, A; Z) = \infty \quad \text{iff} \quad [pB, (1-p)Z] \succcurlyeq A \quad \text{for } 0 < p \leq 1.$$

Figure 3 depicts the situation whereby any gamble between B and Z that offers a positive probability of B is preferred to A. (The 'distance' from Z to B is infinitely greater than the 'distance' from Z to A.)[20]

Case 2: *Zero relative utility* (of A relative to B and Z).

$$U(A, B; Z) = 0 \quad \text{iff} \quad [pB, (1-p)Z] \succcurlyeq A \quad \text{for } 0 < p \leq 1.$$

This is equivalent to *Case 1*, swapping the positions of A and B: any gamble between B and Z that offers a positive probability of B is preferred to A.

Case 3: *Relative utility of 1*.

$$U(A, B; Z) = 1 \quad \text{iff} \quad B \succcurlyeq [pA, (1-p)Z] \text{ and } A \succcurlyeq [pB, (1-p)Z], \text{ for } 0 \leq p < 1.$$

As depicted in *Figure 4*, the agent prefers B to any non-trivial gamble between A and Z, but also prefers A to any non-trivial gamble between B and Z. Even though B is strictly better than A, the agent is unwilling to take a gamble with a positive chance of getting Z if she can have A for sure.

Comparison to lexicographic approach. It is helpful at this point to compare relative utilities to the lexicographic approach. The main difference is that relative utilities represent only partial information about an agent's preferences, a kind of first-order comparison of three outcomes.[21] This gives relative utilities a 'local' feel: we can compare any three outcomes without worrying about all of the other possibilities. As we shall see, this information is typically sufficient for making decisions. The main benefit is that, in contrast to the lexicographic approach, we don't have to stipulate a fixed number of dimensions or indicate what those dimensions mean. This means that RUT does not face the problems of arbitrariness and instability noted at the end of section 6. The representation of relative utilities does not change when we introduce additional outcomes.

To appreciate the contrast between relative utilities and lexicographic utilities more clearly, consider the special case where $U(A, B; Z) = 1$ even though B is strictly preferred to A. This situation is compatible with many different lexicographic utility assignments. For instance, any of the following three sets of lexicographic assignments corresponds to $U(A, B; Z) = 1$, since in each case the agent will prefer A to any non-trivial gamble between B and Z:

[20] This definition of infinite relative utility differs slightly from the one in section 4, by allowing for cases where $Z \approx A$ and/or $A \approx B$. If $Z \approx A$, then $U(B, A; Z) = \infty$ for any $B \succcurlyeq A$; this is a degenerate case of infinite relative utility.

[21] If we have a lexicographic representation, then relative utilities represent information about how three outcomes compare along the dominant dimension where they exhibit any difference.

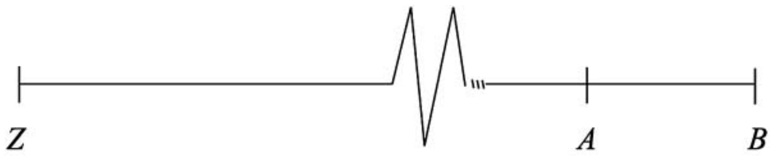

Figure 4: Relative Utility of 1

Assignment 1	Assignment 2	Assignment 3
$u(B) = (1, 5)$	$u(B) = (1, 2, 0)$	$u(B) = (1, 2, 2)$
$u(A) = (1, 4)$	$u(A) = (1, 1, 0)$	$u(A) = (1, 2, 1)$
$u(Z) = (0, 1)$	$u(Z) = (0, 1, 0)$	$u(Z) = (1, 0, 1)$

This contrast may at first seem to favour the lexicographic approach, which lets us distinguish between three cases that 'look the same' on the relative utilities approach (if the only outcomes that we care about are A, B, and Z). As noted, however, the value $U(A, B; Z)$ is typically all that we need to make a decision. In such cases, the problem of choosing a lexicographic representation may be distracting, and any particular choice may seem arbitrary or misleading, as noted in section 6.

Properties of relative utilities. We briefly discuss two important properties of relative utilities that will be useful in the next sub-section.

