Lexitarian Population Ethics:
A Position-Based Axiomatic Approach*

Alyssa Rusonik† ⓡ Larry Samuelson‡ ⓡ and Itzhak Gilboa§

February 5, 2026

Abstract

The paper makes three contributions to the preference aggregation and population ethics literatures. First, we propose the "lexitarian" criterion, according to which the social planner maximizes a vector of lexicographically-ranked utilitarian functions. Second, we suggest modeling the social choice problem using "positions" as a primitive, and axiomatize the lexitarian rule in this context. Third, we argue for an interpretation of the model in terms of the impersonal approach, while prioritizing needs over wants.

1 Introduction

1.1 Motivation

What makes a society "good"? Would we deem it an improvement to impose a modest decrease in the well-being of 99% of society's members in order to attain immense gains for the remaining 1%? Are a dozen starving children worth broad access to education for thousands of others? What if education is replaced by exposure to opera? These problems become even more complicated when we consider populations of different sizes (cf. Parfit, 1976, 1982): do we prefer a population of 10 billion people living in comfort, or 20 billion who

*We thank Loren Fryxell, Margaret Meyer, Jacob Nebel, and Peter Wakker for comments and references. Gilboa gratefully acknowledges Investissements d'Avenir ANR -11- IDEX-0003 / Labex ECODEC No. ANR - 11-LABX-0047, the AXA Chair for Decision Sciences at HEC. Rusonik gratefully acknowledges a state grant managed by the Agence Nationale de la Recherche under the Investissements d'Avenir programme (ANR-18-EURE-0005 / EUR DATA EFM).

†HEC, Paris. alyssa.rusonik@hec.edu

‡Yale University. larry.samuelson@yale.edu

§HEC, Paris. gilboa@hec.fr

suffer recurring hunger? Yet more difficulties arise when discussing future populations. How do we evaluate the welfare of people whose tastes and circumstances are, at best, conjectures?

It is tempting to avoid these difficulties by insisting that individual utilities are not comparable and then seeking refuge in the Pareto ranking, perhaps citing Arrow's impossibility theorem (Arrow, 1950, Maskin and Sen, 2014) and its descendants. This is the path taken by much of economic theory. However, decisions have to be made, even when alternatives are not Pareto ranked. If economists are to have anything to say about public policy, they must have some workable idea of social welfare.

In this paper, we propose a model of welfare aggregation for fixed or variable populations. We make three contributions. First, we propose a "lexitarian" rule, according to which there are several utility functions, and the social planner attempts to maximize their sums in a lexicographic manner. This procedure captures the intuition that trade-offs between individuals' welfare must inevitably be considered in public policy and social planners' problems, but that not all sources of suffering or happiness are created equal. Some trade-offs may be deemed outright unpalatable while others remain a necessary compromise. Second, we introduce the notion of a "position" as the basic primitive, and consider social profiles that describe how many individuals there are in each position, rather than the experiences of each individual per se. Finally, we interpret our model in terms of the impersonal approach to population ethics, while focusing on needs rather than wants. The next three subsections develop each of these points.

1.2 The Aggregation: Lexitarianism

The preference aggregation literature invokes a variety of aggregation rules, such as the utilitarian, Nash, and Rawlsian (or more generally, leximin) social orderings (Osborne, 2025, Section 1.8). The summation of individual measures of well-being is often very intuitively appealing—it makes sense to inconvenience a small number of people for the greater good of a society. At the same time, we also often have the intuition that there are some sacrifices we will not ask, even on the part of a few, for the advantage of the many. To cope with this tension, while retaining the appeal of utilitarian aggregation, we propose a lexicographic version of utilitarianism.

Under the standard utilitarian approach to social welfare, each of the members of a population K is characterized by an outcome drawn from a set P of measures of well-being. Aggregate well-being is given by $\sum_{i \in K} u_i(p_i)$, where u_i is person i's utility function. Section 1.4 explains that we take an impersonal approach in which the social planner uses a single utility function u to evaluate the outcome for every member of the population, so that the utilitarian aggregation takes the form

$$U\left(p_1, ..., p_k\right) = \sum_{i \in K} u\left(p_i\right). \tag{1}$$

Now assume that the social planner has several functions $u_j : P \to \mathbb{R}$ with $j = 1, ..., m$ and consider the lexicographic ordering induced by successively calculating (1) for the functions $(u_1, ..., u_m)$.[1] This enables us to represent preferences that refuse to make trade-offs between certain needs or wants. For example, we can imagine the function u_1 as measuring health and nutrition and u_2 measuring self-fulfillment and personal expression. Such a model would allow one to say that providing medication to sick people may justify lowering the level of nutrition of some others, but that art classes for some individuals should not be traded-off with nutrition of others.

The lexitarian model can generalize leximin (Osborne, 2025, Section 1.8) as well as sufficientarian (Frankfurt, 1987) ideas, without ignoring Pareto improvements. For example, a function u_1 can be an indicator, stating whether an individual has or has not reach a "sufficient" level of a resource, and u_2 may capture other needs or wants the individual might have. Thus, a lexitarian model can respect people's preferences for, say, intellectual stimulation, while insisting that such concerns do not enter the picture until everyone has reached a minimum level of nourishment.

[1]For certain applications, one may wish to consider an infinite hierarchy of utility functions. This will be allowed in our model, as long as each profile obtains non-trivial values only in a finite collection of these.

1.3 The Framework: Positions

When discussing population ethics, in particular when considering the far future, we find it more intuitive to think of the numbers of people who will live in certain conditions rather than of their individual identities. We accordingly propose a model in which the primitive is a set of "positions" people may occupy. The objects of choice are population profiles, which describe how many individuals are at each position. Formally, a profile is a "counter vector" I, where $I(p) \geq 0$ indicates how many individuals are at position p. The overall number of individuals in a profile (ranging over all positions) is finite, but we impose no a priori bounds—neither on the number of positions which may be occupied nor on the number of individuals at each position.

It will be helpful to compare the position-based model to the standard model in the case of utilitarianism. Assume, as above, a population $K = \{1, \ldots, k\}$ of individuals with outcomes drawn from a set P, and an impersonal utility function $u : P \to \mathbb{R}$. The standard model of utilitarianism considers the objects of choice to be functions $f : K \to P$. Axiomatic derivations use a ranking over such functions f as primitive, to derive the maximization of $U(p_1, ..., p_k) = \sum_{i \in K} u(p_i)$. By contrast, in our model the objects of choice are functions whose domain is P and range is the nonnegative integers, i.e., $I : P \to \mathbb{Z}_+$, so that $I(p)$ identifies the number of people characterized by outcome p. Given a utility function $u : P \to \mathbb{R}$, the planner maximizes

$$I \cdot u = \sum_{p \in P} I(p)\, u(p)\,.$$

With enough freedom in the definition of a position, our framework and the standard utilitarian approach are equivalent. Given a function $f : K \to P$, we may define a profile I by $I(p) = |f^{-1}(p)|$ so that $I \cdot u = U(p_1, ..., p_k)$. Conversely, given a profile $I : P \to \mathbb{Z}_+$, we can construct a population $\{1, ..., k\}$ for $k = \sum_{p \in P} I(p)$ and a function $f : K \to P$ such that, for each p with $I(p) > 0$, there are $I(p)$ identical individuals i with $f(i) = p$, so that $U(p_1, ..., p_k) = I \cdot u$.

Despite this formal equivalence, the positions model provides a useful new perspective on social welfare. First, as will be clear below, the axiomatic derivation of the model suggests a new way of conceptualizing utilitarianism and its variants. Specifically, the lexitarian model will be simply axiomatized, with the lexicographic order over utility functions naturally derived from the social planner's preferences. Second, a specific position can also be interpreted as capturing only a part of a person's experience. Thus, in our model each person might be at several positions, reflecting different needs, states of the world, or time periods. This allows us encompass cases in which some competing needs and wants are comparable while others are not. Lastly, the positions model has a built-in anonymity assumption: only the numbers of individuals (in each position) matter, not their identities.

1.4 The Impersonal Approach

A philosophical schism divides the study of population ethics, and the responses to these basic questions, into two main fields of thought: person-affecting views and impersonal approaches (see, e.g., Carlson (1998), Arrhenius (2000), Roberts (2011), Boonin (2014), and Frick (2020); the issue has also been discussed in social choice theory by Blackorby and Donaldson (1984), among others). Person-affecting theories (see, e.g., Navreson, 1973, for a foundational work) maintain that a person's own preferences are the best and indeed only appropriate guide to her welfare, and so it is the individuals' utilities that are to be aggregated. This view appears throughout the economics literature in the form of phrases such as "de gustibus non est disputandum" or "consumer sovereignty", and in models of utilitarianism such as Harsanyi (1955) and Dhillon and Mertens (1999).

Parfit's (1984) discussion of the "non-identity problem" challenges the coherence of person-affecting theories in the context of varying populations. For example, it is not clear we can meaningfully talk about the preferences of people who do not exist. Impersonal approaches accordingly focus on aggregate welfare measures independently of the identities of the individuals. Under this view, the planner assesses the well-being of actual and hypothetical individuals according to criteria that reflect the planner's preferences and only implicitly account for the preferences of the individuals (that is, insofar as they are captured by the planner's preferences).[2]

Our preferred interpretation of the lexitarian model, with its focus on positions rather than individuals, lies squarely within the impersonal approach. Furthermore, we find the model's assumptions more palatable if positions identify needs and the degrees to which they are satisfied rather than preferences (or wants). First, individual preferences are difficult to observe and, given incentives to misrepresent them, difficult to elicit. As a result, attempts to aggregate individuals' "true" subjective utilities are likely to be of limited success. Second, even if these utility functions were directly and unambiguously observable, we argue that maximizing their sums gives rise to unacceptably counterintuitive implications (see subsection 4.1). Third, we expect to have a better idea of the needs of future generations than of their wants. Hence, when considering these generations, and also when trading-off their well-being against that of the present generation, it seems sounder to rely on needs rather than on wants. Finally, if one follows the impersonal approach in allowing a social planner to gauge the utility of individuals based on criteria she deems salient, without soliciting the individuals' subjective assessments of their own positions, then one may wish to limit that social planner's power. Restricting their discourse to perceived needs may be an important step in this direction.

The lexitarian aggregation of preferences is designed to identify pairs of determinants of well-being that can be compared and traded-off, and to separate pairs that cannot. As such, our model relates to several lines of thought in philosophy and in psychology. One of the most prominent contributions is Rawls's (1971) theory of justice, suggesting that a social profile be evaluated by the condition of the individuals who are worst-off. The resulting maxmin

[2]The philosophical dichotomy between the person-affecting and impersonal schools of thought is strongly evocative of a debate at the intersection of logic and linguistics. Narveson (1973) juxtaposes two views. One is Bertrand Russell's claim (see, e.g., Quine (1962, p. 221)) that sentences whose subjects contain references to the non-existent are inherently false (if "they" do not exist, no true assertion can be made about "their" rights, so any statement is a priori false). The other is Strawson's (1950) notion of presupposition: a view that sentences by construction assume the referent's existence (if the "they" in question do exist, the statement makes logical sense and has a truth value, and it is in this context that the statement should be read; and if "they" were, in fact, non-existent, the statement would be assigned no truth value at all).

rule suggests that the central planner should focus on unsatisfied basic needs before considering less-basic ones, let alone desires that are considered to be in the domain of wants, or luxury. Indeed, this approach can be modeled in our framework, using an infinite hierarchy of utility functions, $(u_\alpha)_\alpha$, each indicating whether the individual in question obtains a level of welfare α. In such a model one recovers the leximin criterion, which refines the maxmin one. However, the maxmin (or leximin) criterion is often considered extreme. For example, if the poorest individual has to survive on 400 grams of rice a day, no amount of food given to millions of others would justify dropping this quantity to 390 grams a day. Many would view this as somewhat exaggerated. Our model allows for a utilitarian trade-off between quantities of food, without necessarily trading off space trips against food.

Our notion of a hierarchy of needs is also related to sufficientarianism (see Frankfurt, 1987, Alcantud, Mariotti, and Veneziani, 2022, and Bossert, Cato, and Kamaga, 2023). Sufficientarianism suggests that ethical evaluations should consider the number of individuals who "have enough", without taking into consideration how much more they can have, once they have reached the threshold of "enough". Clearly, this can be captured in our model. In the simplest formulation, one need not resort to lexitarianism: if utility is an indicator function of "having enough", simple utilitarianism would suffice for sufficientianism. However, our model allows for refinements, both below and above the threshold: once the indicator function has been set as the primary criterion, we may proceed to define the degree to which needs are satisfied as a secondary criterion, and then define wants as a tertiary one.

Maslow's (1943) hierarchy of needs can be a source of inspiration for the social planner's lexicographic preferences. In fact, one may view our model as a Maslowian version of Bentham's utilitarian theory.

