1 Introduction
Prioritarianism in Practice

MATTHEW D. ADLER AND
OLE F. NORHEIM*

1.1 Prioritarianism

"Prioritarianism" is a framework for ethical assessment that gives extra weight to the well-being of the worse off. Unlike utilitarianism, which uses a simple summation of well-being, prioritarianism adds up *transformed* well-being numbers: well-being numbers inputted into a concave transformation function.[1] Assume that Andre is worse off than Beatrice. We can increase Andre's well-being by an increment Δw, or increase Beatrice's well-being by the same increment Δw. Utilitarianism is indifferent between these two options. Prioritarianism prefers the first: giving Δw to the worse-off one (Andre) yields a larger ethical improvement than giving Δw to the better-off one (Beatrice).

* *Acknowledgments*. Adler thanks David Levi and Kerry Abrams who, as successive deans of Duke Law School, provided generous financial support for conferences held at Duke Law in connection with this volume. He thanks Carol Abken, Leanna Doty, and Victoria Zellefrow for providing administrative support for these conferences and all other aspects of the "Prioritarianism in Practice" Research Network, and Wick Shreve of the Duke Law Library for research advice and assistance. Norheim thanks the Bill & Melinda Gates Foundation (INV-010174) and the Trond Mohn Foundation (BFS2019TMT02) for financial support to organize the Bergen conference in 2019.

A number of scholars who are not volume authors participated in *Prioritarianism in Practice* conferences, and the volume benefitted from their insights. We thank Mathias Barra, Alexander Cappelen, Jacob Goldin, Ori Heffetz, Anders Herlitz, Miles Kimball, Jeremy Lauer, Paolo Piacquadio, Carl Tollef Solberg, and Bertil Tungodden. Finally, we are grateful to our editor Philip Good, and to everyone else who worked on *Prioritarianism in Practice* at Cambridge University Press and Straive – including Denise Bannerman, Jessica Norman, Sindhuja Sethuraman, and Erika Walsh – for bringing this volume to publication.

[1] To be more precise: the prioritarian transformation function is strictly increasing and strictly concave. Throughout the chapter, we use "concave" in reference to this transformation function as shorthand for strictly increasing and strictly concave.

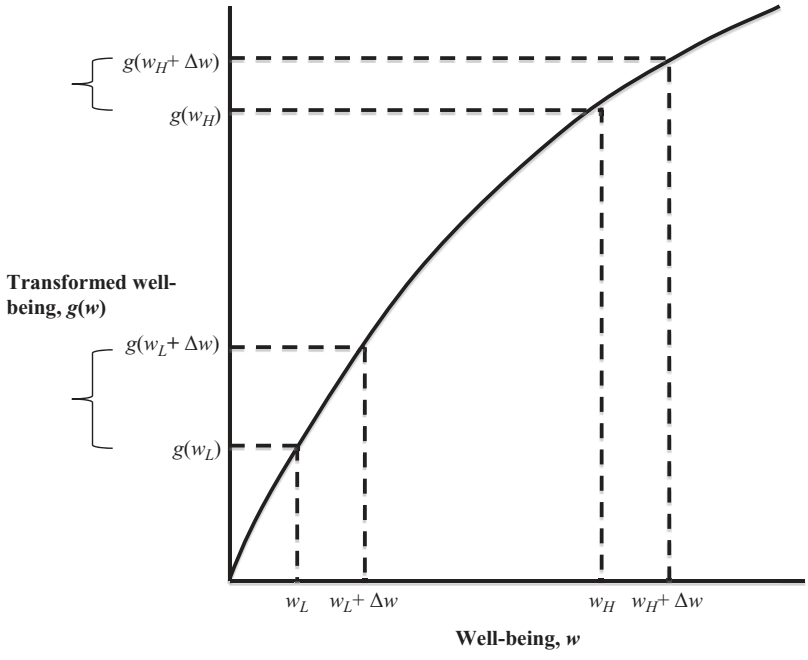

Figure 1.1 Well-being numbers inputted into a concave transformation function

Note: Andre is at lower well-being level w_L and Beatrice at higher well-being w_H. The increase in well-being from w_L to $w_L + \Delta w$ produces a bigger increase in transformed well-being than the increase in well-being from w_H to $w_H + \Delta w$.