(1) $A \succcurlyeq B$ if and only if $U(A, B; Z) \geq 1$.

If $A \approx B$, then $U(A, B; Z) = 1$. But $U(A, B; Z) = 1$ need not imply $A \approx B$, as is clear from *Case 3* above. For a fixed base-point Z, $U(A, B; Z)$ does not always discriminate between distinct outcomes: we may have $U(A, B; Z) = 1$ even though the agent strictly prefers B to A. Still, if the agent strictly prefers B to A, there is *some* base-point Z such that $U(A, B; Z) < 1$.[22] Hence, relative utilities meet the challenge outlined at the end of section 4: they enable us to represent infinite relative utility, and they discriminate between non-equivalent outcomes.

(2) The relative utility function U is linear in the first component.

For any A, A', B, and Z:
$U([pA, (1-p)A'], B; Z) = pU(A, B; Z) + (1-p)U(A', B; Z)$.

In light of linearity, relative utilities of complex gambles can be computed using relative expected utilities, as we show next.

7.2 Working with Relative Utilities: Relative Decision Matrices

A slightly artificial example illustrates how relative utilities allow us to model deontological reasoning about the natural environment. Imagine a debate about the fate of a wilderness area that contains a valuable but non-renewable resource. We consider five options, represented in terms of a *finite* (one-time) value from extracting the non-

[22]In particular, we could take $Z = A$, since $U(A, B; A) = 0$.

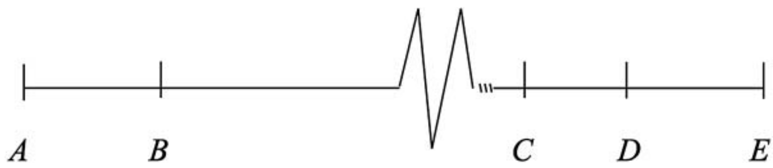

Figure 5: Wilderness example options

renewable resource, plus a *recurring* benefit based on the proportion of unspoiled wilderness.

Option	**Finite value**	**Recurring benefit (% unspoiled)**
A	$0	None
B	$10 billion	None
C	$1 billion	10%
D	$100 million	50%
E	$0	100%

A represents the worst possible outcome: the wilderness area is destroyed and nothing is extracted. On option *B*, maximal value is extracted but the wilderness is destroyed. Options *C* and *D* represent partial extraction; on option E, the wilderness is left untouched.

On the cost-benefit approach, we would assign a dollar value to the recurring benefit (and hence to the total value) for each option. Let's assume, however, that policy-makers want to model the wilderness as infinitely more valuable than the non-renewable resource, while still allowing for comparisons between options *C*, *D*, and *E*. The situation is as depicted in Figure 5.

This situation can be represented with the following relative utilities:

(1) $U(B, C; A) = 0$: *C* is infinitely preferred to *B*. Also, $U(B, D; A) = U(B, E; A) = 0$.
(2) $U(D, E; A) = 1$. Choose *D* over any non-trivial gamble between *E* and *A*.
(3) $U(D, E; C) = 4/9$. Some gambles between *E* and *C* are better than *D*.

In light of the linearity property, it is convenient to represent these values using a *relative decision matrix*. We display the utility of each act/state combination relative to the *best outcome* in the table and a *fixed base-point Z*. For each outcome *O*, the relative decision matrix displays the value $\lambda = U(O, Best; Z)$. This value represents only the dominant dimension of comparison, but the set of these values is often sufficient information to make the best choice.

Consider the wilderness example. Suppose that policy-makers have a choice between guaranteeing *D* and *Coin Toss*, where the latter option is to toss a fair coin and select *E* on *Heads*, *C* on *Tails*. (Think of *Coin Toss* as a metaphor that stands for a policy with a 50% chance of resulting in each of *E* and *C*.) Here, the best outcome is *E* and the worst is *C*, so the relative decision matrix has values $U(E, E; C) = 1$ and $U(C, E; C) = 0$ on the top row, and $U(D, E, C) = 4/9$ on the bottom row:

	Heads	*Tails*
Coin Toss	1	0
Choose D	4/9	4/9

Computing relative expected utilities:

$U(Coin\ toss, E; C) = \frac{1}{2}(1) + \frac{1}{2}(0) = \frac{1}{2}$
$U(D, E; C) = 4/9$

Hence, *Coin Toss* is the better choice. The relative expected utility calculation gives us the preferences of the policy-makers in this situation.