1.5 Planners and Theorists

The social planner we have in mind may be a legislative body who makes decisions on general principles of social and economic policies, or a government agency, who implements such general policies in everyday decisions on budget allocations, project priorities etc. In some cases, one may imagine a planner being a parent making decisions that will affect her children. Alternatively, the planner may be a rhetorical device that a policy advocate uses in arguing for the merits of her policy. The interpretation of the model may well vary with the context and the planner's identity. The main application we have in mind, however, involves a planner who is called upon to make decisions about economic resources. As such, we imagine the planner's choices as an attempt to capture a social consensus, embodying a society's ethos as to how policies are to be evaluated.

The discussion in this paper, as, indeed, most of the literature, is normative in essence. It is important to clarify, however, that we do not believe that we have any privileged access to Truth or Justice. We view the role of theory as clarifying discourse and sharpening arguments, and, to the extent that normative claims are made, they are attempts to model the reader's own preferences. In this context it is useful to bear in mind that there are two types of audiences: fellow economists, who may actually use models to make policy recommendations, and individuals in society who presumably vote on social policy issues. Offering an axiomatic derivation of a formal model—such as lexitarianism—is mostly addressed at professional economists. Discussing the preferred interpretation of the model—such as the priority of basic needs over wants—should have a more general appeal. Naturally, these distinctions are fuzzy. Moreover, some of the claims can be interpreted in either way. Yet, we find the distinction useful to bear in mind, and we emphasize that any normative claim is to be read as an attempt to capture a wide consensus—whether of professionals or of laypeople.

1.6 Preview

Section 2 describes the model and presents the central result, as well as the derivation of the special case of classical utilitarianism in our framework. Section 3 extends the model to encompass time and uncertainty, decomposing the notion of "position" into circumstances, time-periods, and states.

There are several questions that may be raised about our model of lexitarianism and its suggested interpretations. Specifically, a sympathetic reader may ask:

(i) Suppose we accept the axioms and the model. Why should a central planner ignore individual utilities, if we assume that these are given and observable? Why should she be so paternalistic as to assume that she knows what's best for individuals better than they do? In other words, shouldn't the function u in the model be replaced by the actual utilities $(u_i)_i$?[3]

(ii) Next, even if we agree on an impersonal approach, why should the planner focus on needs rather than wants, and, perhaps more fundamentally, what is the difference between these? Is there an observable, scientific way to tell them apart?

(iii) Finally, at a more technical level, what is the right way to capture needs in a position-based model?

Section 4.2 deals with these questions. It argues that the interpretation and modeling choices presented here offer a coherent approach to modeling population ethics, while recognizing that different readers may accept some of these choices but reject others. We conclude with a general discussion in Section 5, while proofs of the representation theorems are given in the appendix.

2 Model and Results

Let there be a set of *positions* P. The set of *profiles* is

$$\mathcal{I} = \left\{ I : P \to \mathbb{Z}_+ \;\middle|\; \sum_{p \in P} I(p) < \infty \right\}.$$

For $I \in \mathcal{I}$, the *support* of I is $supp(I) \equiv \{ p \in P \mid I(p) > 0 \}$. Thus, the set of positions P may be infinite, but each profile I has a finite support. Algebraic operations are performed on $\mathcal{I}$ pointwise so that $nI + mJ$ is well-defined for

[3] This question is particularly relevant in the case of classical utilitarianism. When general lexitarianism is considered, it is not entirely obvious whether individuals have lexicographic preferences and how these should be aggregated, especially when different individuals might well have different numbers of lexicographically-ordered categories.

$n, m \geq 0$ and $I, J \in \mathcal{I}$. $0 \in \mathcal{I}$ has its natural meaning. Similarly, for $I, J \in \mathcal{I}$ the inequality $I \geq J$ is read pointwise. For $p \in P$, $\mathbf{1}_p \in \mathcal{I}$ denotes the indicator function for p. For a subset of positions $A \subset P$, we denote by $I_A \in \mathcal{I}$ the corresponding profile that vanishes outside A, i.e. $I_A = \sum_{p \in A} I(p) \mathbf{1}_p$.

For a function $u : P \to \mathbb{R}$, denote

$$I \cdot u = \sum_{p \in P} I(p) u(p)$$

which is well-defined as I has a finite support. For $u : P \to \mathbb{R}$ and $A \subset P$, $u_A : A \to \mathbb{R}$ is the restriction of u to A.

Let there be given a binary relation $\succsim \subset \mathcal{I} \times \mathcal{I}$. The symmetric and asymmetric parts of $\succsim$ will be denoted, as usual, by $\sim$ and $\succ$, respectively.

We impose four axioms. First,

A1 Weak Order: $\succsim$ is complete and transitive.

The completeness axiom is notoriously problematic, especially in the context of social choice. Its main justification is, as always, that decisions need to be made, and that it is better to bring them forth and see them in the model rather than to leave them unspecified. We return to this point in subsection 4.2.[4]

Transitivity is one of the most basic axioms of choice models, and we do not have anything to say about it that is specific to the current model. The main reason to highlight it is that some of the literature on the "Repugnant Conclusion" suggests dropping this axiom (see Temkin, 1987, Persson, 2004, and Rachels, 2004). We tend to believe that giving up on transitivity is a high price to pay, both theoretically and practically, and that it is worthwhile to explore the possibilities that respect this axiom.

Our next axiom is the main driving force behind the additive representation:

A2 Union: For all $I, J, K \in \mathcal{I}$, $I \succsim J$ iff $I + K \succsim J + K$.

[4]There are papers which opt for quasi-orderings instead, including, for instance Parfit (1982), which introduces indeterminacy via ranges of utility specified for added individuals.

Consider first an interpretation of the model where a position summarizes all the information about an individual, so that each individual is counted in a profile at one position only. In this case, suppose that the social planner can choose between one policy, leading to a population profile I, and another, which will result in a population profile J. There is also another location, where the population profile will be K, irrespective of the current policy choice. The Union axiom states that the preference between I and J should not change if each is combined with the population whose profile is K. Clearly, it is very similar to axioms of separability across populations, and to Savage's P2 (Savage, 1954). It also bears resemblance to von Neumann and Morgenstern's (1944, 1947) Independence axiom, basically stating that a subpopulation that is common to two profiles can be ignored.

The Union axiom is, however, more questionable when a position is interpreted as describing only some aspects of an individual's well-being. For example, if we think food and self-fulfillment as incomparable needs, and if each individual appears in a food-position as well as in a self-fulfillment-position, the axiom states that we ignore inequality in food distribution when we consider inequality in self-fulfillment. We return to discuss these modeling choices in subsection 4.2.

The third axiom is a monotonicity axiom, indicating that the central planner prefers that any population survives rather than dies:

A3 Monotonicity: For all $I \in \mathcal{I}$, $I \succsim 0$.

This axiom is far from trivial. One of the more influential attempts to avoid the repugnant conclusion is *critical level utilitarianism* (Blackorby, Bossert, and Donaldson, 1995; 1997), which posits the existence of a critical threshold below which individuals' well-being does not contribute to aggregate welfare.[5] In our model, this would be akin to dropping A3 and allowing the utility function $u(p)$ to assume negative values. Indeed, one can prove such a version of our main result. In such a model, the critical level would be endogenously

[5] See Broome (2004) for a discussion of the merits and drawbacks of aggregative welfare approaches in general, and non-unique critical thresholds in particular.

derived by the value 0 on the utility scale.[6] However, we take here a more conservative approach according to which the central planner is not supposed to decide that certain lives are not worth living.

For our final axiom, we need the notion of comparability:

Definition 1 *For $p, q \in P$, p and q are* comparable *if there exists $k \geq 1$ such that $k\mathbf{1}_p \succsim \mathbf{1}_q$ and $l \geq 1$ such that $l\mathbf{1}_q \succsim \mathbf{1}_p$.*

Thus, two positions p and q comparable if a profile with sufficiently many people at position p is at least as desirable as a profile with a single person at q, and vice versa. Intuitively, the order $\succsim$ allows trade-offs over comparable positions. Note that the intuition for this definition relies on Monotonicity: it is implicitly assumed that being at each position is a good thing, and thus, unless p and q are incomparable, a large population at one position should eventually be a better society than a single person at the other. Should one believe that a position p describes "life not worth living", a larger population at position p would only make the resulting society worse in the eyes of the central planner.

If positions p and q are not comparable, given completeness, it must be the case that one is infinitely more desirable than another. For example, being healthy may be considered incomparably more important than having a sense of self-fulfillment. Importantly, the notion is endogenous: it is defined by the social planner's preference relation $\succsim$. Our model does not dictate which positions are incomparably more important than others, nor that there be such pairs to begin with. However, with this endogenously-defined notion, we proceed to state the last axiom.

A4 Restricted Archimedeanity: For all $I, J, K, L \in \mathcal{I}$, if $p, q \in supp\,(I) \cup supp\,(J) \cup supp\,(K) \cup supp\,(L)$ are comparable and $K \succ L$, then there exists $n \geq 0$ such that $I + nK \succ J + nL$.

[6] If we deal with the special case of utilitarianism, namely, when the lexicographic ordering is trivially defined by a single function, one only needs to drop A3 and obtain the general result. For the counterpart of the lexicographic result, the definition of "comparability", which follows, would have to be modified.

If we ignore the comparability condition, the axiom is rather standard: it says that, if profile K is strictly preferred to profile L, then taking the union of any profiles I and J with sufficiently many copies of K and L, respectively, would result in a preference for the I-with-K's population over the J-with-L's one. As usual, it is an axiom of commensurability: it states that the difference between J and I cannot be infinitely weightier than that between K and L. The fact that it is restricted to some quadruples (I, J, K, L) is at the heart of lexitarianism: the motivation for our model is precisely the fact that some differences are *not* commensurable. However, the axiom states that incommensurability isn't arbitrary, and it is basically defined by positions: if all the positions involved in (I, J, K, L) are pairwise comparable—according to the planner's own preferences—then the axiom holds.

Next, we define the notion of numerical representation we seek:

Definition 2 *A* lexitarian representation *of* $\succsim$ *is a pair* $\left(\succsim^b, u\right)$ *where* $\succsim^b$ *is a weak order over* P *and* $u : P \to \mathbb{R}_+$ *such that, for all* $I, J \in \mathcal{I}$, $I \succ J$ *iff there exists a* $\succsim^b$*-equivalence class* $A \subset P$ *such that*

$$I_B \cdot u = J_B \cdot u$$

for every $\succsim^b$ *equivalence class for which* $B \succ^b A$ *and*

$$I_A \cdot u > J_A \cdot u,$$

where $\sim^b$ *and* $\succ^b$ *are the symmetric and asymmetric parts of* $\succsim \cdot$*, respectively.*

Note that, for every $I, J \in \mathcal{I}$, we have $I_B = J_B = 0$ for all but finitely many $\sim^b$-equivalence classes B. Hence, if $\succsim$ is complete, this definition also implies that $I \sim J$ iff $I_B \cdot u = J_B \cdot u$ holds for every $\sim^b$-equivalence class B.

A lexitarian representation $\left(\succsim^b, u\right)$ is a rather simple object: it starts with a ranking of positions $\succsim^b$, which can be thought of as "at least as basic as". It also involves a utility function u over positions. The ranking of two profiles I, J by $\left(\succsim^b, u\right)$ is defined by seeking the $\succsim^b$-highest equivalence class that tells them apart, according to their u values. Note that u values over different equivalence classes never get to be compared. Hence we can think of the representation as if there were different utility functions, each defined for a different equivalence class of positions.[7] If it so happens that $\succsim^b$ has finitely many equivalence classes, we can also use the more standard lexicographic representation, with a utility function for each.

We can finally state the main result, proved in Section 6.2:

Theorem 1 *$\succsim$ satisfies A1, A2, A3 and A4 if and only if it has a lexitarian representation $(\succsim^b, u)$. Moreover, for every $\sim^b$-equivalence class B, u_B is unique up to multiplication by a number $\lambda_B > 0$.*

Observe that the statement of the theorem doesn't specify that $\succsim^b$ is unique. It will be clear from the proof that it is unique when restricting attention to the support of u. Positions that are "null" (i.e., positions p such that $\mathbf{1}_p \sim 0$) can be assigned to various $\sim^b$-equivalence classes.

The statements of the axioms clearly require consideration of populations of varying sizes. However, the implied representation can just as clearly be used to evaluate different allocations for a population of fixed size.

In the case of a finite set equivalence classes, the representation can take a more familiar form:

Corollary 1 *Assume that $\succsim^b$ has finitely many equivalence classes. Then $\succsim$ satisfies A1, A2, A3 and A4 if and only if there exists $u : P \to \mathbb{R}_+$ and a partition of P, $(A_1, ..., A_m)$ $(m \geq 1)$ such that $u(p) > 0$ for all $p \notin A_m$, and for all $I, J \in \mathcal{I}$, $I \succsim J$ iff*

$$(I_{A_1} \cdot u, ..., I_{A_m} \cdot u) \geq_L (J_{A_1} \cdot u, ..., J_{A_m} \cdot u),$$

where $\geq_L$ is the familiar lexicographic ordering of elements of $\mathbb{R}^m$.