(*Source*: Figure 1.1 in Matthew D. Adler, *Measuring Social Welfare* (2019), reproduced with permission of Oxford University Press)

See Figure 1.1, illustrating that the prioritarian rule of summing concavely transformed well-being has the effect of giving priority to the worse off.

Utilitarianism has had sweeping influence in academic scholarship (in philosophy, economics, law, public health, and other fields), and in governmental practice, for centuries. (For historical sources on utilitarianism, see Eggleston and Miller [2014]. For contemporary defenses, see, e.g., Brandt [1979], Broome [1991], Goodin [1995], Harsanyi [1977], Singer [2011].) Prioritarianism is a much newer idea. In ethics, prioritarianism is most closely associated with the philosopher Derek

Parfit. Parfit's 1991 Lindley Lecture triggered an intensive philosophical examination of prioritarianism, continuing to the present[2] (Parfit 2000, publishing his 1991 lecture). Although Parfit was not in fact the first philosopher to discuss prioritarianism, it *is* certainly true that wide philosophical investigation and debate about prioritarianism only began with Parfit's work. In short: this is a relatively new idea in ethics, much newer than utilitarianism.

Although the term "prioritarianism" is a piece of philosophical vocabulary that is uncommon in economics, the underlying idea – summing a concave transformation of well-being, and thereby according extra weight to the worse off – has figured in welfare economics since the 1970s. At that time, the concept of the "social welfare function" (SWF) began to play a more central role in several economic literatures: theoretical welfare economics, optimal tax theory, and the measurement of inequality. A variety of types of SWFs have been examined by these literatures, including prioritarian SWFs.[3]

The SWF methodology, in a nutshell, converts each possible outcome of governmental policy choice into a list ("vector") of well-being numbers – measuring the well-being of the individuals in the population of ethical concern. A given outcome is some possible combination of welfare-relevant characteristics (income, health, leisure, happiness, etc.) for each person in the population. Let x be a possible outcome, y a different possible outcome, and so forth. If there are N individuals in the population, then x corresponds to the well-being vector $(w_1(x), w_2(x), \ldots w_N(x))$ – with $w_1(x)$ denoting the well-being number of individual 1 in x, $w_2(x)$ the well-being number of individual 2 in x, and so forth. Similarly, outcome y corresponds to the well-being vector $(w_1(y), w_2(y), \ldots w_N(y))$. The SWF is a rule for ranking well-being vectors and, thereby, the corresponding outcomes. Many different such rules are possible, including the utilitarian SWF, prioritarian SWFs, and others.

In short, the concept of prioritarianism, if not the term "prioritarianism," has figured in various SWF-related economic literatures for

[2] The philosophical literature on prioritarianism is discussed and cited below, in Section 1.2.1.

[3] The literature on SWFs is discussed and cited below, in Section 1.2.2, and in Chapter 2.

the last 50 years. Still, in economics as in ethics, this is a much newer idea than utilitarianism. Further, although a variety of economic concepts and tools have had a large impact on governmental practice, prioritarianism has not (yet) had this influence.

This chapter is the introductory chapter in a volume entitled *Prioritarianism in Practice*. The aim of the volume is to study and showcase the use of prioritarianism as a methodology for evaluating governmental policy, across a variety of policy domains: taxation, health policy, risk regulation, education, climate policy, and the COVID-19 pandemic (a global catastrophe, urgently requiring wise governmental policy, that occurred while the volume was being drafted). The volume also examines prioritarianism as an indicator of social condition (as an alternative to GDP, currently the dominant social-condition metric).

Prioritarianism in Practice is the first book to study the *application* of prioritarianism as a policy-assessment framework. It grows out of an international research network which we founded. The participants in this network are now chapter authors in this volume.

Why study prioritarianism? After all, there are *many* competitors to utilitarianism, including but hardly limited to prioritarianism.

The answer is that prioritarianism retains most of the attractive characteristics of utilitarianism, but drops one feature that many have found problematic: utilitarianism's indifference to the distribution of well-being.[4]

First, prioritarianism, like utilitarianism, is a species of welfare consequentialism. An ethical view is *consequentialist* if it evaluates choices in light of the possible outcomes of those choices. Outcomes are ranked from best to worst; and it is this goodness ranking of outcomes that drives ethical assessment. An ethical view is, more specifically, *welfare* consequentialist if the goodness ranking of outcomes hinges on individuals' well-being. The SWF methodology, in turn, is just a formal procedure for implementing welfare consequentialism.

