

Risk-Free Approaches to the Priority View

David McCarthy

Received: 10 April 2011 / Accepted: 5 April 2012 / Published online: 13 May 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Parfit advertised the priority view as a new and fundamental theory in the ethics of distribution. He never discusses risk, and many writers follow suit when discussing the priority view. This article formalizes two popular arguments for a commonly accepted risk-free definition of the priority view. One is based on a direct attempt to define the priority view, the other is based on a contrast with utilitarianism and egalitarianism. But it argues that neither argument succeeds, and more generally, that it is not possible to make sense of the priority view in a risk-free framework. As a diagnosis, the article suggests that the literature on the priority view has paid insufficient attention to axiomatization and has most likely mistaken the proper role of the use of evaluative measurements in theorizing about the ethics of distribution. Reasons are eventually offered for thinking that despite appearances, approaching the priority view in terms of risk from the outset is quite natural.

1 Introduction

The priority view is widely regarded as one of the most important positions on the ethics of distribution. Closely related views had been discussed before (e.g. Broome 1989; Hurley 1989), but Parfit's 1991 *Lindley Lectures* brought the priority view into prominence (Parfit 2000). Parfit claimed that although many earlier writings on the ethics of distribution were advertised as arguing for egalitarianism, they were better interpreted as advocating the priority view. This mattered for Parfit because he thought that unlike the priority view, egalitarianism faces a serious problem, the so-called leveling down objection.

Broadly speaking, there have been two sorts of reactions to Parfit's work. The first takes Parfit's account of what the priority view is really saying more or less on

D. McCarthy (✉)
Department of Philosophy, University of Hong Kong, Pokfulam, Hong Kong
e-mail: davidmccarthy1@gmail.com

trust, then seeks to extend his discussion or apply it to new problems. For example, it has been asked whether the priority view should be applied to questions about distribution across time but within a single person's lifetime; whether it should be applied to the distribution of some goods but not others; whether it should be combined with egalitarianism; what it implies about problems of population size; and so on (see e.g. McKerlie 2001; Brown 2007).

The second reaction finds Parfit's account of the priority view in need of significant elaboration. Writers in this vein have come to a variety of conclusions about the priority view and its relationship to egalitarianism. For example, Jensen (2003) argues that prioritarian and egalitarian orderings are distinct, Tungodden (2003) argues that the former are strict subset of the latter, while Broome (1991) argues that the priority view is meaningless. It is these kinds of accounts of the priority view I will be concerned with.

One very interesting feature of Parfit's discussion of the priority view is that he never discusses risk. For example, he only addresses the question of when, according to the priority view, one outcome (or history, as I will say) is better than another. Thus I will say that his discussion of the priority view is *risk-free*: it does not address the question of when, according to the priority view, one lottery over histories is better than another.

Much of the subsequent literature follows Parfit in this respect, discussing the priority view in entirely risk-free terms. In other writings Parfit seems to regard the extension of not just the priority view, but any risk-free distributive view, to cover cases involving risk as fairly trivial: an appeal to expected utility theory is all that is needed (Parfit 1984, 2011). But other philosophers see significant issues arising in the extension of even just the priority view to cover risky cases (see e.g. Rabinowicz 2001, 2002; Broome 2004; Fleurbaey 2004; McCarthy 2008). Nevertheless, it is fair to say that all of these writers follow Parfit in regarding the core ideas of the priority view as statable in risk-free terms.

The main claim of this article is that this is a mistake: we cannot make sense of the priority view in a risk-free framework. There is no distributive theory which is statable in risk-free terms, provides an ethically well-motivated alternative to both utilitarianism and egalitarianism, and more or less satisfies Parfit's description of the priority view.

The starting place for this argument is quite simple. There is widespread agreement on which of Parfit's remarks are central to his account of what the priority view is saying. Section 2 rehearses these remarks and discusses how they should be interpreted as providing an implicit definition of the priority view. The discussion is further simplified by the fact that most writers who offer risk-free accounts of the priority view seem to agree on how it should be defined, and Sect. 3 therefore offers what I will call the risk-free definition of the priority view. The complexities begin when we turn to the issue of how the risk-free definition has been defended.

As well as the many clearly distinct risk-free treatments of the priority view mentioned above, some apparently almost identical accounts turn out to be importantly different in crucial but unacknowledged respects. Thus it will help to have some way of organizing the various defenses. One natural division runs

through Parfit's own discussion: some of his remarks seek to explain the priority view directly, while others try to explain it by contrasting it with other distributive views. Although they are often run together, I will therefore distinguish between *direct* and *contrastive* defenses of the risk-free definition. Sections 4 and 5 offer rigorous risk-free versions of these defenses. Sections 6 and 7 then argue that all risk-free direct defenses fail, while Sects. 8 and 9 argue that all risk-free contrastive defenses fail. The conclusion is that we cannot make sense of the priority view in a risk-free framework.

Despite the negative conclusion, Sect. 10 tries to draw three lessons. First, the discussion hints at what I think are promising strategies for trying to define the priority view in terms of risk. I hope to develop this approach elsewhere, but it would seem idiosyncratic without a clear view of the problems with risk-free approaches. Second, it is more or less orthodoxy in moral philosophy to appeal to the existence of various evaluative measures in articulating and defending distributive views. I believe that this is a serious mistake, but while I will not be able to argue here for that general claim, the criticisms of risk-free approaches to the priority view form what I hope is a useful case study. Third, while axiomatic methods are common in discussions of the ethics of distribution in welfare economics, moral philosophy has made much less use of them. But while many risk-free approaches to the priority view seem intuitively compelling, their faultlines only really appear when one asks how they could be axiomatized in a way which yields the needed theorems. Moral philosophy would do well to pay more attention to axiomatization.

2 The Prioritarian Platitudes

The term 'the priority view' was introduced by Parfit, so how the priority view should be understood has to largely defer to what he says about it. This section discusses how his informal remarks could be shaped into an implicit definition. To simplify, I will assume throughout a fixed population of individuals 1, ..., n .

First some terminology. The *betterness relation* holds between two lotteries over histories just in case the first is at least as good as the second. What I will call the *risk-free betterness relation* holds between two histories h_1 and h_2 just in case h_1 is at least as good as h_2 . It is thus a restriction, or special case, of the betterness relation. Roughly speaking, the risk-free betterness relation is *anonymous* just in case the identities of people do not matter. More precisely, suppose that there is some mapping ρ of the population onto itself such that for each member i of the population, h_1 is exactly as good for i as h_2 is for $\rho(i)$. Then the risk-free betterness relation is anonymous iff in all such cases, h_1 is exactly as good as h_2 . The *risk-free principle of personal good* says that if two histories h_1 and h_2 are equally good for each person, they are equally good; and if h_1 is better than h_2 for some member of the population, and at least as good for every member of the population, then h_1 is better than h_2 .

Below are eleven claims Parfit makes about the priority view, directly or implicitly, followed by brief glosses. Apart from the first two, we do not yet aim for much precision. Page references are to Parfit (2000).

P1. According to the priority view, the risk-free betterness relation is anonymous. Parfit does not discuss this idea, but it is obvious from his discussion (348).

P2. According to the priority view, the risk-free principle of personal good is true. What Parfit calls moderate forms of egalitarianism agree that the risk-free principle of personal good is true, although strong forms disagree (377).¹

P3. According to the priority view, benefits have absolute value. How much it matters to benefit someone depends only on how badly off she is, and not on whether others are better off than her (369).

P4. According to the priority view, benefits have declining marginal value. Or in Parfit's words, benefiting people matters more the worse off these people are (366, 369, 372). Parfit writes as if this is the core idea of the priority view. But he also says that this claim does not define a different view as it would be accepted by all egalitarians (368–369).

P5. The priority view emerges from the Divided World example. To be discussed in Sect. 5, this is the only concrete example Parfit offers to illustrate the priority view. Any account of the priority view will have to make sense of it.

P6. The core idea of the priority view is somehow intrapersonal. According to the priority view, what is bad about people being worse off is that they are worse off than they might have been (369).

P7. The priority view has distinctive implications in one-person cases. The gloss on *P4* and *P6* has suggested to commentators, with Parfit's apparent endorsement, that it matters more to benefit someone the worse off she is even when no others are around at all (Rabinowicz 2001, 2002).

P8. The priority view is a new distributive theory and expresses an ethically natural and fundamental idea.

