BOOK REVIEW

Matthew D. Adler: Well-being and fair distribution: beyond cost-benefit analysis

Oxford, xx, 635 p. ISBN: 978-0-19-538499-4

Alex Voorhoeve

1 Introduction

In this extraordinarily thorough and thoughtful work, Matthew Adler proves himself a grandmaster of the moral evaluation of public policy. The core idea of the book is that large-scale public policies should be designed to maximize the expectation of a continuous prioritarian social welfare function (SWF). In other words, they should maximize the expectation of the sum-total of transformed individual utilities, where the transformation gives greater weight to increases in utility the worse off an individual is in absolute terms. [Formally, a continuous prioritarian SWF holds that outcome x is at least as good as outcome y iff, for every utility function $u(.)$ in the set of utility functions $\mathbf{U}$ that represent individual well-being, $\sum_{i=1}^{N} g(u_i(x)) \geq \sum_{i=1}^{N} g(u_i(y))$, where $g(.)$ is a strictly increasing and strictly concave function. If $m(x)$ is a moral value function that evaluates outcomes in this manner, then, on Adler's view, public policy a is at least as good as policy b iff, for a given outcome set $\mathbf{O}$ and a probability function $\pi_a(.)$ linking any given policy a with any outcome, $\sum_{x \in \mathbf{O}} \pi_a(x) m(x) \geq \sum_{x \in \mathbf{O}} \pi_b(x) m(x)$.] The heart of this proposal is simple and close to ideas proposed by others, most notably Parfit (In: Clayton M, Williams A (eds) The ideal of equality, 2002; Utilitas 24:399–440, 2012). The genius of this work lies in the painstaking and creative way it is justified and operationalized. In what follows, I shall summarize the book and critically analyze a few key arguments.

A. Voorhoeve
University Center for Human Values, Princeton University,
Princeton, NJ, USA

A. Voorhoeve (✉)
Department of Philosophy, Logic and Scientific Method, LSE,
Houghton Street, London WC2A 2AE, UK
e-mail: a.e.voorhoeve@lse.ac.uk

2 Welfarism and its alternatives

Chapter 1 delineates the boundaries of, and lays the foundation for, the book's approach. Adler is concerned with fairness in the distribution of well-being within a given population of persons. He recognizes the importance of non-person-related moral considerations, such as the well-being of animals. However, he sets aside these considerations because he believes that while one can act *wrongly* towards such non-persons, one cannot treat them *unfairly* (p. 7). For most of the book, Adler also sets aside two considerations linked to fairness to persons. The first is personal responsibility, which finds its expression in the judgment that those who are badly off due to their free, informed choices have less of a claim on a given increment in well-being than those who are badly off due to no choice of theirs. Adler believes this consideration is important, but makes the simplifying assumption that no one is responsible for his or her level of well-being (in the final chapter, he discusses how to adjust his conclusions when people are responsible for their well-being). The second consideration is that fair treatment of persons requires respecting moral constraints on harming them. Adler outlines why he disagrees with such non-consequentialists (on pp. 24–29), but the book is not an argument against them. Rather, Adler sees his contribution as articulating a particular welfarist, act-consequentialist view. He rightly considers this project of interest even for those who do not share his starting point. After all, even those who believe the moral domain is not exhausted by a concern for good consequences for persons typically hold that these consequences matter.

Chapter 2 compares the SWF approach with familiar alternative forms of policy evaluation, including cost-benefit analysis and criteria based on cost per quality-adjusted life-year. It presents the key objections to these criteria, including that they may violate transitivity and Pareto-optimality.

The discussion of leading alternatives to the SWF approach in this chapter is exhaustive and the theoretical objections offered are persuasive. However, given that many of the objections voiced are well-known, one wonders whether the fact that these methods are still commonly used by governments is not due to lack of familiarity with these objections but rather to practical constraints. Consider, for example, the fixed cost-per-QALY criterion employed by the National Institute for Health and Clinical Excellence in England and Wales to determine whether an intervention is cost-effective (NICE 2008). On theoretical grounds, there is much to be said against such a criterion—it gives no priority to the worst off and, as Adler shows, can lead to Pareto-suboptimal decisions (pp. 148–154). However, it has a practical advantage over a prioritarian SWF approach: it demands less information. To compute it, one need gather only information on the average health-related quality of life improvements due to an intervention. By contrast, Adler's proposal requires information on the entire distribution of affected individuals' quality of life with and without the intervention. An approach based on a prioritarian SWF would therefore be harder to implement. Given that for many health interventions, we lack even information about the average health-related quality of life improvements they generate, this may be a significant obstacle to the implementation of Adler's approach.

