Infinite aggregation and risk

Hayden Wilkinson

Comments welcome: hayden.wilkinson@anu.edu.au

1 Introduction

Each of us ought to make the world better than it would otherwise be, at least somewhat and sometimes.

This need not be because a strict form of consequentialism is true. Perhaps the true moral theory upholds other considerations - constraints, prerogatives, and reasons from other sources. But it is implausible that this theory does not, in a large class of moral decisions, recognise some reasons to bring about the best available outcome.[1] For instance, when one devotes some of one's resources to aid strangers through charity, it seems that one has some reason to use those resources as effectively as possible - to bring about the best available outcome.[2]

On top of this, it is plausible that (at least in some decision contexts) one outcome can be better than another only if the resulting world contains a greater total aggregate of value, impartially construed.[3] Again, this need not be because *only* total value matters; some form of pluralism may be true. A range of the plausible moral theories are *minimally aggregative* in this sense: in a large class of comparisons, they evaluate one state of affairs as better than another only if it contains a greater total aggregate of value; and they entail that we have reason to bring

[1]Rawls (1971:30) himself claims that any plausible deontological theory satisfies this criterion: 'It should be noted that deontological theories are defined as non-teleological ones, not as views that characterise the rightness of institutions and acts independently from their consequences. All ethical doctrines worth our attention take consequences into account in judging rightness. One which did not would simply be irrational, crazy.'

[2]For a defence of that claim, see Pummer (2016).

[3]This definition excludes standard person-affecting views. It also excludes any view with a *monistic* axiology under which value does not admit an additively separable representation (e.g., egalitarianism, maximin, averagism). This exclusion is not because such views escape the infinitarian worries described below - typically, they don't - but for brevity. And it doesn't exclude views with a total utilitarian axiology or a prioritarian axiology in the style of Parfit (1997). Nor does it exclude pluralist views which give any weight at all to either of those axiologies.

about that better world.

Here I am interested in this category of minimally aggregative moral theories. This entire category faces ruinous problems when applied in a physical setting like ours.

Suppose that our universe has infinite spatial volume or temporal duration. Suppose further that it contains infinitely many moral subjects, infinitely many of whom have valuable lives and infinitely many of whom have disvaluable lives (with absolute value greater than some fixed, finite, positive ϵ). In such an infinite universe, the total sum of value minus disvalue is undefined. If this is the case in *all* worlds that our available actions would bring about, then we cannot say that *any* of those worlds contains greater total value than any other. By minimally aggregative views, none are better than any others: they are all incomparable to one another. So such views will give us no reason to bring about any of them over any other. And this extends to a large class of decisions in universes like this: perhaps all pure rescue cases, or all decisions of how to allocate resources for charitable purposes.

Current physical theories suggest that our universe is like this. That the universe has infinite spatial volume is independently implied by the *inflationary view* (Guth, 2007) and the *flat-lambda model* (Wald, 1983; Carroll, 2017), each widely accepted among physicists. And it has an infinite duration, over which life will continue to arise, according to the flat-lambda model. Both views imply that the universe's infinite volume will contain infinitely many tokens of every physically possible small-scale phenomenon, no matter what actions we take (Garriga & Vilenkin, 2001; Linde, 2007; Simone *et al*, 2010; Carroll, 2017). But there will be some physical phenomena which we consider morally valuable (and some disvaluable), e.g., perhaps a human brain experiencing intense pleasure (or pain) for a given duration.[4] So it seems plausible that we are indeed in a physical setting where all minimally aggregative views fail us. We then have reason to be sceptical of them or, if we are particularly confident in either physical theory, it seems that we should reject minimally aggregative views entirely.

But we might replace them with similar, modified views. By tweaking our method of aggregation, and comparing worlds based on something more than a standalone total value - one that is both one-dimensional and real-valued - perhaps something closely resembling a mini-

[4]Specifically, we will have a *countably* infinite number of such tokens. Why? Because they are each positioned in a four-dimensional spacetime. They'll also each occupy some (exclusive) finite region of spacetime - for illustration, for a human brain to experience a given quantity of pleasure, it requires some non-zero spatial volume and some non-zero, finite duration. So we can only fit a countably infinite number of those token events into the world.

mally aggregative view can avoid the problem. Fortunately, we have various proposals for doing this. We have the 'expansionist' methods of Vallentyne & Kagan (1997), Arntzenius (2014), and Wilkinson (2020). We have the weaker proposal of additivism from Lauwers & Vallentyne (2004), described below. We have Bostrom's (2011) hyperreal proposal. And we have Jonsson & Voorneveld's (2018) limit-discounting method, among various others. I'll remain mostly agnostic here on which, if any, of these is the correct method for comparing worlds.

But, whichever of those methods we choose, we face further problems in practice. As mere human agents, we are uncertain of the outcomes of our actions, for *all* actions. Even if we can compare infinite worlds, that alone provides no guidance for action in the relevant decisions. Put differently, even if we had a solution to the *axiological* problem of how to compare worlds, we do not immediately have a solution to the *subjective normative* problem of which risky actions we should take.

For those who hold aggregative views, the standard approach for converting axiology to subjective normative judgements is expected value theory: we evaluate the *expected* total value produced by each action, and are required to choose an action which maximises that expected value (Jackson, 1991). But, for infinite worlds, we often don't have defined total values over which we can take expectations. And by every one of the proposals listed above, we don't obtain any cardinal measure of worlds' value so, even then, we cannot construct expected values. Thus, minimally aggregative views (and even their modified infinitarian relatives) cannot give subjective normative judgements in many decision contexts.

I will propose a solution: the 'expectations of differences' approach. We can adhere to the spirit of expected value reasoning without explicitly assigning expected values to lotteries (the traditional, 'differences of expectations' approach). But in doing so we must satisfy certain crucial desiderata for any plausible ranking of lotteries. This is surprisingly difficult to do. (See Section 5.) And the one solution proposed so far, by Bostrom (2011), Arntzenius (2014), and Meacham (2020), does not satisfy them all.

The paper proceeds as follows. In Section 2, I introduce the necessary formal framework. In Section 3, I describe a minimal aggregation method for comparing *worlds*: Additivity. In Section 4, I describe the Arntzenius-Bostrom-Meacham method for comparing *lotteries*, and hence providing subjective normative judgements, for infinite worlds. In Section 5, I present a damning counterexample to their method, which demonstrates that it violates some basic desiderata for comparing lotteries - desiderata which are on much stronger footing than expected value theory itself. In Section 6, I propose my own solution: the 'expectations of differences' approach. (Readers averse to formalism may wish to skip this section, but will likely still find the rest of the paper illuminating.) Section 7 is the conclusion.

2 Preliminaries

Aggregation is the general method of evaluating worlds based on their total aggregates. The aggregate of a world is some impartial combination of all of the individual instances of value they contain - their *local values*. But how do we identify and demarcate local values? Following others in the literature, I assume that each local value is fundamentally associated with a token entity of some common type which can exist (or have unique counterparts) across different worlds. Those entities might be persons, or person-time-slices, or positions in space and time, or something else. They might have some essential and natural topological structure (as spacetime positions seem to), or they might not (as persons seem not to). There is disagreement in the literature on what the correct type is, and whether they have such structure.[5] Here I will remain agnostic on those questions. But, whatever those tokens are, call them *locations*.

For my purposes, each world W can be represented by an ordered pair $\langle \mathcal{L}, V \rangle$. The plurality of locations it contains is represented by an often-infinite set $\mathcal{L} = \{l_1, l_2, l_3, ...\}$ (in which the subscripts may be entirely arbitrary).And the value at each location is given on a cardinal scale by a function $V : \mathcal{L} \to \mathbb{R}$. I'll often use the subscript of each W_i to identify it with its set of locations $\mathcal{L}_i$ and value function V_i. And where all worlds available in a decision share the same locations, I denote the common location set as $\mathcal{L}$.[6]

I seek an 'at least as good as' relation for comparing these worlds, a binary relation $\succcurlyeq$ on the set $\mathcal{W}$ of (metaphysically) possible worlds. I assume that this relation is both reflexive and transitive[7] over $\mathcal{W}$. (Equivalently, we could say that $\succcurlyeq$ is a *preorder* on $\mathcal{W}$.) Strict betterness is given by the asymmetric component $\succ$, and equality by the symmetric component $\simeq$.

But we don't just want to compare worlds; we want to compare *lotteries* over worlds. For-

[5]For a defence of adopting persons as the appropriate type, see Askell (2018). For arguments in favour of adopting spacetime positions, see Wilkinson (2020; n.d.).

[6]As shorthand, I'll describe worlds as having 'the same locations' whether those locations are really the same or there is merely a unique, bijective counterpart relation between them.

[7]Some reject that a the moral betterness relation between worlds must be transitive (in particular, Temkin 2012). The primary motivation for this rejection is to accommodate some form of pluralism. But here, I am interested *only* in cases in which the only relevant considerations are aggregative ones. So I'll assume that $\succcurlyeq$ must be transitive.

mally, a lottery L is a probability measure on $\mathcal{W}$ - it maps all sets of possible worlds to some probability in the interval $[0,1]$ while obeying the standard probability axioms. The set of all such probability measures on $\mathcal{W}$ is denoted by $\mathcal{P}$. We can also define the *domain* of each lottery L_i by $\mathcal{W}_i = \{W \in \mathcal{W} | L_i(\{W\}) > 0\}$. For two lotteries L_i and L_j, I'll abbreviate the union of their domains $\mathcal{W}_i \cup \mathcal{W}_j$ to $\mathcal{W}_{(i,j)}$.