It would be difficult to uphold the idea that the priority view expresses an ethically natural and fundamental idea if it turned out to rest on an ethically arbitrary or insignificant distinction. To elaborate on the novelty of the priority view we have:

P9. There is an ethically fundamental contrast between the priority view and egalitarianism.

¹ Whenever I speak of egalitarianism I have in mind a pure form of egalitarianism which departs from utilitarianism solely in the direction of being somehow concerned with equality. I have no wish to deny that a concern with equality could be combined with, say, a concern with fairness and a concern with giving priority to the worse off. I have no strong objection to calling such a mixed theory a form of egalitarianism, but that is not how I will use terminology.

Parfit's explication of the priority view is largely based on a contrast with egalitarianism. The final contrast is not that clear however (cf. the gloss on *P2* and *P4*), so this aspect of Parfit's discussion is more suggestive than definite. However:

P10. According to the priority view, and in contrast with egalitarianism, it does not matter how well off people are relative to one another (370).

Now one might think that the “nonrelative/relative concern” distinction is just a paraphrase of the “absolute/nonabsolute value” distinction which features in *P3*. But it will turn out that these two distinctions are naturally interpreted in different ways, and should not be conflated.

P11. There is an ethically fundamental contrast between the priority view and utilitarianism.

Parfit clearly regards the priority view as antiutilitarian, but says little about where, according to the priority view, utilitarianism goes wrong. But with Parfit's apparent endorsement, Rabinowicz (2001, 2002) says that the priority view is motivated by Rawls's famous “separateness of persons” criticism of utilitarianism.

I will call *P1* through *P11* the prioritarian platitudes because they are closer to elaborations on the meaning of ‘the priority view’ than to substantive claims about an independently defined theory. This need not take any controversial view about the analytic/synthetic distinction (Jackson 2000).

Suppose that interpreting some theory as the priority view vindicates the prioritarian platitudes both well enough and better than any other theory. On one view, that theory just is the priority view. On the other main view, it may be vague what the priority view is, but we would nevertheless do best to define the priority view as that theory. I will not try to decide between these views. By contrast, if no theory vindicates the prioritarian platitudes well enough when interpreted as the priority view, ‘the priority view’ is ill-defined.

3 The Risk-Free Definition of the Priority View

Although Parfit does not give a formal definition of the priority view, others have offered one which seems to grow out of his discussion. The definition takes for granted an assumption about evaluative measurement, and we are going to need to look very closely at the role of evaluative measurement in risk-free accounts of the priority view. The assumption in question here is about quantitative measures of how good things are for people, or what I will call individual goodness measures. An *individual goodness measure* is a real-valued function defined on individual/history pairs which purports to measure how good the history is for the individual. For a function g to be an individual goodness measure it will have to satisfy the following (as well as perhaps other constraints).

(1) For any individuals i and j and any histories h_1 and h_2 , h_1 is at least as good for i as h_2 is for j iff $g(i, h_1) \geq g(j, h_2)$.

The literature on the priority view assumes that it makes sense to compare units of individual goodness. This means that it takes for granted the following.

RFI. Individual goodness measures exist and are unique up to positive affine transformation.

A positive affine transformation involves multiplication by a positive constant plus addition of an arbitrary constant. So *RFI* implies that if g is an individual goodness measure, then so too is f just in case $f = ag + b$ for some positive real number a and arbitrary b . Throughout I will take g to be some fixed individual goodness measure. When a particular individual i is fixed I will often write $g(i, h)$ as $g_i(h)$. Then with each history h there is an associated *individual goodness profile*, the n -tuple $[g_1(h), \dots, g_n(h)]$.

It will help to make a standard simplifying assumption. Say that the domain (of the risk-free betterness relation) is *rich* just in case the set of individual goodness profiles is I^n for some interval I of real numbers of positive length. For example, if I is the whole of the real numbers, a rich domain means that for any n -tuple of real numbers, there is some history which has that n -tuple as its individual goodness profile. The assumption of richness is what is often known as a structural assumption. It enables us to provide simple illustrations of many of the distributive ideas we will be discussing, and I will take it for granted.

Since *RFI* implies that g is only one of many individual goodness measures, we always have to check that what we use it to say does not depend on which individual goodness measure it actually is. This is usually straightforward, and will often be left implicit. For example, if the domain is rich relative to one individual goodness measure, *RFI* immediately implies that it is rich relative to any individual goodness measure.

With *RFI* to hand, Broome (1989), Fleurbaey (2004), Jensen (2003) and Tungodden (2003) endorse the following.

The risk-free definition of the priority view. For any population of individuals $1, \dots, n$ there is some strictly increasing and strictly concave function w such that the best available history h containing that population maximizes the value of $w(g_1(h)) + \dots + w(g_n(h))$.²

The discussion will run more smoothly if we generalize slightly. A real-valued function f is said to *represent* an ordering \succsim just in case for all x and y in the set ordered by $\succsim : x \succsim y \iff f(x) \geq f(y)$. Roughly speaking, if a function represents an ordering, it can be regarded as a description of that ordering, with the caveat that one ordering can be described by more than one function. For any real function w , let $P(w)$ be the function on histories defined by

$$P(w)(h) = w(g_1(h)) + \dots + w(g_n(h))$$

² A function f from the reals to the reals is strictly increasing iff for any $x, y \in \mathbb{R}, x > y \implies f(x) > f(y)$. It is strictly concave iff for any distinct $x, y \in \mathbb{R}$ and any $\alpha \in (0, 1), f(\alpha x + (1 - \alpha)y) > \alpha f(x) + (1 - \alpha)f(y)$.

Then with a small generalization, the risk-free definition defines the priority view as saying that there is some strictly increasing and strictly concave w such that $P(w)$ represents the risk-free betterness relation.

But why should we accept the risk-free definition? As already noted, there are two basic approaches both in Parfit's work and in the subsequent literature. The direct defense tries to defend it by appealing the prioritarian platitudes which seem to be solely about the priority view, while the contrastive defense appeals to the platitudes which seem to contrast the priority view with other distributive views. These defenses are often combined, but they involve very different assumptions and problems, so I will present them separately. In keeping with the focus on risk-free approaches to the priority view, the article will say little about attempts to augment these defenses by borrowing ideas which are covertly or overtly about risk.

4 The Direct Defense

Almost everyone who has tried to defend the risk-free definition of the priority view has endorsed the direct defense. It is most explicit in Jensen (2003); see also Rabinowicz (2001, 2002). But I will develop it in my own way and bring out certain important assumptions which are typically left implicit. The direct defense appeals to the prioritarian platitudes which seem to come closest to directly defining the priority view without referring to other distributive views. In particular, the main goal of the direct defense is to argue that we vindicate $P1$ through $P4$ if and only if we accept the risk-free definition.

$P1$ and $P2$ need no further interpretation, so the bulk of the direct defense concerns $P3$ and $P4$, which together say that according to the priority view, benefits have both absolute and declining marginal value. However, how the claim that benefits have both absolute and declining marginal value should be interpreted is not obvious. The claim adopts the language of evaluative measurement, but it is far from clear what exactly it is saying and what assumptions it is taking for granted.

We will therefore need to address three tasks: (i) provide a fully explicit interpretation of $P3$ and $P4$ in the language of evaluative measurement; (ii) connect the risk-free definition with parallel claims made in terms of evaluative measurement; and (iii) use (i) and (ii) to prove that the prioritarian platitudes $P1$ through $P4$ are equivalent to the risk-free definition.

4.1 Interpretation

Suppose the population contains two people, A and B . No matter how we understand benefits, the transition from an individual goodness profile $[0,0]$ to $[1,0]$ provides A with the same benefit as the transition from $[0,10]$ to $[1,10]$. For in each case the transition takes A from 0 to 1. Since B is unaffected by each transition, it is natural to interpret the claim that benefits have absolute value as exemplified by the claim that the two transitions are equally good.

To interpret the idea that benefits have declining marginal value, we need to say more about when two benefits are the same. It is natural to understand this in terms

of arithmetic differences in individual goodness, so that taking A from 0 to 1 provides her with the same benefit as taking her from 9 to 10; see e.g. Jensen (2003). Given *RF1*, this idea is well-defined. The idea that benefits have declining marginal value is then naturally exemplified by the claim that the transition from $[0,10]$ to $[1,10]$ is better than the transition from $[9,10]$ to $[10,10]$. For A receives the same benefit in each case but her starting place is worse in the first transition, while B is unaffected.