3 A measure of welfare

Chapters 3 and 4 discuss, respectively, the theory and practice of interpersonal comparisons of utility. I shall focus my comments on the theoretical proposal in Chapter 3. A core building block, borrowed from Harsanyi, is that in addition to ordinary preferences (such as preferring a meal with friends to an evening of Bible study), individuals have extended preferences over life histories and lotteries of life histories. A life history is being an individual i, with i's beliefs and preferences, in a given outcome (p. 49). An individual k's extended preferences therefore include a comparison of <being individual i, with i's beliefs and preferences, in outcome x> to <being individual j, with j's beliefs and preferences, in outcome y>. (For example, Zelda expresses her extended preferences when she compares <being Arthur, with his beliefs and preferences, in a world in which Arthur is an Epicurean philosopher> to <being Bob, with his beliefs and preferences, when Bob is a Lutheran preacher> and judges the well-being of the latter to be greater.) Harsanyi of course derived a ranking of well-being levels and well-being differences from individuals' self-interested preferences over such life history lotteries when placed behind a veil of ignorance. Adler departs in several ways from Harsanyi, of which I shall discuss two.

3.1 Against using the veil of ignorance for interpersonal comparisons

Adler argues that it is not possible for an individual to do what Harsanyi asks her to do—to evaluate a lottery in which she might be various different persons from a self-interested perspective. According to Adler, Zelda cannot evaluate the prospect of her being Arthur, because taking on Arthur's essential personal characteristics, including his body, beliefs, memories and desires, would, on every leading conception of personal identity, conflict with requirements of personal identity (p. 198). Such a prospect would involve her extinction, so she cannot be asked to evaluate how good it would be for *her* to lead Arthur's life.

Adler instead proposes the following way of making comparisons. Zelda evaluates outcomes and lotteries for herself from the perspective of her rational and informed self-interest. She evaluates outcomes and lotteries for Arthur as his rational and informed guardian would. Finally, Zelda makes interpersonal comparisons directly: she simply ranks a life history of Arthur's against another life history of Bob's in terms of well-being.

Adler's argument against the very possibility of making self-interested evaluations of how good it would be to be different individuals from behind a veil strikes me as novel and important. I do not think it is sound as it stands, but I believe that something close to it may well be.

Contrary to Adler, I do not think the veil of ignorance thought experiment is best understood as asking Zelda to imagine that she will *be* a different person than she *is*. Instead, it is best understood as asking Zelda to imagine the following. An event has just temporarily robbed her of all knowledge of who she is and of all of her ordinary preferences. She retains only her extended preferences. She will soon find out that she is (and has been all along) Arthur or Bob, with Arthur or Bob's preferences, etc. In

this moment of extreme amnesia, she must ask herself who among them she would rather learn that she is.[1]

This thought experiment is consistent with at least one prominent conception of personal identity, on which continuity of body and brain suffices (see Williams 1973). For while Zelda is asked to imagine that she may be any of a number of people with their bodies and brains, she is not asked to imagine that, behind the veil, she occupies a different body and brain than she will occupy when the veil is lifted. In other words, she is not asked to imagine that her essential characteristics (on this view of personal identity) will change; she is merely asked to imagine that she is ignorant of these characteristics.