I'll make a few other abbreviations to keep the notation in check. I'll abbreviate $L(\{W\})$ to $L(W)$ when the input to L is a set $\{W\}$ with only element. And I'll use $L_i(\succcurlyeq W)$ as an abbreviation for the probability that L_i gives to the set of all worlds in $\mathcal{W}_i$ that are *at least as good as* W (or, equivalently, $\{W' \in \mathcal{W}_i | W' \succcurlyeq W\}$). And, last, when a lottery has only one world W in its domain, mapped to probability 1, I'll use W to denote both world and lottery.

You might think that lotteries don't contain all of the information relevant to comparing gambles. If not, we can also talk of the *states of the world* in which a particular world obtains. I have in mind the basic framework of Savage (1954), by which each lottery is associated with *act*: a function from states to their associated worlds (or, in Savage's terminology, *consequences*). For some set of mutually exclusive and exhaustive states $\mathcal{S}$, each act A can be defined as a function $A : \mathcal{S} \to \mathcal{W}$. We will also have some probability measure P on $\mathcal{S}$, independent of A. For each act A_i, the corresponding lottery L_i would be given by $L_i(E) = P(A_i^{-1}(E))$ for any subset E of its domain.

Finally, we need an 'at least as good as' relation for comparing lotteries (or acts): a binary relation $\succcurlyeq_L$ on $\mathcal{P}$. Again, define strict betterness ($\succ_L$) and equality ($\simeq_L$) as the asymmetric and symmetric components, respectively. And again, as basic desiderata, $\succcurlyeq_L$ must be reflexive and transitive. It must also be consistent with $\succcurlyeq$: if $W_1 \succcurlyeq W_2$ then $W_1 \succcurlyeq_L W_2$ too.

3 Comparing worlds: Additivity

The definitions of $\succcurlyeq$ and $\succcurlyeq_L$ above are minimal. We know little more than that they are reflexive and transitive. I'll now strengthen them slightly.

How? Suppose two worlds contain the same locations. Some locations might obtain more value in one world, some more in the other. Sum up all of those increases from one world to the other, minus all of the decreases. If that total sum of all differences is positive, then one world seems clearly better.

For instance, compare worlds W_1 and W_5. Their only differences are that location l_a does better in W_1, and that locations l_b to l_f do better in W_5. Perhaps this comparison represents a decision in which we can either rescue one person from death, or rescue five others (with all else equal).

$$\begin{array}{rcccccccccccc} & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & l_i & l_j & \cdots \\ W_1: & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \cdots \\ W_5: & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots \end{array}$$

We cannot say much about the total value of either world, but we can say something about the subtotal over *any finite subset* of locations. For instance, take the set of all locations which differ between worlds: l_a through l_f. The subtotal over these is 1 for W_1 and 5 for W_5. The worlds have equal values at all other locations, so let us ignore those. And, left with only those first few locations, the difference between W_5 and W_1 is $5 - 1 = 4$. Thus we might claim that W_5 is the better world.

That's *Additivity* in a nutshell. More precisely:

> *Additivity*[8]: For any worlds W_1 and W_2 with the same locations $\mathcal{L}$, $W_1 \succcurlyeq W_2$ if
>
> $$\sum_{l \in \mathcal{L}} V_1(l) - V_2(l) \geq 0$$
>
> (either by converging unconditionally, or by diverging unconditionally to $+\infty$).

Note that Additivity is not biconditional - some worlds may be better than others even though additivity doesn't hold. After all, Additivity by itself is not complete - nowhere near it - nor need it be. I introduce it here only as a weak principle to get us started, which is consistent with almost all of the stronger principles in the literature, but which lacks their controversiality.

But, aside from the lack of controversy, why accept Additivity? As Lauwers & Vallentyne (2004:39) show, any transitive $\succcurlyeq$ which satisfies all three of the following highly plausible principles will also satisfy Additivity.

> *(Strong) Pareto*: For any worlds W_1 and W_2 with the same locations $\mathcal{L}$, if, for all $l \in \mathcal{L}, V_1(l) \geq V_2(l)$, then $W_1 \succcurlyeq W_2$.
>
> If, as well, some $l_i \in \mathcal{L}$ has $V_1(l_i) > V_2(l_i)$, then $W_1 \succ W_2$.

[8]This principle is presented and defended by Vallentyne & Kagan (1997:11), Lauwers & Vallentyne (2004:23), and Basu & Mitra (2007).

This says that, if one world has contains at least as much value as another at *every single* location, then it's at least as good. And if that world has strictly *greater* value at at some locations, then it's strictly better. This sensitivity to changes in local value seems a minimal requirement for any comparison of worlds which stays true to the spirit of aggregation.

> *Separability of Value*: If W_1 and W_2 contain the same locations and $W_1 \succcurlyeq W_2$ then, adding their corresponding local values, $W_1 + W \succcurlyeq W_2 + W$ for all $W \in \mathcal{W}$ with the same locations.[9]

This principle, also called *Translation Scale Invariance* in the literature, ensures that we are *only* sensitive to differences in local value. If between one pair of worlds there is the same pattern of differences as between another pair of worlds, we must rank both pairs the same way. It does not matter what local values we start with in a world; all that matters is what we add or take away; if a certain combination of additions and removals is an improvement to a world, it would count as an improvement to any world. And, in the finite context, this is a key distinguishing feature of aggregative views - such as any which endorse totalism or total prioritarianism. So, moving to the infinite context, Separability of Value seems a necessary condition for staying true to the spirit of aggregation.

And the final condition we need, *Finite Sum*, requires only that $\succcurlyeq$ remains consistent with what our judgements would be if the world contained only finite total value. I take this as a crucial requirement for extensional adequacy.

> *Finite Sum*: If there is a finite total sum of local values in both W_1 and W_2, and the sum in W_1 is at least as great as that in W_2, then $W_1 \succcurlyeq W_2$.

I find all of these principles (and their conjunction) hard to deny. And so Additivity is hard to deny. Given this, it is unsurprising that every one of these principles, and hence Additivity too, is satisfied by each of the plausible stronger proposals in the literature (e.g., Vallentyne & Kagan 1997; Bostrom 2011; Arntzenius 2014; Jonsson & Voorneveld 2018; Wilkinson 2020). Given this broad agreement, I will assume in what follows that Additivity holds for $\succcurlyeq$. But I still remain agnostic as to how we should strengthen it, so my conclusions will still be relevant for all of those stronger views.

[9] Define addition of worlds as follows. For all worlds W_1 and W_2 with the same locations $\mathcal{L}$, the world $W_3 = W_1 + W_2$ is given by $V_3(l) = V_1(l) + V_2(l)$ for all $l \in \mathcal{L}$. Later, I'll define subtraction similarly.

4 Comparing lotteries: Local expectations

In this section, I'll describe a way to extend Additivity to compare lotteries over worlds, which is proposed in some form by both Arntzenius (2014) and Bostrom (2011), and also endorsed by Meacham (2020).

Consider the following two lotteries: W_1, which delivers world W_1 for sure; and L, which brings even odds of W_2 or W_3, as specified below.

$$\begin{array}{rcccccccccccc} & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & l_i & l_j & l_k & \cdots \\ W_1: & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & \cdots \end{array}$$

$$L \left\{ \begin{array}{c|rcccccccccccc} L(W) & & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & l_i & l_j & l_k & \cdots \\ 1/2 & W_2: & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & \cdots \\ 1/2 & W_3: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \end{array} \right.$$

Note that the labelling of locations here may be arbitrary: it need not represent any special structure in our locations (e.g., their position in time). So we can say that W_1 contains infinitely many locations with local value 1, and infinitely many with value 0, but we cannot say anything else about which of those sets of locations is larger.

If Additivity holds, then $W_2 \succ W_1 \succ W_3$. But it says nothing about W_1 versus the risky bet L. Using only Additivity, we cannot straightforwardly construct expected values: the differences between W_1 and each of the others have infinite sums, so there's no clear way to assign a real value to each option.

But here is one way we might still compare the two. In lottery L, each location l has its own prospects: probability $\frac{1}{2}$ of 2, and probability $\frac{1}{2}$ of 0. So we can list the expected *local* value for each l under lottery L.

$$\begin{array}{rcccccccccccc} & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & l_i & l_j & l_k & \cdots \\ \mathbb{E}_L(V(l_i)): & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots \end{array}$$

This defines a new object with the same structure as a world - an 'expected world', with the same set of locations and with a value function given by the *expected* local values under the lottery. Equivalently, $\mathbb{E}W_L = \langle \mathcal{L}, \mathbb{E}_L(V) \rangle$.

This 'expected world' can easily be compared to W_1, via Additivity (indeed, by Pareto alone). And it's strictly better. So we might say that $L \succ_L W_1$.

In effect, this matches the proposals made by Arntzenius[10], Bostrom[11] , and Meacham.To an approximation, they all endorse *Local Expectations*. (See notes for how they differ.)

> *Local Expectations*: For any lotteries L_1 and L_2 such that all $W \in \mathcal{W}_{(1,2)}$ contain the same locations $\mathcal{L}$, $L_1 \succcurlyeq_L L_2$ if $\mathbb{E}W_{L_1} \succcurlyeq \mathbb{E}W_{L_2}$.

If both Local Expectations and Additivity hold, then we can make the comparison between the above lotteries W_1 and L, as demonstrated. And, in general, we then have the result that $L_1 \succeq_L L_2$ if the following sum is (unconditionally) greater than or equal to 0 (or diverges unconditionally to $+\infty$).

$$\sum_{l \in \mathcal{L}} \mathbb{E}_{L_1}(V(l)) - \mathbb{E}_{L_2}(V(l))$$

This seems promising: it appears we can sidestep the whole problem of taking expectations over infinite totals. We just need to lower our expectations, down to the local level.