This example illustrates some of the advantages that relative infinite utilities have over other representations of infinite value. With naïve infinite utilities, calculations yield infinite expected utility for *C*, *D* and *E* (and also for *Coin Toss*), which wrongly suggests indifference among these outcomes. With lexicographic utilities, we can reproduce the result just obtained, but it would take some ingenuity to ensure faithful representation of (2) as well.[23]

7.3 Application to Examples

In this section, we apply RUT to the examples from sections 3 and 5. We briefly discuss *Examples* 1−5 to show that RUT handles each one in a way that parallels the lexicographic approach, although arguably the representation with relative utilities is simpler.

Example 1 (*The Atmosphere*): As before, let $B \equiv$ Allied victory, $A \equiv$ Nazi victory, and $Z \equiv$ destruction of the atmosphere. The picture is the same as in *Figure 4*. We have the following relative utilities:

$U(A, B; Z) = 1$
$U(B, B; Z) = 1$
$U(Z, A; Z) = 0$

The relative decision matrix for this decision is the following:

	Combustible atmosphere		*Non-combustible atmosphere*	
	A	*B*	*A*	*B*
Build the bomb	0	0	1	1
Don't build the bomb	1	1	1	1

If $Pr(Z\ /\ Build\ Bomb) = p > 0$, then the relative expected utilities are

$U(Build\ bomb, B; Z) = 1-p$
$U(Don't\ build\ bomb, B; Z) = 1$

and hence the bomb should not be built.

Example 2 (*Northern Gateway Pipeline*): Suppose, for the sake of argument, that the best possible outcome is $B \equiv$ *the Pipeline is built and there is no oil spill*. Let $Z \equiv$ *Disastrous oil spill* and let $A \equiv$ *Status quo (no pipeline)*. The picture is again as in *Figure 4*, and, relative to the best outcome *B* and base-point *Z*, we have the following relative decision matrix:

	Oil spill	*No spill*
Build Northern Gateway	0	1
Don't build it	—	1

[23] One natural lexicographic representation is as follows: $u(A) = (0, 0)$, $u(B) = (0, 10000)$, $u(C) = (0.1, 1000)$, $u(D) = (0.5, 100)$ and $u(E) = (1, 0)$. This yields the preference for *Coin Toss* over *D*, but fails to yield a preference for *D* over all gambles between *A* and *E*.

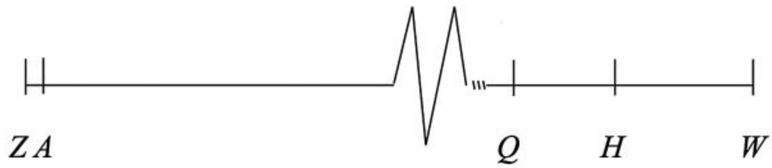

Figure 6: Mangrove forest

The key assumption in this reasoning is that $U(A, B; Z) = 1$: these are the preferences of somebody who thinks that no gamble with any positive chance of a spill is preferable to the status quo. In this case, we maximize relative expected utility by *not* building the pipeline.

Example 3 (*Mangrove forest*): As before, let $H \equiv$ *Half the forest is saved*, $W \equiv$ *The whole forest is saved*, and let $A \equiv$ *The forest is destroyed*. Let Z be any base-point worse than A. We model the assumption that both H and W are infinitely better than A by

$$U(H, A; Z) = U(W, A; Z) = \infty.$$

We can also model the assumption that we are unwilling to trade H for any gamble that might result in its destruction:

$$U(H, W; Z) = U(H, W; A) = 1.$$

In order to discriminate between H and W, consider a different base-point, $Q \equiv$ *One-quarter of the forest is saved*. Imagine that our preferences are as pictured in Figure 6. In this case, we have the following:

$$0 < U(H, W; Q) < 1.$$

There is some non-trivial gamble between W and Q that is preferred to H.