However, on at least one other conception of personal identity, problems arise. Lewis (1983), for example, argues that "psychological connectedness," or a sufficient degree of similarity in memories, beliefs, values, and aims, is a necessary condition for personal identity. Note, however, that there may be a complete lack of such connectedness between the person behind the veil and the people she must imagine she might be. Behind the veil, Zelda has no memories and no ordinary preferences, so in this respect she is very dissimilar to each of the persons with whom she is asked to identify. She does, of course, keep her extended preferences, since without them she would have no basis on which to evaluate life histories. These extended preferences will depend on her view of the good life (e.g. they will be very different if she is an Epicurean than if she is a Lutheran). And since views of the good life differ thoroughly, the people whose lives she is evaluating may have radically different extended preferences than she has behind the veil. This means that the degree of connectedness between the person behind the veil and the people she is meant to imagine she might be may be too low to satisfy Lewis' criterion of personal identity. On Lewis' view, the person behind the veil may therefore be unable to imagine herself to be identical to the persons with whom she is asked to identify.[2]

Even if this lack of psychological connectedness does not violate the conditions of personal identity, there is still a problem. Parfit (1987), for example, argues that while psychological connectedness is not *required* for personal identity, it *is* "what matters in survival." If this is correct, and if there is little connection between the person behind the veil and the people she must imagine she might be, then this veiled person may have, at best, only a much attenuated self-concern for these people.

In sum, it seems that if psychological connectedness matters to personal identity or to self-concern, then Adler's conclusion stands. Self-interested veil of ignorance reasoning cannot provide a good interpersonally comparable measure of utility.

---

[1] I am grateful to Michael Otsuka for discussion of Adler's objection to veil of ignorance reasoning.

[2] Yet another leading conception of personal identity requires psychological continuity, defined as an overlapping chain of strong connectedness (see Parfit 1987). Since psychological connectedness between Zelda behind the veil and Arthur and Bob is low, the required identification may be problematic on this view also.

3.2 Differences in extended preferences

Adler's second point of departure from Harsanyi is to drop the latter's assumption that everyone's rational extended preferences are identical. Adler holds that Harsanyi's assumption is indefensible, because individuals' extended preferences reflect their views of the good life, which will differ. Of course, this means that there may be as many different rankings of life histories as there are individuals—indeed, given that individuals' preferences might change, there may be even more. Adler's response to this variability is to propose that the social planner evaluate life histories as follows. Life history <*i*; *x*> is better than (as good as) life history <*j*; *y*> just in case the former is better (equally good) on *every* person's actual or possible rational, informed, extended preferences (p. 226). Otherwise, the two life histories are incomparable. He believes this proposal has an "appealing liberal flavor," because it draws on convergent preferences to rank life histories.

Adler is surely right that individuals' extended preferences differ. But his proposal has an obvious drawback. An Epicurean will rank life histories by the pleasure they yield; someone who takes seriously Luther's dictum that the point of life is "suffering and the cross"[3] will have a very different ranking. And so on. Given the diversity of views of the good life, Adler's approach may well result in incomparability on a scale that threatens the project of ranking policies. It therefore seems to me that we should explore liberal approaches that allow for full comparability. One such approach is to construct a measure that is publicly *used* as a standard of well-being, but not *endorsed* as the correct view of the good life. To be liberal, such a measure should also be as accommodating as possible to different views of the good life. An example in a limited domain is the preference-based measure of utility employed in health economics, in which the value 0 is assigned to death at a young age, the value 1 to a 90-year life in good health, and the value of health states in between is set equal to the probability p that would render a person indifferent (if she were exclusively self-interested and rational) between a life of 90 years in that state and a lottery with chance $(1-p)$ of death at a young age and p of 90 years in good health. Such a measure mixes "objective" elements (assigning the values of 0 and 1 to the same objective conditions for each person) with "subjective" elements (allowing individuals to assign personalized values to all other conditions). On this measure, anhedonia might be judged a terrible condition for an Epicurean and a minor harm to a Lutheran. But the use of this measure by a public authority would not involve the assertion by that authority that it really is better to be a Lutheran than an Epicurean with this condition.

4 Prioritarianism

Chapters 5, 6, and 7 justify the prioritarian evaluation of outcomes and policies. Chapter 8 remarks on the extension of the analysis to variable population cases and cases in which individuals bear some responsibility for their well-being.

---

[3] Quoted in Williams (1972, p. 76).