[10] Arntzenius (2014:55-6) proposes the following.

> *Weak Location Criterion*: For lotteries A and B, $A \succ B$ iff $\sum_{l \in \mathcal{L}} \mathbb{E}_A(V(l)) - \mathbb{E}_B(V(l))$ "...is absolutely convergent and > 0, where we are summing over all (epistemically possible) [locations]..." $l \in \mathcal{L}$.

This is *almost* equivalent to the conjunction of Local Expectations and Additivity. Except: 1) it defines only a strict betterness relation $\succ_L$, so does not imply that $L_1 \simeq_L L_2$ for any lotteries; 2) it remains silent if the sum diverges unconditionally to $+\infty$, even in cases of certainty in which Additivity gives a verdict. And 3) it allows worlds in both lotteries to contain different locations (e.g., the same people) summing value instead over all epistemically possible locations, which requires that we assign some value to the non-existent lives. I don't want to make a stand on such cases here, and my modifications to (1) and (2) should be uncontroversial.

[11] Bostrom's (2011:20-4) proposal is more complicated. He describes (but doesn't explicitly endorse) an approach by which we represent the total value of each world with a *hyperreal number*: a vector of (countably) infinite length which consists of the cumulative sums of local values, summed in some common order. In the example above, if we chose to sum in the order l_a, l_b, l_c, etc, then we would represent the total value of W_1 by the hyperreal $(1, 1, 2, 2, 3, 3, ...)$ and of W_2 by $(2, 4, 6, 8, 10, ...)$. And we can say that W_2 is better: if one hyperreal has entries larger than another at 'sufficiently many' positions, then it is the larger number. W_2 has larger entries than W_1 at *all* positions, so it has the larger total.

The standard of 'sufficiently many' can vary here but one very minimal condition is that, if the entries of one hyperreal are greater than another in all but finitely many positions, then that the former is the larger hyperreal. In effect, this condition is equivalent to Additivity. (There are stronger conditions we might apply too, but those can be chosen to make the hyperreal approach equivalent to any plausible aggregation rule we want - see Pivato 2014.)

We can also sum hyperreals and multiply them by real probabilities just as we can with vectors, and so they can give us expected values in roughly the old-fashioned way. L would then have expected value $(1, 2, 3, 4, 5, 6, ...)$, which is identical to the hyperreal total value of $\mathbb{E}W_L$ from above. Since this is larger than the total value of W_1, we could say that $L \succ_L W_1$. In effect, this is equivalent to applying Local Expectations as above, but under different formalism.

5 The problem

But there's a problem with Local Expectations. Consider the example of *Egregious Energy*.

> **Example: Egregious Energy**
>
> A new energy source has been discovered, and you must decide whether humanity takes advantage of it.
>
> If we do use it, once we get it working there we will obtain enormous amounts of energy and many lives will be improved. But there are a few downsides. One is that the fuel needed is limited - we will only reap its benefits for a short time. Another is that it may take a while to get it working but, the longer it takes, the longer that cornucopia of energy will last. But the greatest problem is pollution - this energy source produces a novel form of pollution which will badly harm human health. That pollution and its effects will decrease over time, but we will never be able to eradicate it. If we use this source of energy, that pollution will continue to harm human well-being for the entire future of humanity, which I'll assume will be infinitely long. In short, using this energy source will produce some finite benefit, but also cause an infinite total amount of harm. (The relevant probabilities and values are given below.)
>
> And, for simplicity, exactly the same persons (and person-time-slices) will exist at exactly the same physical positions whether or not we adopt this new energy source.

Since the same persons exist at the same physical positions either way, we can treat the resulting worlds as having the same locations. Then we can represent your options as W_0 and the lottery L. If you forego the energy source, you produce W_0, a world with constant cardinal value 0 at all locations, simply representing the baseline of what would have happened otherwise.[12] And if you choose to have humanity adopt it, you produce L, a lottery over infinitely many worlds $\{W_1, W_2, ...W_j, ...\}$ (each with probability $\frac{1}{2^j}$). You're uncertain of how long it takes the energy source to start working, during which time everyone just obtains the same value 0 as in W_0. Once

[12] Note that local values of 0 here do not imply that the lives in question are right on the boundary of not worth living. They might be extremely valuable lives. But these local values are here represented cardinally - the numbers only capture the relative size of the differences between them. So these same representations of W_0 and L could describe worlds in which everyone has a blissful life and suffers only a slight reduction in quality of life due to the pollution. Or they could describe worlds in which everyone suffers terribly and that pollution makes life even worse.

it's working, some number of people obtain some greater value, represented by 2. And then, once the fuel runs out, every person obtains less value than the baseline, represented by some negative number. Note that, in every single one of these worlds in L, the total value diverges unconditionally to $-\infty$. So they're all infinitely worse than W_0.[13] Note also that the sequence of locations $(l_1, l_2, ...)$ may be chronological, but that need not have any moral significance.

$$\begin{array}{rccccccccc} & l_1 & l_2 & l_3 & l_4 & l_5 & l_6 & l_7 & l_8 & \cdots \\ W_0: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \end{array}$$

$$L \left\{ \begin{array}{c|rcccccccccc} L(W) & & l_1 & l_2 & l_3 & l_4 & l_5 & l_6 & l_7 & l_8 & \cdots \\ 1/2 & W_1: & 2 & -1/2 & -1/2 & -1/4 & -1/4 & -1/4 & -1/4 & -1/8 & \cdots \\ 1/4 & W_2: & 0 & 2 & 2 & -1/4 & -1/4 & -1/4 & -1/4 & -1/8 & \cdots \\ 1/8 & W_3: & 0 & 0 & 0 & 2 & 2 & 2 & 2 & -1/8 & \cdots \\ 1/16 & W_4: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{array} \right.$$

We can use Local Expectations to compare L to W_0. As above, we take L's expected value for each location. We obtain $\mathbb{E}W_L$ as below.

$$\begin{array}{rccccccccc} & l_1 & l_2 & l_3 & l_4 & l_5 & l_6 & l_7 & l_8 & \cdots \\ \mathbb{E}W_L: & 1 & 1/4 & 1/4 & 1/16 & 1/16 & 1/16 & 1/16 & 1/64 & \cdots \end{array}$$

The expected local values are all greater than 0, so Additivity (or just Pareto alone) implies that $\mathbb{E}W_L \succ W_0$. And, together with Local Expectations, that implies that $L \succ_L W_0$. It is allegedly better to adopt the energy source, even though that is guaranteed to leave infinitely many people worse off.

But this is implausible. *Every single world* in L is worse than W_0 (by Additivity). So L guarantees us a worse outcome. Yet Local Expectations still says that it is the better choice. This means that Local Expectations clashes with *Guaranteed Betterness.*

> *Guaranteed Betterness*: If *every* world in $\mathcal{W}_1$ (the domain of L_1) is better than *every* world in $\mathcal{W}_2$, then $L_1 \succ_L L_2$.

[13] In each world W_i, the local value at l_j is given by:

$$V_i(l_j) = \begin{array}{cc} 0 & \text{for } i < 2^i \\ 2 & \text{for } 2^i \leq j < 2^{i+1} \\ -\frac{1}{2^k} & \text{for } 2^k \leq j < 2^{k+1}, \forall k > i \end{array}$$

This conflict can be stated more formally as Theorem 1.

> **Theorem 1**: For any reflexive, transitive relation $\succcurlyeq_L$ on $\mathcal{P}$, if $\succcurlyeq_L$ satisfies Local Expectations then it cannot satisfy both Guaranteed Betterness and Additivity.

To prove this would be straightforward. We would need only compare W_0 and L from Egregious Energy. *Any* $\succcurlyeq_L$ on the set of all lotteries which satisfies both Guaranteed Betterness and Additivity must give the verdict that W_0 is better. And any such relation which satisfies Local Expectations and Additivity (or just Pareto) must disagree: L is better. So such relations can *only* satisfy Guaranteed Betterness and Additivity if they do not satisfy Local Expectations.

Is this so bad? Yes. Although Local Expectations seems plausible, Guaranteed Betterness seems *undeniable*. Why? One of the most basic claims of decision theory is: if we must decide between two outcomes of which one is strictly better, and there is no risk involved, *of course* we ought to choose the better one. Likewise for minimally aggregative views in ethics - in the ethical cases we're interested in, we ought to choose the better outcome. Intuitively, this shouldn't change if we add a little uncertainty to the outcomes of each, but little enough that *every* outcome of one option L_2 is still worse than every outcome of the other, L_1. No matter the result of L_1 or of L_2, no matter the state of the world, no matter if we swap those states around, no matter how lucky or unlucky we are, we are *certain* that L_2 will turn out *worse*. For goodness' sake, we had better not choose it!

Further, the motivation of expected value reasoning (at least in contexts in which it is appropriate to use) is to help us obtain what we ultimately care about. As Schoenfield (2014:268) forcefully points out in another context, we ultimately care about obtaining outcomes that are valuable, not outcomes that are merely *expectedly valuable*. We should only be interested "...in *expected* value insofar as it helps [us] obtain what is actually important: value. ... Any theory of expected value that makes demands that don't make sense given our concern with *value* can't do what expected value theory is meant to do." (*ibid.*) With the goal of promoting value, a demand that we abandon Guaranteed Betterness make no sense all.

In my view, this goal of promoting value also places stronger requirements on our rules for comparing acts and lotteries. For one, if one lottery is guaranteed to bring about a better world than another *no matter which state arises*, then it is the better lottery. So says *Statewise Dominance*.