In the probabilistic version of the example, the choice is between the two gambles $G_W = [pW, (1-p)A]$ and $G_H = [pH, (1-p)A]$. If the base-point is A, then $U(G_H, G_W; A) = 1$: we fail to discriminate between the two gambles, since both are infinitely better than A. But if instead the base-point is $G_Q \equiv [pQ, (1-p)A]$, then

$$U(G_H, G_W; G_Q) = U(H, W, Q),$$

a value between 0 and 1. Given a suitable choice of the base-point, RUT discriminates between the two gambles and clearly prescribes G_W over G_H.

Example 4 (*Endangered species*): Let $S \equiv$ *The species is saved*, so that $\sim S$ represents the outcome where the species is not saved. As before, P represents the *passive act*, while I represents *active intervention* to save the species, and we assume that $Pr(S \mid P) = 0.01$ and $Pr(S \mid I) = 0.9$. We want to represent that S is infinitely better than $\sim S$, and also to prefer the act I with a higher probability of bringing about S. To represent the situation, we may suppose that the best outcome is $B \equiv S \,\&\, P$, while the worst is $Z \equiv \sim S \,\&\, I$. Figure 7 provides the picture that goes with the intuitions about relative infinite value. The relative utilities here are as follows:

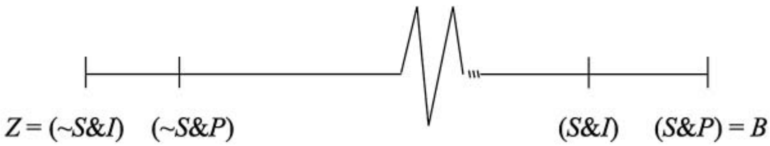

Figure 7: Endangered species

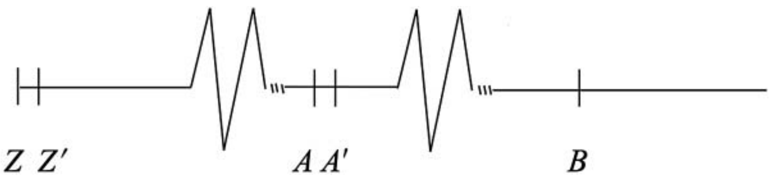

Figure 8: Multiple endangered species

$$U(S \,\&\, P, B; Z) = 1$$
$$U(S \,\&\, I, B; Z) = 1$$
$$U(\sim S \,\&\, P, B; Z) = 0$$
$$U(\sim S \,\&\, I, B; Z) = 0$$

Hence, the relative decision matrix is

	S **(*saved*)**	**$\sim$*S*** **(*not saved*)**
P **(*passive*)**	1	0
I **(*active*)**	1	0

Computing relative utilities, we have

$$U(P, B; Z) = 0.01 \text{ and } U(I, B; Z) = 0.9.$$

In agreement with the obvious intuition, we should prefer active intervention.

Example 5 (*Multiple endangered species*): Let $B \equiv$ *Two species are saved*, $A \equiv$ *One species is saved*, and $Z \equiv$ *No species is saved*. If our intuition is that B is infinitely preferable to A, which in turn is infinitely preferable to Z, then the picture is as shown in Figure 8.

In order to represent this situation with relative utilities, we need at least two outcomes Z, Z' where no species are saved, and two outcomes A, A' where one species is saved, as shown. We then have

$$U(A, Z'; Z) = U(B, A'; A) = \infty.$$

At the same time, we are unwilling to exchange A for a tiny chance to achieve B:

$$U(B, A; Z) = 1.$$

At the end of section 6, we achieved the same result with a complicated lexicographic representation, but we noted that introducing additional outcomes (such as $C \equiv$ *Three species are saved*) requires a completely different representation. No such revision is needed with relative utilities. We simply add $U(C, B; Z) = U(C, A; Z) = 1$, with no change to existing relative utility assignments. In this respect, relative utility behaves like a scalar utility function.

8. Conclusion

Some decisions require us to think about actions that may cause irreversible damage to, or loss of, a precious part of the natural world. The lexicographic and relative-utilities frameworks provide helpful ways of modelling the idea that such a part of nature is infinitely valuable *relative* to something else. As we have seen, these approaches don't suffer from the defects alleged by critics of infinite utility. Nevertheless, in this final section, we identify two significant concerns that apply to relative infinite utilities.