Chapter 5 makes the case for a continuous prioritarian SWF and in particular for the Atkinson SWF, which says that outcome x is at least as good as y iff $(1-\gamma)^{-1} \sum_{i=1}^{N} u_i(x)^{1-\gamma} \geq (1-\gamma)^{-1} \sum_{i=1}^{N} u_i(y)^{1-\gamma}$, for all $u(.)$ in the relevant set **U** of utility functions. (The parameter γ is an inequality-aversion parameter that takes some value larger than zero.)

Adler's defense is axiomatic: he carefully outlines the case for the Pigou–Dalton axiom, an axiom of separability, a continuity axiom, and ratio-rescaling invariance and shows that they together yield the Atkinson SWF. (Ratio-rescaling invariance allows for a fixed set of well-being ratios between life histories. Adler can employ it because he introduces a zero for all utility functions at the point at which the evaluator would be indifferent, on behalf of the person whose life history is being evaluated, between that person living the life history in question and that person's non-existence.)

The philosophical heart of the chapter is its articulation of a claim-based notion of fairness that draws on Thomas Nagel's work (see Nagel 1979). Adler argues that in justifying a distribution, we must consider what can be said in favor of it from each individual's perspective. In a choice between two outcomes x and y, he holds that this leads us to the following ideas:

(i) an individual has a claim in favor of x iff it makes him better off than y;
(ii) his claim is stronger the larger the increment of well-being that x offers him over y;
(iii) his claim is stronger the lower the absolute level of utility from which the increment occurs;
(iv) x is fairer than y iff the sum-total of strength-weighted claims in favor of x is greater than the sum-total of these claims in favor of y.

Element (i) distinguishes Adler's view from the kind of egalitarianism defended by Temkin (1996). The latter speaks of claims of unfairness on behalf of individuals who are worse off than others, even if nothing can be done to improve their lot. Adler dismisses this view on grounds of its purported conflict with Pareto optimality (pp. 335–337). Adler knows, of course, that all sensible egalitarians care about improving well-being as well as about reducing inequality. But he claims that such pluralist egalitarianism will conflict with Pareto-optimality, because it would be implausible for egalitarians to hold that the value of a Pareto-superior improvement in well-being for the best off always trumps the disvalue of the resulting increase in inequality.

Element (iii) distinguishes Adler's view from utilitarianism. Adler justifies it by an appeal to the "separateness of persons." Respect for this separateness, he argues, requires that we care about the distribution of well-being among individuals (pp. 316–317). He illustrates this with a case in which individuals' interests conflict. Suppose we must choose between x, which yields 8 utils for Arthur and 3 for Bob, and y, which yields 9 for Arthur and 2 for Bob. For both individuals, 1 util is at stake in this choice. Surely, Adler writes, Bob's claim to this util is stronger than Arthur's just because he is worse off.

Chapter 6 discusses the question whether the object of such prioritarian concern should be low lifetime utility or low utility in some part of a life. In a nutshell, Adler's

main argument is this. The unity of an individual across time allows for intrapersonal compensation of losses in one period by gains in another. For example, for a given individual, a loss in an early period may be compensated by a slightly greater utility gain in a later period, even if the individual is worse off in the earlier period than in the later period. A combination of such an earlier loss and a later greater gain would therefore be good for an individual and should therefore be judged morally good. But if we applied prioritarian weighting to sub-lifetime utilities, then we might judge this combination to be morally bad. So lifetime utility, rather than sub-lifetime utility, should be the object of prioritarian concern.

Chapter 7 examines whether prioritarians should give extra weight to gains in utility for individuals with low *expected* utility, or to those with low utility *outcomes*. The former respects the ex ante Pareto principle in all cases, whereas the latter does not. Adler argues that the former view ought to be rejected along with the ex ante Pareto principle because it may lead to a choice of an alternative with worse outcomes in every state of the world. Consider, for example, Case 1, in which ε is a small amount of utility.