> *Statewise Dominance*: Let $A_1, A_2 : \mathcal{S} \to \mathcal{W}$ be some acts with corresponding lotteries

L_1, L_2. If $A_1(S) \succcurlyeq A_2(S)$ for all $S \in \mathcal{S}$, then $L_1 \succcurlyeq_L L_2$. If, as well, $A_1(S') \succ A_2(S')$ for some $S' \in \mathcal{S}$, then $L_1 \succ_L L_2$.

This too seems undeniable. We know that *some* state S will obtain and, no matter which it is, L_1 turns out as good or better. If we want to promote value, it makes little sense to not consider L_1 as good as or better than L_2.

But it doesn't matter in which specific state a world obtains; all that matters is its probability of obtaining. If we take two equally likely states, and swap around the worlds associated with those states in a lottery, that shouldn't make the lottery better or worse. As both Easwaran (2014) and Bader (2018) show, if we accept that such permutations cannot make a lottery better or worse and we accept Statewise Dominance, then we must also accept *Stochastic Dominance*.

> *Stochastic Dominance*: Let L_1 or L_2 be any lotteries on $\mathcal{W}$ with $\mathcal{W}_{(1,2)}$ totally ordered by $\succcurlyeq$. If $L_1({\succcurlyeq}W) \geq L_2({\succcurlyeq}W)$ for all $W \in \mathcal{W}_{(1,2)}$, then $L_1 \succcurlyeq_L L_2$.
>
> If, as well, $L_1({\succcurlyeq}W) > L_2({\succcurlyeq}W)$ for some $W \in \mathcal{W}$, then $L_1 \succ_L L_2$.

This is a strengthening of Statewise Dominance and, in turn, of Guaranteed Betterness. But it is not a radical one. When our decision theory deals with finite payoffs, all three principles are consistent with, but weaker than, standard expected value theory - for instance, they do not rule out risk aversion (see Buchak 2013). Accept expected value theory and you must accept Stochastic Dominance. But, while expected value theory is somewhat controversial in the existing literature, Stochastic Dominance and its kin are not. To the best of my knowledge, no normative decision theory which violates this form of Stochastic Dominance has been seriously proposed by any philosopher.[14]

But Local Expectations violates Stochastic Dominance, along with its weaker siblings. So we must reject it. But how then can we compare lotteries over infinite worlds?

6 My proposal: Expectations of differences

The challenge is to provide a rule for comparing lotteries over infinite worlds which avoids the above problems, and which satisfies Stochastic Dominance. This challenge is greater in some

[14] Schoenfield (2014) rejects a similar, but distinct, form of Stochastic Dominance which gives verdicts even when outcomes in the domains aren't totally ordered by the betterness relation. She raises no objection to weaker formulations like mine.

cases than in others. I'll start by giving a weak version of my rule which can handle easy cases. Then, I'll give a stronger version which can handle harder cases too.

6.1 Version 1

We start with a class of easy cases: comparing lotteries in which each pair of worlds differ by at most a finite sum of local differences. L_1 and L_2 are two such lotteries.

$$L_1 \begin{cases} L_1(W) \\ 1/2 \\ 1/2 \end{cases} \begin{array}{|rccccccccc} & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & \cdots \\ W_0: & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & \cdots \\ W_6: & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots \end{array}$$

$$L_2 \begin{cases} L_2(W) \\ 1/2 \\ 1/2 \end{cases} \begin{array}{|rccccccccc} & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & \cdots \\ W_2: & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & \cdots \\ W_3: & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & \cdots \end{array}$$

Here's how we might proceed. First, with just Additivity, we can rank the worlds $W_6 \succ W_3 \succ W_2 \succ W_0$.

Next, we might pick a world at random - say, W_0 - and represent these worlds by their sum of differences with that baseline world. For W_0, that's 0. For W_6, it's 6. For W_2, it's 2. And for W_3, it's 3. So we have a nice finite, cardinal 'total' to represent each world. And that gives us all we need to represent each lottery with an expected 'total' value.

$$L_1(W_0) \cdot 0 + L_1(W_0) \cdot 6 = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 6 = 3$$

$$L_2(W_2) \cdot 2 + L_2(W_3) \cdot 3 = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 3 = 2\frac{1}{2}$$

Calculated this way, the expected 'total' of L_2 is less than that of L_1, so we might claim that L_1 is the better lottery. Putting this approach more precisely, we have *Expectations of Differences 1.*

> *Expectations of Differences 1 (ED1)*: Let L_1 and L_2 be any lotteries on $\mathcal{W}$ for which all worlds in $\mathcal{W}_{(1,2)}$ contain the same locations $\mathcal{L}$. If there exists a world $W_* \in \mathcal{W}_{(1,2)}$ such that the sum
>
> $$\sum_{W_i \in \mathcal{W}_1} L_1(W_i)\big(\sum_{l \in \mathcal{L}} V_i(l) - V_*(l)\big)$$

is greater than or equal to the corresponding sum for L_2 (mutatis mutandis), then $L_1 \succcurlyeq_L L_2$.

ED1 resembles the way we calculate finite expectations, but with crucial tweaks. In the finite setting, we take the expectation of the *total* value in each world and see how they differ. Here, instead, we represent the value of each world by the sum of all its differences from some 'baseline' world W_*. And then, for each lottery, we take the expectation of that sum.

The standard approach compares expected totals to see which is larger - effectively, it relies on the *differences of expectations.* My approach is to instead use the *expectations of differences.* If we wanted, we could do this in finite cases. We would reach the same verdict as we get with expected total values - in finite cases, the approaches are equivalent. But in infinite cases like this, they come apart, and one of the two clearly does better.

But this approach won't always succeed. For some pairs of lotteries, the expectation of differences will not exist for *any* W_*. Then my rule often falls silent. But it won't *always* fall silent in such cases. Consider a further class of easy cases: when one of those sums diverges unconditionally to (positive or negative) infinity. As long as *only* one of them does, then we can still say which lottery is better. For instance, we can do so in the case of Egregious Energy. Recall W_0 and L.

$$\begin{array}{rccccccccc} & l_1 & l_2 & l_3 & l_4 & l_5 & l_6 & l_7 & l_8 & \cdots \\ W_0: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \end{array}$$

$$L \left\{ \begin{array}{c|rccccccccc} L(W) & & l_1 & l_2 & l_3 & l_4 & l_5 & l_6 & l_7 & l_8 & \cdots \\ 1/2 & W_1: & 2 & -1/2 & -1/2 & -1/4 & -1/4 & -1/4 & -1/4 & -1/8 & \cdots \\ 1/4 & W_2: & 0 & 2 & 2 & -1/4 & -1/4 & -1/4 & -1/4 & -1/8 & \cdots \\ 1/8 & W_3: & 0 & 0 & 0 & 2 & 2 & 2 & 2 & -1/8 & \cdots \\ 1/16 & W_4: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{array} \right.$$

Unlike the Arntzenius-Bostrom-Meacham approach, Expectations of Differences implies that W_0 is better than L. How? First, let our 'baseline world' be W_0. Then the sum of differences between W_0 and W_0 is simply 0. Meanwhile, between each world in the domain of L and W_0, the sum diverges unconditionally to $-\infty$.[15]

We have unconditional divergence for one lottery (L) but not the other (W_0), so we face the sort of situation mentioned above. And ED1 can still say, thankfully, that $W_0 \succcurlyeq L$. After all, $-\infty$ is a lot less than 0. So we seem to be doing better than the Arntzenius-Bostrom-Meacham approach.

ED1 has other arguments in its favour too. Intuitively, it seems a plausible and natural way to judge lotteries. It's a rough analogue of Additivity for this new setting of comparing lotteries. In fact, in cases of certainty, it *implies* Additivity. And, like Additivity, the most compelling reason to accept it is that it is entailed by some highly plausible conditions. (All proofs are in the appendix.)

> **Theorem 2**: For any reflexive, transitive relation $\succcurlyeq_L$ on $\mathcal{P}$, if $\succcurlyeq_L$ satisfies Statewise Dominance, Finite Expectations, and Separability of Value (for Lotteries) then it satisfies ED1.

Statewise Dominance will be familiar from the previous section. Recall that it is consistent not just with standard expected value theory but also with risk aversion. To impose risk neutrality, we can use *Finite Expectations*. This is the analogue of Finite Sum (from Section 3) for comparing lotteries. It requires only that $\succcurlyeq_L$ remains consistent with what our judgements would be if our lotteries contained only finite expected total value. If $\succcurlyeq_L$ doesn't satisfy this, then it's not an adequate extension of finite expected value theory.

> *Finite Expectations*: Let L_1 and L_2 be any lotteries such that the expected total sum of local values converges unconditionally to finite values k_1 and k_2, respectively. $L_1 \succcurlyeq_L L_2$ if and only if $k_1 \geq k_2$.

Then there's the highly plausible *Separability of Value (for Lotteries)*. As did Separability of Value above for comparisons of worlds, this says that we can take any two lotteries and add

[15] The calculation is:

$$\begin{aligned}
&\sum_{W_i \in \mathcal{W}} L(W_i)\Big(\sum_{l \in \mathcal{L}} V_i(l) - V_0(l)\Big) \\
&= \sum_{W_i \in \mathcal{W}} \frac{1}{2^i}\Big(\sum_{l \in \mathcal{L}} V_i(l) - 0\Big) \\
&= \sum_{W_i \in \mathcal{W}} \frac{1}{2^i}\Big(0 + 0 + \ldots + 2 \cdot 2^{i-1} + \sum_{k=1}^{\infty} -1\Big) \\
&\to -\infty
\end{aligned}$$

whatever string of local values we want to (all of the worlds in the domain of) them, and the resulting lotteries are ranked the same way. This implies that comparisons of lotteries are *only* sensitive to differences in the local values of the worlds in their domains, and that local values are additively separable. This is a direct analogue of Separability of Value for the lottery context, so much so that I'll abbreviate it to 'Separability of Value' in what follows.