The first concern revisits a point raised in section 4: any type of infinite utility induces paralysis into day-to-day decision-making. Although we have found ways to avoid the absurd consequences of naïve infinite-value models, a decision to model a part of nature as infinitely valuable in a *relative* sense may still have stark implications. Large deposits of natural gas have been discovered in wilderness areas of British Columbia. The development of the LNG (liquid natural gas) industry could add billions of dollars to the provincial economy, but suppose that extraction poses a miniscule risk of irreversible damage to these wilderness areas. Many people would be willing to take that risk. But if we treat the wilderness as infinitely valuable relative to the development of the LNG industry, our hands are tied. We should not take any chances. This example is structurally identical to *Examples 1* and *2*.

More generally, everyday use of infinite-value models threatens to lead to a radically impoverished quality of life. I would like to go to the movies this weekend. The short drive to the movies entails a slight but definite risk of a fatal car crash. I value my life infinitely, relative to watching a movie. According to RUT, it seems that I should stay home.[24] This is a very counter-intuitive result. We take daily risks with our lives and with the lives of our loved ones.

The second concern is closely related. Each example that we have considered pits the protection of some part of nature against some other good. Perhaps it is plausible that this part of nature is infinitely valuable relative to that other good. In complex cases, however, human livelihoods and possibly human lives are at stake, no matter which option we choose. This was clearly the case in deciding whether to halt research on the atomic bomb. It is arguably the case for *Northern Gateway*. In these decisions, we are pitting priceless outcomes against each other. The concern here is that, when we evaluate the environment as relatively infinite in value, we may be underestimating the value of non-environmental goods.

We offer the following considerations in response to the first concern. First, infinite-value models are not always appropriate. People need not take their lives, the environment, or anything else to be infinitely valuable relative to some other objective. If we are willing to take chances with the environment, then we don't regard it as having relatively infinite value. In such cases, the cost-benefit approach is just fine (and is consistent with our approach).

Second, we suggest that a fully developed model of decision-making should take into account the context of each decision. Two similar decisions might be modelled with relatively infinite utilities in one case but not the other. In case A, I am presented with a stark choice between a dull afternoon at home and a roll of the dice to determine whether I enjoy a pleasant experience at the movies (99.99% chance) or face an

[24] Suppose that *U*(*Stay Home*, *Movies*; *Car Crash*) = 1 and *U*(*Car Crash*, *Movies*; *Car Crash*) = 0. It follows that *U*([p *Movies*, $(1-p)$ *Car Crash*], *Movies*; *Car Crash*) = p. If $p < 1$, then the gamble is inferior to *Stay Home*.

immediate painful death (0.01% chance). I reject the gamble. In case B, just as I am stepping into my car to drive to the theatre, somebody reminds me of the risk of a crash (0.01% chance). I dismiss this worry and go to the movies.

Roughly speaking, the explanation might be that I prefer a dull risk-free existence only *finitely* over death, but I prefer an interesting life *infinitely* over a dull existence. If my decision to drive to the theatre in case B is part of a pattern of choices that make up a full life, then I don't view it as analogous to case A, which is an isolated decision. Perhaps the same point applies to many day-to-day decisions involving the environment. For a major policy decision affecting the environment, however, this line of thinking is much less plausible. Such decisions are vastly important one-of-a-kind choices that require careful consideration. RUT remains a valuable modelling tool for this type of case. We acknowledge, however, that we need a systematic account of how to combine infinite-value models with everyday decision-making.[25]

As for the second concern, there are indeed cases where lives and fundamental values are at stake, no matter what we decide to do. We suggest that RUT offers the flexibility to develop different models for thinking about such cases. If the survival of an ecosystem is weighed against extracting a valuable natural resource that could make the difference between a living wage and starvation for workers, we may model the ecosystem as infinitely valuable relative to some objectives but on par with or less valuable than others. In reflecting on such cases, RUT offers advantages over the lexicographic utilities approach. A representation with relative utilities is not tied to a fixed set of dimensions or a rigid hierarchy of outcomes. RUT thus has the kind of flexibility associated with the cost-benefit approach, yet it faithfully represents deontological reasoning. We conclude that RUT holds considerable promise for environmental decision-makers, as a framework within which both deontological and utilitarian considerations can find a place. [26]

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Ackerman, Frank and Lisa Heinzerling 2002. Pricing the Priceless: Cost-Benefit Analysis of Environmental Protection, *University of Pennsylvania Law Review* 150/5: 1553−84.