These ideas need to be generalized, but they rely on the idea that we understand what it means to say that one transition is at least as good as another. But those who accept the direct defense plainly have a view about this; see e.g. Broome (1991) p. 179 and Jensen (2003) p. 99. Roughly speaking, the idea is that one transition is better than another just in case the first produces a greater increase in goodness than the second.

More precisely, *goodness measures* are real-valued functions defined on histories which, perhaps among other things, represent the risk-free betterness relation. For talk about comparisons of increases in goodness to be well-defined, we need to assume

RF2. Goodness measures exist and are unique up to positive affine transformation.

Taking *RF2* for granted, we can now be more precise. Let G be a goodness measure. Then the idea is that the transition from a history k_1 to a history k_2 is at least as good as the transition from k_3 to k_4 just in case $G(k_2) - G(k_1) \geq G(k_4) - G(k_3)$. *RF2* guarantees that this is well-defined; that is, it is independent of the particular choice of goodness measure.

We have seen natural exemplars of the claim that benefits have absolute value and the claim that benefits have declining marginal value. With the understanding of when one transition is better than another to hand, the following are natural generalizations of the exemplars.

RF3. *Benefits have absolute value* iff for any histories h_1, \dots, h_4 containing a population $1, \dots, n$ any individual k , and any goodness measure G : if (i) for each member i of the population apart from k , h_1 is exactly as good for i as h_2 , and h_3 is exactly as good for i as h_4 , and (ii) h_1 is exactly as good for k as h_3 , and h_2 is exactly as good for k as h_4 , then (iii) $G(h_2) - G(h_1) = G(h_4) - G(h_3)$.

RF4. *Benefits have declining marginal value* iff for any histories h_1, \dots, h_4 containing a population $1, \dots, n$ any individual k , and any goodness measure G : (i) for each member i of the population apart from k , h_1, \dots, h_4 are all equally good for i , and (ii) $g_k(h_2) - g_k(h_1) = g_k(h_4) - g_k(h_3) > 0$ and $g_k(h_3) > g_k(h_1)$, then (iii) $G(h_2) - G(h_1) > G(h_4) - G(h_3)$.

These two claims provide interpretations of the two italicized statements. *RF1* makes explicit a presupposition which is vital to *RF4*, while *RF2* makes explicit a presupposition which is vital to both *RF3* and *RF4*. The interpretation of the prioritarian platitudes *P3* and *P4*, with all their presuppositions fully spelled out, is therefore provided by *RF1* through *RF4*.

4.2 The Risk-Free Definition and Goodness

We have just interpreted the crucial platitudes $P3$ and $P4$ in terms of goodness measures. The risk-free definition of the priority view says nothing overt about goodness measures. So to use the platitudes to defend the risk-free definition, we need to somehow connect the risk-free definition with suitable claims involving goodness measures. But a natural answer is to hand, for the following seems to be taken for granted by many friends of the risk-free definition (e.g. Jensen 2003 p. 99).

RF5. $P(w)$ represents the risk-free betterness relation for some strictly increasing and strictly concave function $w \Rightarrow P(w)$ is a goodness measure.

4.3 Deriving the Risk-Free Definition

RF1 through *RF5* together make fully explicit all the assumptions needed for the direct defense. We now just need to show that with them, the appropriate prioritarian platitudes lead to the risk-free definition. One direction of the following result is implicit in the literature.

Proposition 1 *Assume RF1 through RF4, and a rich domain. Then the risk-free betterness relation is anonymous, the risk-free principle of personal good is true, and benefits have absolute and declining marginal value \iff for some strictly increasing and strictly concave w , $P(w)$ is a goodness measure, unique up to positive affine transformation.*

Assuming *RF1* through *RF4*, versions of the direct defense I am familiar with make a case for the “ \Leftarrow ” part of the biconditional, which is obvious. But they need the “ \Rightarrow ” part as well. This is more important and harder to show, but Appendix A provides a proof. But Proposition 1 yields the result we are seeking.

Theorem 1 *Assume RF1 through RF5, and a rich domain. Then $P1$ through $P4$ are true \iff the risk-free definition of the priority view is correct.³*

In summary, the direct defense relies heavily on claims made in the language of evaluative measurement. But these claims need to be interpreted and their presuppositions made fully explicit. Together *RF1* through *RF5* do that. But with *RF1* through *RF5* to hand, it follows that we vindicate the core prioritarian platitudes $P1$ through $P4$ if and only if we accept the risk-free definition.

³ *Proof.* Assume *RF1* through *RF5*. For “ \Rightarrow ”, suppose $P1$ through $P4$ are true. These together imply that according to the priority view, the left-hand side of the biconditional in Proposition 1 is true. It then follows that according to the priority view, the right-hand side of the biconditional is true. But if $P(w)$ is a goodness measure for some strictly increasing and strictly concave w , $P(w)$ certainly represents the risk-free betterness relation. Hence $P1$ through $P4$ jointly imply the risk-free definition. For “ \Leftarrow ”, suppose the risk-free definition is correct. Given *RF2* and *RF5* it then follows that according to the priority view, the right-hand side of the biconditional in Proposition 1 is true. Given *RF1*, *RF3* and *RF4*, that then implies that according to the priority view, the left-hand side of the biconditional is true, which implies $P1$ through $P4$, as required.

Sections 6 and 7 will offer some criticisms of the direct defense. But first it will help to understand the other main way of trying to defend the risk-free definition, the contrastive defense.

5 The Contrastive Defense

The contrastive defense also aims to establish that we vindicate certain prioritarian platitudes if and only if we accept the risk-free definition. Like the direct defense, it regards some of these platitudes as needing further interpretation; unlike the direct defense, it avoids making use of goodness measures and appeals instead to contrasts with other distributive views. The literature mostly focuses on egalitarianism, but we will need to look at utilitarianism as well.

5.1 The Contrast with Egalitarianism

Egalitarians are concerned with inequality, but to be concerned with inequality is to be concerned with certain patterns between people. This is crucial to Parfit's understanding of the contrast between egalitarianism and the priority view. As he puts it, egalitarians are concerned with relativities, but prioritarrians are not. But how should the concern with patterns be understood?

Suppose two histories are such that that for each member i of some subgroup of the population, the two histories are equally good for i . In trying to rank the two histories, we might be tempted to ignore the members of the subgroup, for none of them is affected by the choice between the two histories. However, it has seemed to many people that the presence of these people might well make a difference to the patterns of inequality in the two histories, in one case for the better and in the other case for the worse. It has therefore seemed to many people that a defining feature of egalitarianism is the denial that it is always safe to ignore such subpopulations (Sen 1973; Broome 1989; Temkin 1993).

To illustrate with an example modeled on the Allais paradox (cf. Sen 1973; Broome 1989), let k_1 through k_4 have respective individual goodness profiles $[1,1,1,1]$, $[3,0,1,1]$, $[1,1,0,0]$ and $[3,0,0,0]$. Someone impressed by the equality in k_1 might think it better than k_2 , but think that there is not much difference in the badness of the inequality in k_3 and k_4 and think k_4 better than k_3 because of its greater total. But suppose we ignore persons 3 and 4 in the comparison between k_1 and k_2 on the grounds that they are unaffected by the choice, and also ignore them in the comparison between k_3 and k_4 for the same reason. Then claiming that k_1 is better than k_2 forces us to conclude that k_3 is better than k_4 .

To formalize, suppose we can partition the population into two groups I and J such that (i) for all i in I , h_1 is exactly as good for i as h_3 , and h_2 is exactly as good for i as h_4 ; and (ii) for all j in J , h_1 is exactly as good for j as h_2 , and h_3 is exactly as good for j as h_4 . The risk-free betterness relation is *strongly separable* if and only if in all such cases: h_1 is at least as good as h_2 iff h_3 is at least as good as h_4 . For short, I will often refer to this idea as 'risk-free strong separability', and will usually drop

the ‘risk-free’. Then a popular way of at least partially characterizing egalitarianism is to claim that it denies strong separability.

This argument generalizes. Being concerned with inequality is one way of being concerned with how well off people are in relation to one another, but the argument seems to apply to any kind of concern with how well off people are relative to one another. This suggests that the “nonrelative versus relative” distinction should be interpreted in terms of the “strongly separable versus not strongly separable” distinction. Notice how different this interpretation of “nonrelative versus relative” is from the earlier interpretation of the superficially similar “absolute versus nonabsolute” distinction. The latter was interpreted by assuming a lot about the quantitative structure of goodness; the former has just been interpreted solely in terms of qualitative features of the risk-free betterness relation.