> *Separability of Value (for Lotteries)*: For any lotteries L_1 and L_2 and world $W' \in \mathcal{W}$ with the same locations as all worlds in $\mathcal{W}_{(1,2)}$, let lottery L_1' on $\mathcal{W}$ be defined as $L_1'(W) = L_1(W - W')$ for all $W \in \mathcal{W}_1 + W'$, and similarly for L_2'. If $L_1 \succcurlyeq_L L_2$ then $L_1' \succcurlyeq_L L_2'$.[16]

I find Statewise Dominance, Finite Expectations, and Separability of Value, as well as their conjunction, hard to deny. So ED1 is on firm ground as a minimal principle for comparing lotteries, just as Additivity is for comparing worlds. And it seems to me to be on strictly *firmer* ground than Local Expectations, which implies Finite Expectations and Separability of Value but violates all three of the dominance principles we saw above.

6.2 Version 2

But that first principle doesn't get us far. Consider L_1 and L_2.

$$L_1 \left\{ \begin{array}{c|cccccccc} L_1(W) & & l_a & l_b & l_c & l_d & l_e & \cdots \\ 0.1 & W_4: & 4 & 4 & 4 & 4 & 4 & \cdots \\ 0.9 & W_1: & 1 & 1 & 1 & 1 & 1 & \cdots \end{array} \right.$$

$$L_2 \left\{ \begin{array}{c|cccccccc} L_2(W) & & l_a & l_b & l_c & l_d & l_e & \cdots \\ 0.2 & W_{2,1}: & 2 & 1 & 2 & 1 & 2 & \cdots \\ 0.8 & W_0: & 0 & 0 & 0 & 0 & 0 & \cdots \end{array} \right.$$

Take any pair of those worlds and sum their local differences; the result is infinite. If we apply ED1 here, no matter which baseline-world we pick, the expected sums are undefined. Using ED1 alone, we cannot compare these lotteries. So we need a stronger rule.

Before I present that rule, consider how little we need to compare two lotteries over finite payoffs. We can make judgements without knowing the probabilities in each lottery - we just need

[16] This also implies the converse since, for all W', the condition also applies to $-W'$.

to know the *differences* in probability of each outcome between one lottery and the other. And we can make judgements without knowing the values of the payoffs - we just need to know the (scale of the) differences between them. We can even make judgements often without knowing the *precise* differences in probabilities and values - upper and lower bounds may be enough.

Here are two lotteries over finite payoffs which we can compare with incomplete information.

L_1': value $v+u+1$ with probability p; value v with probability $1-p$.

L_2': value $v+1$ with probability $2p$; value 0 with probability $1-2p$.

Here, v and u are some real numbers such that $v>0$ and $u>2$. And p is some probability in $(0,\frac{1}{2})$. We cannot assign precise values to any of them, perhaps because those values are vague or indeterminate. But we can still judge which lottery is better. Take their expected values: L_1' has expectation $v+p(u+1)$, which is greater than $v+3p$; and L_2' has expectation $2p(v+1)$, which is less than $v+2p$. So L_1' is better. Imprecision is no problem here.

We have similar information when comparing the above lotteries L_1 and L_2 over infinite worlds - it turns out that they are structurally equivalent to the finite lotteries L_1' and L_2'. So we can take a fairly similar approach.

First, take the difference in probabilities of obtaining a world at least as good as each W_i between L_1 and L_2, given by $\Delta p_i = L_1(\succcurlyeq W_i) - L_2(\succcurlyeq W_i)$. Where this is positive, L_1 has a greater probability of delivering W_i or better; where it's negative, L_2 has the greater probability.

$$\begin{aligned}
\Delta p_4 &= L_1(\succcurlyeq W_4) - L_2(\succcurlyeq W_4) = 0.1\\
\Delta p_{2,1} &= L_1(\succcurlyeq W_{2,1}) - L_2(\succcurlyeq W_{2,1}) = -0.1\\
\Delta p_1 &= L_1(\succcurlyeq W_1) - L_2(\succcurlyeq W_1) = 0.8\\
\Delta p_0 &= L_1(\succcurlyeq W_0) - L_2(\succcurlyeq W_0) = 0
\end{aligned}$$

Second, take the difference between each world and the next best world in the domain. If our worlds had finite total values, we would represent those differences with finite values. But, with infinite worlds, we must represent differences as worlds themselves - worlds given by the differences in local values. For each world W_i which has some distinct next best world W_j in $\mathcal{W}_{(1,2)}$, the difference between them is given by $D_i = \langle \mathcal{L}, V_i - V_j \rangle$, with local values $V_i(l) - V_j(l)$ at every location. (Note that if there is no world in $\mathcal{W}_{(1,2)}$ that is strictly worse than W_i, then D_i is undefined.)

$$\begin{array}{rcccccc} & l_a & l_b & l_c & l_d & l_e & \cdots \\ D_4: & 2 & 3 & 2 & 3 & 2 & \cdots \\ D_{2,1}: & 1 & 0 & 1 & 0 & 1 & \cdots \\ D_1: & 1 & 1 & 1 & 1 & 1 & \cdots \end{array}$$

Third, compare how large those differences are relative to one another. Again, this would be straightforward with real, finite differences, but less so when the differences are infinite worlds. Fortunately, we *can* compare their size. To start with, we can use the same old $\succcurlyeq$ relation that we would use to compare the size of worlds. By Additivity (or just Pareto), we'll have $D_4 \succ D_1 \succ D_{(2,1)}$.

And, further, we can judge their relative size by comparing them to scalar multiples of one another.[17] A scalar multiple can be defined as $k \cdot W = \langle \mathcal{L}, k \times V_i \rangle$ for any real k, with local values each k times that in W. Then we have:

$$2 \cdot D_1 \prec D_4 \prec 3 \cdot D_1$$

$$0 \cdot D_1 \prec D_{2,1} \prec 1 \cdot D_1$$

So we can say that D_4 is more than twice, but less than three times, as great as D_1. And $D_{2,1}$ is less than D_1, but greater than 0 times D_1. And that's all we need. The information we have here is analogous to what we had for L_1' and L_2' above, and the lotteries L_1 and L_2 are analogous to them as well. D_1 is analogous to v, D_4 analogous to u, and $D_{2,1}$ analogous to the value 1. We know that $D_4 \succ 2 \cdot D_{2,1}$, much like we knew that $u > 2$ before. And so, if these lotteries can be compared in an analogous way, we can say $L_1 \succ_L L_2$.

And we can make judgements much more generally, even for lotteries which do not resemble L_1' and L_2' specifically. To see why, note that, if we were dealing with *any* lotteries L_1 and L_2 over finitely-valued outcomes $w \in \mathcal{W}_{(1,2)}$, we would say that $L_1 \succcurlyeq_L L_2$ if and only if L_1 had at least as great an expected value:

$$\mathbb{E}(L_1) - \mathbb{E}(L_2) = \sum_{w \in \mathcal{W}_{(1,2)}} (L_1(w) - L_2(w)) \times V(w) \geq 0$$

And this equation rearranges to the following, where the outcomes $(w_1, w_2, w_3, ...)$ are ordered from least to most value. Here, we use the probability of each lottery turning out at least as

[17] Since the local values in the original worlds W_i and W_j are represented on a common interval scale, the local values in D_i can be represented on a *ratio scale* - they have an absolute zero, 0. So it makes sense to compare their absolute size via scalar multiplication, and likewise for the difference-worlds at large.

good as each w_i, rather than the probability of w_i alone.

$$\mathbb{E}(L_1) - \mathbb{E}(L_2) = \sum_{w_i \in \mathcal{W}} \Big(L_1(\succcurlyeq w_i) - L_2(\succcurlyeq w_i) \Big) \Big(V(w_i) - V(w_{i-1}) \Big) \geq 0$$

That difference in probabilities between L_1 and L_2 is just Δp_i from above. And the difference in values is equivalent to D_i. If we can represent the D_is with some real values k_i which measure their relative size, then we can use this equation as is. But we don't even need their precise relative size. We can get by with k_is which capture an upper or lower bound on the relative size of each D_i.

In effect, we can apply *Expectations of Differences 2*, which goes like this.

Expectations of Differences 2 (ED2): Let L_1 and L_2 be any lotteries on $\mathcal{W}$ such that all $W \in \mathcal{W}_{(1,2)}$ have the same locations.

If, for some $W_* \in \mathcal{W}_{(1,2)}$, and for all $W_i \in \mathcal{W}_{(1,2)}$ such that $\Delta p_i \neq 0$, there exists $k_i \in \mathbb{R}$ such that

$$k_i \Delta p_i \cdot D_* \preccurlyeq \Delta p_i \cdot D_i \qquad \text{and} \qquad \sum_{W_i \in \mathcal{W}_{(1,2)}} k_i \Delta p_i \geq 0$$

then $L_1 \succcurlyeq_L L_2$.

Here, W_* is just some world in the domain, and D_* the difference between it and the next best world, against which the size of each D_i will be compared.

The first equation constrains the k_i that we can use as a stand-in for each D_i. It represents an upper or lower bound on how many times greater D_i is than D_*. To say that $L_1 \succcurlyeq_L L_2$, we need k_i to give an *upper* bound on the size of D_i when W_i (or better) is *more* likely under L_1 - when $\Delta p_i > 0$. And we need k_i to give a *lower* bound when it's more likely under L_2 - when $\Delta p_i < 0$. That's why Δp_i appears on both sides of that equation - it allows the inequality to switch direction when $\Delta p_i < 0$.