For those who endorse welfare consequentialism as the appropriate architecture for ethical assessment and, specifically, the assessment of

[4] The normative case for prioritarianism, summarized in the next several paragraphs, is presented in much greater detail in Chapter 2.

governmental policies, the fact that prioritarianism – like utilitarianism – has this structure is a welcome feature of prioritarianism.[5]

Further, both the utilitarian SWF and prioritarian SWFs are "score-based"; in each case, the SWF operates by assigning a numerical score to a given vector and then ranking vectors in the order of these scores.[6] (The utilitarian score is just the sum of well-being; the prioritarian score is the sum of concavely transformed well-being.) By contrast, for example, the "leximin" SWF – roughly corresponding to John Rawls' notion of "maximin" (Rawls 1999) – employs an algorithm for ranking well-being vectors that cannot be summarized by a numerical score.[7]

Score-based SWFs are especially tractable – a virtue for policy assessment – and can readily be generalized to policy choice under uncertainty. If we conceptualize a governmental policy as a probability distribution across outcomes, then the utilitarian SWF can be used to rank policies by assigning each policy its expected score (namely, the expected sum of well-being). The prioritarian SWF can be applied under uncertainty in an analogous way (now, policies are ranked according to the expected sum of concavely transformed well-being).[8]

Prioritarianism shares yet another desirable feature with utilitarianism: it is *separable*. If a group of individuals is unaffected by policy choice – whichever policy might be chosen, the well-being of individuals within the group doesn't change – then utilitarian assessment can

[5] One important critique of welfare consequentialism is that it fails to take account of individual responsibility. How to refine prioritarianism so as to incorporate individual responsibility is the topic of Chapter 11. A different critique is that welfare consequentialism ignores deontological constraints and other non-consequentialist factors. For those who believe that ethics is a hybrid of consequentialism and non-consequentialist factors, prioritarian policy analysis can still be seen as capturing one *part* of ethical assessment, namely the consequentialist component (Adler 2019, pp. 24–27).

[6] We use the term "SWF" to mean what is, strictly, a social welfare ordering: a rule for ranking well-being vectors. Some (not all) such rules can be represented by a real-valued function $S(\cdot)$, such that: vector $\mathbf{w}$ is at least as good as vector $\mathbf{v}$ if and only if $S(\mathbf{w}) \geq S(\mathbf{v})$. In our terminology, these SWFs are "score-based."

[7] Leximin orders two well-being vectors according to the levels of the worst-off individuals in each; if the worst-off individuals are equally well off in the two vectors, according to the levels of the second-worst-off individuals; and so forth.

[8] This formula (the expected sum of concavely transformed well-being) is one methodology for applying prioritarianism under uncertainty, so-called "ex post prioritarianism" (EPP). Chapter 2, Section 2.7, discusses both EPP and other possible formulas.

simply ignore the individuals.[9] The utilitarian comparison of policies is wholly determined by the well-being of *affected* individuals. The same is true of prioritarianism. Separability, like the feature of being score-based, increases the tractability of both utilitarianism and prioritarianism.

The critical *difference* between the two approaches is that prioritarianism embodies a concern for the distribution of well-being that is absent from utilitarianism. The utilitarian SWF is such that: if total well-being is greater in outcome y than outcome x, y is better than x regardless of how well-being is distributed in the two outcomes. In particular, imagine that Cedric is better off than Dahlia in x. Cedric's well-being in y increases by Δw^*, while Dahlia's decreases by Δw. As long as the increment to Cedric's well-being is larger than the loss to Dahlia (that is, as long as $\Delta w^* > \Delta w$), then utilitarianism ranks y better than x *regardless* of the well-being levels of the two individuals, and regardless of the magnitudes of Δw^* and Δw (even if Cedric's gain is only slightly more than Dahlia's loss). This is problematic. By contrast, prioritarianism may well prefer x; whether it does depends upon the well-being levels of the two individuals, the comparative magnitudes of Cedric's gain and Dahlia's loss, and the concave transformation function.