Bartha, Paul 2007. Taking Stock of Infinite Value: Pascal's Wager and Relative Utilities, *Synthese* 154/1: 5−52.

Bartha, Paul forthcoming. Making Do Without Expectations, *Mind.*

Bentham, Jeremy 1789 (1843). *An Introduction to the Principles of Morals and Legislation*, in *The Collected Works of Jeremy Bentham*, Vol. I, ed. John Bowring, London: William Tait, 1−154.

Broome, John 2004. *Weighing Lives*, Oxford: Clarendon Press.

Broome, John 2008. What is Your Life Worth? *Daedalus* 137/1: 49−56.

Bulte, Erwin H. and G. Cornelius van Kooten 2000. Economic Science, Endangered Species, and Biodiversity Loss, *Conservation Biology* 14/1: 113−19.

[25] It must also be conceded that the lack of any such substantive account is potentially a very serious drawback to such models.

[26] Versions of this paper were presented to the UBC Philosophy Department, to CLMPS 2015 in Helsinki, and to INEM at the University of Cape Town. We have benefitted greatly from comments from these audiences, and from suggestions by two anonymous referees for this journal.

Colyvan, Mark 2008. Relative Expectation Theory, *The Journal of Philosophy* 105/1: 37–44.
Colyvan, Mark, Damian Cox, and Katie Steele 2010b. Modelling the Moral Dimension of Decisions, *Noûs* 44/3: 503–29.
Colyvan, Mark and Alan Hájek forthcoming. Making Ado Without Expectations, *Mind*.
Colyvan, Mark, James Justus, and Helen M. Regan 2010a. The Natural Environment Is Valuable but Not Infinitely Valuable, *Conservation Letters* 3/4: 224–28.
Compton, Arthur Holley 1956. *Atomic Quest: A Personal Narrative*, Oxford: Oxford University Press.
Costanza, Robert, Ralph d'Arge, Rudolf de Groot, Stephen Farber, Monica Grasso, Bruce Hannon, Karin Limburg, Shahid Naeem, Robert V. O'Neill, Jose Paruelo, Robert G. Raskin, Paul Sutton, and Marjan van den Belt 1997. The Value of the World's Ecosystems Services and Natural Capital, *Nature* 387/6230: 253–60.
Daily, Gretchen C. (ed.) 1997. *Nature's Services: Societal Dependence on Natural Ecosystems*, Washington, D.C.: Island Press.
Diamond, Peter A. and Jerry A. Hausman 1994. Contingent Valuation: Is Some Number Better Than No Number?, *The Journal of Economic Perspectives* 8/4: 45–64.
Dreier, Jamie 1996. Rational Preference: Decision Theory as a Theory of Practical Rationality, *Theory and Decision* 40/3: 249–76.
Easwaran, Kenny. 2008. Strong and Weak Expectations, *Mind* 117/467: 633–41.
Easwaran, Kenny. 2014. Regularity and Hyperreal Credences, *The Philosophical Review* 123/1: 1–41.
Farber, Stephen C., Robert Costanza, and Matthew A. Wilson 2002. Economic and Ecological Concepts for Valuing Ecosystem Services, *Ecological Economics* 41/3: 375–92.
Feldman, Allan 1994. The Value of Life, Working Paper No. 94–21, *Department of Economics, Brown University*, URL=http://www.econ.brown.edu/fac/Allan_Feldman/efiles/PDFs/The%20Value%20of%20Life.pdf
Fishburn, Peter C. 1974. Exceptional Paper—Lexicographic Orders, Utilities and Decision Rules: A Survey, *Management Science* 20/11: 1442–71.
Hájek, Alan 1997. The Illogic of Pascal's Wager, in *Proceedings of the 10th Logica International Symposium*, Liblice, ed. Timothy Childers et al., Filosophia, The Institute of Philosophy of the Academy of Sciences of the Czech Republic: 239–49.