Once we have those k_is to represent the size of the differences, the second equation runs the same probability-weighted sum we saw above. Like above, when it's positive, L_1 is better. And that's ED2.

But should we accept ED2? I think so, and not just based on that analogy with imprecise values. For one, it implies all of the (very plausible) judgements of ED1, as long as $\succcurlyeq$ obeys Additivity.

Theorem 3: For any reflexive, transitive relation $\succcurlyeq_L$ on $\mathcal{P}$, if $\succcurlyeq_L$ satisfies Expectations of Differences 2 and Additivity, then it satisfies Expectations of Differences 1.

Given that compatibility, we already know that this version of Expectations of Differences will deal with cases like Egregious Energy better than rival views. But it's also stronger than the previous version, as demonstrated in the example above.

Also in its favour, the rule is implied by the conjunction of several highly plausible principles.

Theorem 4: For any reflexive, transitive relation $\succcurlyeq_L$ on $\mathcal{P}$, if $\succcurlyeq_L$ satisfies Stochastic Dominance, Separability of Value (for Lotteries), Independence, and Extrapolated Expectations, then it satisfies Expectations of Differences 2.

You'll recognise Stochastic Dominance and Separability of Value from above. Then we have two newcomers.

First, *Independence* is the same principle that is often used to axiomatise expected utility theory. Suppose we evaluate L_1 is at least as good as L_2. Independence says that we could mix each lottery with some third lottery L_3, whatever it might be, and the resulting mixed lotteries would be ranked the same way. Equivalently, take a mixed lottery which gives you probability p of running lottery L_1 as normal, and a probability $1-p$ of running L_3 instead; Independence says that that mixed lottery is still at least as good as the mixed lottery in which L_2 replaces L_1. More formally, it can be stated as follows.

Independence: Define lotteries $L_{1\vee 3}, L_{1\vee 3}$ on $\mathcal{W}$ by

$$L_{1\vee 3}(W) = p \times L_1(W) + (1-p) \times L_3(W) \qquad \text{for all } W \in \mathcal{W}$$

$$\text{and } L_{2\vee 3}(W) = p \times L_2(W) + (1-p) \times L_3(W) \qquad \text{for all } W \in \mathcal{W}$$

For any lotteries L_1, L_2, and L_3 on $\mathcal{W}$ and for any $p \in [0,1]$, $L_{1\vee 3} \succcurlyeq_L L_{2\vee 3}$ if and only if $L_1 \succcurlyeq L_2$.

We then have *Extrapolated Expectations*, which should also be uncontroversial for those who accept expected value theory in the finite context. It merely says that, for any world $W' \succ \mathbf{0}$, a lottery with probability p of W' (and $\mathbf{0}$ otherwise) is precisely as good as $p \cdot W'$.

Extrapolated Expectations: For any world $W' \in \mathcal{W}$ such that $W' \succ \mathbf{0}$ and any $p \in (0,1]$ the following lottery $L \simeq_L p \cdot W'$.

$$L(W) = \begin{cases} p & \text{for } W = W' \\ 1-p & \text{for } \mathbf{0} \\ 0 & \text{for } W \notin \{W', \mathbf{0}\} \end{cases}$$

If we faced a similar lottery over finite payoffs, w and 0, we'd assign it expected value $p \cdot w$ without hesitation. Extrapolated Expectations requires that we treat infinite worlds in the same way. And it's on firm footing of its own - if our $\succcurlyeq$ for comparing worlds obeys Additivity, as any plausible $\succcurlyeq$ does, then Extrapolated Expectations follows straightforwardly from Finite Expectations and Statewise Dominance. It is also worth noting that it follows from Local Expectations - if one is tempted to adopt Local Expectations instead of ED2, Extrapolated Expectations must hold either way.

And if you accept Extrapolated Expectations, Independence, Stochastic Dominance, and Separability of Value, then you must also accept Expectations of Differences 2, by Theorem 4.

6.3 Weaknesses of Expectations of Differences

You might notice in the definition of Expectations of Differences 2 that the antecedent of the conditional is awfully strong. The rule, then, is awfully weak. It can only provide judgements when:

- 1) *Every* world in the domain of either lottery contains precisely the same locations;
- 2) *Every* pair of such worlds is comparable;
- 3) *Many* pairs of such worlds have a difference which is also comparable to many of the other differences; and
- 4) Those differences are *also* comparable with certain scalar multiples of other differences (depending on the probabilities in the lotteries).

To demonstrate my view's weakness, here is a pair of lotteries similar to those I used earlier to demonstrate Local Expectations in action, but with *slightly* different probabilities. (Here, ϵ is some small positive number.)

$$
\begin{array}{rccccccccccccc}
 & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & l_i & l_j & l_k & \cdots \\
W_1: & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & \cdots
\end{array}
$$

$$
L \begin{cases} L(W) \\ 1/2 - \epsilon \\ 1/2 + \epsilon \end{cases}
\left|
\begin{array}{rcccccccccccc}
 & l_a & l_b & l_c & l_d & l_e & l_f & l_g & l_h & l_i & l_j & l_k & \cdots \\
W_2: & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & \cdots \\
W_3: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots
\end{array}
\right.
$$

What does ED2 say? Well, here are our differences-below, and a relevant scalar multiple.

$$
\begin{array}{rcccccc}
 & l_a & l_b & l_c & l_d & l_e & \cdots \\
D(W_2, W_1): & 1 & 2 & 1 & 2 & 1 & \cdots \\
D(W_1, W_0) = W_1: & 1 & 0 & 1 & 0 & 1 & \cdots \\
(1+\epsilon') \cdot D(W_1, W_0): & 1+\epsilon' & 0 & 1+\epsilon' & 0 & 1+\epsilon' & \cdots
\end{array}
$$

By Additivity alone, we cannot compare $D(W_2, W_1)$ to the scalar multiple $(1+\epsilon') \cdot D(W_1, W_0)$, since the former does better at locations $l_b, l_d, l_f, ...$ and the latter does better at the rest. The sum of their local differences is undefined. And given the probabilities here, there is no way to generate k_is both of the equations needed for ED2.

Thus, ED2 and Additivity together say nothing about how we should compare W_1 to L. And this is *even though* every pair of worlds at play is comparable and, on top of that, so are the differences between each of them. Even worse, we only made the smallest of changes to the lottery - all we did was give a mere $+\epsilon$ of probability mass to one outcome, and now we cannot say a thing!

My first response to this silence is simply: *tu quoque*. Suppose we adopt the rival view of Arntzenius, Bostrom, and Meacham, and we swap out ED2 for Local Expectations. Then we still cannot compare these two lotteries. So I am doing no worse here. But that is little comfort.

My second response is that this is no shortcoming of Expectations of Differences; it is a shortcoming of Additivity. Additivity is a very weak constraint on betterness. I have used it so far *because* it is so weak - this weakness makes it uncontroversial and, indeed, all of the plausible stronger principles in the literature are consistent with it. We can treat the conjunction of ED2 and Additivity similarly: it is just a weak and (hopefully) uncontroversial condition for comparing lotteries. So it is fitting that it makes no judgement in this case. The correct judgement is not obvious, so it should remain silent.

My third response is that, actually, ED2 allows us to do even better. It can be combined with almost any betterness relation we choose - the definition above makes no reference to Additivity, but instead to some unspecified $\succcurlyeq$ relation. We might adopt the $\succcurlyeq$ relation from Vallentyne & Kagan (1997:19), Jonsson & Voorneveld (2018), or Wilkinson (2020), each of which is much stronger than Additivity. Personally, I favour the betterness relation described in Wilkinson (*ibid.*) - unlike the others, it can be shown to give comparisons for every pair of physically possible worlds that could occur as a result of our actions (see Wilkinson n.d.).

And even without a betterness relation stronger than Additivity, ED2 does just fine in the examples in the previous sections, including Egregious Energy. And so this approach is not so weak after all.

7 Conclusion

If you hold a minimally aggregative ethical view, you may be dismayed to discover that the universe is infinite. Your moral theory seems to fall silent in all cases where it must rely on facts of betterness. You may want to advise others to make the world better, but your axiology tells you that this is impossible.

It may be a relief to then come across the existing proposals for betterness relations which resemble finite aggregationism but which fare better in infinite worlds (e.g., Vallentyne & Kagan 1997; Lauwers & Vallentyne 2004; Bostrom 2011; Arntzenius 2014; Jonsson & Voorneveld 2018; Wilkinson 2020). But then you recall that you are a limited epistemic agent - you're uncertain about the effects of your actions - and you may be dismayed once more. After all, most of these betterness relations offer no clear way to compare lotteries over infinite outcomes, and so no way to make subjective normative judgements. And those which do offer this have implausible implications, as we saw in Section 5.

If you have ridden that emotional rollercoaster, you have my sympathy. I hope that I've allowed you to disembark. We now have a plausible method for comparing lotteries over infinite outcomes, which avoids the problems of the previous proposals. With a sufficiently complete betterness relation, this proposal can restore the subjective normative judgements of minimally aggregative views. We may well be able to make decisions based on what will promote the good after all.

8 Appendix

Theorem 2: For any reflexive, transitive relation $\succcurlyeq_L$ on $\mathcal{P}$, if $\succcurlyeq_L$ satisfies Statewise Dominance, Finite Expectations, and Separability of Value (for Lotteries) then it satisfies ED1.

