Meeting of the Aristotelian Society held online on 24 May 2021 at 5:30 p.m.

XII—A NEW METHOD FOR VALUE AGGREGATION

KENNY EASWARAN

Many axiological theories ground the goodness or badness of options in the aggregate of the goodness or the badness of these options for individuals. Most commonly, this works by summing (or averaging), and taking the expectation of this result if there is uncertainty. Such theories face problems dealing with infinite populations, for which sums or averages are infinite or undefined. They fetishize certain mathematical operations, in a subject that is not inherently mathematical. The fact that the sum is the target of maximization is said to mean that the theory ignores the separateness of persons. I propose an extension of my 2014 paper, 'Decision Theory without Representation Theorems'. I illustrate how the resulting theory can account for a case involving an infinite population, dealing with the objections. I connect this to the theory of measurement, and explain why the mathematical operations are co-extensive with the result without grounding it. Because there is no sum in these infinite populations, it avoids the worry that aggregative theories treat individuals as secondary to some aggregate.

I

Introduction. Many axiological theories say that betterness and worseness of options[1] is in some sense grounded in the aggregate of the goodness or badness of these options for distinct individuals. This is most obviously true for certain consequentialist theories like utilitarianism, but at least some non-consequentialist theories have aggregative axiological components as well. Many such theories aggregate, not only over the actual goodness and badness of these options for agents, but also over a range of possible outcomes of these options, to deal with the normative significance of (subjective or objective) uncertainty.

[1] I use the word 'option' to remain neutral about whether the theory is about acts or social policies or some other sort of option. In this paper I only address the questions of axiology, about which options are better or worse, and not the questions of deontology, about which options are permissible or obligatory.

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

The most common method for aggregating these goodnesses and badnesses is first to represent each numerically, then within each possible outcome to sum these values for all agents, and then to take the 'expected value' of these totals, which is an average of the totals, weighted by the probabilities of the possible outcomes (using a subjective or objective probability function). In this paper, I develop a new method, extending the decision theory of Easwaran (2014a) to a method for aggregating across multiple agents.[2] In §II, I describe the general techniques I propose for developing such an ordering of options. In §III, I use these techniques to create one example ordering. In §IV, I illustrate how this method works for a case with an infinite population, showing that this method applies to cases for which traditional methods do not apply. I also show that when there is a finite set of agents and a precise, numerical probability for each outcome, it yields the same verdicts as calculating the expected value of the sum of the goodnesses for the agents. In §V, I give some suggestions of ways to begin to vary this formal method to accommodate alternatives to addition (for instance, to put greater importance on the well-being of those worse off) and to expected value (for instance, to put some importance on the difference between risk and certainty), as well as to aggregate across multiple dimensions of value that might be only partially commensurable, such as health, wealth and happiness.

One important limitation of the method I describe is that it requires that the set of agents is identical for each state of the world and for each option. Thus, as specified, it can't be used to address questions of 'population ethics' (such as whether it is better for there to be a large population with moderately good lives or a small population with exceedingly good lives). Extending the method to allow for varying sets of agents is an important project for future research.

The remainder of this section gives several motivations for the philosophical interest of the case with an infinite population. While some motivations depend on the possibility or actuality of an infinite population of agents that are affected by decisions, others merely use the infinite case to better clarify the metaphysical dependencies of

[2] Many authors use the word 'aggregate' to refer only to theories where there is some mathematical representation such as the sum of individual goodnesses that determines overall betterness and worseness. As will become clear later, my theory has no natural such representation, but it is still 'aggregative' in the sense that overall betterness and worseness is determined entirely by individual goodness and badness.

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

parts of the method even in finite cases. I claim that these motivations show some advantages of the new method I propose over either the simple method of adding and calculating expected value or the more general method of social welfare functions as in Adler (2019).

Infinite Ethics. The first set of motivations for considering these cases involving infinity is that various ethicists have thought that these cases are in fact of direct relevance to ethical theories. Aggregative theories that sum the goodnesses for individual agents are often said to suffer from 'infinitary paralysis' in such cases, because the total goodness in any such world will be infinite, and thus all options will be on a par. Some theorists use this possibility to motivate non-additive theories of aggregation, while others use it as a challenge to aggregative ethics, or to challenge the assumptions that led to the infinite population.

Vallentyne and Kagan (1997) just take for granted that there might be a world with infinitely many agents, and develop an ethical theory that could apply to such a world. Nick Bostrom (2011) supplements this with an appeal to contemporary cosmology to argue that it is a very real possibility that the *actual* world is infinite in extent, and contains infinitely many agents throughout its regions. Mark Johnston (2017) goes further and argues that four-dimensionalist metaphysics, or perhaps even just naturalism, implies that there are infinitely many 'personites' that occupy overlapping temporal extents of a single human body, each with its own moral consideration, so that even when only a single person is involved, there is already a threat of infinitary paralysis.

In the economics literature—for instance, Ramsey (1928), Diamond (1965), and Basu and Mitra (2003)—this infinite population is thought of as extending through time. A policymaker making decisions regarding infrastructure investment, climate change and contemporary economic consumption is thought of as making decisions that affect the welfare of both current and future generations. Since there is no known upper bound on the number of generations that society will last, these decisions are thought of as affecting infinitely many generations. In this context, infinitary paralysis is sometimes addressed by applying a 'discount rate' to the interests of future generations; but Ramsey notes that this is 'a practice which is ethically indefensible and arises merely from the weakness of the imagination', and his successors generally agree.

doi: 10.1093/arisoc/aoab008

Some of this literature focuses on showing the impossibility of finding a 'social welfare function' for an infinite population satisfying both a Pareto condition and an equity condition. A 'social welfare function' is a function that determines a single number to summarize the overall social welfare for any particular distribution of welfare of individuals. Like Vallentyne and Kagan, I avoid the problem by creating an aggregative method that results in a merely partial ordering rather than a numerical representation (which implicitly carries a total ordering). Bostrom and Arntzenius both challenge the method of Vallentyne and Kagan, saying that a partial ordering for social welfare is not sufficient to deal with the sort of aggregation across possibilities that occurs when there is uncertainty about the outcome. Bostrom proposes a method where persons are aggregated first, and uncertainty second. Arntzenius proposes a method where uncertainty is aggregated first, and persons second. I use an analogy proposed by Harsanyi (1953) (and later associated also with Rawls) to suggest that either asymmetry is unappealing. The method I propose is provably commutative, so that there is no fundamental asymmetry between the two. But the end result is still a partial ordering, so it avoids the impossibility results of the economists.

Like the methods proposed by other authors in this literature, my method depends essentially on the ordering of the agents in an infinite sequence. As these authors note, this is an unappealing feature of the method. None of these methods can be considered a proper solution to these sorts of cases unless the normative significance of this ordering can be justified, which may be easier for the economic motivation involving a sequence of generations than for Johnston's motivation involving 'personites'.

Mathematical Representation. Another motivation for this method is to give a proper explanation for the appearance of various mathematical entities and operations in an aggregative ethical theory. Many standard aggregative theories assume there is such a thing as a numerical representation of individual goodness, and then say that a state of affairs is better if and only if the sum (or average) of these numerical values is greater. However, this often leaves unexplained why *addition* (or averaging) is the relevant operation, rather than multiplication, weighted averaging, or some other social welfare function. One aim of the method I propose is to show what sort of

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

philosophical work needs to be done to ground the particular mathematical operation that is used.