Note that this representation coincides with the lexicographic representation of $\succsim$ by $(u_1, ..., u_m)$ where each u_i vanishes outside A_i. Conversely, given a lexicographic representation by finitely many functions $(u_1, ..., u_m)$ over a population $K = 1, ..., k$ of individuals, their utility profiles can be embedded

[7] For notational simplicity, we use but one function u.

in our model by introducing m equivalence classes of positions, each reflecting the range of u_i for a given i, as explained above for the utilitarian case ($m = 1$).

A simple and familiar special case is utilitarianism, namely the case in which all positions are comparable. In that case we may assume a stronger version of the Archimedeanity axiom:

A4* Strong Archimedeanity: For all $I, J, K, L \in \mathcal{I}$, if $K \succ L$, then there exists $n \geq 0$ such that $I + nK \succ J + nL$.

For the statement of this axiom, one need not define comparability of positions. However, it is obvious that (given the other axioms) any two positions p, q will be comparable. In other words, given A1-A4, we could also impose the condition that any two positions are comparable and obtain A4*.

Section 6.1 proves the following:

Theorem 2 $\succsim$ *satisfies A1, A2, and A4* if and only if there exists a function*

$$u : P \to \mathbb{R}$$

such that, for all $I, J \in \mathcal{I}$

$$\begin{array}{rcl} I & \succsim & J \\ & \textit{iff} & \\ I \cdot u & \geq & J \cdot u \end{array} \tag{2}$$

Moreover, in this case u is unique up to multiplication by a positive number.

Furthermore, if the above hold, then A3 holds iff $u \geq 0$.

Theorem 1 in Chapter 3 of Krantz, Luce, Suppes, and Tversky (1971, p. 74) encompasses the main part of this result (i.e., without A3 and nonnegativity) within a more general and abstract framework.[8] Our setting is more concrete, and, as a result, the axioms and proof are commensurately simpler.[9] Indeed, from a mathematical viewpoint, our argument is similar to de Finetti's (1931,

[8]We thank Jacob Nebel for pointing this out to us. The result is also cited as Proposition 1 in Nebel (2024).

[9]Krantz et al. (1971) deal with an abstract "extensive structure", involving a binary operation that need not be commutative, nor even associative. (An axiom named "weak

1937) derivation of maximization of expected value relative to a subjective probability (see the version described in Gilboa, 2009). In that model, P is a set of states of the world, and a real-valued function from the states to the real line describes a bet. Similar axioms are used to derive a separating hyperplane between the bets that are at least as desirable as zero and those that are not, and an additional monotonicity axiom guarantees that (apart from the trivial case of indifference), the gradient of the hyperplane can be interpreted as a probability over the states. The major conceptual difference between the results is that in de Finetti's model the utility function is assumed known (represented by the values of each function), and the derived weights are probabilities—whereas in the present model the given values are numbers of occurrences of positions, and the utility is derived as the weights assigned to these positions. The difference in interpretation also results in a mathematical difference: in the present model the values of the functions compared are only non-negative integers, rather than any real numbers. Consequently, the proof of this result is more involved than it would be in the real-valued case.

In sum, the four axioms are equivalent to the existence of a number $u(p)$ for each position p, such that profiles are evaluated by taking the summation of $u(p)$ over all positions within each equivalence class, and then lexicographically comparing the resulting vectors of sums. The numbers specified by u are unique up to a positive multiplicative constant.[10] There is no assumption about the utility u as a function of p; the mathematical derivation uses the space of counter vectors (the $I, J, ...$) and not the utility values or even the

associativity"guarantees, however, that in case associativity fails, the order of operations does not affect preferences. No such axiom is needed in case commutativity fails.) Thus, one of their axioms isn't needed in our setup, and two others can be simplified. More importantly, the additive operation induces a simple geometric structure, and the proof is basically a separating hyperplane argument. The corresponding proof in Krantz et al. is, naturally, more involved, and relies on representations constructed by "standard sequences" in less structured spaces (including Theorems 4 and 5 in their Chapter 2).

[10]While the set of positions may be infinite, every profile considered has only finitely many individuals, so that the summation of their utilities is well-defined. It is possible, for example, that a position is defined by a level of wealth, and that the set of positions is a continuum. Yet, a given population will only have finitely many individuals, and therefore only finitely many positions p such that $I(p) > 0$.

positions.

3 Time and Uncertainty

The model as discussed so far implicitly assumes that there is no uncertainty regarding the outcomes of social policies, and that all occur contemporaneously. This is clearly a very restricted setup. People live in different periods of time, and there is considerable uncertainty about the number of people alive at any time, let alone about the position(s) they will be at. Indeed, some of the most prominent questions of population ethics have to do with the effect of current policies on the uncertainties faced by future generations.

In this section we re-interpret the notion of a "position" to allow us to deal with such questions based on the analysis above. Specifically, we will assume that there is a set of "circumstances", playing the role of positions in a one-stage model under certainty, and a position will now be a circumstance experienced at a given state of the world and a given period.

Formally, let C be a set of circumstances, let $T = \{0, 1, 2, ...\}$ be a set of time periods, and let Ω be a finite set of states of the world. Assume that $P = C \times T \times \Omega$. It will be convenient to denote subsets of $\mathcal{I}$ defined by supports that are restricted to states or time periods. With a slight abuse of notation that is unlikely to cause any confusion, we define

$$\begin{aligned} \mathcal{I}_t &= \{ I \in \mathcal{I} \mid I(c, t', \omega) = 0 \qquad \forall t' \neq t \} \\ \mathcal{I}_\omega &= \{ I \in \mathcal{I} \mid I(c, t, \omega') = 0 \qquad \forall \omega' \neq \omega \} \\ \mathcal{I}_{t,t'} &= \{ I \in \mathcal{I} \mid I(c, t'', \omega) = 0 \qquad \forall t'' \neq t, t' \} \end{aligned}$$

Define a state ω to be *null* if the value of $I(c, t, \omega)$ does not affect preferences. We impose three additional axioms:

A5 State Invariance: For all non-null $\omega, \omega' \in \Omega$, all $I, J \in \mathcal{I}_\omega$, and $I', J' \in \mathcal{I}_{\omega'}$, if for all $(c, t) \in C \times T$, $I(c, t, \omega) = I'(c, t, \omega')$ and $J(c, t, \omega) = J'(c, t, \omega')$, then $I \succsim J$ iff $I' \succsim J'$.

A6 Time Invariance: For all $t, t' \in T$, all $I, J \in \mathcal{I}_t$, and $I', J' \in \mathcal{I}_{t'}$, if for all $(c,\omega) \in C \times \Omega$, $I(c,t,\omega) = I'(c,t',\omega)$ and $J(c,t,\omega) = J'(c,t',\omega)$, then $I \succsim J$ iff $I' \succsim J'$.

Axioms A5 and A6 are the natural conditions required to have a utility function that is state- and time-independent. They can be stated in other ways as well: first, instead of limiting attention to profiles whose support only involves one state (or one period t), they can be stated for any pairs of profiles I, J, I', J' such that I' and J' are obtained from I and J, respectively, by permuting the corresponding states (or periods). This seemingly stronger requirement will obviously be equivalent to the current formulation in light of the additive structure guaranteed by the results of Section 2.

Clearly, Axioms A5 and A6 consider profiles that do not have the same number of individuals in all states and/or in all time periods. This is in line with our general setup, assuming that profiles with different numbers of individuals can be compared. Indeed, this flexibility of the model seems necessary to discuss some of the problems of population ethics: when considering the long-term risks of climate change or of nuclear wars, the number of individuals alive will typically differ across time periods and across states of the world.

While A5 and A6 treat uncertainty and time symmetrically, the next axiom is specific to the time dimension:

A7 Dynamic Consistency: For all $t \in T$, all $I, J \in \mathcal{I}_{t,t+1}$, and $I', J' \in \mathcal{I}_{t+1,t+2}$, if for all $(c,\omega) \in C \times \Omega$ and $t' \in T$, $I(c,t'+1,\omega) = I'(c,t',\omega)$ and $J(c,t'+1,\omega) = J'(c,t',\omega)$, then $I \succsim J$ iff $I' \succsim J'$.

We start with the (simpler) case of utilitarianism, that is, when all positions are comparable.

Proposition 1 $\succsim$ *satisfies A1, A2, A4* as well as A5, A6, and A7 if and only if there exists a utility function defined on circumstances,*

$$\hat{u} : C \to \mathbb{R}$$

a probability vector $pr \in \Delta(\Omega)$ and a factor $\delta > 0$ such that, for all $I, J \in \mathcal{I}$

$$\begin{aligned} I \ \succsim \ J \quad &\Leftrightarrow \quad V(I) \geq V(J) \\ &\textit{where} \\ V(I) \ &= \ \sum_{\omega \in \Omega} pr(\omega) \sum_{t \in T} \delta^t \sum_{c \in C} I(c, t, \omega) \hat{u}(c) \end{aligned} \tag{3}$$

Moreover, in this case $\hat{u}$ is unique up to multiplication by a positive number, whereas pr and δ are unique if $\hat{u}$ is not identically zero.[11]

Corollary 2 *Finally, under the above condition, A3 holds iff u is nonnegative.*

Note that the proposition does not ensure that $\delta < 1$. Since only vectors I with finite support are considered, there is no mathematical difficulty in allowing $\delta \geq 1$. It is not difficult to state an assumption that would guarantee $\delta < 1$.[12]

For the general case of lexitarian preferences, a few adaptations are needed. First, we have to ask which pairs of positions are reasonably comparable. Our basic intuition is that comparability is a matter of the circumstances, not of time or uncertainty. For example, luxury goods may be incomparable to life-saving medication, but there is no reason that luxury goods ten years from now would be incomparable to the same goods today, or that disease in the state of global warming would be incomparable to disease in the state of no climate change. We therefore impose the following:

A8 Comparability: For all $c \in C$, $(t, \omega), (t', \omega') \in T \times \Omega$, (c, t, ω) and (c, t', ω') are comparable.

We can now state

Proposition 2 $\succsim$ *satisfies A1-A8 iff it has a lexitarian representation $\left(\succsim^b, u\right)$ such that:*

[11]A constant zero-valued u corresponds to the case of a trivial relation for which $I \sim J$ for all I, J.

[12]Such an assumption could compare a vector I in $\mathcal{I}_t$ with the corresponding vector in $\mathcal{I}_{t+1}$—but the direction of preference between them would have to depend on the preference between I and 0.

(i) There are $\hat{u} : C \to \mathbb{R}_+$, and, for each $\sim^b$- equivalence class A, a probability vector $pr_A \in \Delta(\Omega)$ and $\delta_A > 0$ such that, for all $(c, t, \omega) \in A$,

$$u(c, t, \omega) = pr_A(\omega)\, \delta_A^t \hat{u}(c)$$

(ii) For every $c \in C$ and all $(t, \omega), (t', \omega') \in T \times \Omega$, $(c, t, \omega) \sim^b (c, t', \omega')$.

Moreover, in this case $\hat{u}$ is unique up to multiplication by a positive number, whereas pr_A and δ_A are unique if $\hat{u}$ is not identically zero.

Thus, retaining the weaker version of Archimedeanity (A4 rather than A4*), Proposition 2 allows for lexitarian preferences in the more elaborate model, reflecting dynamic and uncertainty considerations. However, axiom A8—clearly reflected in condition (ii) of the representation—restricts the lexicographic nature of preferences to the comparison of basic circumstances. Hence, other reasons for incomparabilities, such as lexicographic probabilities as in Blume, Brandenburger, and Dekel (1991), are ruled out by A8.

Observe that the axioms, as stated, allow the probability and the discount factor to depend on the equivalence class. One may strengthen A5 and A6 to relate preferences across equivalence classes to obtain a tighter representation where pr_A and δ_A are independent of A.

4 Wants, Needs, and Positions

This section addresses the three interpretational questions raised above: (i) Should we insist on the impersonal approach? (ii) Should the model reflect wants as well as needs? (iii) How should needs be described by positions?

4.1 Wariness of Person Affecting Aggregation

The standard person-affecting approach to social welfare aggregates a vector of utilities, representing the well-being of the various individuals. We are skeptical of such person-affecting approaches to welfare, even with a fixed population and directly observable utility functions. We use utilitarianism as a laboratory for exploring this issue, ending with a brief discussion of other aggregation rules.

Suppose first that the planner must choose the quantity of a single good $x \in \mathbb{R}_+$. There are two individuals, with utility functions

$$
\begin{aligned}
u_1(x) &= x \\
u_2(x) &= -\ln(1+x).
\end{aligned}
$$

No matter how the utilities of the two individuals are weighted, the weighted sum of utilities grows without bound as x increases, and hence the planner will set x to be as large as allowed by the feasible set. Examples of this type are often interpreted as person 1 being a sadist who enjoys tormenting person 2, who prefers not to be tormented. It is counterintuitive that the planner could recommend letting the sadist run amok.