Proof: Define the binary relation $\succcurlyeq_{\text{ED1}}$ on $\mathcal{P}$ by:

$L_1 \succcurlyeq_{\text{ED1}} L_2$ if and only if there is some $W_* \in \mathcal{W}$ such that

$$\sum_{W_j \in \mathcal{W}_1} L_1(W_j)\big(\sum_{l \in \mathcal{L}} V_j(l) - V_*(l)\big) \geq \sum_{W_j \in \mathcal{W}_2} L_2(W_j)\big(\sum_{l \in \mathcal{L}} V_j(l) - V_*(l)\big) \qquad \text{(i)}$$

This is the relation given by ED1 alone. To prove Theorem 2, it suffices to show that, for any L_1 and L_2 such that $L_1 \succcurlyeq_{\text{ED1}} L_2$, it must also hold that $L_1 \succcurlyeq_L L_2$ if $\succcurlyeq_L$ abides by Statewise Dominance, Finite Expectations, and Separability of Value (for Lotteries).

Let $L_1, L_2 \in \mathcal{P}$ be some such lotteries. (We already know that such lotteries exist - see Section 6.) Then there exists some $W_* \in \mathcal{W}$ such that (i) holds. So either 1) both sides of the inequality converge unconditionally to some real values, with the LHS greater than or equal to the RHS, or 2) the LHS diverges unconditionally to $+\infty$, the RHS diverges unconditionally to $-\infty$, or both.

1) If both sides of (i) converge:

Define two additional lotteries L_1^* and L_2^* by $L_1^*(W - W_*) = L_1(W)$ and $L_2^*(W - W_*) = L_2(W)$ for all $W \in \mathcal{W}_{(1,2)}$ (and $L_1^*(W), L_2^*(W) = 0$ otherwise). Since both sides of (i) converge unconditionally to real values, the expected total sums of local value in L_1^* and L_2^* will be finite, and the expected total sum for L_1^* will be greater than or equal to that for L_2^*. So $L_1^* \succcurlyeq_L L_2^*$, by Finite Expectations. Then, by Separability of Value (for Lotteries), $L_1 \succcurlyeq_L L_2$, as required.

2) If either or both sides of (i) diverge:

The LHS diverges to $+\infty$, or the RHS diverges to $-\infty$, or both. If both, then define L_1^* as above and $L_2^*(W_0) = 1$. If not both, then define L_1^*, L_2^* as above.

If the LHS diverges to $+\infty$, then there is some lottery L_1^{**} with finite expected sum such that $L_1^* \succcurlyeq_L L_1^{**}$ by Statewise Dominance and $L_1^{**} \succcurlyeq_L L_2^*$ by Finite Expectations. We can obtain L_1^{**} from L_1^* by replacing any worlds in its domain that have total sums of local value greater than some chosen large finite bound with some other worlds with totals below that bound.

If the LHS does not diverge but the RHS does diverge to $-\infty$, then there is some lottery L_2^{**}

such that $L_2^* \preccurlyeq_L L_2^{**}$ by Statewise Dominance and $L_2^{**} \preccurlyeq_L L_1^*$ by Finite Expectations.

Either way, $L_1^* \succcurlyeq_L L_2^*$. Separability of Value (for Lotteries) then implies that $L_1 \succcurlyeq_L L_2$. $\square$

Theorem 3: For any reflexive, transitive relation $\succcurlyeq_L$ on $\mathcal{P}$, if $\succcurlyeq_L$ satisfies ED2 and Additivity, then it satisfies ED1.

Proof:

Define a binary relation $\succcurlyeq_{\text{ED2}}$ on $\mathcal{P}$ by:

> $L_1 \succcurlyeq_{\text{ED2}} L_2$ if and only if 1) for all worlds $W_a, W_b \in \mathcal{W}_{(1,2)}$, $\mathcal{L}_a = \mathcal{L}_b$, and 2) there is some $W_* \in \mathcal{W}_{(1,2)}$ such that, for each $W_i \in \mathcal{W}_{(1,2)}$, there is some $k_i \in \mathbb{R}$ such that
>
> $$k_i \Delta p_i \cdot D_* \preccurlyeq \Delta p_i \cdot D_i \qquad \text{and} \qquad \sum_{W_i \in \mathcal{W}_{(1,2)}} k_i \Delta p_i \geq 0$$

This is the relation given by ED2 alone. Define $\succcurlyeq_{\text{ED1}}$ as above, and let $L_1, L_2 \in \mathcal{P}$ be any lotteries such that $L_1 \succcurlyeq_{\text{ED1}} L_2$. And assume that $\succcurlyeq$ satisfies Additivity. Then, to prove Theorem 3, it suffices to show that $L_1 \succcurlyeq_{\text{ED2}} L_2$.

Since $L_1 \succcurlyeq_{\text{ED1}} L_2$, for some $W' \in \mathcal{W}_{(1,2)}$,

$$\sum_{W_j \in \mathcal{W}_1} L_1(W_j)\big(\sum_{l \in \mathcal{L}} V_j(l) - V'(l)\big) \geq \sum_{W_j \in \mathcal{W}_2} L_2(W_j)\big(\sum_{l \in \mathcal{L}} V_j(l) - V'(l)\big)$$
$$\Rightarrow \sum_{W_j \in \mathcal{W}_{(1,2)}} (L_1(W_j) - L_2(W_j))\big(\sum_{l \in \mathcal{L}} V_j(l) - V'(l)\big) \geq 0 \qquad \text{(i)}$$

Assume that the signs of $L_1(W_j) - L_2(W_j)$ and $\sum_{l \in \mathcal{L}} V_j(l) - V'(l)$ differ for some $W_j \in \mathcal{W}_{(1,2)}$, and hence also that the signs are the same for some other W_i. (ii) (If not, then we immediately have $L_1 \succcurlyeq_{\text{ED2}} L_2$.)

Given (i), either 1) for at least one W_j, $\sum_{l \in \mathcal{L}} V_j(l) - V'(l)$ diverges unconditionally to $+\infty$, or 2) all of those sums are finite.

1) For at least one W_j, $\sum_{l \in \mathcal{L}} V_j(l) - V'(l)$ diverges unconditionally to $+\infty$:

There will be at least one D_i with an infinite sum of local values and $\Delta p_i > 0$, but none with $\Delta p_i > 0$. Given (ii), we also have at least one D_* with negative Δp_*. Given that the sum diverges unconditionally, that D_* must only have a finite total sum of local values. For all finite-sum D_i, set $k_i = 0$. For all infinite-sum D_i, set $k_i = 1$. Then we have $k_i \Delta p_i \cdot D_* \preccurlyeq \Delta p_i \cdot D_i$ for all D_i.

And $\sum_{W_i \in \mathcal{W}_{(1,2)}} k_i \Delta p_i \geq 0$, since it will simply be the sum of the (at least one) infinite-sum D_i with positive Δp_i. Thus, both conditions are thus satisfied, and we have $L_1 \succcurlyeq_{\text{ED1}} L_2$, as required.

2) For all W_j, $\sum_{l \in \mathcal{L}} V_j(l) - V'(l)$ is finite:

Then each D_i will also have a finite total sum of local value, $S_i = \sum_{l \in \mathcal{L}} D_i(l)$.

Let $D_* = D'$. Then let each $k_i = \frac{S_i}{S_*}$. Since all S_i are finite and all $D_i \succcurlyeq 0$, Additivity says that $k_i \Delta p_i \cdot D_* \simeq \Delta p_i D_i$ if and only if $k_i \Delta p_i \times S_* = \Delta p_i S_i$. Since $k_i = \frac{S_i}{S_*}$, this holds for all D_i, as required.

We now seek the second condition. First, note that $P(W) \cdot V + (1 - P(W)) \cdot V' = V' + P(W) \cdot (V - V')$. By iterating that rearrangement, we can obtain the following from (i).

$$\sum_{W_i \in \mathcal{W}_{(1,2)}} (P_{L_1}(W_i \text{ or better}) - P_{L_2}(W_i \text{ or better})\big(\sum_{l \in \mathcal{L}} V_i(l) - V_j(l)\big) \geq 0$$

$$(\text{ where } W_j = \max\{W \in \mathcal{W}_{(1,2)} | W_j \prec W_i\})$$

$$\Leftrightarrow \sum_{W_i \in \mathcal{W}_{(1,2)}} \Delta p_i \sum_{l \in \mathcal{L}} D_i(l) \geq 0$$

$$\Leftrightarrow \sum_{W_i \in \mathcal{W}_{(1,2)}} \Delta p_i S_i \geq 0$$

$$\Leftrightarrow \frac{1}{S_*} \sum_{W_i \in \mathcal{W}_{(1,2)}} \Delta p_i S_i \geq 0$$

$$(\text{ since } S_* > 0)$$

$$\Leftrightarrow \sum_{W_i \in \mathcal{W}_{(1,2)}} \Delta p_i k_i \geq 0 \quad \square$$

To prove Theorem 4 below, it will help to first establish Lemma 1. If Independence and Extrapolated Expectations hold, then this lemma effectively implies that we can evaluate any lottery in which every outcome is some scalar multiple of some world $W \succ \mathbf{0}$. And we evaluate it as equally good as the probability-weighted sum of the scalar multiples of W. Effectively, the lottery is valued at its expected value, as some multiple of W.

Lemma 1: Let $L \in \mathcal{P}$ be any finitely-supported lottery such that, for some $W \in \mathcal{W}$, $L(W_j) > 0$ if and only if W_j such that $W_j = k_j \cdot W$ for some positive, real k_j. If $\succcurlyeq_L$ is a reflexive, transitive relation on $\mathcal{P}$ which satisfies Independence and Extrapolated Expectations, then:

$$L \simeq_L (\sum_{W_j \in \mathcal{W}} L(W_j) \times k_j) \cdot W$$

Proof:

Let $k_{\max}$ be the greatest such k_j (which exists and is unique, since L is finitely-supported). For each k_j, $0 < \frac{k_j}{k_{\max}} < 1$.