Case 1. Utilities of two alternatives

Alternative	Person	State of the world (equiprobable)	
		s_1	s_2
a	Arthur	0.8	0.8
	Bob	0.8	0.8
b	Arthur	1	$0.6 + \varepsilon$
	Bob	$0.6 + \varepsilon$	1

In this choice, b is ex ante Pareto-superior to a. However, for a small enough ε, a prioritarian will regard the gain in utility from $0.6+\varepsilon$ to 0.8 as morally more important than the gain in utility from 0.8 to 1. She will therefore regard a as offering a better outcome than b no matter whether the state of the world is s_1 or s_2. She can therefore infer that if she were to learn the true state of the world, she would invariably prefer a. Adler (p. 510) believes it is a basic principle of rationality that when she knows that she would prefer a with full information, she should decide in line with this preference.

One might wonder how this conclusion meshes with Adler's derivation of his preferred form of prioritarianism from consideration of what can be said in favor of a distribution or policy from each person's perspective. After all, ex ante, both Arthur and Bob would prefer policy b. However, Adler can respond that this ex ante unanimity is merely the result of lack of information. There is a conflict of interests in final well-being between Arthur and Bob, a conflict that would be expressed if they had full information about how the policies would affect them. It is therefore plausible for a decision-maker to regard this as an interpersonal tradeoff case in which she must decide between one person's improvement from 0.8 to 1 and another's improvement from $0.6+\varepsilon$ to 0.8. And, for a small ε, a prioritarian decision-maker should resolve this conflict by choosing the latter improvement.

I am sympathetic to many of Adler's points in these chapters. However, I believe it is doubtful whether Adler's favored principle follows from his starting points.

To begin with, consider Case 2, in which $0 < \lambda < 0.2$, so that only c is ex ante Pareto-optimal.

Case 2. Utilities of two alternatives

Alternative	Person	State of the world (equiprobable)	
		s_1	s_2
a	Arthur	0.8	0.8
	Bob	0.8	0.8
c	Arthur	1	$0.8 - \lambda$
	Bob	0.8	0.8

Only Arthur's well-being is affected by this choice. On Adler's view about individual claims, it follows that only Arthur has a claim. And given that c is a better prospect for Arthur, it would seem natural to hold that Arthur has a stronger claim to c than to a. After all, even if things go badly for him because s_2 is the state of the world, the decision-maker can say that she chose c for his sake. Claims-based reasoning, which focuses on what can be said in favor of a policy on behalf of each individual, therefore naturally favors c. However, for some λ in the specified range, Adler's prioritarianism would choose a.

Adler might attempt a defense of this choice by noting that he rejected the appeal to ex ante Pareto in Case 1. However, the reasons for rejecting ex ante Pareto in Case 1 do not apply in Case 2. First, in Case 1, the decision-maker can infer that if she knew the state of the world, she would invariably favor the ex ante Pareto-sub-optimal a. But she can make no such inference in Case 2: if s_1 is the state of the world, then c is better, while if s_2 is the state of the world, then a is better. Second, in Case 1, there is a conflict of individuals' interests in final well-being, which one ought arguably to resolve in favor of the worst off. By contrast, in Case 2, there is no such conflict.

The conclusion I draw from these cases is that for a view that takes individuals' claims as its starting point, and that eschews an appeal to the unfairness of inequality, ex ante Pareto is an attractive principle when there are no interpersonal conflicts of interest, either ex ante or ex post. There is therefore a tension between Adler's starting point and his favored form of prioritarianism, which violates ex ante Pareto in such no-conflict cases.[4]

A contrast between the following Case 3 and Case 2 offers a further problem for Adler (Voorhoeve and Fleurbaey 2012).

---

[4] Adler (pp. 518–524) remarks that he finds ex ante Pareto attractive in cases in which there is no outcome inequality because all individuals' outcomes are perfectly correlated. He therefore expresses sympathy for Fleurbaey (2010) proposal to evaluate an outcome by its "equally-distributed equivalent"—the level of utility that, if equally distributed, would be as valuable as that outcome. Fleurbaey's SWF respects ex ante Pareto in cases in which there is no outcome inequality, but violates ex ante Pareto in Case 2, in which there is outcome inequality. The difficulty for Adler is that since he does not regard outcome inequality as in itself unfair, he cannot countenance this distinction between these cases. It would therefore be inconsistent of him to adopt Fleurbaey's proposed SWF.