A large literature has responded to similar criticisms, with rule utilitarianism being perhaps the most prominent response to address such concerns. Our view is that this challenge is not fatal for basic utilitarianism itself. The prospect that maximizing utility may call for individual 1 to torture individual 2 is unsettling because people now commonly find torture abhorrent, even if they are only impartial observers. If so, this should be reflected in the utility functions of the members of the society. A straightforward utilitarian calculation will then lead the planner to restrain the sadist.[13]

A second concern is that if a planner is supposed to maximize a (weighted) sum of utilities, she has to choose for each individual a utility function, from a set of utility functions that represent the individual's preferences, but that are unique only up to an increasing affine transformation. The choice of a utility function from this set boils down to determining an individual's weight in the social aggregation. Asking individuals to provide their utility functions is tantamount to asking them how important they think they are. The result is likely to be an arms race: each agent faces unlimited incentives to multiply their utility function by ever-larger constants in an attempt to increase their

[13]Pinker (2011, Chapter 4) presents evidence that attitudes toward torture have evolved. For much of human history, torture was considered unexceptional, and was widely practiced. A utilitarian living in such a time might have had personal preferences that precluded torture, but would not have argued that utilitarianism precluded it. Preferences have now shifted, and with them the implications of utilitarianism.

allocation.

Harsanyi (1955) is silent on this question, devoting a careful argument to establishing that the planner must maximize some weighted sum of utilities while saying nothing about where the weights come from. Dhillon and Mertens (1999) resolve this issue by assuming that every individual's utility function has infimum zero and supremum one, and then invoking an anonymity axiom to assume that every individual receives the same weight in the summation. One problem with this 0-1 normalization is that individuals can differ in their concepts of "best" and "worst" outcomes. For example, a person of faith who believes in the afterlife can conceive of Heaven and Hell as possible outcomes, which might be beyond the scope of an atheist's mental space. As a result, the atheist will have a higher weight in the social welfare function, required to normalize her mundane best- and worst-cases to the same scale as the believer's. Should the atheist promote maximization of such a sum, she could be thought of as telling the believer, "Well, you anyway care more about the afterlife, so why don't you give me all the material goods and live in poverty in the meantime?" This type of reasoning isn't completely foreign to human thought; indeed, believers who join monasteries might be viewed as following such trade-offs between earthly and heavenly payoffs. Yet, it seems problematic to endorse such a policy for a social planner, who would be penalizing the believer for her faith.[14]

Third, person-affecting aggregation gives rise to counterintuitive implications even in the simplest of settings. Suppose there is a single private good x and three individuals, with utility functions

$$\begin{aligned} u_1(x_1) &= \sqrt{x_1} \\ u_2(x_2) &= x_2 \\ u_3(x_3) &= (x_3)^2. \end{aligned}$$

[14]One could instead suggest that the atheist can conceive of Heaven and Hell, but that he attaches zero probabilities to experiencing these outcomes after his death. But this raises the measurability problem again: it might be difficult to design a procedure that would tell us—and the atheist himself—what should be his utility for such extreme outcomes, given that he does not believe any act he takes would result in such outcomes with any positive probability.

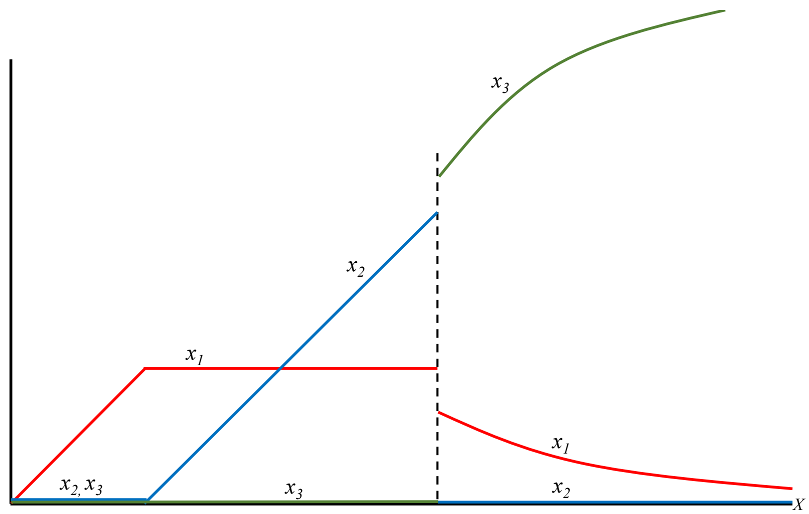

Figure 1: Quantity of resource allocated to person 1 (x_1, red), person 2 (x_2, blue when positive), and person 3 (x_3, green when positive) as a function of total resource X.

Let there be a total quantity X, and assume that we divide it among the three individuals so as to maximize the utilitarian criterion $\sum_{i=1}^{3} \alpha_i u_i(x_i)$, for the case $\alpha_1 = \alpha_2 = \alpha_3$. Our focus is on the way the optimal allocation (x_1, x_2, x_3) varies as X varies. Figure 1 presents the qualitative solution. Never do all three people receive a positive quantity of resource, and there are both non-monotonicities and discontinuities of the allocations with respect to the total amount to be shared. Individual 2 initially receives nothing, then experiences a period in which all of any additional resource is given to her, and is then entirely deprived of the resource. Individual 3 initially receives nothing, then jumps to having the largest allocation, and eventually consumes the entire supply. We might seek refuge from this outcome by attaching different weights to the people, but no assignment of weights will resolve all issues.

While this example seems bizarre, it is not unique. One may construct sufficiently many similarly bizarre examples to populate a zoo. Moreover, the counterintuitive conclusions do not depend on utilitarian aggregation. Suppose one adopted the Rawlsian order, leading the planner to equalize utility levels across individuals. Should individual 1 receive a meager allocation, simply because she is apparently more easily satisfied with smaller rations than either individuals 2 or 3? Should individual 3 receive the bulk of the resource, simply because he is morose and can be placated only by a massive allocation?

If we set out to avoid such quandaries, then the aggregation must not depend on the individual's utilities alone—the aggregation must not be purely person-affecting. The planner must step in, at least supplementing the information contained in preferences with the planner's own view on how to aggregate the agents' well being. This brings us to an impersonal approach.

4.2 Wants vs. Needs

Allowing the planner to invoke her own assessment of individuals' well-being raises challenges for the planner and raises the concern of the planner abusing her discretion. Both difficulties are mitigated by having the planner base her assessment of social welfare on needs rather than wants.

If the distinction between wants and needs is to make sense, we must be able to tell them apart. A cold beer on a hot summer day might be pleasurable, as well as solving the pain of thirst. A concert may be a pleasure to attend, and it may be solving a problem of boredom. Are we dealing with wants or needs? Can there be any pleasure that doesn't satisfy a need of some sort?

We find an analogy between the want-need distinction and the question of interpersonal comparisons of utility. In both cases, a general, theoretically-neat resolution of the problem is elusive. In both, one might then seek refuge in the dry, clean bastion of logical positivism, avoiding all assessments of needs or wants, rather than trod the marshes of shady psychological interpretations. However, doing so leaves not only the marshes but also the seas to politicians who fear no theoretical confusion. We believe there are sufficiently large bodies of clear water out there for economists to introduce these distinctions into their formal models. For instance, it appears evident that owning a private jet isn't a need in the sense that having a dinner is. It seems equally obvious that, given the choice, a social planner should opt for feeding the hungry before supplying private jets to the well-nourished. Our preferred interpretation of the model reflects this distinction.

Can the planner form a reasonable idea of needs and wants? A positive answer is much more likely in the case of needs. The measurement of well-being by subjective self-reports of the extent to which wants are satisfied, used in psychological studies for decades, gave rise to Easterlin's Paradox (Easterlin 1973, 1974, Diener, 1984). The basic argument is that self-reported well-being depends on one's aspirations and expectations, which change with experience as in adaptation level theory (Helson 1947, 1948). The implication is that, if we rely on self reports to determine well-being, it is doubtful that improving well-being is a worthy or even feasible goal.[15] This rather nihilistic view of well-being has been contested, and arguably the only clear conclusion to be drawn from this debate is that research in the social sciences has not reached a conclusion about the definition and measurement of happiness.[16] By contrast, there seems to be little doubt about misery. Suffering hunger and cold, slavery and disease make people miserable. This suggests that public policies that go beyond Pareto domination might do well to focus on minimizing misery rather than maximizing happiness. Translating this discussion to our terms, our view is that wants are irretrievably elusive, while the planner can form a reasonable conception of individuals' needs. The need to focus on needs rather than wants

[15]This point was made by Brickman and Campbell (1971), who suggested that there is no way to attain happiness apart from "stepping off the hedonic treadmill": if more resources raise one's aspirations, so that yet more resources are needed, the pursuit of happiness via material well-being is hopeless. In a similar vein, the famous study by Brickman, Coates, and Janoff-Bulman (1978) compared the subjective well being of lottery winners and accident victims, and found that the dramatic life events had little effect on long-term well-being.

[16]See Lucas, Dyrenforth, and Diener, 2008, Veenhoven and Vergunst, 2014, Easterlin, 2017 on Easterlin's paradox and Mancini, Bonanno, and Clark, 2011, on the effect of major life events on well-being. Moreover, the psychological literature has also shown that subjective well-being can be manipulated (see Schwarz and Clore,1983, as well as Strack, Martin, and Schwarz, 1988). To many economists, the mere findings of Brickman, Coates, and Janoff-Bulman (1978) were sufficient to conclude that subjective well-being isn't a measure one should take seriously. Casual observation suggests that, presenting the findings about lottery winners and accident victims in an economics department invariably raises the question, "Yes, but would they switch?"In order to cope with these difficulties, Kahneman, Krueger, Schkade, Schwarz, and Stone (2004) suggested the Day Reconstruction Method. However, this method has also been criticized for ignoring any subjective inputs. Moreover, it has been argued that meaning is also crucial to happiness (Kauppinen, 2013, and Li, Jonah, Wong, and Chao, 2019), that is, that well-being has both a hedonic and an eudaimonic component.

becomes all the stronger when considering, for example, humanity 200 years hence. What will be their wants? Will they enjoy Beethoven more than techno music? Watching movies or playing computer games? Or are they more likely to enjoy pastimes we can't even conceive of? It seems hard to tell. By contrast, it doesn't seem to be far-fetched to assume that, like us, they will need food and shelter. Again, there is much less uncertainty (or ambiguity) about needs than about wants.

The function u appearing in (2), capturing the preferences of the planner, or the advocate motivating a social policy, or the moralist opining on the nature of a good society, may be informed by the preferences of the individuals in the society, but is not mechanically determined by the latter. The planner may recognize that all of the individuals in Section 4.1 value the good x, but may not respond to the fact that one person is easily satisfied with small quantities of x while another is satisfied only with large dollops of x, allowing the planner to escape the difficulties raised in Section 4.1.[17]

This impersonal approach may appear to cede a great deal of arbitrary power to the planner. What would constrain her? What would prevent a government from imposing its own values on an entire population in an undemocratic way? We believe that the focus on needs, as opposed to wants, limits the risks of abuse of the social aggregation rule. To this end, the specification of needs should reflect a consensus within the society or polity to which the analysis applies.[18] To allay fears that a reliance on consensus dooms us to incompleteness, we note that we do not seek consensus over social policies, but only the set of needs relevant for evaluating those policies. We would hope to reach a wide consensus over questions of the type, "Is food a basic need?"

[17]Paul Bloom (2016) argued that sympathy is good for public policy, but empathy is not. A social planner can appropriately understand and make use of people's priorities without adopting their preferences as his own. Adam Smith (1759) and David Hume (1939-40, 1951) have similarly argued for limited empathy.

[18]The notion of a consensus is related to Pareto domination and, naturally, to incomplete orders. Indeed, Aumann (1962) suggested a model of incomplete preferences that can be viewed as the intersection of weak (complete) orders defined by different utility functions. Bewley (2002) adopted a similar approach to probabilities, whereas Galaabaatar and Karni (2013) combined the two.

while expecting disagreement over "Should we tax the rich in order to feed the poor?" We envision a person advocating for a particular view of what constitutes a social good as appealing to our axioms in arguing that support for her argument is appropriately couched in terms of the lexitarian aggregation of needs. For support of her vision of what these needs are, she must then appeal to the consensus implicit in the structure and workings of the society around her.[19]

At a practical level, we may begin with some needs on which a consensus exists—say, food—and ask, which other needs are comparable to these. For example, a person might be asked, Is education a need? We suggest that such a question be operationalized by the question, "Is there a number n such that depriving n children of education is worse than having one child being hungry?" Clearly, we do not purport to suggest any answers to such questions, and we don't think it is the task of theorists to provide such answers. However, a theoretical model can render such a question more concrete, and help people decide what needs they would like to be included in the social planner's considerations.

The interpretation of the model and the notion of needs clearly depend on the context. While our prime application is for policy decisions at the level of an economy, one can also think of social choice problems in other contexts, where "needs" can be differently interpreted. Consider, for example, parents who have two children, Ada and Bobby. Ada loves mathematics, having obtained a PhD in the field and a tenured academic position. Bobby loves chess. He participates in chess clubs and tournaments, but supports himself by working in an unrelated and unrewarding job. The parents may well consider self-fulfillment as a need, and provide more financial support to

[19]Views as to what constitutes a need and which things are desires that go beyond needs differ across societies and have evolved over time. Many European societies consider health care and education to be needs while this is not the case in the United States. Some societies consider that ability to participate in unhindered public speech a need; others do not. Most societies now take freedom from slavery and torture to be needs, though this is not quite universal. Internet access would recently have been unimaginable even as a want, while some now view it as a need. One might view the United Nations' Universal Declaration of Human Rights an attempt to make explicit the social consensus on needs.