Work of this sort for mathematical operations in other applications has been developed for decades in the field of 'measurement theory' (with Krantz et al. 1971 as a classic reference). One important result is that there is some conventionality of the choice of operation: any physical operation that can be represented by the mathematical operation of addition can also be represented by the mathematical operation of multiplication, if the numerical representation is sufficiently altered. (This is the principle by which a slide rule works: it has at least two sets of numbers on it, one combining additively and the other combining multiplicatively, enabling easy conversion between different mathematical operations.) Field (1980) shows how these methods can even be extended to develop a version of Newtonian gravitational mechanics without any appeal to mathematical entities, grounding all explanations in the physical. The particular mathematical operations are then shown to yield results that are co-extensive with the physical relationships. My aim is to do something similar for value theory, showing what sort of postulates are needed to ground the meaningfulness of various mathematical operations. Even if one disagrees with the particular postulates I propose, they provide an illustration of the sort of postulates that are needed.

Easwaran (2014a) shows that a particular preference structure on acts with uncertain outcomes gives rise to an overall structure that agrees with expected value decision theory when expected values are defined, but still yields a consistent theory of preference among acts for which expected values are not defined. The method developed later in this paper extends this method to cases of interpersonal decision, and shows that in cases where they are well-defined, additive interpersonal aggregation and expected value decision theory agree with the results of this method, but the method yields results even in cases where these are undefined. Even the use of numbers to represent goodnesses can to some extent be abstracted away from, as discussed in §v.

Separateness of Persons. Another point of my infinite examples is to refute a certain objection to aggregative theories. This objection suggests that aggregative theories ignore the 'separateness of persons', and perhaps treat individuals as mere 'receptacles of value'

Proceedings of the Aristotelian Society, *Vol.* cxxi, *Part* 3
doi: 10.1093/arisoc/aoab008

rather than as having value of their own. Norcross (2009) gives both a metaphysical interpretation of this objection (alleging that utilitarianism is committed to all individuals being nothing more than parts of a single 'super-person', whose interests are the ones that matter) and a moral interpretation (questioning the legitimacy of trading off the interests of one against the interest of another), and argues that neither is a proper objection. I don't have anything to add in response to these two versions of the objection, but I think there is a third, subtly different version of the objection between the two which my method does address. As Chappell (2015) puts this objection, it is about treating individuals 'as constitutive means to the aggregate welfare, rather than as distinct ends in themselves'. Chappell goes on to argue that utilitarianism is not committed to this sort of treatment, but can in fact see individuals as ends in themselves.

Gustafsson (2021) makes this argument more explicitly. He shows that additive utilitarianism for a fixed finite population is extensionally equivalent to a theory generated by four axioms that each satisfy an individualist constraint. However, it remains open to an opponent to say that these axioms are somehow *grounded* in the sum. If two accounts are extensionally equivalent, it is sometimes hard to say which is prior and explains the other.

One role that the infinite case plays in this paper is to help make this response to this particular interpretation of the challenge mathematically very clear. In the infinite case, there is *no* numerical representation of the aggregate welfare, but only representations of the welfare of each individual. One outcome or option will be better than another *because* the option is good *for* particular individuals in particular states of the world, rather than *because* this outcome has some greater aggregate value. My method works exactly the same in the infinite case as in the finite case, which illustrates that even in the finite case, the method doesn't *depend* on the aggregate value. This purpose is served even if the infinite cases aren't actually possible—they still help illustrate the dependence among the aspects of the method, and we don't have to worry about whether the dependence on the ordering of the agents makes the theory get the comparisons 'wrong'.

I do not address the more fundamental objection to trading off goodness for one individual against goodness for another, which Norcross and Chappell aim to address. On some level, I think there

doi: 10.1093/arisoc/aoab008

is no non-question-begging response to this objection, because on some level the objection just is the denial of aggregation. However, because of the parallel my method draws between agents and states, my method suggests a parallel objection. I say that trading off goodness for one agent against goodness for another is analogous to trading off goodness in one possible state (say, if the coin comes up heads) against goodness in another (if the coin comes up tails). In §v I outline a few ways that alternative trade-off equivalences could be defined. There are some formal similarities between the ways that one might accommodate risk aversion in trade-offs among possible outcomes and inequality aversion in trade-offs among persons. This suggests that further work on these analogies might help understand the full force of the 'separateness of persons' worry, and whether an analogous worry should be considered regarding the 'separateness of possible states'. Perhaps the actual world ought to enter differently into moral theories that deal with uncertainty from the way in which other possible outcomes do. Modal realists who insist that every possible outcome is in some sense real might further insist that it is wrong to trade off the interests of people in one possibility against the interests of people in another.

II

General Formal Tools. I assume throughout that there is a fixed set $\mathcal{A}$ of *agents* and a fixed set $\mathcal{S}$ of possible *states*. I assume that each *option* yields a well-defined goodness for each agent in each state, and that these goodnesses can be represented by real numbers. (This assumption is justified, and alternatives are considered, in §v.) Thus each option can be represented by a function assigning a real number to each agent-state pair. I let $\mathcal{O}$ be the set of all functions from agent-state pairs to real numbers. The primary aim of this paper is to develop a method for defining a betterness relation on $\mathcal{O}$. Even though many members of $\mathcal{O}$ aren't actually options that anyone has available, their positions show where options *would* be if there *had been* options with those distributions of goodness, and I claim that they can help us understand why the actual options are compared as they are.

If O is a specific member of $\mathcal{O}$, there are two natural classes of entities to associate with it. For any agent A in $\mathcal{A}$, there is a *prospect*

Proceedings of the Aristotelian Society, *Vol.* cxxi, *Part 3*
doi: 10.1093/arisoc/aoab008

for that agent associated with O, which is the amount of goodness that option will yield for that agent in each state. I represent this prospect by the function $O(A, \cdot)$, which I will also write as $O(A)$. I let the set of prospects $\mathcal{P}$ be the set of all functions from states to real numbers.

Similarly, for any state S in $\mathcal{S}$, there is a *world* for that state associated with the option, which is the amount of goodness that option will yield for each agent in that state. I represent a world by the function $O(\cdot, S)$, which I will also write as $O(S)$. (Since agents and states are distinct entities, there is no ambiguity about whether a notation with an option followed by an entity in parentheses denotes a prospect or a world, though this is strictly an abuse of notation.) I let the set of worlds $\mathcal{W}$ be the set of all functions from agents to real numbers.

The central formal tools I use are partial orderings[3] and equivalence relations. These are relations, and each relation has associated with it a *domain*, which is the set of entities that are apt to bear the relation to each other—usually either $\mathcal{O}$, $\mathcal{P}$ or $\mathcal{W}$.

A *weak partial ordering*, which I will always notate with '$\geqslant$', is a relation that is reflexive (for all x in its domain, $x \geqslant x$) and transitive (for all x, y, z in its domain, if $x \geqslant y$ *and* $y \geqslant z$, then $x \geqslant z$). A *strict partial ordering*, which I will always notate with '$>$', is a relation that is irreflexive (for all x in its domain, $\neg(x > x)$) and transitive (for all x, y, z in its domain, if $x > y$ and $y > z$, then $x > z$). An *equivalence relation*, which I will always notate with '$=$', is a relation that is reflexive (for all x in its domain, $x = x$), transitive (for all x, y, z in its domain, if $x = y$ and $y = z$, then $x = z$), and symmetric (for all x, y in its domain, if $x = y$, then $y = x$).