Case 3. Utilities of two alternatives

Alternative	Person	State of the world (equiprobable)	
		s_1	s_2
a	Arthur	0.8	0.8
	Bob	0.8	0.8
d	Arthur	0.8	$0.8 - \lambda$
	Bob	1	0.8

The anonymized outcome in each state of the world in Case 3 is identical to Case 2. On Adler's prioritarianism, it is therefore true that if λ is such that c is preferable to a, then d is preferable to a; and if λ is such that a is preferable to c, then a is preferable to d. Adler's view therefore fails to respond to the fact that Case 2 and Case 3 differ in morally relevant respects. Case 2 involves only an intrapersonal tradeoff. For Arthur, the prospect of the gain from 0.8 to 1 is the desirable flipside of his exposure to the risk of a smaller loss from 0.8 to $0.8 - \lambda$ (recall that $0 < \lambda < 0.2$). By contrast, Case 3 involves only an interpersonal tradeoff. If the decision-maker chooses d, Bob will face just a prospect of a gain and Arthur will face just a risk of a loss. Moreover, Arthur will be worse off than Bob no matter what happens.

These differences make a clear difference to the justifiability of choosing the risky option. In Case 2, the decision-maker can justify c to Arthur on the grounds that in taking a risk, she was doing the best she could for him, given the knowledge she had at the time. By contrast, in Case 3, she must justify exposing Arthur to a risk of a loss, thereby making him worse off than Bob, by appealing to the equal chance of a greater gain accruing to Bob. For a given, equally likely, loss and greater gain, exposing a person to the risk of that loss in order to give that very same person the chance at the gain is easier to justify than exposing a person to the risk of that loss in order to give someone *else* (who will be better off no matter what happens) the chance at the gain. There will therefore be some λ in the specified range for which c is permissible in Case 2, but for which d is impermissible in Case 3. This difference in the justifiability of the risky option is a consequence of the difference between the unity of the individual (which gives us reason to make pure intrapersonal tradeoffs for a person's sake) and the separateness of persons (which requires that in interpersonal tradeoffs, we give greater weight to the interests of the worse off). Adler's starting points—respect for the unity of the individual and the separateness of persons—therefore naturally lead to conclusions that are at odds with his favored form of prioritarianism.

5 Conclusion

Well-Being and Fair Distribution displays a breathtaking knowledge of philosophy, law, decision theory, and social choice. On each of the topics it discusses, it represents competing views accurately and in detail and assesses them judiciously. Individual chapters will therefore be valuable as critical surveys for researchers. They also ought to figure on reading lists for graduate courses on policy evaluation, well-being, and distributive justice. In addition, the book contains novel arguments and proposals,

especially on the measure of well-being and the philosophical foundations of priori-tarianism. Though I have argued that some of Adler's arguments don't succeed, I also believe that many do. In all cases, I have learned a great deal from grappling with them. I predict that engagement with the book will prove similarly rewarding for anyone in the field.[5]

References

Fleurbaey M (2010) Assessing risky social situations. J Political Econ 118:649–680
Lewis D (1983) Survival and identity. In: Philosophical papers, vol I. Oxford University Press, Oxford, pp 55–77
Nagel T (1979) Equality. In: Mortal questions. Cambridge University Press, Cambridge, pp 106–27
NICE (2008) Social value judgments, 2nd edn. NICE, London
Parfit D (1987) Reasons and persons. Clarendon Press, Oxford
Parfit D (2002) Equality or priority. In: Clayton M, Williams A (eds) The ideal of equality. Palgrave, Basingstoke, pp 81–125
Parfit D (2012) Another defence of the priority view. Utilitas 24:399–440
Temkin L (1996) Equality. Oxford University Press, Oxford
Voorhoeve A, Fleurbaey M (2012) Egalitarianism and the separateness of persons. Utilitas 24:381–398
Williams B (1972) Morality. Cambridge University Press, Cambridge
Williams B (1973) The self and the future. In: Williams B (ed) Problems of the self. Cambridge University Press, Cambridge, pp 46–63