Further, prioritarianism is quite flexible regarding the degree of ethical priority to the well-being of the worse off. Here, a contrast not only with utilitarianism, but also leximin, is instructive. The utilitarian SWF is a specific such rule (a specific rule for ranking well-being vectors), as is the leximin SWF. The utilitarian SWF gives zero priority to the worse off; the leximin SWF gives absolute priority to the worse off. Prioritarianism is an entire *family* of SWFs; the choice of concave transformation function determines a particular such rule. A more concave function means that, as between two individuals at two given well-being levels, the worse-off one has a greater degree of priority. The prioritarian, by means of her choice of transformation function, can specify whatever degree of priority for the worse off she judges

[9] To be more precise, the utilitarian SWF and prioritarian SWFs both satisfy an axiom of Separability with respect to the ranking of well-being vectors. Further, they both can be applied under uncertainty in a manner that satisfies an analogous axiom with respect to policies. Only SWFs that satisfy Separability with respect to the vector ranking can be separable with respect to policies. See Chapter 2, Section 2.4.

reasonable – approaching utilitarianism at one extreme and the absolute priority of leximin at the other.

In the remainder of this introductory chapter to *Prioritarianism in Practice*, we do the following. First, we briefly survey existing scholarship on prioritarianism in philosophy, economics, and public health (the main academic literatures in which prioritarianism has, to date, played a role), so as to give the reader a sense of the intellectual backdrop for the volume. Second, we provide an overview of the volume.

1.2 Scholarship on Prioritarianism: A Brief Survey

1.2.1 Philosophy

As mentioned, it was Derek Parfit's 1991 Lindley Lecture that triggered widespread philosophical attention to prioritarianism[10] (Parfit 2000, publishing his 1991 lecture). There is now a substantial body of work in academic philosophy on the topic.

In the lecture, Parfit introduces prioritarianism – contrasting it with egalitarianism – with reference to a hypothetical case previously described by the philosopher Thomas Nagel in his article, "Equality" (Nagel 1979). Nagel imagines that he has two children, the first who is healthy and the second who has a serious health condition and is worse off than the first; and that he faces a choice between moving to a city or a suburb. In the city, the second child would have access to treatment for her condition, and so she would be somewhat better off than in the suburb. In the suburb, the first child would flourish, so much so that the difference for him between moving to the suburb and moving to the city is larger than the benefit the second child would reap from moving to the city.

Parfit uses well-being numbers to present Nagel's case. The first child would be at well-being level 20 in the city, and would gain 5 units moving to the suburb. The second child, who would be worse off than the first regardless of where the family lives, would be at level 9 in the suburb and 10 in the city. See Table 1.1 (from Parfit [2000, p. 83]).

[10] There is some philosophical work on the concept of prioritarianism (not using that term) that slightly predates Parfit. See McKerlie (1984); Temkin (1983); Weirich (1983).

Table 1.1. ***Nagel's two-child case, as presented by Parfit***

	First child	Second child
Move to the city	20	10
Move to the suburb	25	9

Utilitarianism recommends moving to the suburb: the sum total of the children's well-being is greater in the suburb (34) than the city (30). Nagel, discussing the case, writes that the decision to move instead to the city would be an "egalitarian" decision (Nagel 1979, p. 124). Parfit suggests that the rationale for moving to the city *might* be egalitarian, but it might instead be prioritarian. The aim of the Lindley Lecture is to differentiate these two distinct species of non-utilitarian reasoning.

Parfit observes that the most plausible version of egalitarianism is *pluralist*. Egalitarians endorse "The Principle of Equality": "It is in itself bad if some people are worse off than others" (Parfit 2000, p. 84). But they also should endorse "The Principle of Utility": "It is in itself better if people are better off" (ibid.). An egalitarian who endorsed only the first principle, not the second, would see no difference between everyone being equally well off at a low level of well-being, and everyone being equally well off at a high level of well-being.

Within the space of pluralist egalitarianism, Parfit differentiates between "strong" and "moderate" egalitarians. Strong pluralist egalitarians believe that the "Principle of Equality" may outweigh the "Principle of Utility" even in cases where the more equal outcome is worse for some and better for none. As between an outcome in which everyone is at well-being level w, and a second outcome in which some are at w and others are better off than w, the strong pluralist egalitarian *might* choose the first outcome. By contrast, moderate pluralist egalitarians accept that, in such a case, the two principles always balance against each other so as to favor the second outcome. In the language of welfare economics, moderate pluralist egalitarians accept the Pareto axiom: if some are better off in outcome x than y, and none are worse off, then x is better than y.[11]

[11] "Pareto" is used in Chapter 2 to mean the combination of two axioms: Strong Pareto (if some are better off in one outcome as compared to a second, and none are worse off, then the first outcome is better), and Pareto Indifference (if each

In the two-child case, the Pareto axiom does not come into play: the first child is better off in the suburb, the second in the city. A moderate pluralist egalitarian who gives little weight to equality will recommend the suburb; if she gives more weight to equality, she will recommend the city.