Given the partial characterization of egalitarianism in terms of the denial of strong separability, and given the way Parfit contrasts the priority view with egalitarianism, it seems mandatory to follow Broome (1989), Jensen (2003) and McCarthy (2008) and accept

RF6. The fundamental contrast between the priority view and egalitarianism is that the priority view accepts risk-free strong separability, while egalitarianism rejects it.

This takes us some way towards a definition of the priority view, but not all of the way. Utilitarians claim that one history is at least as good as another if and only if it contains at least as great a sum of individual goodness. That immediately implies that utilitarians accept strong separability, so we cannot define the priority view solely in terms of strong separability. We need to look at how it contrasts with utilitarianism. In the spirit of the way the contrastive defense has so far been developed, I will avoid trying to do this by appealing to goodness measures.

5.2 The Contrast with Utilitarianism

As I will understand it, there are two parts to the contrast with utilitarianism. The first seeks to articulate a principle which provides the content of the antiutilitarian component of the priority view. The second tries to show that the motivation for this principle arises from a natural criticism of utilitarianism.

A *Pigou–Dalton transfer* takes a unit of individual goodness from a better off person and gives it to a worse off person while leaving the first person no worse off than the second person. Everyone else is left unaffected. In speaking of units of individual goodness, this definition presupposes *RF1*. Consider

The Pigou–Dalton principle. If a history h_1 can be obtained from a history h_2 by a Pigou–Dalton transfer, then h_1 is better than h_2 .

There are three reasons for thinking that the Pigou–Dalton principle spells out the antiutilitarian content of the priority view. First, several philosophers have thought that even if the interpretation of the claim that benefits have declining marginal value is not entirely clear, that claim at least obviously entails the Pigou–Dalton principle (see e.g. Jensen (2003)). Second, consider a minor variation on what Parfit

calls the Divided World. Two communities of the same size are unaware of each other's existence. In h_1 everyone is at individual goodness level 150. In h_2 one community is at 200 and the other is at 100. Parfit suggests that h_1 is better than h_2 , and that this claim expresses the heart of the priority view. But this conclusion can be derived from a series of applications of the Pigou–Dalton principle. Third, Parfit clearly implies that the priority view departs from utilitarianism in the direction of giving more priority to the worse off. But the Pigou–Dalton seems to be the weakest natural principle one could accept which gives more priority to the worse off than utilitarianism does. Thus perhaps because of the combination of these three reasons, Tungodden (2003) and McCarthy (2008) suggest that the Pigou–Dalton principle just is the antiutilitarian component of the priority view.

Turning now to motivation, Rabinowicz (2001, 2002) claimed, with Parfit's apparent endorsement, that the antiutilitarian aspect of the priority view is motivated by the famous separateness of persons criticism of utilitarianism in Rawls (1971).

Rawls's starting point was the impartial spectator argument. According to this argument, two histories h_1 and h_2 can be compared as follows (Rawls 1971; Nagel 1970). The lives of each member of the population in h_1 are concatenated into a single life via something like the serial-life construction of Lewis (1946); likewise for h_2 . The impartial spectator imagines the lives in full detail, and ranks the histories by ranking their associated serial-lives. Rawls claimed that transfers of units of individual goodness from one part of an (ordinary) life to another part are unproblematic. When this idea is applied to serial-lives, the impartial spectator construction implies that transfers of units of individual goodness from one person's life to a different person's life are equally unproblematic, and the result is utilitarianism. But Rawls objected. He claimed that because the impartial spectator construction conflates the lives of different people into the life of a single person, utilitarianism "does not take seriously the distinction between persons" (Rawls 1971 p. 28).

If one is moved by this objection, one will look for some natural principle which restricts transfers between different people more than utilitarianism does in a way which somehow takes the separateness of persons into account. The Pigou–Dalton principle seems to be the obvious candidate, for in effect it says that it is bad to take a unit of individual goodness from a worse off person and give it to a better off person. Thus the Pigou–Dalton principle can be seen as being motivated by a concern with the separateness of persons. Combining the contrastive and motivational remarks, it then becomes natural to accept

RF7. The fundamental contrast between the priority view and utilitarianism is that the priority view accepts the Pigou–Dalton principle, while utilitarianism rejects it.

I stress that I am not endorsing the claim that utilitarianism improperly ignores the separateness of persons. Rather, I have sketched an influential view about the way various ideas are taken to be connected, and have suggested that *RF7* makes some of these connections explicit in a natural way.

5.3 Deriving the Risk-Free Definition, Again

We can now conclude the contrastive defense. Taking *RF1* for granted we have

Proposition 2 *Assume a population of at least three people and a rich domain. Then the risk-free betterness relation can be represented by $P(w)$ for some strictly increasing and strictly concave $w \iff$ the risk-free betterness relation is anonymous and continuous, the risk-free principle of personal good is true, the risk-free betterness relation is strongly separable, and the Pigou–Dalton principle is true.*

This follows from what Wakker (1989) calls the central theorem for additive representation. The claim that the risk-free betterness relation is continuous is usually regarded as a technical condition and often ignored. Just to illustrate, suppose one individual goodness profile is better than another. Then continuity implies that some sufficiently small change to the former will still leave it better than the latter; see Wakker for details. But Proposition 2 entails a conclusion of the desired form.

Theorem 2 *Assume *RF1*, *RF6*, *RF7*, at least three people, and a rich domain. Then *P1*, *P2*, *P9* and *P11* are true and the risk-free betterness relation is continuous \iff the risk-free definition of the priority view is correct.*

In summary, *RF1* is commonly assumed in discussions of the ethics of distribution, while *RF6* and *RF7* provide a natural interpretation of the way the priority view contrasts with egalitarianism and utilitarianism. But if we take continuity on trust then provided there are at least three people,⁴ it follows that we vindicate the prioritarian platitudes *P1*, *P2*, *P9* and *P11* if and only if we accept the risk-free definition.

5.4 Summary

It should be stressed that both the direct and the contrastive defenses of the risk-free definition do not mechanically exploit the prioritarian platitudes, but depend on further interpreting some of them. The direct defense makes assumptions about goodness measures to interpret the idea that benefits have both absolute and declining marginal value. The contrastive defense depends on views about the separateness of persons and about concerns with patterns to interpret the way the priority view is supposed to contrast with utilitarianism and egalitarianism. Both

⁴ The case where the population contains exactly two people is tricky. In such cases, strong separability implies nothing beyond what is already implied by the risk-free principle of personal good, and does not guarantee an additive representation. It is therefore not much use in distinguishing the priority view from egalitarianism in such cases, so *RF6* is at least slightly exaggerated and should be restricted to the case where there are at least three people. To guarantee an additive representation in the $n = 2$ case one needs to appeal to the so-called hexagon condition; see Wakker (1989) for more details. However, the hexagon condition is extremely complicated and it is hard to see that it expresses a natural ethical idea, let alone one which naturally separates the priority view from egalitarianism. I will not say anything more about this, but it points to a significant gap in the literature and one more difficulty for the contrastive defense beyond those to later be discussed.

defenses make assumptions about the existence and uniqueness of individual goodness measures.

Friends of the risk-free definition will no doubt be encouraged by the fact that two rather different forms of argument lead to that definition. Like having two independent pieces of evidence for the same hypothesis, two different arguments for the risk-free definition will seem to provide greater support for that definition than each taken separately. I will refer to the conjunction of these defenses as the risk-free defense.

We have now seen rigorous versions of the two main approaches to defending the risk-free definition of the priority view. We will soon note some variants on these, but now it is time to start criticizing them. The next two sections will argue that the direct defense is a failure, and the two after that will argue that the contrastive defense also fails.

6 Criticisms of The Direct Defense: The Structure of Goodness

Goodness measures play a crucial role in the direct defense. Assumptions about them are often left implicit or taken for granted in defenses of the risk-free definition of the priority view. The direct defense was formulated to make these assumptions fully explicit, but we will now see that the assumptions make the direct defense highly problematic.