Hájek, Alan 2003. Waging War on Pascal's Wager, *The Philosophical Review* 112/1: 27–56.
Hausman, Jerry 2012. Contingent Valuation: From Dubious to Hopeless, *The Journal of Economic Perspectives* 26/4: 43–56.
Hausner, Melvin 1954. Multidimensional Utilities, in *Decision Processes*, ed. Robert M. Thrall, Clyde H. Coombs and Richard L. Davis, New York: John Wiley & Sons: 167–80.
Howarth, Richard B. 1995. Sustainability under Uncertainty: A Deontological Approach, *Land Economics* 71/4: 417–27.
Justus, James, Mark Colyvan, Helen Regan, and Lynn Maguire 2009. Buying Into Conservation: Intrinsic Versus Instrumental Value, *Trends in Ecology and Evolution* 24/4: 187–91.
Kareiva, Peter 2010. Conservation Science: Trade-in to Trade-up, *Nature* 466/7304: 322–23.
Kareiva, Peter and Michelle Marvier 2012. What is Conservation Science? *BioScience* 62/11: 962–69.
Luce, R. Duncan and Howard Raiffa 1957. *Games and Decisions: Introduction and Critical Survey*, New York: John Wiley & Sons.
McCauley, Douglas J. 2006. Selling Out on Nature, *Nature* 443/7107: 27–8.
Mill, John Stuart 1863 (1906). *Utilitarianism*, Chicago: University of Chicago Press.
Mill, John Stuart 1874 (2006). *Nature*, in *Collected Works of John Stuart Mill*, Vol. X, ed. John M. Robson, Indianapolis: Liberty Fund, 373–402.
Miller, Brian, Michael E. Soulé, and John Terborgh 2014. 'New Conservation' or Surrender to Development?, *Animal Conservation* 17/6: 509–15.
Pearce, D. 1998. Cost Benefit Analysis and Environmental Policy, *Oxford Review of Economic Policy* 14/4: 84–100.
Pearce, David W., Giles Atkinson, and Susana Mourato 2006. *Cost Benefit Analysis and the Environment: Recent Developments*, Paris: Organisation for Economic Co-operation and Development Publishing.
Resnik, Michael 1987. *Choices: An Introduction to Decision Theory*, Minneapolis: University of Minnesota Press.
Rhodes, Richard 1986. *The Making of the Atomic Bomb*, New York: Simon and Schuster.
Rudin, Walter 1976. *Principles of Mathematical Analysis*, 3rd edn., New York: McGraw Hill.

Sagoff, Mark 2004. *Price, Principle, and the Environment*, Cambridge: Cambridge University Press.
Sen, Amartya 2000. The Discipline of Cost-Benefit Analysis, *The Journal of Legal Studies* 29/S2: 931−52.
Skala, Heniz 1975. *Non-Archimedean Utility Theory*, Dordrecht: D. Reidel Publishing.
Soulé, Michael 1985. What Is Conservation Biology? A New Synthetic Discipline Addresses the Dynamics and Problems of Perturbed Species, Communities, and Ecosystems, *BioScience* 35/11: 727−34.
Soulé, Michael 2013. The 'New Conservation', *Conservation Biology* 27/5: 895−97.
Spash, Clive L., and Nick Hanley 1995. Preferences, Information, Biodiversity Preservation, *Ecological Economics* 12/3: 191−208.
Swift, Anthony, Nathan Lemphers, Susan Casey-Lefkowitz, Katie Terhune, and Danielle Droitsch 2011. Pipeline and Tanker Trouble: The Impact to British Columbia's Communities, Rivers, and Pacific Coastline from Tar Sands Oil Transport, New York: Natural Resources Defense Council, URL=http://www.livingoceans.org/media/reports-publications/pipeline-and-tanker-trouble
von Neumann, John and Oskar Morgenstern 1953. *Theory of Games and Economic Behaviour*, Princeton: Princeton University Press.