A distinct rationale for moving to the city, Parfit observes, is prioritarianism. Actually, Parfit doesn't use the term "prioritarianism." Instead he speaks of "the Priority View" or "prioritarians." Philosophical nomenclature thereafter shifted slightly, with "prioritarianism" becoming the standard name for what the Lindley Lecture denotes the "Priority View."

Parfit characterizes prioritarianism ("the Priority View") as follows: "Benefitting people matters more the worse off these people are." He continues:

> For Utilitarians, the moral importance of each benefit depends only on how great this benefit would be. For *Prioritarians*, it also depends on how well off the person is to whom this benefit comes. We should not give equal weight to equal benefits, whoever receives them. Benefits to the worse off should be given more weight. (Ibid., p. 101)

Depending on how it assigns moral weights to well-being gains at various levels, prioritarianism may recommend moving to the city in the two-child case. It may conclude that the moral impact of a well-being improvement for the second child from level 9 to level 10 is greater than the moral impact of a well-being improvement for the first child from level 20 to level 25.

The crucial difference between egalitarianism and prioritarianism, according to Parfit, is that egalitarians care about "relativities" while prioritarians do not.

> [O]n the Priority View, we do not believe in equality. We do not think it in itself bad, or unjust, that some people are worse off than others. This claim can be misunderstood. We do of course think it bad that some people are

person is just as well off in one outcome as a second, then the two outcomes are equally good). For purposes of Parfit's discussion of egalitarianism and prioritarianism, the relevant part of the Pareto axiom combination is Strong Pareto: moderate pluralist egalitarians and prioritarians accept, while strong pluralist egalitarians reject, the Strong Pareto axiom. In this chapter, therefore, we use "Pareto" as shorthand for "Strong Pareto."

worse off. But what is bad is not that these people are worse off than *others*. It is rather that they are worse off than *they* might have been.

Consider next the central claim of the Priority View: benefits to the worse off matter more. . . . In this view, if I am worse off than you, benefits to me are more important. Is this *because* I am worse off than you? In one sense, yes. But this has nothing to do with my relation to you.

It may help to use this analogy. People at higher altitudes find it harder to breathe. Is this because they are higher up than other people? In one sense, yes. But they would find it just as hard to breathe even if there were no other people who were lower down. In the same way, on the Priority View, benefits to the worse off matter more, but that is only because these people are at a lower *absolute* level. It is irrelevant that these people are worse off *than others*. . . .

The chief difference is, then, this. Egalitarians are concerned with *relativities*: with how each person's level compares with the level of other people. In the Priority View, we are concerned only with people's absolute levels. (Ibid., p. 104)

Parfit does not use axioms to describe prioritarianism. However, his discussion of the difference between prioritarianism and egalitarianism can readily be translated into axiomatic language. The Separability axiom says: the ranking of any two outcomes is invariant to the well-being levels of individuals who are equally well off in both. More precisely: Let x and y be two outcomes such that some individuals are not equally well off in the two (these individuals are "affected" as between x and y); and other individuals *are* equally well off in the two (these individuals are "unaffected" as between x and y). Further, let x^* and y^* be two other outcomes which are related to x and y as follows: each individual who is affected as between x and y is at the same well-being level in x^* as she is in x, and the same well-being level in y^* as she is in y (thus also affected as between x^* and y^*); and each individual who is unaffected as between x and y is unaffected as between x^* and y^* (although perhaps at a different well-being level in the x^*/y^* pair than the x/y pair). In other words, the x^*/y^* pair is the same as the x/y pair with respect to everyone's well-being *except* for the well-being levels of the unaffected. Separability requires that, for any two such outcome pairs, the x^*/y^* ranking must be the same as the x/y ranking.[12]

A prioritarian outcome ranking satisfies Separability, while an egalitarian outcome ranking does not. Separability is a parsimonious

[12] See Chapter 2, Table 2.4.