When '$>$' or '$\geqslant$' appear without a subscript, they will stand for the underlying relation on goodness for a particular individual in a particular state, and I will always use a subscript when they are meant to stand for other partial orderings, including any orderings on worlds, prospects or options. When '$=$' appears without a subscript, it will stand for the relation of strict identity, and I will always use a subscript when it is meant to stand for another equivalence relation.

Many other methods (notably, anything that falls under the rubric

[3] Technically, mathematicians will call these *pre*-orderings, because distinct objects can both bear the weak relation to each other.

doi: 10.1093/arisoc/aoab008

of 'social welfare functions', as in Adler 2019) require that betterness be a *complete* ordering rather than a partial ordering; that is, for every x, y in its domain, either $x \geqslant y$ or $y \geqslant x$. My method does not. Following my work in Easwaran (2014a), I start with a very limited pair of partial orderings $>_{OD}$ and $\geqslant_{OD}$ on O, and find a set of equivalence relations $=_i$, and use them to extend these partial orderings to more detailed partial orderings $>_i$ and $\geqslant_i$ that represent the candidate for overall betterness. I don't claim that the particular partial orderings I end up with tell the full story about betterness, just that they give part of it. It is plausible that further extensions should be required, even if all the extensions I have made so far are correct.

However, in addition to allowing the method to proceed sequentially in this way, the partiality of these orderings is not merely because of the possibility of further extension. It is plausible that any normatively significant ordering here *must* be partial. As proved by Dubey (2011), there is no way to construct a total ordering satisfying both a 'weak Pareto condition' and an 'anonymity condition' without using the axiom of choice. As I have said in earlier work (Easwaran 2014b), I find it implausible that any ordering depending essentially on the axiom of choice could be normative. Since the weak Pareto condition and anonymity are parts of any plausible aggregative theory—I show later where they arise in my method—I think this means we should expect any normative ordering on infinite populations to be partial.

The main equivalence relations I consider are generated by transformation groups on $\mathcal{O}$, $\mathcal{P}$ or $\mathcal{W}$. A function from a set to itself is called a *permutation* of that set if and only if it is one-to-one, total, and onto. A set G of permutations of a set is said to be a *transformation group* on that set if and only if the identity transformation is in the set, the inverse of any function in the set is also in the set, and the composition of any two functions in the set is also in the set. For any collection of permutations of a set, the transformation group *generated by* those permutations is the set of all permutations that result from composing any finite sequence of permutations or their inverses from that collection. This is the smallest transformation group containing all those permutations.

If G is a transformation group on a set, then we can define an equivalence relation $=_G$ on that set by saying $x{=_G}y$ if and only if there is a g in G with $g(x) = y$. This relation is reflexive because the identity transformation is in the group, symmetric because the

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

inverse of a transformation is in the group, and transitive because the composition of any two transformations in the group is in the group. Intuitively, I will use transformation groups which represent ways of transforming options into other options that should be considered equally good (though as I suggest later, at least some might better be interpreted as 'almost equally good').

I say a transformation group G on a set *commutes* with a strict (or weak) partial ordering $>_i$ (or $\geqslant_i$) on the same set if and only if for every x, y in the set, and every g in G, $x >_i y$ (or $x \geqslant_i y$) if and only if $g(x) >_i g(y)$ (or $g(x) \geqslant_i g(y)$).

I say a transformation group G on a set is *compatible* with a strict partial ordering $>_i$ on the same set if and only if there is no x in the set and g in G, with $x >_i g(x)$. Intuitively, if a transformation group is supposed to represent ways of transforming options into other options that are *equally* good, then it shouldn't transform one option into one that is strictly *worse* than it.

The basic tool I use for extending one pair of partial orderings to a broader one is to use a transformation group that is compatible with the strict ordering and commutes with both. If $>_\mathrm{i}$ and $\geqslant_\mathrm{i}$ are the initial orderings, and G is the group, then I define $x >_G y$ if and only if there is a g in G with $x >_i g(y)$, and $x \geqslant_G y$ if and only if there is a g in G with $x \geqslant_i g(y)$. The central theorem of Easwaran (2014a) shows that these new relations are also a strict and weak partial ordering. The basic idea is that if $>$ tells us that x is better than y, and G tells us that y is equally good as $g(y)$, then x should be better than $g(y)$.

III

A Worked-Out Example Ordering. This section uses the tools given above to create a partial ordering of options. I start with a simple partial ordering of options involving strict dominance, and extend it with two equivalence relations given by transformation groups—one giving trade-offs between states and the other giving trade-offs between persons. The particular ordering I end up constructing by these means is not a plausible candidate for a complete and correct axiological theory, but as I show in §IV, it yields plausible results on a particular space of options, and illustrates the features I mentioned in §I.

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

Dominance. The partial orderings I begin with are based on the concept of dominance, also known in the economics literature as a 'Pareto condition'. Versions of the partial orderings can be defined on $\mathcal{O}$, $\mathcal{P}$ and $\mathcal{W}$. If O_1 and O_2 are options, then I define $O_1 >_{OD} O_2$ if and only if there is some $\epsilon > 0$ such that for every agent A and state S, $O_1(A, S) > \epsilon + O_2(A, S)$, and I define $O_1 \geq_{OD} O_2$ if and only if for every agent A and state S, $O_1(A, S) \geq O_2(A, S)$. ($>_{PD}$ and $\geq_{PD}$ are defined similarly on $\mathcal{P}$, and $>_{WD}$ and $\geq_{WD}$ are defined similarly on $\mathcal{W}$. When referring to a generic relation in these families I will refer to '$>_{xD}$' or '$\geq_{xD}$'.)

There are two slightly weaker strict dominance relations that one might naturally consider in place of my $>_{OD}$, and it is important to note that the one I have defined is distinct from either. One is the relation $>_{OD1}$ that holds between O_1 and O_2 if and only if, for every A and S, $O_1(A, S) > O_2(A, S)$. (This is the one Dubey calls the 'weak Pareto condition'.) Another is the relation $>_{OD2}$ that holds between O_1 and O_2 if and only if, for every A and S, $O_1(A, S) \geq O_2(A, S)$ and additionally, there exists A_0 and S_0 with $O_1(A_0, S_0) > O_2(A_0, S_0)$. (This is the one Dubey calls the 'strong Pareto condition'.) Both of these are plausible candidates for inclusion in a partial ordering of the sort I would like to construct, but the latter is not compatible with the particular equivalence relations I include in my example, and I haven't determined whether the former is. It is in some ways better to think of the ordering I produce as representing one option being 'much better than' another, while these other two are better candidates for building a relation that captures 'slightly better than' as well.

World Shifts. To define the first transformation group, I need to assume a structure on the set $\mathcal{S}$ of states. I assume that there is a numerical probability function P defined on an algebra of subsets of $\mathcal{S}$. That is, $P(\mathcal{S}) = 1$, and if s_1 and s_2 are disjoint subsets of $\mathcal{S}$, then $P(s_1) \geq 0$, $P(s_2) \geq 0$, and $P(s_1 \cup s_2) = P(s_1) + P(s_2)$.

Now, for any set X of agents, any two disjoint sets s_1 and s_2 of states, and any real number k, I can define a permutation $g_{X,s_1,s_2,k}$ on $\mathcal{O}$ as follows. If O is any element of $\mathcal{O}$, then I define

Proceedings of the Aristotelian Society, Vol. CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

$$g_{X,s_1,s_2,k}(O)(A,S) = \begin{cases} O(A,S) - k & S \text{ in } s_1 \text{ and } A \text{ in } X \\ O(A,S) + k & S \text{ in } s_2 \text{ and } A \text{ in } X \\ O(A,S) & \text{otherwise} \end{cases}$$

That is, this transformation involves taking away k units of goodness from members of X in states in s_1 and giving k units of goodness to members of X in states in s_2.