Bobby than to Ada, so that Bobby can satisfy the self-fulfillment need as does Ada. In contrast, self-fulfillment may not qualify as a need the social planner takes into account when devising the income tax code.

4.3 Modeling Needs by Positions

We use the notion of a position to describe the extent to which needs are satisfied. Suppose that there are R needs, and that for each person the extent to which need $r \in \{1, \ldots, R\}$ is met can be described by a number drawn from the set N_k. There are two obvious ways in which the R needs can be captured by positions in our model: (i) we might think of a position as an element of $\prod_{r=1}^{R} N_r$, interpreted as a vector specifying the extent to which each of the R needs is met. Each member of the population would then appear once in a population profile. (ii) We might also think of positions as elements of $P = \bigcup_{r=1}^{R} N_r$. Here, each position gives information only about a single need, and each individual appears R times in each profile, so as to collect information about the extent to which each need is satisfied for this individual.

For example, suppose there are two needs, food and space trips. Each is classified as unsatisfied or wholly satisfied, so that $N_1 = N_2 = \{0, 1\}$. The first modeling option would have $P = \{(0,0), (0,1), (1,0), (1,1)\}$ and a profile I would indicate that there are $I(0,0)$ people who are starving and make no space trips, $I(1,0)$ people who have enough to eat but don't travel in space, and so on. A lexitarian model $\left(\succsim^b, u\right)$ with a trivial $\succsim^b$ (i.e., $\succsim^b = P \times P$ and the strong Archimedean axiom holds) would face the social planner with two choices: either u ignores the second coordinate of the positions, or else one is driven to the conclusion that sufficiently many space trips for some individuals compensates for the starvation of others. To avoid this (repugnant) conclusion, one may wish to use the additional freedom of the lexitarian model. Suppose that $(1,0) \sim^b (1,1) \succ^b (0,0) \sim^b (0,1)$. That is, the relation $\succsim^b$ has two equivalence classes, one in which the individual is well-fed, another in which she starves, with the former ranked strictly above the latter. Within each indifference class, the function u may still represent a preference for space travel, so that $u(1,1) > u(1,0)$. However, there will then exist an number $n \geq$

1 such that $n\left(u\left(1,1\right)-u\left(1,0\right)\right)>u\left(1,0\right)-u\left(0,0\right)$, meaning that providing space travel to n well-fed individuals can compensate for moving one well-fed individual into the starving $\succsim^b$ equivalence class.

By contrast, the second option defines $P=\{nf,f,ns,s\}$ where "f" and "nf" stand for "food" and "no-food", respectively, and "s" and "ns" are similar for space travel. With a trivial $\succsim^b$ one faces the same problem: either $u\left(ns\right)=u\left(s\right)$, and the social planner ignores preferences for space trips, or else there is a number of space trips that would justify starvation. However, if $f\sim^b nf\succ^b s\sim^b ns$, and $u\left(f\right)>u\left(nf\right)$, $u\left(s\right)>u\left(ns\right)$, then one obtains a ranking that first considers food, and only when the number of hungry individuals is identical across two profiles, does it compare space trips. This allows the planner to recognize that food and space trips both have value, but to reject trade-offs between the two. We thus have a model in which needs are the carriers of the planner's utility, while individuals are collections of needs. Observe, however, that the lexitarian model retains the additive nature of utilitarianism. Therefore, in this modeling of needs, they are evaluated in an additive way: the degree to which space trips are preferred to their absence $\left(u\left(s\right)-u\left(ns\right)\right)$ is independent of the individual's hunger. Clearly, one may generalize the model to be non-additive across needs, even if it retains additivity across individuals.[20]

5 Discussion

5.1 Common Practice and Government Budgets

The lexitarian model captures several features of the common practice of government policy. First, economic policy goes beyond Pareto optimality, and, in practice, makes interpersonal comparisons of utility despite the official position of microeconomic textbooks. For example, progressive taxation is typi-

[20]It should be noted that the axiomatic derivation in this paper assumes that all profiles I are conceivable. This means that one should be able to think of a profile that has $I\left(f\right)=1$ and vanishes elsewhere. Such a profile does not provide a full description of an individual's circumstances. It may be interpreted as a situation in which the social planner does not know the degree to which some needs are satisfied.

cally motivated by the intuition that it is easier for the rich than for the poor to part with a given amount of money. Second, choices are made in ways that seem to be intuitively in line with the need/want distinction. Governments allocate budgets to cope with hunger and disease, but not for space trips as a recreational activity.

Beyond these extremes, people and polities have different preferences. For instance, are cultural activities a need that is on par with food and medication? Should the government finance, say, a state opera? Or, to make the question more concrete, is one child's hunger worth n people enjoying an opera? Our model is designed to deal with such problems, allowing one to say that space tourism isn't considered a "need" at all, while opera might be a need which is in a lower category than food.

Actual government decisions necessarily make trade-offs between all of the programs the government funds. Should a government allocate some of its budget to a state opera, it invariably provides less funds to food programs and to life-saving medical treatments. It is fair to ask, therefore, what is the use of a lexitarian model: in the final analysis, are there any policy decisions in which one finds equality at levels $(u_1, ..., u_\ell)$ and hence the decision is determined by differences in the function $u_{\ell+1}$?

Our answer two-fold. First, as is the case with any model, our model involves various idealizations. While perfect equality of $(u_1, ..., u_\ell)$ is indeed unlikely, in reality our ability to measure degrees of need-satisfaction is limited, and utilities can be measured up to just-noticeable-differences. A government may decide to allocate some funds to the opera, considering amounts that would not make a noticeable difference in the degree of satisfaction of more basic needs of the population at large. Second, we imagine expecting a government to motivate a policy decisions by exhibiting a model rationalizing the decision. Faced with a proposal to allocate funds to food assistance, the government might identify a list of priority classes reflecting the levels in our ordering $\succsim^b$ and a description of the trade-offs within classes under which the policy constitutes an improvement.[21] Importantly, various feasibility constraints may

[21]Gilboa and Samuelson (2022) offer a similar criterion for an agent, such as a portfolio

render it impossible to satisfy more highly-regarded needs, whereas less crucial ones can be satisfied to higher degrees. The lexitarian model provides the government with a convenient language for expressing these priorities and trade-offs.

5.2 The Repugnant Conclusion

Parfit (1984) raises the "Repugnant Conclusion", frequently implied by aggregative welfare approaches: namely, that a sufficiently-large, marginally-subsistent society would always be mechanically preferred to a small, flourishing one.[22] Several directions have emerged in the population ethics literature to address this critique, including variants of average utilitarianism[23] (Harsanyi, 1955; Sikora, 1975); rank-dependent approaches (Zuber and Asheim, 2012); prioritarianism (Adler, 2012; Parfit, 1997); and various threshold-based modifications, including leximin principles (Hammond, 1976), utilitarianism dampened by population (Hurka, 1983), and maximin welfare orderings (Bossert, 1990), and critical thresholds below which individuals' well-being does not contribute to aggregate welfare (Blackorby, Bossert, and Donaldson, 1995; 1997). Recent contributions that deal with the repugnant conclusion and related problems in a utilitarian framework include Pivato (2020) and Nebel (2024).

Our model nests the generally "prioritarian" impulse of focusing on the most-suffering and the threshold-based intuition, common to these approaches, and extends it by including a rank-ordering of priorities on which the social planner lexicographically judges a society.

manager or a government, making decisions under uncertainty on behalf of others.

[22] Also known as the "Mere Addition Paradox", the paradoxical nature of the problem is contested: some philosophers have disputed the "repugnancy" of the conclusion (Huemer, 2008; Tännsjö, 2002; Gustaffson, 2022), while others have considered this as revelatory of a fundamental flaw to aggregation of welfare in ethics and utilitarianism (Temkin, 2012).

[23] Which, nonetheless, are exposed to critiques analogous to the repugnant conclusion; see also Parfit (1984) and Arrhenius (2000).

6 Proofs

The proof of Theorem 1 builds on the proof of Theorem 2, and so we present the latter first.

6.1 Proof of Theorem 2

It is straightforward that the representation (2) implies that $\succsim$ satisfies A1-A3. We thus turn to prove sufficiency of the axioms, from which the uniqueness of the representation will also follow. Naturally, the idea is to extend the relation, given for natural-number vectors, to real-valued vectors and apply a separation theorem. More specifically, we wish to successively extend $\succsim$ to vectors of integer and rational numbers, and then apply a separation theorem to the convex hull of the rational-valued vectors. To this end we define the following sets of vectors:

$$\begin{aligned} \mathcal{I}_{\mathbb{Z}} &= \{ I : P \to \mathbb{Z} \mid \#supp(I) < \infty \} \\ \mathcal{I}_{\mathbb{Q}} &= \{ I : P \to \mathbb{Q} \mid \#supp(I) < \infty \} \end{aligned}$$

For two sets $A \subset B$ and two binary relations on them, $\succsim \subset A \times A$ and $\succsim' \subset B \times B$, we say that $\succsim'$ is an *extension* of $\succsim$ if it extends both the weak and strict preference, that is, if, $\succsim \subset \succsim'$ and $\succ \subset \succ'$. We will also need to refer to the versions of axioms A1-A3 where the vectors are in $\mathcal{I}_{\mathbb{Z}}$ and in $\mathcal{I}_{\mathbb{Q}}$ (rather than only in $\mathcal{I}$). To simplify notation, we will use the same terms "A1", "A2", "A3", rather than redefine them for these domains.

Assume, then, that $\succsim \subset \mathcal{I} \times \mathcal{I}$ satisfies A1-A3. We first extend A2 to apply also for vectors K that are integer-valued but not necessarily nonnegative.

Lemma 1 *There exists a unique $\succsim_{\mathbb{Z}}$ on $\mathcal{I}_{\mathbb{Z}}$ satisfying A1-A3 and extending $\succsim$ (from $\mathcal{I}$ to $\mathcal{I}_{\mathbb{Z}}$).*

Proof: Let there be given $I, J \in \mathcal{I}_{\mathbb{Z}}$. Define $I \succsim_{\mathbb{Z}} J$ if there exists $K \in \mathcal{I}$ such that $I + K \succsim J + K$ (here and in the sequel, the claim that $\succsim$ holds between two vectors is taken to mean that they are both in $\mathcal{I}$ and that the relation holds between them). Clearly, $\succsim_{\mathbb{Z}}$ extends $\succsim$. Completeness of $\succsim$ on

$\mathcal{I}$ implies completeness of $\succsim_{\mathbb{Z}}$ on $\mathcal{I}_{\mathbb{Z}}$. To see that $\succsim_{\mathbb{Z}}$ is transitive, assume that for $I, J, L \in \mathcal{I}_{\mathbb{Z}}$ we have $I \succsim_{\mathbb{Z}} J$ and $J \succsim_{\mathbb{Z}} L$. Let $K, K' \in \mathcal{I}$ be such that (i) $I + K \succsim J + K$ and (ii) $J + K' \succsim L + K'$. Applying A2 to (i) and K' and then to (ii) and K, we have $I + K + K' \succsim J + K + K' \succsim L + K + K'$ and thus $I \succsim_{\mathbb{Z}} L$.

Next we wish to show that $\succsim_{\mathbb{Z}}$ satisfies A2 on $\mathcal{I}_{\mathbb{Z}}$. Let there be given $I, J, K \in \mathcal{I}_{\mathbb{Z}}$, with $I \succsim_{\mathbb{Z}} J$. This means that there exists $K' \in \mathcal{I}$ such that $I + K' \succsim J + K'$. Assume first that $K \in \mathcal{I}$. By A2 $I + K' + K \succsim J + K' + K$ so that $I + K \succsim_{\mathbb{Z}} J + K$. Similarly, $I + K \succsim_{\mathbb{Z}} J + K$ implies $I \succsim_{\mathbb{Z}} J$ (using $(K + K') \in \mathcal{I}$).

For a general (not necessarily nonnegative) $K \in \mathcal{I}_{\mathbb{Z}}$, let $K^+, K^- \in \mathcal{I}$ be such that $K = K^+ - K^-$. Then $I + K = I + K^+ - K^-$. Denote $I' = I - K^- \in \mathcal{I}_{\mathbb{Z}}$. Similarly, let $J' = J - K^-$. By the restricted A2 (applied to the nonnegative $K^- \in \mathcal{I}$), $I' \succsim_{\mathbb{Z}} J'$, iff $I' + K^- \succsim_{\mathbb{Z}} J' + K^-$, that is iff $I \succsim_{\mathbb{Z}} J$. Applying the restricted A2 again to I', J', this time with the non-negative $K^+ \in \mathcal{I}$, we get $I' \succsim_{\mathbb{Z}} J'$ iff $I' + K^+ \succsim_{\mathbb{Z}} J' + K^+$. Thus, $I \succsim_{\mathbb{Z}} J$ iff $I' + K^+ \succsim_{\mathbb{Z}} J' + K^+$. It is left to note that $I' + K^+ = I - K^- + K^+ = I + K$ and similarly $J' + K^+ = J + K$, so that $I \succsim_{\mathbb{Z}} J$ iff $I + K \succsim_{\mathbb{Z}} J + K$.