Let L_j be the lottery such that $L_j(k_{\max}) = \frac{k_j}{k_{\max}}$ and $L_j(\mathbf{0}) = 1 - \frac{k_j}{k_{\max}}$. (So the remaining probability mass goes to $\mathbf{0}$.) Since $k_j \cdot W = \frac{k_j}{k_{\max}} \cdot (k_{\max} \cdot W)$, Extrapolated Expectations implies that:

$$k_j \cdot W \simeq_L L_j \qquad \text{(i)}$$

By Independence, we can replace each world $k_j \cdot W$ in L with the lottery L_j. In other words, by (i) and Independence, $L \simeq_L L'$, where:

$$L'(W') = \begin{cases} \sum_{k_j \in \mathbb{R}} L(k_j \cdot W) \times \frac{k_j}{k_{\max}} & \text{for } W' = k_{\max} \cdot W \\ 1 - L'(k_{\max} \cdot W) & \text{for } \mathbf{0} \\ 0 & \text{for } W' \neq k_{\max}, \mathbf{0} \end{cases}$$

By Extrapolated Expectations, $L' \simeq_L (\sum_{k_j \in \mathbb{R}} L(k_j \cdot W) \times \frac{k_j}{k_{\max}}) \cdot (k_{\max} \cdot W)$.

$\therefore L \simeq_L (\sum_{W_j \in \mathcal{W}} L(W_j) \times k_j) \cdot W$, as required. $\square$

Theorem 4: For any reflexive, transitive relation $\succcurlyeq_L$ on $\mathcal{P}$, if $\succcurlyeq_L$ satisfies Stochastic Dominance, Separability of Value, Independence, and Extrapolated Expectations, then it satisfies ED2.

Proof:

Define $\succcurlyeq_{\text{ED2}}$ as above, and let L_1, L_2 be any lotteries such that $L_1 \succcurlyeq_{\text{ED2}} L_2$. To prove Theorem 4, it suffices to show that $L_1 \succcurlyeq_L L_2$.

If $\Delta p_i \geq 0$ for all $W_i \in \mathcal{W}_{(1,2)}$, then $L_1 \succcurlyeq_L L_2$, by Stochastic Dominance, as required.

If not, there are some $W_i, W_j \in \mathcal{W}_{(1,2)}$ such that $\Delta p_i < 0$ and $\Delta p_j > 0$. Assume from here on that such W_i and W_j exist.

From the definition of $\succcurlyeq_{\text{ED2}}$, there is some $W_* \in \mathcal{W}_{(1,2)}$ such that, for each $W_i \in \mathcal{W}_{(1,2)}$ there is $k_i \in \mathbb{R}$ such that:

$$k_i \Delta p_i \cdot D_* \preccurlyeq \Delta p_i \cdot D_i \text{ (i)}$$

$$\text{and} \qquad \sum_{W_i \in \mathcal{W}_{(1,2)}} k_i \Delta p_i \geq 0 \text{ (ii)}$$

For some such k_is and W_*, define lotteries $L_1^D, L_2^D \in \mathcal{P}$ by

$$L_1^D((\sum_{\{W_j \in \mathcal{W}_{(1,2)} | W_j \preccurlyeq W_i\}} k_j) \cdot D_*) = L_1(W_i) \qquad \text{for all } W_i \in \mathcal{W}_1$$

and likewise for L_2^D, mutatis mutandis.

And $D_* \succ \mathbf{0}$ since it is the difference between some world and a worse world, by Separability of Value. So Lemma 1 applies to both L_1^D and L_2^D.

$$\therefore L_1^D \simeq_L (\sum_{W_i \in \mathcal{W}_1} L_1(W_i)(\sum_{\{W_j \in \mathcal{W}_{(1,2)} | W_j \preccurlyeq W_i\}} k_j)) \cdot D_* \qquad \text{(by Lemma 1)}$$

and likewise for L_2^D (mutatis mutandis).

So $L_1^D \succcurlyeq_L L_2^D$ iff that scalar multiple of D_* is as good or better than the corresponding world for L_2^D. Since $D_* \succ \mathbf{0}$, that holds if and only if:

$$\sum_{W_i \in \mathcal{W}_1} L_1(W_i)(\sum_{\{W_j \in \mathcal{W}_{(1,2)} | W_j \preccurlyeq W_i\}} k_j) \geq \sum_{W_i \in \mathcal{W}_2} L_2(W_i)(\sum_{\{W_j \in \mathcal{W}_{(1,2)} | W_j \preccurlyeq W_i\}} k_j)$$

$$\Leftrightarrow k_{\min} + \sum_{W_i \in \mathcal{W}_{(1,2)}} L_1(\succcurlyeq W_i) \times k_i \geq k_{\min} + \sum_{W_i \in \mathcal{W}_{(1,2)}} L_2(\succcurlyeq W_i) \times k_i$$

$$\text{for } k_{\min} = \min\{k_i | W_i \in \mathcal{W}_{(1,2)}\}$$

$$\Leftrightarrow \sum_{W_i \in \mathcal{W}_{(1,2)}} k_i \Delta p_i \geq 0, \text{ which is given by (ii).}$$

Therefore, $L_1^D \succcurlyeq_L L_2^D$.

For some such k_is and W_* as above, define lotteries L_1^*, L_2^* on $\mathcal{W}$ by

$$L_1^*(W - W_*) = L_1(W) \qquad \text{for all } W \in \mathcal{W}_1$$

and likewise for L_2^*(mutatis mutandis).

These lotteries resemble L_1 and L_2; they have the same probabilities, but W_* is subtracted from each outcome. As a result, each outcome $W - W_*$ can be represented as a sum of differences

$$W - W_* = \sum_{\{W_j \in \mathcal{W}_{(1,2)} | W_* \preccurlyeq W_j \preccurlyeq W\}} D_j$$

$$(\text{or, for } W \prec W_*, W - W_* = - \sum_{\{W_j \in \mathcal{W}_{(1,2)} | W \preccurlyeq W_j \preccurlyeq W_*\}} D_j).$$

But, by (i), whenever $\Delta p_i > 0$,

$$W_i - W_* = \sum_{\{W_j \in \mathcal{W}_{(1,2)} | W_* \preccurlyeq W_j \preccurlyeq W_i\}} D_j \succcurlyeq (\sum_{\{W_j \in \mathcal{W}_{(1,2)} | W_j \preccurlyeq W_i\}} k_j)) \cdot D_*.$$

And, whenever $\Delta p_i < 0$, the inequality is reversed.

Therefore, by Stochastic Dominance and (i), $L_1^* \succcurlyeq_L L_1^D$ and $L_2^* \preccurlyeq_L L_2^D$. And, since $L_1^D \succcurlyeq_L L_2^D$, transitivity implies that $L_1^* \succcurlyeq_L L_2^*$.

By Separability of Value, $L_1 \succcurlyeq L_2$ if and only if $L_1^* \succcurlyeq_L L_2^*$. Therefore, $L_1 \succcurlyeq L_2$, as required. $\square$

References

[1] Arntzenius, Frank 2014, 'Utilitarianism, decision theory and eternity', *Philosophical Perspectives* 28, no. 1:31-58.

[2] Askell, Amanda 2018, *Pareto Principles in Infinite Ethics*, PhD thesis (New York University).

[3] Bader, Ralf 2018, 'Stochastic dominance and opaque sweetening', *Australasian Journal of Philosophy* 96, no. 3:498-507.

[4] Bostrom, Nick 2011, 'Infinite ethics', *Analysis and Metaphysics* 10:9.

[5] Buchak, Lara 2013, *Risk and rationality* (Oxford: Oxford University Press).

[6] Carroll, Sean M. 2017, 'Why Boltzmann brains are bad', *arXiv preprint* arXiv:1702.00850.

[7] Easwaran, Kenny 2014, 'Decision theory without representation theorems', *Philosopher's Imprint* 14, no. 27:1-30.

[8] Garriga, Jaume, & Alexander Vilenkin 2001, 'Many worlds in one', *Physical Review D* 64, no. 4:043511.

[9] Guth, Alan H. 2007, 'Eternal inflation and its implications', *Journal of Physics A: Mathematical and Theoretical* 40, no. 25:6811.

[10] Jackson, Frank 1991, 'Decision-theoretic consequentialism and the nearest and dearest objection', *Ethics* 101, no. 3:461-82.

[11] Jonsson, Adam, & Mark Voorneveld 2018, 'The limit of discounted utilitarianism', *Theoretical Economics* 13, no. 1:19-37.

[12] Lauwers, Luc & Peter Vallentyne 2004, 'Infinite utilitarianism: more is always better', *Economics & Philosophy* 20, no. 2:307-30.

[13] Meacham, Christopher J.G. 2020, 'Too much of a good thing: Decision-making in cases with infinitely many utility contributions', *Synthese* 2020:1-41.

[14] Parfit, Derek 1997, 'Equality and priority', *Ratio* 10:202-21.

[15] Pummer, Theron 2016, 'Whether and where to give', *Philosophy & Public Affairs* 44, no. 1:77-95.

[16] Rawls, John 1971, *A Theory of Justice* (Cambridge: Harvard University Press).

[17] Schoenfield, Miriam. "Decision making in the face of parity." *Philosophical Perspectives* 28 (2014).

[18] Tarsney, Christian, 'Exceeding expectations: Stochastic dominance as a general decision theory', unpublished manuscript.

[19] Vallentyne, Peter, & Kagan, Shelly 1997, 'Infinite value and finitely additive value theory', *Journal of Philosophy* 94, no. 1:5-26.

[20] Wilkinson, Hayden 2020. "Infinite aggregation: Expanded addition." *Philosophical Studies* (forthcoming).

[21] Wilkinson, Hayden. "Chaos, add infinitum." Unpublished manuscript.