I define G_{WS} to be the group generated by the set of all permutations $g_{X,s_1,s_2,k}$, where X is any set of agents, s_1 and s_2 are any two sets of states with $P(s_1) = P(s_2)$, and k is any difference in basic goodness.[4] It is straightforward to see that G_{WS} commutes with $>_{OD}$ and $\geq_{OD}$. Proving that G_{WS} is compatible with $>_{OD}$ takes a bit more work.[5]

It turns out that G_{WS} is not compatible with the strongest version of dominance, $>_{OD2}$. If O_1 is any option, and O_2 is identical except that in a set of worlds of probability 0 some set of agents are better off, then $O_2 >_{OD2} O_1$, but $O_2 =_{WS} O_1$. However, the proof given for compatibility with $>_{OD}$ in fact shows compatibility with $>_{OD1}$ as well.

Person Shifts. There are several normative problems for the second transformation group I consider, so this should be taken even more strongly as a placeholder in a proof-of-concept, rather than a candidate for a correct normative theory. The first problem is that to define this group, I need to assume a particular structure on the set $\mathcal{A}$ of agents. For now, I assume the set of agents is countable, and assume a fixed numbering of the agents $A_0, A_1, A_2, \ldots$ by the natural numbers. The dependence on this structure is shared by other aggregative approaches to infinite populations (such as Vallentyne and Kagan 1997, Basu and Mitra 2003, Arntzenius 2014, and Bostrom

[4] It is also possible to define a very similar transformation group on $\mathcal{P}$, the set of prospects, if X is just the set consisting of the single agent for whom this prospect applies. This is effectively what is done in Easwaran (2014a, §3.5).

[5] Proof: Assume that O_0 is related to O_n by a sequence of permutations $g_1 = g_{X_1,s_{11},s_{21},k_1}, g_2 = g_{X_2,s_{12},s_{22},k_2}, \ldots g_n = g_{X_n,s_{1n},s_{2n},k_n}$, with $O_{i+1} = g_{i+1}(O_i)$, for $0 \leq i < n$. Fix some particular agent A and consider the prospects $O_i(A)$. Although the prospects $O_i(A)$ may not have well-defined expected values, it is clear that the *difference* between prospects $O_i(A)$ and $O_{i+1}(A)$ must have expected value 0, and in fact must only take on the finitely many values $+k$, $-k$, and 0. Since the expected value of a sum is the sum of the expected values, the difference between prospects $O_n(A)$ and $O_0(A)$ must have expected value 0 as well, and must also take on only finitely many values. If a function takes on only finitely many values and has expected value 0, then not all of those values can be positive. Thus it is not the case that $O_0(A) >_{PD} O_n(A)$, and so it is not the case that $O_0 >_{OD} O_n$, QED.

Proceedings of the Aristotelian Society, Vol. CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

2011), so I see this as a shared problem for all methods (like the problem of interpersonal utility comparison) rather than a distinctive problem for my method.

The second problem, which I see as more significant, is that with the new group H_{PS}, it will turn out that although $=_{H_{PS}}$ is compatible with $>_{OD}$, it is *not* compatible with $>_{OD2}$. In particular, if option O_1 makes everyone better off in every world than option O_2 by at least ϵ, then we will have $\neg(O_1 =_{H_{PS}} O_2)$; but if there are finitely many people that are better off in O_1 than O_2, and everyone else is equal, then $O_1 =_{H_{PS}} O_2$. I think this makes it plausible to think that $=_{H_{PS}}$ doesn't mean that the two options are *equally* good, but perhaps suggests they are *almost equally* good. If we let $>_F$ and $\geq_F$ be the relations we get by composing $>_{OD}$ and $\geq_{OD}$ with $=_{G_{WS}}$ and $=_{H_{PS}}$, then this means that if $O_1 >_F O_2$, we should trust this judgement, but if $O_1 \geq_F O_2$ and $O_2 \geq_F O_1$, then we shouldn't necessarily trust this judgement that the two are actually equally good. (I suspect my proof of compatibility extends to compatibility with $>_{OD1}$, but I'm not certain.)

For any set X of agents, and any set s of states, natural number n, and difference of goodnesses k, I define

$$h_{X,s,n,k}(O)(A_i, S) = \begin{cases} O(A_i, S) - k & S \text{ in } s,\ A_{i+n} \text{ not in } X, \text{ and } A_i \text{ in } X \\ O(A_i, S) + k & S \text{ in } s,\ A_{i+n} \text{ in } X, \text{ and } A_i \text{ not in } X \\ O(S) & \text{otherwise} \end{cases}$$

That is, this transformation involves taking away k units of goodness from any member of X in states in s, and giving k units of goodness to any person 'n spots to the right' of a member of X in states in s. For a person that is both in X *and* n spots to the right of a member of X, the net result is no change.

I define H_{PS} to be the group generated by the set of all permutations $h_{X,s,n,k}$.[6] It is straightforward to see that H_{PS} commutes with $>_{OD}$ and $\geq_{OD}$. Proving that H_{PS} is compatible with $>_{OD}$ takes a bit more work.[7]

[6] It is also possible to define a very similar transformation group on W, the set of worlds, if we just ignore s. When X is just a single person, this formalizes the 'anonymity' condition of Dubey.

[7] Lemma: If $h_{X,s,n,k}(O_1) = O_2$, then for any S in s, $\left(\sum_{i=a}^{i=b} O_1(A_i, S)\right) - \left(\sum_{i=a}^{i=b} O_2(A_i, S)\right) \leq nk$.

We can prove the lemma by noting that the difference between the sum of O_1 and the sum of O_2 over a particular interval will ignore any shift in goodness that occurs

doi: 10.1093/arisoc/aoab008

It is also straightforward to see that H_{PS} commutes with G_{WS}. (The generators of H_{PS} commute with each other, and when commuting with a generator of G_{WS}, one just applies the element of G_{WS} to the set s of states used in that generator.) The proof that the two groups together are compatible with $>_{OD}$ just puts together the ideas of the proof of the compatibility of each individually.

To see that H_{PS} is *not* compatible with $>_{OD2}$, consider the element $h_{\mathcal{A},\mathcal{S},n,k}$. This element takes k units of goodness away from *every* agent (since every agent is in $\mathcal{A}$), and gives k units to every agent who is n spaces to the right of anyone. That is, this takes k units away from agents $A_0, \ldots, A_{n-1}$, and leaves everyone else unchanged. As a result, this particular equivalence relation treats options as equally good if they differ only for finitely many people. H_{PS} is therefore not a good candidate for telling us when two options are precisely equally good. But since $>_F$, the composition of $>_{OD}$ with G_{WS} and H_{PS}, is a strict partial ordering that extends $>_{OD}$, it is plausible that its comparisons are correct as judgements that one option is *infinitely* better than another.

The two transformation groups I propose both work in the same sort of way. G_{WS} uses probability to determine when two sets of states are 'equal', and says that two options are equal if and only if they differ by the same amount for some fixed set of agents in those two sets of states. H_{PS} uses a translation by n spaces to determine when two sets of agents are 'equal', and says that two options are equal if and only if they differ by the same amount in some fixed set of states for the agents in those two sets. Different specifications of when two sets are 'equal' will yield natural alternatives to these two transformation groups. There may also naturally be some transformations to consider that involve interactions between agents and

entirely within the interval. It tells us how much goodness is shifted into the interval (which can only be at most k units of goodness per agent in the top n positions in the interval) and how much goodness is shifted out of the interval (which can only be at most k units of goodness per agent in the bottom n positions in the interval). Thus, the total difference can't have absolute value greater than nk.