We now turn to A3. Assume that $I, J, K, L \in \mathcal{I}_{\mathbb{Z}}$, and $K \succ_{\mathbb{Z}} L$. Let $K' \in \mathcal{I}$ be such that $I + K', J + K' \in \mathcal{I}$. As $K \succ_{\mathbb{Z}} L$, there exists $K'' \in \mathcal{I}$ such that $K + K'' \succ L + K''$. Applying A3 (for $\succ$ on $\mathcal{I}$) to $I + K', J + K'$ and the pair $K + K'' \succ L + K''$, we conclude that for some $n \geq 0$ we have $I + K' + n\,(K + K'') \succ J + K' + n\,(L + K'')$. This implies $I + nK + (K' + nK'') \succ_{\mathbb{Z}} J + nL + (K' + nK'')$ and, by A2, this is equivalent to $I + nK \succ_{\mathbb{Z}} J + nL$.

By similar arguments, $\succsim_{\mathbb{Z}}$ is the unique extension of $\succsim$ to $\mathcal{I}_{\mathbb{Z}}$ that satisfies the three axioms. $\square$

Lemma 2 *For all $I, J \in \mathcal{I}_{\mathbb{Z}}$, (i) $I \succsim_{\mathbb{Z}} J$ iff $I - J \succsim_{\mathbb{Z}} 0$; (ii) for all $n > 1$, $I \succsim_{\mathbb{Z}} J$ iff $nI \succsim_{\mathbb{Z}} nJ$.*

Proof: Follows from A2 and transitivity. $\square$

We can now make the next step into the rationals.

Lemma 3 *There exists a unique* $\succsim_{\mathbb{Q}}$ *on* $\mathcal{I}_{\mathbb{Q}}$ *satisfying A1-A3 and extending* $\succsim_{\mathbb{Z}}$ *(from* $\mathcal{I}_{\mathbb{Z}}$ *to* $\mathcal{I}_{\mathbb{Q}}$*).*

Proof: Let there be given $I, J \in \mathcal{I}_{\mathbb{Q}}$. Define $I \succsim_{\mathbb{Q}} J$ if there exists $n \geq 1$ such that $nI \succsim_{\mathbb{Z}} nJ$. Clearly, $\succsim_{\mathbb{Q}}$ extends $\succsim_{\mathbb{Z}}$. Completeness of $\succsim_{\mathbb{Q}}$ follows from completeness of $\succsim_{\mathbb{Z}}$ on $\mathcal{I}_{\mathbb{Z}}$. To see that transitivity holds as well, assume that $I, J, K \in \mathcal{I}_{\mathbb{Q}}$ with $I \succsim_{\mathbb{Q}} J$ and $J \succsim_{\mathbb{Q}} K$. Let n be such that $nI \succsim_{\mathbb{Z}} nJ$ and let m be such that $mJ \succsim_{\mathbb{Z}} mK$. Applying Lemma 2 to both, we conclude that $nmI \succsim_{\mathbb{Z}} nmJ \succsim_{\mathbb{Z}} nmK$ transitivity of $\succsim_{\mathbb{Z}}$ implies that of $\succsim_{\mathbb{Q}}$.

Next consider A2. It suffices to show that $I \succsim_{\mathbb{Q}} J$ implies $I + K \succsim_{\mathbb{Q}} J + K$ (as the converse argument is symmetric). Let there be given $I, J, K \in \mathcal{I}_{\mathbb{Q}}$, with $I \succsim_{\mathbb{Q}} J$. Let $n \geq 1$ be such that $nI \succsim_{\mathbb{Z}} nJ$. Let $m \geq 1$ satisfy $mK \in \mathcal{I}_{\mathbb{Z}}$. By Lemma 2, $nmI \succsim_{\mathbb{Z}} nmJ$ and A2 yields $nmI + nmK \succsim_{\mathbb{Z}} nmJ + nmK$, that is, $I + K \succsim_{\mathbb{Q}} J + K$.

We now turn to A3. Assume that $I, J, K, L \in \mathcal{I}_{\mathbb{Q}}$, and $K \succ_{\mathbb{Q}} L$. Let $m \geq 1$ satisfy $mJ, mI \in \mathcal{I}_{\mathbb{Z}}$. For some $m' \geq 1$, $m'K \succ_{\mathbb{Z}} m'L$. By Lemma 2, $mm'K \succ_{\mathbb{Z}} mm'L$. Applying A3 (on $\succsim_{\mathbb{Z}}$ on $\mathcal{I}_{\mathbb{Z}}$) to $mm'I, mm'J, mm'K, mm'L$ we infer that there exists $n \geq 0$ such that $mm'I + nmm'K \succ_{\mathbb{Z}} mm'J + nmm'L$, which, by definition, means that $I + nK \succ_{\mathbb{Q}} J + nL$.

Finally, similar arguments imply that the extension $\succsim_{\mathbb{Q}}$ is unique. □

Lemma 4 *For all* $I, J \in \mathcal{I}_{\mathbb{Q}}$*, (i)* $I \succsim_{\mathbb{Q}} J \iff I - J \succsim_{\mathbb{Q}} 0$*; (ii) for all* $n > 1$*,* $I \succsim_{\mathbb{Q}} J \iff nI \succsim_{\mathbb{Q}} nJ$*.*

Proof: Identical to the proof of Lemma 2. □

It is worth mentioning explicitly the following:

Lemma 5 *For all* $I, J \in \mathcal{I}_{\mathbb{Q}}$*, and every rational number* $\alpha > 0$*,* $I \succsim_{\mathbb{Q}} J$ *iff* $\alpha I \succsim_{\mathbb{Q}} \alpha J$ *(and thus* $I \succ_{\mathbb{Q}} J$ *iff* $\alpha I \succ_{\mathbb{Q}} \alpha J$*).*

Proof: Follows from Lemma 4. □

We turn to the definition of the function u. If $I \sim J$ for all $I, J \in \mathcal{I}$, select $u(\cdot) \equiv 0$. Otherwise, assume that $I^* \succ J^*$ for some $I^*, J^* \in \mathcal{I}$. Consider a finite $P_0 \subset P$ such that $P^* \equiv supp(I^*) \cup supp(J^*) \subset P_0$. We will define u

for such a subset P_0 and then show how these definitions can be "patched" together to a definition of u on all of P. Let $\mathcal{I}_{\mathbb{Q}_0}$ be the set of vectors in $\mathcal{I}_{\mathbb{Q}}$ that vanish outside of P_0. Define

$$\begin{aligned} A &= \{ I \in \mathcal{I}_{\mathbb{Q}_0} \mid I \succ_{\mathbb{Q}} 0 \} \\ B &= \{ I \in \mathcal{I}_{\mathbb{Q}_0} \mid 0 \succ_{\mathbb{Q}} I \} \\ E &= \{ I \in \mathcal{I}_{\mathbb{Q}_0} \mid I \sim_{\mathbb{Q}} 0 \} \end{aligned}$$

We consider A, B, E as subsets of the finite-dimensional Euclidean space $\mathbb{R}^{|P_0|}$, and observe the following: (I) A, B, E are pairwise disjoint; (II) $A \cup B \cup E = \mathcal{I}_{\mathbb{Q}_0}$; (III) $A = -B$; (IV) $0 \in E$; (V) for all $I, J \in \mathcal{I}_{\mathbb{Q}_0}$, $I \succ_{\mathbb{Q}} J \iff I - J \in A$, $J \succ_{\mathbb{Q}} I \iff I - J \in B$, and $I \sim_{\mathbb{Q}} J \iff I - J \in E$; and (VI) $A, B \neq \varnothing$.

Lemma 6 *Suppose that $I, J \in A(B, E)$ and $\alpha \in [0,1] \cap \mathbb{Q}$. Then $\alpha I + (1-\alpha) J \in A(B, E)$.*

Proof: Follows from Lemma 5, A2, and transitivity. $\square$

In the following, for $I \in \mathbb{R}^{|P_0|}$ and $\varepsilon > 0$, $N_\varepsilon(I)$ denotes the ε-neighborhood of I in $\mathbb{R}^{|P_0|}$.

Lemma 7 *Let $I \in A(B)$. There exists $\varepsilon > 0$ such that for every $J \in N_\varepsilon(I) \cap \mathcal{I}_{\mathbb{Q}_0}$, we have $J \in A(B)$.*

Proof: Let there be given $I \in A$. For $1 \leq j \leq |P_0|$, let e_j be the j-th unit vector in $\mathbb{R}^{|P_0|}$. We wish to find an $\varepsilon_j > 0$ such that $[I - \varepsilon_j e_j, I + \varepsilon_j e_j] \cap \mathcal{I}_{\mathbb{Q}_0} \subset A$. Consider $K^+ = I + e_j \in \mathcal{I}_{\mathbb{Q}_0}$. By A3 there exists $n \geq 1$ such that $K^+ + nI \succ_{\mathbb{Q}} 0$. Hence by Lemma 4), $\frac{1}{n+1}[K^+ + nI] \succ_{\mathbb{Q}} 0$. Thus, for $\varepsilon_j^+ = \frac{1}{n+1}$ we have $\left[I, I + \varepsilon_j^+ e_j\right] \cap \mathcal{I}_{\mathbb{Q}_0} \subset A$. Similarly, there exists $\varepsilon_j^- > 0$ such that $\left[I - \varepsilon_j^- e_j, I\right] \cap \mathcal{I}_{\mathbb{Q}_0} \subset A$—and we define $\varepsilon_j = \min\left(\varepsilon_j^+, \varepsilon_j^-\right)$.

Let $\bar{\varepsilon} = \min_j \varepsilon_j$ and set $\varepsilon = \frac{\bar{\varepsilon}}{\sqrt{|P_0|}} > 0$. It follows that the ε-ball around I is included in the convex hull of $\{I - \bar{\varepsilon} e_j, I + \bar{\varepsilon} e_j\}_{j \leq |P_0|}$. Consider a rational point $J \in N_\varepsilon(I) \cap \mathcal{I}_{\mathbb{Q}_0}$. By Carathodory's Theorem, there are $l \leq |P_0| + 1$ points $\{I_r\}_{r \leq l}$ in $\{I - \bar{\varepsilon} e_j, I + \bar{\varepsilon} e_j\}_{j \leq |P_0|}$ and nonnegative numbers $\{\alpha_r\}_{r \leq l}$ adding up to 1 such that $J=\sum_{r\leq l}\alpha_r I_r$, and, furthermore, these $\{\alpha_r\}_{r\leq l}$ are unique (for the given set $\{I_r\}_{r\leq l}$). This implies that $\{\alpha_r\}_{r\leq l}$ are rational (as the unique solution to a system of equations with rational numbers). It follows, from iterated applications of of Lemma 6, that $J=\sum_{r\leq l}\alpha_r I_r\in A$.

The argument for B is symmetric. $\square$

Lemma 8 *For every $I\in E$ there are arbitrarily close points $J^+\in A$ and $J^-\in B$.*

Proof: Follows from Lemma 4, A2, and transitivity. $\square$

Lemma 9 *There exists $u:P_0\to\mathbb{R}$ such that, for all $I,J\in\mathcal{I}_{\mathbb{Q}_0}$*

$$
\begin{array}{rcl}
I & \succsim & {}_{\mathbb{Q}_0}J \\
 & & \textit{iff} \\
\sum_{p\in P_0} I(p)\,u(p) & \geq & \sum_{p\in P_0} J(p)\,u(p)
\end{array}
$$

Further, u is unique up to multiplication by a positive number.

Proof: We know that A and B are closed under rational convex combinations. Consider $A^0=int\,(conv\,(A))$ and $B^0=int\,(conv\,(B))$, where $conv$ refers to the convex hull in $\mathbb{R}^{|P_0|}$ (including not-necessarily rational points, and not-necessarily rational convex combinations). By Lemma 7 the sets $A^0,B^0\subset\mathbb{R}^{|P_0|}$, which are open and convex, are non-empty (and full-dimensional). By Lemma 7 we also know that $A\subset A^0$ and $B\subset B^0$. Moreover, for any $I\in A^0(B^0)$ there exists $\varepsilon>0$ such that $N_\varepsilon(I)\cap\mathcal{I}_{\mathbb{Q}_0}\subset A(B)$. In particular, this means that $A^0\cap B^0=\varnothing$: if there were $x\in A^0\cap B^0$, we would be able to find two neighborhoods thereof, such that in one all rational points have to be in A and in the other—in B, in contradiction to $A\cap B=\varnothing$.