The direct defense makes two assumptions about goodness measures. In *RF2* it assumes that goodness measures exist and are unique up to positive affine transformation. In *RF5* it assumes that if $P(w)$ represents the risk-free betterness relation for some strictly increasing and strictly concave function w , then $P(w)$ is a goodness measure. Together, these imply that according to the risk-free definition of the priority view, the priority view implies not just that for some strictly increasing and strictly concave function w , $P(w)$ represents the risk-free betterness relation, but also that $P(w)$ is a goodness measure, unique up to positive affine transformation. We need to ask whether the two assumptions are true, and whether the direct defense really needs them.

Let's begin with their truth. It is far from obvious. The only obvious constraint that a goodness measure G has to satisfy is that

$$(2) \text{ For all histories } h_1 \text{ and } h_2, G(h_1) \geq G(h_2) \text{ iff } h_1 \text{ is at least as good as } h_2.$$

In other words, goodness measures have to represent the risk-free betterness relation. However, for any function h and any strictly increasing function f we have

$$\text{For all } x \text{ and } y : h(x) \geq h(y) \iff f(h(x)) \geq f(h(y)).$$

Thus when a function represents an ordering, so does any increasing transformation of the function. But this means that if (2) is the only constraint which a goodness measure has to satisfy, then goodness measures are unique only up to increasing transformation. But there are vastly many more increasing transformations than positive affine transformations, so if (2) is indeed the only constraint on goodness measures, then *RF2* is false.

The point can be put another way. If for some reason we are convinced of *RF2*, then *RF5* becomes questionable. Even if it is given that $P(w)$ represents the risk-free betterness relation, all that follows from *RF2* is that the goodness measures are the positive affine transformations of some increasing transformation of $P(w)$, not of $P(w)$ itself. Thus the combined package of *RF2* and *RF5* is highly problematic.

Let's now ask how troubling this is for the direct defense. On the face of it, it is very troubling. The direct defense needs both *RF2* and *RF5*. It certainly needs *RF2*, for as already noted, the inequalities which *RF3* and *RF4* use to interpret the idea that benefits have absolute and declining marginal value are well-defined if and only if goodness measures are unique up to positive affine transformation. But *RF5* is needed as well, as a simple example will illustrate.

Without *RF5*, *RF2* only tells that the risk-free definition of the priority view tells us that according to the priority view, for some strictly increasing and strictly concave w , some increasing transformation of $P(w)$ is a goodness measure. But let $Q(w)$ be the fifth power of $P(w)$, an increasing transformation of $P(w)$, let w be the square root function, and consider what happens if $Q(w)$ is a goodness measure.

Assume a population of two people and let h_1 through h_6 have individual goodness profiles $[0,0]$, $[1,0]$, $[0,10]$, $[1,10]$, $[9,10]$ and $[10, 10]$. According to the direct defense, the idea that benefits have absolute value was exemplified by the claim that the transition from h_1 to h_2 produces the same increase in goodness as the transition from h_3 to h_4 . Likewise, the idea that benefits have declining marginal value was taken to be exemplified by the claim that the transition from h_1 to h_2 produces a greater increase in goodness than the transition from h_5 to h_6 . However, if $Q(w)$ is a goodness measure, a quick calculation shows that both of these claims are false. Thus if the natural default position that goodness measures are unique only up to increasing transformation is correct, the risk-free definition of the priority view certainly does not imply that according to the priority view, the value of benefits is absolute, nor does it imply that according to the priority view, benefits have declining marginal value. But that implies that the risk-free definition does not vindicate the relevant prioritarian platitudes.

7 Criticisms of The Direct Defense: Ethical Significance

Nevertheless, the direct defense is natural enough, and we should not be too quick to dismiss it. This section therefore looks at various ways in which it might be modified in the light of the preceding criticisms. But to evaluate these modifications we first need to say a little more about what a successful defense of a definition of the priority view would achieve.

Much of the excitement about Parfit's discussion of the priority view came from the impression that he had uncovered a fundamental distributive theory which was genuinely different from others on the market. So vindicating *P8*, part of which says that the priority view expresses an ethically natural and fundamental idea, is central to the project of providing an explicit definition of the priority view.

Parfit's vision seems convincing in part because the idea that benefits have absolute and declining marginal value seems ethically natural and fundamental. But

the direct defense interpreted this idea in terms of goodness measures, and that ran into the threat of presupposition failure. This section will examine ways in which that problem with the direct defense might be avoided, but we will see that these attempts come with costs. In particular, they leave us with no reason to believe that the idea that benefits have absolute and declining marginal value has any real ethical significance.

I see little hope for defending *RF2* in a risk-free framework in a way which does not make disputed assumptions about which theory about the ethics of distribution is correct. But an argument for the correctness of a definition of the priority view should rest on a more neutral basis than that. Instead, this section examines three attempts to patch up the direct defense without relying on *RF2*.

The first attempt starts by considering what I will call the *strong* risk-free definition of the priority view, proposed by Rabinowicz (2001), which says that the following is a definition of the priority view.

- (3) According to the priority view, for some strictly increasing and strictly concave w , $P(w)$ is a goodness measure, unique up to positive affine transformation.

We could then drop *RF2* and still run one direction of the direct defense. Without going through the details, we could conclude that the strong risk-free definition of the priority view vindicates the prioritarian platitudes the direct defense appeals to, in particular the platitudes concerning the absolute and declining marginal value of benefits.

We cannot claim that those platitudes imply the strong risk-free definition of the priority view. The basic problem is that in the absence of *RF2*, the interpretations offered by *RF3* and *RF4* of what it is for benefits to have absolute and declining marginal value are ill-defined.⁵ Nevertheless, the fact that the strong risk-free definition vindicates the platitudes in question may seem to provide support for the strong risk-free definition.

However, the one-way nature of this defense should arouse our suspicion. For if the prioritarian platitudes do not imply the strong risk-free definition, it is unclear what the ethical content of the priority view is according to the strong risk-free definition beyond what was already implied by the original risk-free definition. The original definition tells us that according to the priority view, $P(w)$ represents the risk-free betterness relation for some strictly increasing and strictly concave w . This is consistent with the claim that according to the priority view, the goodness measures are the positive affine transformations of the cube root of $P(w)$; the claim that they are the positive affine transformations of the cube of $P(w)$; or the claim that they are the positive affine transformations of $P(w)$ itself, as the strong risk-free definition claims. But each of these pairwise-inconsistent claims makes the priority view imply exactly the same claims about the extension of the risk-free betterness relation as the original risk-free definition. We are therefore left with little

⁵ I am being slightly casual. A bit more precisely, *RF3* and *RF4* suffer from presupposition failure if we abandon *RF2*. But there are different views about what to say about presupposition failure, and I sidestep the question of whether it means that *RF3* and *RF4* are false, meaningless, truth-valueless or whatever. Whatever we say, it is obvious that this direction of the direct defense is in trouble.

understanding of what the genuine ethical content of the priority view is according to the strong risk-free definition which is not already built into the original risk-free definition. But without this understanding, we cannot claim that on the strong risk-free definition, the priority view expresses a natural and basic ethical idea.

The second attempt starts by trying to fix the problem with the first. Instead of proposing to make (3) come out true by definition, we might try to derive (3) by trying to argue for

(4) $P(w)$ represents the risk-free betterness relation for some strictly increasing and strictly concave function $w \implies P(w)$ is a goodness measure unique up to positive affine transformation.

I believe (4) is tacitly accepted by several writers (e.g. Jensen 2003 p. 99), and in fact, there is an interesting argument for accepting it. Suppose that $P(w)$ represents the risk-free betterness relation. Because of its additive form, it is appealingly simple. But of all its increasing transformations, the only ones which preserve the appealingly simple additive form are the positive affine transformations. All the others result in functions which are intuitively more mathematically complicated.

Now suppose we wish to define goodness measures in a way which enable us to talk meaningfully about units of goodness. After all, philosophers do sometimes talk quantitatively about goodness and are likely to continue to, so it would be useful if we could arrive at a convention which fixed the meaning of this talk. That means fixing goodness measures up to positive affine transformation. It would be foolish to make our measures needlessly complicated, and therefore sensible to adopt (4) as a convention. In effect, we now have a pragmatic argument for adopting (4).