Now consider what happens if $O_\circ$ is related to O_m by a set m of individual generators to get from one to the other, and $O_\circ$ dominates O_m by at least ϵ. If n is the maximum of the distance of the shifts involved, k is the maximum of the absolute values of the amount of goodness shifted, and m is the number of steps, then consider $\left(\sum_{i=a}^{i=a+1+mnk/\epsilon} O_0(A_i, S)\right) - \left(\sum_{i=a}^{i=a+1+mnk/\epsilon} O_m(A_i, S)\right)$. By m applications of the lemma, this difference is at most mnk. But if one of them ϵ-dominated the other, then the difference would be at least $mnk + \epsilon$. Thus, the group is compatible with the ordering, QED.

Proceedings of the Aristotelian Society, *Vol.* cxxi, *Part* 3
doi: 10.1093/arisoc/aoab008

states that I do not consider here. I do not claim that the particular dominance relations and particular transformation groups I consider here are the *right* ones to use, but just claim that they are *plausible*, and illustrate some important features of the particular aggregative method I propose, as well as of aggregation generally.

IV

An Example Comparison of Options. Nick Bostrom asks us to consider worlds W_1, W_2, W_3 with the following goodnesses for all persons:

$$W_1(A_i) = i+2$$
$$W_2(A_i) = \begin{cases} i+3 & i \text{ is divisible by } 3 \\ i & \text{otherwise} \end{cases}$$
$$W_3(A_i) = 1$$

Bostrom notes that the methods of Vallentyne and Kagan (1997) allow us to say that W_1 is better than W_2 and that W_2 is better than W_3, but also notes that we want to be able to compare *options* and not just worlds. We want to know whether an option that yields W_2 for certain is better or worse than one with a chance of yielding W_1 and a chance of yielding W_3.

I will consider a situation in which there are $n+1$ states $\mathcal{S} = \{S_0, S_1, \ldots, S_n\}$, all of equal probability, with O_1 yielding W_3 in S_0 and W_1 otherwise, and O_2 yielding W_2 for certain. I will show that in $>_F$, the final ordering achieved by composing $>_{OD}$ with G_{PS} and H_{WS}, we have $O_2 >_F O_1$. That is, even though W_1 is infinitely better than W_2, and W_2 is infinitely better than W_3, W_2 is infinitely better than any non-extreme gamble between W_1 and W_3. (Thus this sample ordering violates a continuity principle that some authors endorse.)

First, let $W_2'(A_i) = i+1$ for all i. Let $X = \{A_{3i}\}$ be the set of all agents whose index is a multiple of 3. Then $h_{X,s,2,1}$ is the person shift that shifts 1 unit of goodness from each person whose index is a multiple of 3 to the person two spaces to their right, and $h_{X,s,1,1}$ is the person shift that shifts 1 unit of goodness from each person whose index is a multiple of 3 to the person one space to their right. Putting

Proceedings of the Aristotelian Society, Vol. CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

these two shifts together converts W_2 to W_2'. Since $W_1 >_{WD} W_2' >_{WD} W_3$, we can recover the initial judgement of Vallentyne and Kagan. Furthermore, if O_2' is the option that yields W_2' for certain, then $O_2 =_{PS} O_2'$.

Now let $W_1'(A_i) = i$ for all i, and let $W_3'(A_i) = 2n+1$. Recall that there are $n+1$ states, $S_0, \ldots, S_n$, and that O_1 is the option that yields W_3 in S_0 and W_1 in all other states. Let O_1' be the option that yields W_3' in state 0 and W_1' in all other states. Let $g_m = g_{\mathcal{A},\{m\},\{0\},2}$ be the world shift that takes 2 units of goodness from everyone in state m and gives 2 units of goodness to everyone in state 0. Then applying $g_1, \ldots, g_n$ successively to O_1 turns it into O_1'. Thus, $O_1' =_{G_{WS}} O_1$.

Under O_1', only if $i < 2n+1$ does agent A_i ever receive more than i units of goodness (and only in S_0). Recall that O_2' is the option that yields W_2 in every state, so that A_i receives $i+1$ units in every state. Thus, O_2' yields at least a full unit more goodness than O_1' in every state for all but the first $2n+1$ agents. From the observation above that person shifts can change the goodness for any finite set of agents, we see that $O_2' >_{H_{PS}} O_1'$.

Putting all these observations together, we see that $O_2 >_F O_1$. Thus this method is able to provide a comparison of options involving infinitely many persons, as well as uncertainty, without ever directly computing a sum or an expectation. We just observe that some options are equivalent to other options under person shifts or world shifts, and eventually find options equivalent to the original ones, where one dominates the other.

Infinite Ethics. Bostrom and Arntzenius both consider this case, and give examples of methods that can yield the same result that my method yields. However, both of them describe methods that have an unpleasant asymmetry, unless one brings in tools from advanced set theory.

Both authors evaluate options in a three-step process. Arntzenius (2014) tells us to identify the prospects $O(A)$ for each agent, calculate the expected value of each prospect, and then find a way to compare the resulting infinite sequences (either the method of Vallentyne and Kagan, or another method described later). Bostrom (2011), by contrast, tells us to identify the worlds $O(S)$ for each state, find a way to assign a value to these infinite sequences of goodnesses, and

Proceedings of the Aristotelian Society, *Vol.* cxxi, *Part* 3
doi: 10.1093/arisoc/aoab008

then calculate the expectation of these values using the probabilities of these worlds. Each considers a few ways to aggregate across persons, but makes an arbitrary choice to do this either before or after applying expected value to aggregate across states.

Harsanyi makes statements suggesting that we should treat states and agents symmetrically, rather than prioritizing one over the other: '[A] value judgment on the distribution of income would show the required impersonality to the highest degree if the person who made this judgment had to choose a particular income distribution in complete ignorance of what his own relative position (and the position of those near to his heart) would be within the system chosen. This would be the case if he had exactly the same chance of obtaining the first position (corresponding to the highest income) or the second or the third, etc., up to the last position (corresponding to the lowest income) available within that scheme' (Harsanyi 1953, pp. 434–5). That is, when comparing worlds with different distributions of goodness across persons, a person should think of the different agents as possibilities for who the person will be, and evaluate these worlds the same way they would evaluate the corresponding prospects. Versions of this idea have also become famous among philosophers through the Rawlsian 'veil of ignorance', which involves treating the different agents in the world as if they were different possible states for a person facing an individual decision problem.

In my method, the world-shift and person-shift trade-offs commute with each other, so we treat states and agents symmetrically. Neither aggregation has priority over the other, even though they aren't done 'simultaneously'. Arntzenius and Bostrom both observe that there is a way to make their methods have the same sort of commutativity.

The way to do this is to use a mathematical construct known as an 'ultrafilter' to assign values from a mathematical structure known as the 'hyperreals' to infinite sequences of real numbers. Three details of this construction matter here. One is that any form of calculation that can be done on standard reals, such as expected value, also has a counterpart that can be done on hyperreals. The second detail is that doing a calculation on the hyperreals resulting from some infinite sequences always yields the same result as doing that same calculation on the individual real numbers in the sequence and then finding the hyperreal corresponding to the resulting sequence,

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

so that this restores the commutativity that is lost by applying the Vallentyne and Kagan method to the expectations. The third, and most significant for my purposes, is that constructing the hyperreals, and creating an ultrafilter generally, requires the set-theoretic axiom of choice, and involves an infinite set of arbitrary choices.