Thus we can apply a separating hyperplane theorem to conclude that there exists $u:P_0\to\mathbb{R}$ and a number $c\in\mathbb{R}$ such that

$$
\begin{array}{rclcl}
I\in & A^0 & \Longrightarrow & I\cdot u>c & \text{(i)} \\
I\in & B^0 & \Longrightarrow & I\cdot u<c & \text{(ii)}
\end{array}
$$

We first note that the implications in (i) and in (ii) can also be reversed. To see this, assume that $I \cdot u > c$. Choose $\varepsilon > 0$ such that for every $J \in N_\varepsilon(I)$, we have $J \cdot u > c$. We argue that, if $J \in \mathcal{I}_{\mathbb{Q}_0}$, we have $J \in A$. Indeed, if $J \in B$ then $J \in B^0$ and (ii) would imply $J \cdot u < c$. If, however, $J \in E$, then, by Lemma 8, there are points $J' \in B$ arbitrarily close to J, so that we can pick $J' \in N_\varepsilon(I) \cap B$. This would contradict (ii) again. Thus all rational points in $N_\varepsilon(I)$ are in A and $I \in A^0$ follows. The argument is symmetric for B (and the converse of (ii)).

Next, observe that for $I \in E$ we have to have $I \cdot u = c$, because an inequality would imply that I is in A or in B. Finally we observe that $c = 0$ because $0 \in E$.

We conclude that for all $I \in \mathcal{I}_{\mathbb{Q}}$

$$
\begin{aligned}
I \in A &\iff I \cdot u > 0 \\
I \in E &\iff I \cdot u = 0 \\
I \in B &\iff I \cdot u < 0
\end{aligned}
$$

which means that, for all $I, J \in \mathcal{I}_{\mathbb{Q}_0}$, $I \succsim_{\mathbb{Q}_0} J$ ($I \succ_{\mathbb{Q}_0} J$) iff $I - J \in A \cup E$ ($I - J \in A$) iff $(I - J) \cdot u \geq 0$ (> 0)—and u is unique up to multiplication by a positive constant $\lambda > 0$. □

To conclude the proof of the theorem, we now define u for all p. We apply Lemma 9 and fix one $u_0 : P_0 \to \mathbb{R}$ that represents $\succsim_{\mathbb{Q}_0}$ on $\mathcal{I}_{\mathbb{Q}_0}$. For every $p \notin P_0$, we consider $P_p = P \cup \{p\}$ and the corresponding $\mathcal{I}_{\mathbb{Q}_p}$. There exists a unique function u_p that extends u_0 to P_p and represents $\succsim_{\mathbb{Q}_p}$. Let $u(p) = u_p(p)$. Thus u is defined for all of P.

To see that u so defined represents $\succsim_{\mathbb{Q}}$ on all of $\mathcal{I}_{\mathbb{Q}}$, consider $K, L \in \mathcal{I}_{\mathbb{Q}}$. Limit attention to $P' = P^* \cup supp(K) \cup supp(L)$. By the same arguments, there exists a $u' : P' \to \mathbb{R}$ that represents $\succsim_{\mathbb{Q}'}$. It thus also represents $\succsim_{\mathbb{Q}_0}$ on $\mathcal{I}_{\mathbb{Q}_0}$ and $\succsim_{\mathbb{Q}_p}$ on $\mathcal{I}_{\mathbb{Q}_p}$ for every $p \in P'$. Hence it has to equal u up to multiplication by a positive constant. □

6.2 Proof of Theorem 1

We begin with sufficiency of the axioms. We first note that

Lemma 10 *For all $I, J, K, L \in \mathcal{I}$ and all $m \geq 1$,*

(i) $I \succsim J$ and $K \succsim (\succ) L$ imply $I + K \succsim (\succ) J + L$;

(ii) If $K \sim L$ then $I \succsim J$ iff $I + K \succsim J + L$;

(iii) $I \succsim (\succ) J$ implies $mI \succsim (\succ) mJ$.

Proof: (i) follows from two applications of A2 and transitivity: $I \succsim J$ iff $I + K \succsim J + K$ and $K \succsim L$ iff $J + K \succsim J + L$.

(ii) follows from two applications of (i) if $I \succsim J$ then $K \succsim L$ implies $I + K \succsim J + L$. Conversely, if $I + K \succsim J + L$, had $J \succ I$ been the case, (i) would yield $J + L \succ I + K$, a contradiction.

(iii) follows from iterated applications of (i). $\square$

We now turn to define the relation $\succsim \cdot$: For $p, q \in P$, $p \succsim \cdot q$ if $\exists k \geq 1$ such that $k\mathbf{1}_p \succsim \mathbf{1}_q$.

Lemma 11 *(i) $\succsim \cdot$ is complete;*

(ii) $\succsim \cdot$ is transitive;

(iii) $\sim \cdot$ coincides with comparability.

Proof: (i) follows from completeness of $\succsim$, say, applied to $\mathbf{1}_p$ and $\mathbf{1}_q$.

(ii) follows from transitivity of $\succsim$ and Lemma 10: assume that $p \succsim \cdot q$ and $q \succsim \cdot r$. Let $k, l \geq 1$ be such that $k\mathbf{1}_p \succsim \mathbf{1}_q$ and $l\mathbf{1}_q \succsim \mathbf{1}_r$. Part (iii) of Lemma 10 implies that $kl\mathbf{1}_p \succsim l\mathbf{1}_q$ and transitivity—that $kl\mathbf{1}_p \succsim \mathbf{1}_r$, hence $p \succsim \cdot r$.

Finally, (iii) is immediate. $\square$

Let us define the $\sim \cdot$-equivalence classes to be $\mathcal{A} \equiv P/ \sim \cdot$. We use $\succsim \cdot$ over $\mathcal{A}$ with the obvious meaning. We can write, for each I,

$$I = \sum_{A \in \mathcal{A}} I_A$$

and note that $I_A = 0$ for all but finitely many A's. Equivalently, for every I there is a finite set of $\sim \cdot$-equivalence classes, $(A_1, ..., A_m)$ such that $A_i \succ \cdot A_{i+1}$ (for $1 \leq i < m$) and

$$I = \sum_{i \leq m} I_{A_i}$$

(and $I_B = 0$ for all $B \in \mathcal{A} \backslash \{A_1, ..., A_m\}$).

The key step in the proof is to show that, if two profiles I, J can be ranked on a given $A \in \mathcal{A}$, anything that happens in lower indifference classes of positions does not matter. Formally,

Lemma 12 *Let there be given $I, J \in \mathcal{I}$ and $A \in \mathcal{A}$ such that*

(i) $I_A \succ J_A$

(ii) $I_B \sim J_B$ for all $B \in \mathcal{A}$ with $B \succ \cdot A$;

Then $I \succ J$.

Proof: We first write

$$\begin{aligned} I &= \sum_{\substack{B\in\mathcal{A}\\ B\succ\cdot A}} I_B + I_A + \sum_{\substack{C\in\mathcal{A}\\ A\succ\cdot C}} I_C \\ J &= \sum_{\substack{B\in\mathcal{A}\\ B\succ\cdot A}} J_B + J_A + \sum_{\substack{C\in\mathcal{A}\\ A\succ\cdot C}} J_C \end{aligned}$$

and it will be useful to look also at

$$\begin{aligned} I^* &= \sum_{\substack{B\in\mathcal{A}\\ B\succ\cdot A}} I_B + I_A \\ J^* &= \sum_{\substack{B\in\mathcal{A}\\ B\succ\cdot A}} J_B + J_A \end{aligned}$$

Because $I_A \succ J_A$, we can Lemma 10 consecutively for every $B \succ \cdot A$ such that $I_B \neq 0$ or $J_B \neq 0$ (and there are finitely many such B's) to conclude that $I^* \succ J^*$. We now wish to prove, that we can change I^* and J^* on any position p in C with $A \succ \cdot C$, without changing the strict preference. As there are finitely many positions that distinguish (I, J) from (I^*, J^*), the argument can be used inductively to show that $I^* \succ J^*$ does indeed imply $I \succ J$.

Let there be given such a p, so that $q \succ \cdot p$ for every $q \in A$. Consider the modified profiles

$$\begin{aligned} I_p^* &= I^* + I(p)\,\mathbf{1}_p \\ J_p^* &= J^* + J(p)\,\mathbf{1}_p \end{aligned}$$

We wish to show that $I_p^* \succ J_p^*$. We will use A4 to conclude that $I_p^* \succsim I^*$, and thus it would suffice to prove that $I^* \succ J_p^*$.[24] Assume not. Then

$$J_p^* = J^* + J(p)\,\mathbf{1}_p \succsim I^*$$

Consider any $q \in A$. Apply the Archimedean axioms to the four profiles $0, \mathbf{1}_q, I_A, J_A$ whose supports are all included in A. Since $I_A \succ J_A$, there exists $k \geq 1$ such that $(0+)\, kI_A \succ \mathbf{1}_q + kJ_A$. (That is, adding one at position q does not suffice to reverse the strict preference $kI_A \succ kJ_A$.) However, $J^* + J(p)\,\mathbf{1}_p \succsim I^*$ implies (by Lemma 10 (iii)) that

$$kJ^* + kJ(p)\,\mathbf{1}_p \succsim kI^*$$

Using Lemma 10 (ii), when applied to all B with $B \succ \cdot A$ (for which $I_B \sim J_B$ and $kI_B \sim kJ_B$), we also have

$$kJ_A + kJ(p)\,\mathbf{1}_p \succsim kI_A \succ \mathbf{1}_q + kJ_A$$

The preference $kJ_A + kJ(p)\,\mathbf{1}_p \succ \ \mathbf{1}_q + kJ_A$ implies, using Lemma 10 (i), $kJ(p)\,\mathbf{1}_p \succ \mathbf{1}_q$ which means that $p \succsim \cdot q$—a contradiction.

Thus we have $I_p^* \succ J_p^*$ and, by repeating the argument inductively for every p in $supp(I) \cup supp(J)$ with $q \succ \cdot p$ for $q \in A$, we conclude that $I \succ J$. $\square$

We now complete the proof of the theorem. For every $A \in \mathcal{A}$, we restrict attention to

$$\mathcal{I}_A = \{\, I_A \mid I \in \mathcal{I} \}$$

and notice that $\succsim$ satisfies A1-A4 on this space, where A3, the unrestricted version of Archimedeanity, follows from the fact that all I_A vanish outside of the equivalence class A. Theorem 2 guarantees that existence of a nonnegative u_A that represents $\succsim$ on $\mathcal{I}_A$ via the product operation. The function u has to vanish outside of A. We define $u = \sum\limits_{A \in \mathcal{A}} u_A$ which is well-defined because each p belongs to a unique A.[25] It is straightforward to verify that $(\succsim \cdot, u)$ is a lexitarian representation of $\succsim$.

[24] If A4 is not assumed, the argument is similar.

[25] Alternatively, one could consider the restrictions of profiles to equivalence classes, obtain a function u_A that is defined only over A , and define u as their union.

The proof of necessity of the axioms is straightforward and therefore omitted.

Clearly, we obtain a function u that can be separately re-scaled on each $A \in \mathcal{A}$, but is otherwise unique: there can be at most one $A \in \mathcal{A}$ consisting of p with $u(p) = 0$, where for $q \in A^c$ we have $u(q) > 0$. Further, for $q_1, q_2 \in A^c$, $u(q_1)/u(q_2)$ is uniquely defined. If we wish the relation $\succsim \cdot$ to have maximal equivalence classes, as in our construction, then it is unique. However, one may have an equivalent representation with other relations $\succsim' \cdot$, which agree on the equivalence classes over which $u(q) > 0$ but may not consider all positions p with $u(p) = 0$ to be $\sim' \cdot$ equivalent. $\square$

6.3 Proof of Proposition 1

It is straightforward to see that the axioms are necessary: given a representation of $\succsim$ by $V(I) = \sum_{\omega \in \Omega} pr(\omega) \sum_{t \in T} \delta^t \sum_{c \in C} I(c, t, \omega) \hat{u}(c)$, we have a representation as in Theorem 2 with $u(c, t, \omega) = pr(\omega) \delta^t \hat{u}(c)$. This means that, restricting attention to profiles in $\mathcal{I}_\omega$ or in $\mathcal{I}_{\omega'}$ preferences are identical if both ω, ω' are non-null (or if both are null, which means that $u(c, t, \omega) = u(c, t, \omega') = 0$ for all $(c, t) \in C \times T$). Similarly, A6 and A7 follow from the representation.

Conversely, assume that A5-A7 hold. Consider non-null $\omega, \omega' \in \Omega$ and their corresponding $\mathcal{I}_\omega$,$\mathcal{I}_{\omega'}$. Applying Theorem 2 to each of $\mathcal{I}_\omega$,$\mathcal{I}_{\omega'}$ separately, we conclude that there exist functions $u_\omega, u_{\omega'} : C \times T \to \mathbb{R}$ that represent $\succsim$ on these restricted spaces via the inner products (i.e., for $I, J \in \mathcal{I}_\omega$, $I \succsim J$ iff $I \cdot u_\omega \geq J \cdot u_\omega$, and similarly for $I, J \in \mathcal{I}_{\omega'}$). By A5, $\succsim$ on $\mathcal{I}_\omega$ is identical to its restriction to $\mathcal{I}_{\omega'}$. Thus, u_ω also represents $\succsim$ on $\mathcal{I}_{\omega'}$ (and vice versa). By the uniqueness result, $u_{\omega'}$ is a positive multiple of u_ω. Fix one non-null state ω_0 (if such exists) and for every ω' choose $\lambda_{\omega'} > 0$ such that $u_{\omega'} = \lambda_{\omega'} u_{\omega_0}$. That is, $u(c, t, \omega') = \lambda_{\omega'} u(c, t, \omega_0)$. If all states are null, u vanishes. Otherwise, normalize $(\lambda_{\omega'})_{\omega'}$ to be a probability vector pr.