Unfortunately, however, this pragmatic argument is self-defeating. Suppose that the betterness relation is represented by taking the expected value of $f(P(h))$ for some strictly increasing but nonlinear function f . In note 7 in the next section we will see that there are plenty of important views about the ethics of distribution which say just that. But then it is arguable that the simplest overall way of measuring goodness is by taking the goodness measures to be the positive affine transformations of $f(P(h))$, not of $P(h)$. For that means that one lottery is better than another just in case its expected goodness is greater, an appealingly simple representation. In effect, this means we have a counterexample to the pragmatic argument for adopting (4).

The third attempt starts by trying to fix the problem with the second. Consider what I will call the *risk-involving* definition of the priority view, which says that the following is a definition of the priority view.

(5) According to the priority view, for some strictly increasing and strictly concave w , the betterness relation is represented by taking expectations of $P(w)$.⁶

I believe Broome (1991) was the first to propose this definition as an amendment to the risk-free definition, a definition he had proposed earlier in the same book. It clearly abandons a risk-free approach to the priority view, but let's set that aside.

⁶ Thus for any lotteries L_1 and L_2 , L_1 is at least as good as L_2 iff the expected value of $P(w)(h)$ under L_1 is at least as great as the expected value of $P(w)(h)$ under L_2 .

Given (5), the reasoning behind the counterexample to the pragmatic argument just discussed now motivates adopting (3). But as we have already seen, (3) implies that the prioritarian platitudes featuring in the direct defense, particularly those concerning absolute and declining marginal value, are vindicated.

However, it is important to understand the limitations of this conclusion. Suppose we adopt the risk-involving definition. Then relative to some conventions for speaking quantitatively about goodness, the priority view implies that benefits have both absolute and declining marginal value. Relative to other conventions, the priority view implies neither.

In the absence of any obvious constraints on goodness measures beyond (2), it seems that if we are going to speak quantitatively about goodness, we should adopt a convention on pragmatic grounds. Given (5), (3) is then the best convention to adopt on the grounds that it makes the description in terms of goodness of what the priority theory says appealingly mathematically simple. This turns out to vindicate the idea that the priority view implies that benefits have both absolute and declining marginal value. However, this is nothing more than a feature of a convention chosen on the grounds of mathematical simplicity. But there is surely no a priori connection between mathematical simplicity and ethical significance. Thus this approach to goodness measures gives us no reason to believe that the priority view as defined by the risk-involving definition rests on or embodies ideas concerning the absolute and declining marginal value of benefits which are ethically natural and fundamental. In retrospect, this objection would have applied to the second attempt as well.

I cannot claim to have surveyed every way in which the meaning of quantitative talk about goodness might be fixed. I will say more in Sect. 10 where I will suggest that many moral philosophers have made a methodological mistake about the proper role of goodness measures in ethical theorizing. But for now the arguments of this section suggest that it is going to be difficult to patch up the direct defense in a way which both fends off the threat of presupposition failure and shows that the claim that benefits have absolute and declining marginal value expresses an ethically natural and fundamental idea.

However, since the direct and contrastive defenses are largely independent, it is possible that despite the failure of the direct defense, the contrastive defense succeeds. That would be enough to vindicate the risk-free definition, but the next two sections will argue that the contrastive defense also fails.

8 Criticisms of The Contrastive Defense: Egalitarianism

The substantive claim the contrast with egalitarianism rested upon was *RF6*, the idea that the boundary between the priority view and egalitarianism should be understood in terms of strong separability. In Sect. 8.1 we will look at this appeal to strong separability more closely, and find it problematic. However, one major feature of Parfit's discussion has not yet been properly discussed. One of Parfit's main reasons for finding the priority view appealing was his claim that egalitarianism suffers from a serious problem, the so-called leveling down

objection. I will later say how we might be able to treat this objection as an implicit partial definition of the priority view. This might lead to a new defense of characterizing the boundary between the priority view and egalitarianism in terms of strong separability. But it might instead lead to a different way of drawing the boundary, so we will treat the leveling down objection separately in Sect. 8.2.

8.1 Strong Separability

This subsection will argue that as so far presented, the contrastive defense does a poor job at vindicating various platitudes connected with egalitarianism, in particular *P8*, *P9* and *P10*. I begin with two preliminary remarks about strong separability before turning to the platitudes.

First, although it has often been said that egalitarianism and strong separability are incompatible, to my knowledge no one has offered a simple example which shows that egalitarians are forced to reject strong separability. In Sect. 5.1 we noted that a modification of the Allais paradox which might seem to play this role was proposed by Sen (1973) and Broome (1989). But as Broome (2004) admits (and perhaps Sen intended all along), the example only shows that egalitarians can reject strong separability, not that they have to. Moreover, the argument given in 5.1 for the claim that egalitarianism should be characterized as denying strong separability is overstated. It too only makes a case for the claim that egalitarians can reject strong separability, not that they have to.

Second, recall that Sect. 5.1 presented an argument to the effect that the “nonrelative versus relative” distinction should be formalized via the “strongly separable versus not strongly separable” distinction. But this conclusion is mistaken. Consider the following.

The weak equality principle. Assume a constant population. Suppose that two lotteries L_1 and L_2 are equally good for each member of the population, but L_1 guarantees equality of outcome while L_2 guarantees inequality of outcome. Then L_1 is better than L_2 .

To illustrate with an example due to Myerson (1981), the weak equality principle implies that $[\frac{1}{2}, [4, 4]; \frac{1}{2}, [0, 0]]$ is better than $[\frac{1}{2}, [4, 0]; \frac{1}{2}, [0, 4]]$.

The weak equality principle naturally expresses the idea that equality is valuable, and is plainly concerned with how well off people are relative to others. But it turns out to be consistent with strong separability of the risk-free betterness relation.⁷ Hence we cannot say that the “strongly separable versus not strongly separable”

⁷ *Sketch proof.* Assume that one lottery is at least as good for an individual as another lottery just in case the first gives the individual at least as much expected individual goodness. Let w be any strictly increasing and strictly concave function, assume a fixed population $1, \dots, n$ and let $E(w)$ be the function defined on histories by $E(w)(h) = w^{-1}(\frac{1}{n} \sum_{i=1}^n w(g_i(h)))$. The claim that the betterness relation is represented by taking expectations of $E(w)(h)$ is easily seen to entail both the weak equality principle and that the risk-free betterness relation is strongly separable. Whether or not such accounts of the betterness relation are genuine forms of egalitarianism involves much more complicated philosophical and mathematical issues than I can go into here. For what it is worth, I argue in McCarthy (forthcoming) that they are, but nothing here hinges on that conclusion.

distinction captures the “nonrelative versus relative” distinction as there are natural extensions of strongly separable orderings of histories which are concerned with relativities. At best, strong separability is necessary for a nonrelative concern with people, but not sufficient.

With these remarks to hand, what can we say about the claim that the contrastive defense succeeds in vindicating the platitudes associated with egalitarianism? Consider first the question of whether there is an ethically fundamental contrast between egalitarianism and the priority view. In light of the first remark, the case for characterizing egalitarianism in terms of the denial of strong separability is far from convincing.⁸ There have been two reactions to this from friends of the risk-free definition.

The first position continues to advocate characterizing egalitarianism as denying strong separability, and thus makes a simple contrastive case for characterizing the priority view as accepting strong separability. This guarantees that no egalitarian ordering can be a prioritarian ordering, and vice versa. And as Jensen (2003) p. 92 puts it, it has the advantage that “a definition leading to practical differences in the ordering of outcomes is more interesting”. However, we lack a simple example which shows that a concern with equality forces one to reject strong separability, and the “nonrelative versus relative” distinction does not pick out strong separability as a fundamental property. So while this position draws a clear boundary between egalitarianism and the priority view, the cost is that it gives us no reason for thinking it is a boundary which has fundamental ethical significance. It therefore does a poor job at vindicating *P8* and *P9*.