However, as mentioned above and in my Easwaran (2014b), I claim that no normative theory should depend on any mathematical entity proof of whose existence requires the axiom of choice. Although such entities are perfectly legitimate members of whatever sort of realm mathematical entities occupy, they are too complex to be relevant to agents in the real world. These entities also typically have uncountably infinitely many counterparts, each yielding different results for the application at hand. If there were some non-arbitrary way to choose among these counterparts, then the construction would not require the axiom of choice. Thus it seems unclear how any one of them rather than any other could become the one that is normative.

Like the methods of Bostrom and Arntzenius, but unlike that of Vallentyne and Kagan, my method can account for betterness relations among options with infinite populations as well as uncertainty. Unlike the methods of Bostrom and Arntzenius, my method can treat states and persons symmetrically without requiring the axiom of choice. Thus, I claim, my method yields an account of infinite ethics superior to these competitors.

Mathematical Representation. Both Bostrom and Arntzenius take the normativity of certain mathematical operations for granted. In the case of the expected value calculation, there are various justifications of it in the literature, but for the use of the ultrafilter to get a hyperreal representation, there is very little justification. But I claim that even in the finite case, where instead of an ultrafilter we appeal to addition, the use of this particular mathematical operation needs to be justified. I claim that the transformation groups I appeal to can explain why expectation and addition yield the right comparisons in finite cases.

If there are only finitely many agents, then $>_{OD}$ and G_{WS} can apply exactly as in the case with infinitely many agents. However, H_{PS} needs to be slightly modified to remove permutations moving goodness from agents near the end of the list to positions that are beyond the end of the list. The natural permutation group is in fact generated

doi: 10.1093/arisoc/aoab008

by the set of shifts from any individual agent to any other, or by the set of shifts from any finite set of individual agents to any other set of the same cardinality.[8] Under this finite version of H_{PS}, it turns out that every world is equivalent to a world that yields zero goodness for all agents but the first, who receives goodness equal to the sum of goodnesses for all agents in the original world.

Thus this person-shift trade-off group explains the use of the mathematical operation of addition in finite cases. It is because a world in which one person gets $x+k$ units of goodness and another gets y units is just as good as one in which the first person gets x units of goodness and the other gets $y+k$ units that overall betterness tracks the sum of the goodnesses. (This is similar to the argument of Gustafsson 2021 that his axioms yield a betterness relation co-extensive with that of additive utilitarianism.) In §V, I give the beginnings of an exploration of how alternative approaches to trade-offs might be able to yield results co-extensive with non-additive theories.

Results in Easwaran (2014a) show that if there is only a single agent, then if the set of states is divisible—that is, for every set of states of positive probability, there are two disjoint subsets of positive probability—then whenever the expected value of two options are distinct real numbers, the one with the higher expected value is preferred to the other. In that paper, I argued that this showed that the world-shift trade-off group explains the use of expected value for options in which it is defined. The method also yields distinct comparisons for options without well-defined expected values, which may be of further interest if one thinks that a normative theory should be able to apply to such examples. But even if one thinks that only finite options with well-defined expected values are relevant, this method shows that expected value follows from a commitment to certain trade-offs and doesn't take this calculation either as primitive, or as a consequence of an abstract mathematical representation theorem.

The only reason these calculations of addition across agents or expected value across worlds matter, on this account, is because they

[8] As shown by Jonsson (2018), the group I consider in the infinite case is not generated by either the set of shifts among finite sets of agents or shifts among infinite sets of the same cardinality. Figuring out precisely which sets of agents to treat as 'equal' in the infinite case is one of the biggest difficulties for identifying a properly normative example of the technique I develop.

Proceedings of the Aristotelian Society, Vol. CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

happen to yield results that agree with the method proposed here. But because this method yields results in cases where sums or expected values aren't well-defined, either because the population is infinite or because the probabilities line up in a way that defies the calculation of expected value, we can see that the explanation can't go the other way. The work I have done here is not as complete as the work done in Krantz et al. (1971) to explain the conventional choice of addition as an operation on distances or masses, but I claim that it is a step towards this. Like Field (1980), I show that the use of the mathematical operations is a conservative way to reason about the underlying facts (normative in my case, physical in his), but the fundamental facts do not themselves involve mathematical operations.

Separateness of Persons. The fact that this method works just as well in the infinite case as in the finite case also enables a response to one version of the 'separateness of persons' objection to aggregative moral theories. This version of the objection alleges that such theories take the aggregate to matter primarily, and individuals matter only in terms of their contribution to that aggregate. However, for the objection to arise, there must *be* some aggregate. In this case it is straightforward to prove that there is no function $f : \mathcal{O} \to \mathbb{R}$ such that $O_1 >_F O_2$ only if $f(O_1) > f(O_2)$.[9] Thus the theory I describe at least does not work by saying there is some such thing as the aggregate, having a real-number value, such that goodness for individuals is secondary to this fundamental aggregate goodness. (Note that by the definitions of Hirose 2013 or Gustafsson 2021, this means that my theory is not 'aggregative'. But my point is that a theory can be aggregative in the sense of saying that overall betterness depends on only the goodness and badness for each individual, without there being an aggregate of this sort.)

In the case with a finite population, we can represent this aggregate as just the sum of the goodnesses for the individuals. But I claim that whatever motivates the aggregation method in the finite case

[9] If there were such a function, then there would be a countable set of options such that no option is strictly preferred to all of them. However, for any countable set of options, it is straightforward to come up with an option that is strictly preferred to all of them. If $O_1, O_2, O_3\ldots$ are the countable set of options, then we define a new option O by $O(A_n, S) = \epsilon + max_{i<n}(O_i(A_n, S))$. For any O_i, O is better than O_i by at least ϵ in every world for all but finitely many agents, and thus by person-shift equivalence, O is preferred to O_i overall.

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

must be the same as whatever motivates the same method in the infinite case. Since there is, in general, no sum of the goodnesses in the infinite case, this must not be what motivates the method. Thus I claim that the importance of the sum of goodnesses for agents in a finite case is grounded in the importance we see for trade-offs between persons, and not in the aggregate.

One can of course construct *some* mathematical set X to interpret as the goodness of the aggregate, and a function $f : \mathcal{O} \to X$ such that $O_1 >_F O_2$ if and only if $f(O_1) > f(O_2)$. (For instance, the objects could be the equivalence classes of options under $=_F$.) However, without any appealing intrinsic characterization of X, it's implausible to say that X is more fundamental than the individuals. This would be especially implausible if in fact the final aggregate ordering is a partial ordering, while goodness for individuals can be represented more simply in an ordering like the real numbers.

As argued by Chappell (2015) and Gustafsson (2021), everything has been defined in terms of goodness for individuals. While some versions of the 'separateness of persons' objection reject even the possibility of trading off benefits to some against harms to others, the type of aggregative theory I describe at least still treats the *persons* as fundamental, and fundamentally distinct, even though their interests can be traded off against each other.