Next consider two periods $t, t' \in T$ and, using similar reasoning applied to A6, conclude that, for every $t > 0$ there exists $\mu_t > 0$ such that $u(c, t, \omega_0) = \mu_t u(c, 0, \omega_0)$. Define $\hat{u}(c) = u(c, 0, \omega_0)$ and observe that $u(c, t, \omega) = pr(\omega) \mu_t \hat{u}(c)$.

Finally, invoke A7 to conclude that, for $\delta \equiv \mu_1 > 0$ we have to have $\mu_t = \delta^t$ for all t. $\square$

6.4 Proof of Proposition 2

We start with necessity of the axioms. A8 follows from condition (ii). We first note that, for each $\sim^b$-equivalence class A, A5-A7 hold for profiles whose supports are contained in it, $\mathcal{I}_A$. This follows from the necessity of the axioms in Proposition 1. To see that these axioms hold in general, it remains to follow the lexicographic representation. For example, consider A5. Let there be given two non-null $\omega, \omega' \in \Omega$ and $I, J \in \mathcal{I}_\omega$, and $I', J' \in \mathcal{I}_{\omega'}$ with $I(c,t,\omega) = I'(c,t,\omega')$ and $J(c,t,\omega) = J'(c,t,\omega')$ for all $(c,t) \in C \times T$. We need to show that $I \succsim J$ iff $I' \succsim J'$. Assume that $I \succ J$. Let A be the $\succsim^b$-maximal equivalence class such that $I_A \nsim J_A$ (that is, $I_A \succ J_A$). Then, for any equivalence class $A' \succ^b A$ we have $I_{A'} \sim J_{A'}$ and hence $I'_{A'} \sim J'_{A'}$. Similarly, for A we have $I'_A \succ J'_A$ and thus $I' \succ J'$. If, however, $I \sim J$, we have to have $I_A \sim J_A$ for all A and thus also $I'_A \sim J'_A$ and $I \sim J$ follows. Similar reasoning applied to the necessity of A6 and A7.

To prove sufficiency of the axioms, we apply the sufficiency part of Proposition 1 for each $\sim^b$-equivalence class A separately, and derive the representation as in (i) with $u(c,t,\omega) = pr_A(\omega)\, \delta_A^t \hat{u}(c)$ for all (c,t,ω) (where pr_A and δ_A are derived from Proposition 1). Finally, (ii) follows from A8. $\square$

7 Bibliography

Adler, Matthew D. (2012), *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis.* Oxford University Press.

Alcantud, Jose-Carlos R., Marco Mariotti, and Roberto Veneziani (2022), "Sufficientarianism". *Theoretical Economics*, **17**: 1529-1557.

Arrhenius, Gustaf (2000), "An Impossibility Theorem for Welfarist Axiologies", *Economics and Philosophy*, **16**: 247-266.

Arrow, Kenneth J. (1950), “A Difficulty in the Concept of Social Welfare”, *Journal of Political Economy*, **58**: 328-346.

Aumann, Robert J. (1962), “Utility Theory Without the Completeness Axiom”, *Econometrica*, **30**: 445-462.

Bewley, Truman (2002), “Knightian Decision Theory, Part I”, *Decisions in Economics and Finance*, **25**: 79-110.

Blackorby, Charles, Walter Bossert, and David Donaldson (1995), “Intertemporal Population Ethics: Critical-Level Utilitarian Principles”, *Econometrica*, **65**: 1303-20.

Blackorby, Charles, Walter Bossert, and David Donaldson (1997), “Critical-Level Utilitarianism and the Population Ethics Dilemma”, *Economics and Philosophy*, **13**: 197-230.

Blackorby, Charles and David Donaldson (1984), “Social Criteria for Evaluation Population Change”, *Journal of Public Economics*, **25**: 13-33.

Bloom, Paul (2016), *Against Empathy: The Case for Rational Compassion.* Ecco/Harper Collins.

Blume, Lawrence, Adam Brandenburger, and Eddie Dekel (1991), “Lexicographic Probabilities and Choice under Uncertainty”, *Econometrica*, **59**: 61-79.

Boonin, David (2014), *The Non-Identity Problem and the Ethics of Future People*, Oxford University Press.

Bossert, Walter (1990), “Maximin Welfare Orderings with Variable Population Size”, *Social Choice and Welfare*, **7**: 39-45.

Bossert, Walter, Susumu Cato, and Kohei Kamaga (2023), “Thresholds, Critical Levels, and Generalized Sufficientarian Principles”, *Economic Theory*, **75**: 1099-1139.

Broome, John (2004), *Weighing Lives*, Oxford University Press.

Brickman, Philip, Donald T. Campbell (1971), "Hedonic Relativism and Planning the Good Society", in M. H. Appley (Ed.), *Adaptation Level Theory: A Symposium.* Academic Press.

Brickman, Philip, Dan Coates, and Ronnie Janoff-Bulman (1978), "Lottery Winners and Accident Victims: Is Happiness Relative?" *Journal of Personality and Social Psychology*, **36**: 917-927.

Carlson, Erik (1998), "Mere Addition and the Two Trilemmas of Population Ethics", *Economics & Philosophy*, **14**: 283-306.

de Finetti, Bruno (1931), "Sul Significato Soggettivo della Probabilitá ", *Fundamenta Mathematicae*, **17**: 298-329.

de Finetti, Bruno (1937), "La Prévision: Ses Lois Logiques, Ses Sources Subjectives", *Annales de l'Institut Henri Poincaré*, **7**: 1-68.

Dhillon, Amrita, and Jean-François Mertens (1999), "Relative utilitarianism", *Econometrica* , **67**(3), 471-498.

Diener, Ed (1984), "Subjective Well-being", *Psychological Bulletin,* **95**: 542-575.

Easterlin, Richard A. (1973), "Does money buy happiness?" *The Public Interest*, **30**: 3-10.

Easterlin, Richard A. (1974), "Does Economic Growth Improve the Human Lot? Some Empirical Evidence", in *Nations and Households in Economic Growth.* P. A. David and M. W. Reder, Academic Press: 89-125.

Easterlin, Richard A. (2017), "Paradox Lost?", *Review of Behavioral Economics*, **4**: 311-339.

Frankfurt, Harry G. (1987), "Equality as an Ideal", *Ethics*, **98**: 21-43.

Frick, Johann (2020), "Conditional Reasons and the Procreation Asymmetry", *Philosophical Perspectives*, **34**: 53-87.

Galaabaatar, Tsogbadral and Edi Karni (2013), "Subjective Expected Utility With Incomplete Preferences", *Econometrica*, **81**: 255-284.

Gilboa, Itzhak (2009), *Theory of Decision under Uncertainty.* Econometric Society Monograph Series Vol. No. 45. Cambridge: Cambridge University Press.

Gilboa, Itzhak and Larry Samuelson (2022), "What Were You Thinking? Revealed Preference Theory as Coherence Test", *Theoretical Economics*, **17**, 507-519.

Gustafsson, Johan E. (2022), "Our intuitive grasp of the repugnant conclusion." *The Oxford Handbook of Population Ethics*, 371-89.

Hammond, Peter J. (1976), "Equity, Arrow's Conditions, and Rawls' Difference Principle", *Econometrica,* **44**: 793-804.

Harsanyi, John C. (1955), "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, **63**: 309-321.

Helson, Harry (1947), "Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data", *American Journal of Psychology*, **60**: 1-29.

Helson, Harry (1948), "Adaptation-Level as a Basis for a Quantitative Theory of Frames of Reference", *Psychological Review*, **55**: 297-313.

Huemer, Michael (2008), "In Defence of Repugnance", *Mind*, **117**(468): 899-933.

Hume, David (1739-40), *A Treatise of Human Nature.* Edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Clarendon Press, 1978.

Hume, David (1751), *An Enquiry Concerning the Principles of Morals.* Edited by Tom L. Beauchamp. Oxford: Oxford University Press, 1998.

Hurka, Thomas (1983), "Value and Population Size", *Ethics*, **93**: 496-507.

Kahneman, Daniel, Alan B. Krueger, David A. Schkade, Norbert Schwarz, and Arthur A. Stone (2004), "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method", *Science*, **306**: 1776-1780.

Kauppinen, Antti (2013), "Meaning and Happiness", *Philosophical Topics*, **41**: 161-185.

Krantz, David H., R. Duncan Luce, Patrick Suppes, and Amos Tversky (1971), *Foundations of Measurement: vol. I.* Academic Press, New York.

Li, P. F. Jonah, Y. Joel Wong, and Ruth C.-L. Chao (2019), "Happiness and Meaning in Life: Unique, Differential, and Indirect Associations with Mental Health", *Counselling Psychology Quarterly*, **32**: 1-19.

Lucas, Richard E., Portia S. Dyrenforth, and Ed Diener (2008), "Four Myths about Subjective Well-Being", *Social and Personality Compass*, **2**: 2001-2015.

Mancini, Anthony D., George A. Bonanno, and Andrew E. Clark (2011), "Stepping Off the Hedonic Treadmill: Individual Differences in Response to Major Life Events", *Journal of Individual Differences*, **32**: 144-152.

Maslow, Abraham H. (1943), "A Theory of Human Motivation", *Psychological Review*, **50**: 370-396.

Maskin, Eric and Amartya Sen (2014), *The Arrow Impossibility Theorem.* Columbia University Press, New York.

Narveson, Jan (1973), "Moral Problems of Population", *The Monist*, **57**: 62-86.

Nebel, Jacob (2024), "Extensive Measurement in Social Choice", *Theoretical Economics*, **19**: 1581-1618.

Osborne, Martin J. (2025), *Models in Political Economy.* OpenBook Publishers, Cambridge, UK.

Parfit, Derek (1976), "On Doing the Best for Our Children", in: M. Bayles. ed.. *Ethics and Population* (Schenkman, Cambridge).

Parfit, Derek (1982), "Future generations: Further problems", *Philosophy and Public Affairs*, **11**: 113-172.

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

Parfit, Derek (1997), "Equality and Priority", *Ratio,* **10**: 202-221.

Persson, Ingmar (2004), "The Root of the Repugnant Conclusion and its Rebuttal", in J. Ryberg and T. Tannsjö (eds.) *The Repugnant Conclusion*, Library of Ethics and Applied Philosophy: 187-200.

Pinker, Steven (2012), *The Better Angels of Our Nature: Why Violence Has Declined*. Penguin Books, London.

Pivato, Marcus (2020), "Rank-additive population ethics", *Economic Theory,* **69**: 861-918.

Quine, Willard Van Orman (1962), *Methods of Logic*, London: Routledge & Kegan Paul.

Rachels, Stuart (2004), "Repugnance or Intransitivity: A Repugnant but Forced Choice", in J. Ryberg and T. Tannsjö (eds.) *The Repugnant Conclusion*, Library of Ethics and Applied Philosophy: 163-186.

Rawls, John A. (1971), *A Theory of Justice*. Harvard University Press.

Roberts, Melinda A. (2011), "The Asymmetry: A Solution ", *Theoria,* **77**: 333-367.

Savage, Leonard J. (1954), *The Foundations of Statistics*. New York: John Wiley and Sons. (Second addition in 1972, Dover)

Schwarz, Norbert and Gerald L. Clore (1983), "Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States", *Journal of Personality and Social Psychology*, **45**: 513-523.

Sikora, Richard I. (1975), "Utilitarianism: The Classical Principle and the Average Principle ". *Canadian Journal of Philosophy,* **5**: 409-419.

Smith, Adam (1759), *The Theory of Moral Sentiments*. Edited by D. D. Raphael and A. L. Macfie. Indianapolis: Liberty Fund, 1982.

Strack, Fritz, Leonard L. Martin, and Norbert Schwarz (1988), "Priming and Communication: Social Determinants of Information Use in Judgments of Life Satisfaction", *European Journal of Social Psychology*, **18**: 429-442.

Strawson, Sir Peter Frederick (1950), "On Referring ", *Mind*, **59**: 320-344.

Tännsjö, Torbjörn (2002), "Why We Ought to Accept the Repugnant Conclusion", *Utilitas*, **14**: 339-350.

Temkin, Larry S. (1987), "Intransitivity and the Mere Addition Paradox", *Philosophy and Public Affairs*, **16**: 138-187.

Temkin, Larry S. (2012), *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning.* Oxford University Press.

Veenhoven, Ruut and Floris Vergunst (2014), "The Easterlin Illusion: Economic Growth Does Go with Greater Happiness", *International Journal of Happiness and Development*, **1**: 311-343.

von Neumann, John and Oskar Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton, N.J.: Princeton University Press.

Zuber, Stéphane and Geir B. Asheim (2012), "Justifying Social Discounting: The Rank-Discounted Utilitarian Approach", *Journal of Economic Theory,* **147**: 1572-1601.