The second position, endorsed by Fleurbaey (2004) and Tungodden (2003), says that denying strong separability is an option for egalitarians, but not a necessity. Since the Pigou–Dalton principle is one of the main egalitarian ideas, this position will regard prioritarian orderings as a strict subset of egalitarian orderings. But it still maintains that there is some sort of contrastive case for defining the priority view in this way. In making the priority view a specialization of an already known basic distributive position, this position appears to undermine the idea that there is something importantly new about the priority view. Thus it is not clear how it vindicates *P8* and *P9*. But Fleurbaey and Tungodden have a response. What is distinctive about the priority view is not what it says about the way histories should be ranked, but why it says it. So they would suggest that fundamental distributive theories, such as egalitarianism and the priority view, should be identified not just with accounts of the extension of the risk-free betterness relation, but should also include justifications of those accounts. This makes room for the position that the

⁸ Notice that my claim is that it is hard to see a good reason to characterize egalitarianism in terms of the denial of strong separability. For all I have argued, it might turn out to be the case that egalitarianism entails the denial of strong separability. For example, it is arguable that any form of egalitarianism is going to have to capture the weak equality principle. In conjunction with other ideas egalitarianism should accept, that is likely to make egalitarianism a complicated theory. Perhaps that theory entails the denial of strong separability. But this does not mean that egalitarianism should be characterized in terms of strong separability. The characterization of a theory should be done in terms of axioms whose meaning and motivation is fairly close to being self-evident; I have simply been claiming that we lack such a motivation for the idea that egalitarians have to deny strong separability.

priority view is a new and fundamental distributive view even if prioritarian orderings are a subset of egalitarian orderings.

This position is going to have to locate all contrasts between the priority view and egalitarianism at the level of justification, but it gives little indication of how to do that. But this creates a worry. It is certainly possible to use different words at the level of justification: one view solely talks about giving priority to the worse off and nothing else, while the other talks about reducing inequality. But when a concern with reducing inequality is combined with strong separability, it is far from obvious that the slogan of giving priority to the worse off which is associated with the priority view does not simply summarize this form of egalitarianism. It is often said that a concern with equality can be expressed in a variety of ways (Jensen 2003), and it is far from clear that different words used at the level of justification really do indicate two fundamentally different kinds of positions.

Parfit does not discuss fairness, but to fully vindicate *P8*, the priority view would have to be shown to be distinct from a concern with fairness. I do not have the space to discuss fairness, but it is worth noting that on the view of Broome (1990–1991), fairness requires goods to be distributed in proportion to how much people would benefit from them. This idea has various interpretations, but the most natural turn out to be versions of the priority view according to the risk-free definition (Segal 2006). This puts further pressure on the idea that the priority view is a new theory in the ethics of distribution.

In summary, as it stands, the contrastive defense fails to give an account of the priority view and egalitarianism which shows that there is a clear and ethically important distinction between the priority view and egalitarianism, and thereby fails to vindicate *P8* and *P9*. In addition, no matter how it is defended, the risk-free definition of the priority view fails to vindicate *P10*.

8.2 The Leveling Down Objection

Parfit (360) says that according to egalitarianism,

E. Equality is intrinsically valuable.

He then uses this claim to try to motivate the priority view by claiming that egalitarianism is subject to what he called the leveling-down objection. The objection claims that whenever there is inequality, bringing everyone down to the level of the worst off member(s) of the population is in one way better according to egalitarianism since it brings about perfect equality. But Parfit suggested that many people will find this conclusion “absurd” (364).

Since Parfit does not offer an explicit definition of the priority view, it is tempting to use these remarks as the basis for an implicit partial definition of the priority view: the priority view should in part be characterized as making whatever claim it is whose denial makes egalitarianism subject to the leveling down objection.

This proposal has a potential advantage. The defenses we discussed in Sect. 8.1 of characterizing the boundary between the priority view and egalitarianism in terms of strong separability faced difficulties in explaining why that boundary has any great ethical significance. But if egalitarianism is subject to a significant objection in

virtue of denying some claim, then characterizing the boundary between the priority view and egalitarianism in terms of that claim will pick out an important ethical issue.

Since the priority view rejects *E*, the most appealing version of this proposal would be to interpret *E* in a way which yields in a clear disagreement between egalitarianism and the priority view and which shows why the egalitarian position on the disagreement results in a leveling down objection. However, if we remain within a risk-free framework, it is difficult to see how to do this. For example, suppose we interpret *E* as

- (6) An equal distribution is better than an unequal distribution of the same sum of individual goodness.

According to the risk-free definition, the priority view also implies (6). So we will have failed to have established that the priority view is a new and fundamental distributive view. We could try to separate prioritarianism from egalitarianism by distinguishing between the reasons for which they accept (6), egalitarianism because it accepts *E*, prioritarianism for some other reason. But this leads us back to the problems of locating the contrasts at the level of justification discussed in the previous subsection.

In my opinion (McCarthy forthcoming), the only really satisfactory way of interpreting the claim that equality is intrinsically valuable is to move to a risk-involving framework. Consider the weak equality principle. That principle says that when there is nothing to choose between two lotteries from the point of view of each person, if one lottery guarantees equality while another guarantees inequality, then the one which guarantees equality is always better. It is better because, to use Parfit's phraseology, the histories which could result from the first lottery are, from the point of view of egalitarianism, "in one way better" than the histories which could result from the second. It is hard to think of a more direct expression of the claim that equality is intrinsically valuable than the weak equality principle, and moreover, there is plainly no reason for the priority view to accept it.

We have been entertaining the conjecture that the priority view should be characterized as making some claim whose denial makes egalitarianism subject to a leveling down objection. Suppose that interpreting *E* via the weak equality principle does not result in a significant leveling down objection. Then this route to understanding the priority view fails at the outset. So suppose instead that it does lead to a leveling down objection. It would still not follow that characterizing the priority view as accepting strong separability avoids the leveling down objection. This is established by note 7, which shows that a sensible theory containing the weak equality principle is consistent with strong separability. But more, note 7 entails that according to the risk-free definition, the priority view can consistently be extended to a theory which is vulnerable to a leveling down objection. Therefore, obviously, the risk-free definition of the priority view cannot be defended as articulating a theory which avoids the leveling down objection. Thus on either supposition, appealing to the leveling down objection does not provide support for the risk-free definition.

9 Criticisms of the Contrastive Defense: Utilitarianism

This section will argue that the contrastive defense does a poor job at vindicating various platitudes connected with utilitarianism, in particular *P6*, *P7*, *P8* and *P11*. These claims are all connected with the content and motivation of the Pigou–Dalton principle.

Consider first *P7*, the idea that the priority view has distinctive implications in one-person cases. Utilitarianism implies that in such cases, a history h_1 is at least as good as h_2 just in case h_1 is at least as good for the sole inhabitant as h_2 . So does egalitarianism, for in one-person cases inequality is never an issue. But the risk-free definition of the priority view implies that the priority view has exactly the same implications in one-person worlds as both these views. This is not particularly surprising. Since inequality is never an issue in one-person worlds, the antiegalitarian component of the priority view, whatever exactly it is, is not by itself going to have distinctive one-person implications. But if the Pigou–Dalton principle is the antiutilitarian component of the priority view, that aspect of the priority view will by itself have no one-person implications at all, for it only applies when at least two people are around.

Now one might think that this is not too devastating for the risk-free definition. Parfit does not explicitly discuss one-person worlds, and an adequate account of the priority view need not vindicate all of the prioritarian platitudes: it just has to vindicate enough of them well enough. But the fact that the Pigou–Dalton principle essentially involves two people leads to deeper trouble.

P6 tells us that the core of the priority view is somehow intrapersonal. So it is odd that it should be interpreted by a principle which essentially involves at least two people. The contrastive defense rested in part on the claim that the Pigou–Dalton principle is obviously entailed by the idea that benefits have declining marginal value. Since the idea that benefits have declining marginal value is not clearly well-defined, this point would be better put by claiming that the Pigou–Dalton principle provides a natural, although perhaps not complete, interpretation of the idea that benefits have declining marginal value. So one might think that there is nevertheless a very close connection between the Pigou–Dalton principle and the core of the priority view.

But this claim is problematic. There is no obvious reason why one could not accept that benefits have declining marginal value by accepting that benefiting men matters more the worse off they are and that benefiting women matters more the worse off they are, but nevertheless hold that benefiting women matters much more than benefiting men. After all, Parfit takes care to emphasize that the slogan that benefiting the worse off matters more is to be given a strictly intrapersonal interpretation: “benefiting people matters more the worse off these people are” (366, 369, 372). But on the view just mentioned, it will sometimes be better to give a unit of individual goodness to a better off woman than to a worse off man, contra the Pigou–Dalton principle.