Even if one objects to the particular way trade-offs have been implemented in the infinite case, the fact that this method yields consistent comparisons without a numerical representation of aggregate value indicates that the comparisons do not depend on this aggregate value. Getting the details of the transformation groups and the partial orderings right is important for the use of this method in actual infinite ethics. But if one thinks the finite cases are the ones that actually matter ethically, the fact that *some* consistent version of the infinite case exists still makes the same point about the dependence of the axiology on the individuals rather than on some aggregate value.

V

Alternative Formal Settings. Although I have described one particular example of how to fill out my formal method for aggregation, the method is quite modular. In this section I describe several considerations that might motivate alternative ways of filling in these parts.

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

I have assumed that each option yields, for each agent and state, a precise level of goodness. Furthermore, I have assumed that these goodnesses can be understood via real numbers. I then define the world-shift transformation group G_{WS} and the person-shift transformation group H_{PS} in terms of these real numbers. One might worry about the amount of mathematics that has already crept in at this stage, even before I mention addition or expectation. Savage (1954) argues that the use of real numbers for goodness only makes sense when combined with probability. In the absence of probability, classical utility theory only yields a meaningful ordinal scale of values.

However, what is essential for Savage in probability theory is that it gives a way to say when two differences of goodness between outcomes are the same size. Ramsey (1926) is more explicit about this. Let α be some outcome better than β, and γ be some outcome better than δ. Let S_1 and S_2 be two states. Ramsey first defines S_1 and S_2 as equally probable if and only if an option that yields α in S_1 and β in S_2 is just as good as an option that yields β in S_1 and α in S_2. He then says that the difference between α and β is equal to the difference between γ and δ if and only if an option yielding α in S_1 and δ in S_2 is just as good as an otherwise identical option yielding β in S_1 and γ in S_2. That is, the differences are the same if and only if receiving α rather than β in one state is a sufficient trade-off against receiving δ rather than γ in an equally probable state.

Once this concept of 'equal difference' of goodnesses is defined, the claim that goodness has the structure of the real numbers follows from the assumption that goodnesses are dense (between any two goodnesses there is another) and unbounded (for any goodness and any difference, there is another goodness that has that difference from it), that they have the Archimedean property (that is, for any two differences of goodness, there is some finite number of copies of one that are enough to outweigh the other), and that there are limits for the sequences that 'should' converge. (Ramsey is not explicit about this, but presumably he means Cauchy's condition.)

If we give up density, unboundedness, or convergence, then goodnesses can still be represented by real numbers—the only loss is that not every real number corresponds to some particular realizable level of goodness. This might make it harder to compare some specific pairs of options using my methods. For instance, to compare an option in which three people each receive goodness 4 and a fourth person receives goodness 0 to another option in which all four people

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

receive goodness 2, we transform the first into an option in which all four people receive goodness 3, and then observe that it dominates the second option. However, if there is no goodness that corresponds to the real number 3 on this scale, then this is no option at all, and the comparison is blocked.

But the method already requires the use of 'options' that aren't actually available. As Dreier (1996) notes, standard decision-theoretic representation theorems require the consideration of options that can't actually be brought about. In addition to the option that yields an unencumbered wet walk in the rainy state of the world and an unencumbered dry walk in the sunny state of the world, and the option that yields an umbrella-encumbered dry walk in both states of the world, there must be an option that yields an unencumbered dry walk in the rainy state of the world and an unencumbered wet walk in the sunny state of the world. I think it is not such a problem either for theories based on these theorems, or for my method, that they depend on these 'impractical preferences' involving impossible 'options'. On my account, what is essential is just that some options are better than others, and if one actual option is better than such an impossible option, and some other actual option is worse, then the first actual option is better overall, even if we couldn't establish this by direct comparison. Similarly, if the 'option' that is required for this comparison requires the use of some goodness level that can't actually be realized by any outcome, but occupies a distinct place in the ordering of goodnesses, I think this reveals the same thing.

Giving up the Archimedean property yields a somewhat more significant change in the numerical representation. There could be some agent for whom each additional dollar of wealth adds the same non-zero amount of goodness to an outcome, while no finite number of dollars of wealth adds as much goodness as a single additional lifelong friendship. If we tried to represent goodness for such an agent on a real-valued scale, we would need to represent either the value of a friendship as infinitely large or the value of a dollar as infinitesimally small. There are mathematical structures that can be used to represent such sorts of goodness. For some of these non-Archimedean fields, enough of the mathematics of the real numbers is doable that the proofs of compatibility of G_{WS} still go through, but for H_{PS} we might require some modification. Since some objections to aggregative theories focus on the Archimedean nature of the additive real numbers (so that if a headache and a death are both

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part* 3
doi: 10.1093/arisoc/aoab008

represented by real values, then some finite number of headaches must be as bad as one death), giving a fuller account of a non-Archimedean version of this method would help address some other objections to aggregation.

But all of this depends on the 'equal difference' concept. Ramsey and Savage cash this out in terms of trade-offs between different states of equal probability. But in the setting with multiple agents, this could also be cashed out in terms of trade-offs between different persons. In the example version of the theory, I took for granted that there was a pre-existing real-valued representation of goodness for each agent in each possible state, and both G_{WS} and H_{PS} expressed the same trade-offs involving differences of the same size, among corresponding sets of states or corresponding sets of persons. These relations require an interpersonal (and inter-state) comparison of differences of goodness. However, they don't require an interpersonal identification of goodness zero. This might be needed for plausible extensions of my method to options on which populations might differ, but my method yields the same comparisons between options if we change our numerical scale of goodness by shifting the zero point of the goodness scale for some particular set of persons, or some particular set of states.

However, if we have a single absolute scale of goodnesses fixed for all persons and states, there are other trade-off concepts we could use. Prioritarian aggregative ethicists say that differences of goodness matter more when they occur at lower absolute levels. A world where two people are both at goodness 100 might be just as good as a world where one is at 50 and the other 200 because the interests of someone at goodness 50 matter more than the interests of someone at goodness 100 or 200. Certain economic theories of risk-sensitivity in decision theory say the same thing. A prospect of a guaranteed goodness 100 might be just as good as a prospect with a $1/2$ chance of goodness 50 and a $1/2$ chance of goodness 200 because the increase from 50 to 100 matters as much as the increase from 100 to 200.

The use of G_{WS} and H_{PS} assumed that an increase of from goodness x to $x+k$ in one set of states or for one set of persons can always trade off against a decrease from goodness y to $y-k$ for any other corresponding set of states or persons. But we could in principle instead assume that an increase from x to $x+k$ trades off against a decrease from y to $y-k'$, where k and k' are possibly distinct, but

doi: 10.1093/arisoc/aoab008

related to each other on some schedule dependent on x and y. These trade-off schedules for world shift and person shift need to be the same in order for the relevant shifts to be compatible. If not, then there will be some levels of goodness $a > b, c > d$, with b, c distinct, such that a shift from goodness a to b in one set of worlds trades off with a shift from goodness d to b in another equal set of worlds, while a shift from goodness a to c for one person trades off with a shift from goodness d to c for another person. But then an option where one person gets goodness a if heads and d if tails, and another person gets goodness d if heads and a if tails, would be world-shift-equivalent to an option where both people get goodness b with certainty, and person-shift-equivalent to an option where both people get goodness c with certainty. If b and c are distinct, then one of these would dominate the other, so the transformation groups would not be compatible.

This suggestion that the differences that trade off depend on the absolute level of goodness for a person in a state, rather than being fixed at all levels, ends up being less structurally significant than one might think. Ramsey's method of declaring two differences equal if they can trade off against each other yields a way to renumber the goodnesses so that we basically have the original proposal. If the change from 1 unit of goodness to 0 units trades off against the change from 1 unit of goodness up to x units, then we can call x '2'. If the change from 1 unit to 0 also trades off against the increase from x to y, then we can call y '3'. As long as the trade-offs satisfy the principles Ramsey assumed, we will end up with a unique renumbering of the whole scale. Any fixed trade-off schedule between different goodnesses in equal sets of persons or states yields something very much like the original proposal, with the main variation effectively coming down to which sets of persons and which sets of states are considered 'equal', and how strong a dominance ordering we begin with. This is a kind of 'diminishing marginal value' for goodness, and just indicates that the scale we use for aggregation is ordinally equivalent but cardinally distinct from the scale we used for individual goodness.

However, egalitarians about aggregation of goodness, and Lara Buchak's (2013) view of risk aversion in decision theory, ask for something somewhat different which yields a more significant change. Instead of trading off k units of goodness at level x for k'

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

units of goodness at level y, egalitarians trade off k units of goodness for the best-off person against k' units of goodness for the nth best-off person, regardless of the absolute level of goodness each person is at. Similarly, Buchak has a trade-off schedule that depends on the probability of being in a better-off state. There may be transformation group approaches to defining these types of trade-offs, but it seems less plausible that such groups can be commutative and compatible with each other.

In addition to taking absolute goodness or positional goodness into account when determining trade-offs between persons and states, one other natural consideration is that goodness might depend on some third (or fourth, or more) parameter as well. For instance, Cain (1995) describes some examples in which, in addition to an infinite population, there is an infinite set of times. To deal appropriately with these cases, we would have to think of options as functions assigning goodnesses to *three* (or more) parameters rather than just two. There would need to be a transformation group giving trade-offs among those parameters (as well as perhaps more relating these parameters). There may be reasons to continue to treat the parameters symmetrically, or reasons why their aggregation should not commute. This might end up being a better setting to address the questions arising in economics, where the infinite population is really a finite population evolving through unbounded time.

Finally, there might be dimensions of goodness that can't be compared directly, as some say about health, wealth and happiness. Again, it seems reasonable to start with dominance, and extend it with judgements about some trade-offs being equivalences, even if the resulting ordering is still a partial ordering.

VI

Conclusion. Starting with a strict form of dominance, and extending with equivalences derived from transformation groups expressing various trade-offs, I have developed a partial ordering among options for an infinite population under uncertainty. In previous work, I suggested that this method can better help explain why expected value is appropriate for deciding among options for one individual, by showing how the use of expected value (where it is defined) is grounded in these trade-offs. In this paper, I have suggested

doi: 10.1093/arisoc/aoab008

that it can also help understand the way aggregative ethical theories are grounded in individual goodness. In finite cases, it explains why addition is appropriate, but because it extends just as well to infinite cases, it shows that the grounding is asymmetric—aggregate goodness is grounded in individual goodness, and not vice versa.[10]

Department of Philosophy, Texas A&M University
301 YMCA Building, College Station
TX 77843-4237, USA
easwaran@tamu.edu

References

Adler, Matthew D. 2019: *Measuring Social Welfare: An Introduction*. Oxford: Oxford University Press.

Arntzenius, Frank 2014: 'Utilitarianism, Decision Theory, and Eternity'. *Philosophical Perspectives*, 28(1), pp. 31–58.

Basu, Kaushik, and Tapan Mitra 2003: 'Aggregating Infinite Utility Streams with Intergenerational Equity: The Impossibility of Being Paretian'. *Econometrica*, 71(5), pp. 1557–63.

Bostrom, Nick 2011: 'Infinite Ethics'. *Analysis and Metaphysics*, 10, pp. 9–59.

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

Cain, James 1995: 'Infinite Utility'. *Australasian Journal of Philosophy*, 73(3), pp. 401–4.

Chappell, Richard Yetter 2015: 'Value Receptacles'. *Noûs*, 49(2), pp. 322–32.

Diamond, Peter A. 1965: 'The Evaluation of Infinite Utility Streams'. *Econometrica*, 33(1), pp. 170–7.

Dreier, Jamie 1996: 'Rational Preference: Decision Theory as a Theory of Practical Rationality'. *Theory and Decision*, 40(3), pp. 249–76.

[10] I have been very fortunate to have many opportunities to present work related to this project, and to discuss it with friends and colleagues. I would like to thank audiences at (2015) the Formal Ethics Workshop at Universität Bayreuth, the Munich Center for Mathematical Philosophy, the Princeton Workshop on Infinite Value, (2016) Carnegie Mellon University, Columbia University, the Society for Exact Philosophy meeting in Miami, the Decisions Games and Logic conference at the University of Michigan, (2018) the International Rationality Summer Institute, (2019) the University of Missouri, Leeds University (2020) Cornell University, University of North Carolina, (2021) UC San Diego, the Formal Value Theory Workshop at the University of Colorado, and the Aristotelian Society, and in particular Peter Vallentyne for several rounds of discussion and feedback on this topic over many years.

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008

Dubey, Ram Sewak 2011: 'Fleurbaey–Michel Conjecture on Equitable Weak Paretian Social Welfare Order'. *Journal of Mathematical Economics*, 47(4–5), pp. 434–9.
Easwaran, Kenny 2014a: 'Decision Theory without Representation Theorems'. *Philosophers' Imprint*, 14(27), pp. 1–30.
——2014b: 'Regularity and Hyperreal Credences'. *Philosophical Review*, 123(1), pp. 1–41.
Field, Hartry 1980: *Science without Numbers*. Princeton, nj: Princeton University Press.
Gustafsson, Johan 2021: 'Utilitarianism without Moral Aggregation'. Forthcoming in *Canadian Journal of Philosophy*.
Harsanyi, John C. 1953: 'Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking'. *Journal of Political Economy*, 61(5), pp. 434–5.
Hirose, Iwao 2013: 'Aggregation and the Separateness of Persons'. *Utilitas*, 25(2), pp. 182–205.
Johnston, Mark 2017: 'The Personite Problem: Should Practical Reason Be Tabled?' *Noûs*, 51(3), pp. 617–44.
Jonsson, Adam 2018: '*On Permutations and Equity among Generations*'. *Social Science Research Network (SSRN)*, 11 June 2018. https://ssrn.com/abstract=2810869.
Krantz, David, Duncan Luce, Patrick Suppes, and Amos Tversky 1971: *Foundations of Measurement*, vol. 1. New York: Academic Press.
Norcross, Alastair 2009: 'Two Dogmas of Deontology: Aggregation, Rights, and the Separateness of Persons'. *Social Philosophy and Policy*, 26(1), pp. 76–95.
Ramsey, Frank Plumpton 1926: 'Truth and Probability'. *Proceedings of the Aristotelian Society*, 7(1), pp. 153–70. Reprinted in his *The Foundations of Mathematics and Other Logical Essays*, pp. 156–98. Edited by R. B. Braithwaite, with a preface byG. E. Moore. New York: Harcourt, Brace and Co., 1931.
——1928: 'A Mathematical Theory of Saving'. *Economic Journal*, 38(152), pp. 543–59.
Savage, Leonard J. 1954: *The Foundations of Statistics*. New York: Dover Books.
Vallentyne, Peter, and Shelley Kagan 1997: 'Infinite Value and Finitely Additive Value Theory'. *Journal of Philosophy*, 94(1), pp. 5–26.

Proceedings of the Aristotelian Society, *Vol.* CXXI, *Part 3*
doi: 10.1093/arisoc/aoab008