In effect, the claim that the Pigou–Dalton principle provides a natural interpretation of the idea that benefits have declining marginal value takes anonymity for granted. But while anonymity is plausible, the men and women

example shows that it does not belong to the principle Parfit was trying to articulate which he summarized via the slogan that benefiting the worse off matters more. More precisely, let P stand for the core antiutilitarian idea behind the priority view, an intrapersonal principle which has at least some important connection with the idea that benefits have declining marginal value. Then our task is to figure out what P is given that it has to satisfy: [P and anonymity] \Rightarrow Pigou–Dalton, despite the fact that $P \not\Rightarrow$ Pigou–Dalton. But it is very hard to see what P could be. The only idea which looks at all promising is to interpret P just as the claim that benefits have declining marginal value. But this leads us straight back to the direct defense and the problems with goodness measures already rehearsed.

A related worry about the Pigou–Dalton principle is that it presupposes that individual goodness measures are unique up to positive affine transformation. But it is far from obvious that they are. And if they are, it is far from obvious that a particular principle expressed in terms of them, such as the Pigou–Dalton principle, has any natural ethical significance. But the issues here parallel the worries about goodness measures already discussed, so I will not discuss them any further.

Turning now to motivation, there is a logical peculiarity in the contrastive defense. The contrastive defense claimed that the Pigou–Dalton principle is well motivated because utilitarianism ignores the separateness of persons, and that moving to the Pigou–Dalton principle is needed to respect the separateness of persons. Utilitarianism was said to ignore the separateness of persons because it rests on the impartial spectator argument, and the premises of that argument ignore the separateness of persons. But just because some bad argument implies utilitarianism, it would be a logical blunder to conclude merely from that fact that there is something wrong with utilitarianism. Perhaps some good argument implies utilitarianism.

There is growing acknowledgement that by far the best defense of utilitarianism is provided by a theorem due to Harsanyi (1955).⁹ Aside from some idealizing assumptions to do with comparability which can most likely eventually be dispensed with, there is a strong prima facie case for its premises. This puts those who wish to defend distributive alternatives to utilitarianism under a perhaps helpful obligation: say which of the Harsanyi premises you disagree with, and why. But this has at least two implications which friends of the priority view will have to take seriously. First, the theorem does not make any assumptions about evaluative measurements. So as it stands, it is far from clear where the Pigou–Dalton principle even takes issue with the Harsanyi premises. So the contrastive defense does a poor job at vindicating $P8$. Second, the theorem does not use anything like an impartial spectator style construction and conflate persons in the way the impartial spectator argument is said to. So the separateness of persons objection has no obvious target, and anyone who wishes to use something like it to articulate or defend a distinctively prioritarian critique of utilitarianism has much work to do. But without a clear, nonegalitarian motivation for the Pigou–Dalton principle, the contrastive defense also does a poor job at vindicating $P8$.

⁹ See e.g. Broome (1991) and Gibbard (2008). For an accessible discussion of the technicalities of the theorem, see Resnik (1987).

10 Conclusions

Given the restriction to a risk-free framework, I have argued for three things. First, the direct defense of the risk-free definition of the priority view may well suffer from presupposition failure, and even if it does not, it fails to show that the ideas it takes to define the priority view have any fundamental ethical significance. Second, the egalitarian half of the contrastive defense rests on an unconvincing characterization of egalitarianism, and natural responses to this problem involve either drawing a somewhat arbitrary boundary between egalitarianism and the priority view, or else obscurely claiming that the differences between the two views are all located at the level of justification. Third, the utilitarian half of the contrastive defense leaves it unclear where the priority view takes issue with the best defense of utilitarianism, and whether the priority view rests on a well-motivated but nonegalitarian criticism of utilitarianism. In addition, the risk-free definition of the priority view fails to vindicate some of the prioritarian platitudes no matter how it is defended, in particular the platitudes connected with the intrapersonal nature of the priority view. These failures are so extensive that I conclude that given the restriction to a risk-free framework, no theory vindicates the prioritarian platitudes well enough for the priority view to be well-defined. But despite this negative conclusion, three positive methodological lessons emerge, to do with risk, measurement, and axiomatization.

Consider first axiomatization. Thanks to Harsanyi and others, we have a gold standard for developing and defending views about the ethics of distribution: the informal articulation and defense of natural and appealing ethical ideas; the formalization of such ideas into claims suitable to be premises of formal theory, with any problematic presuppositions made fully explicit; and proof of a representation theorem characterizing the theory in question. It is unfortunate that much work in philosophy about the ethics of distribution does not look beyond the first stage. In the case of discussions of the priority view, this has had three bad effects. First, the characterization of egalitarianism as, for example, accepting that equality is intrinsically valuable, falls far short of being a suitable premise of a formal theory. It needs to be interpreted in a way which has concrete and unambiguous implications about betterness. The problems we have seen with the egalitarian half of the contrastive defense come from neglecting that task. Second, the utilitarian half of the contrastive defense suffers from neglecting the fact that we do have a good axiomatic defense of utilitarianism, the one provided by Harsanyi. Third, the direct defense rests on several claims couched in terms of evaluative measures. But when we make explicit the presuppositions needed for these claims to be well-defined and to deliver the theorems the direct defense requires, it becomes clear that the direct defense is highly problematic.

Consider next measurement. In the physical and social sciences, there is a fairly standard view about measurement: at the more fundamental level there are just various sorts of qualitative structures which can be described in measurement-free terms, while quantitative measurements emerge at a less fundamental level as useful descriptions of those qualitative structures. Which measurements are meaningful and the extent to which they are unique depends on the nature of the underlying

qualitative structures (see e.g. Krantz et al 1971; Narens 2007; Roberts 2009). It is odd that much theorizing about the ethics of distribution takes the reverse approach. Evaluative measures (e.g. goodness measures and individual goodness measures) are regarded as exogenously supplied and tacitly assumed to be reasonably unique, then put to work in the articulation and defense of distributive theories whose ultimate output is qualitative, namely an account of when one situation is better than another. The direct defense of the risk-free definition of the priority view provides a good example. I believe that there is no good reason for the reversal of direction taken in moral philosophy.¹⁰ While it would take a separate paper to argue that this reversal is a mistake, the difficulties the direct defense faces in virtue of the assumptions it makes about evaluative measurement may serve as a useful case study.

Consider finally risk. Given that Parfit and much of the subsequent literature on the priority view never mentions risk, trying to define the priority view in terms of risk from the outset would have seemed a fool's errand. But apart from the difficulties we have seen with risk-free approaches to the priority view, we have seen some positive reasons for approaching the priority view in terms of risk. The essentially risk-involving weak equality principle is most likely still some distance from characterizing the core of egalitarianism; roughly speaking, it only offers an interpretation of "equality is intrinsically valuable", but egalitarians will also want to say "less inequality is intrinsically better than more inequality". Nevertheless, the weak equality principle already appears to be something the priority view should reject, so pursuing this may lead to the antiegalitarian component of the priority view. In turn, articulating the antiutilitarian component of the priority view is going to have to pay attention to the premises of Harsanyi's theorem, but those too are deeply to do with risk.

Now it might be objected: even if we eventually have to consider risk when developing an account of the priority view, it does not follow that we cannot make a lot of progress by first considering the risk-free case. But in response, I think it is quite surprising how little progress we can make in a risk-free framework.

The only concrete risk-free claim about the priority view we have examined which survives all the various criticisms of risk-free approaches is as follows. Say that a function f on individual/history pairs represents the risk-free extended individual betterness relation just in case for any individuals i and j and any histories h_1 and h_2 , $f(i, h_1) \geq f(j, h_2)$ iff h_1 is at least as good for i as h_2 is for j . Now although it is far from obvious whether egalitarianism has to reject strong separability, it seems reasonably clear that the priority view should accept it. The claim which then survives is that according to the priority view,

- (7) Given a population of individuals $1, \dots, n$ of least three people, for some function f which represents the risk-free extended individual betterness relation, the risk-free betterness relation is represented by $f_1 + \dots + f_n$, where $f_i(h) := f(i, h)$.

¹⁰ A broadly similar view is expressed in Broome (1991) and Hammond (1991).

Assuming a rich domain, this follows along the lines of the derivation of Theorem 2 from the assumption that according to the priority view, the risk-free principle of personal good is true, and the risk-free betterness is anonymous, continuous, and strongly separable. However, this achieves very little. For (7) is accepted by utilitarianism, and is not obviously not accepted by egalitarianism. Moreover, there is something potentially misleading about characterizing the priority view in terms of strong separability. Utilitarianism implies strong separability, but strong separability is not one of the Harsanyi premises, nor is it implied by any single one of those premises. Thus for utilitarians, strong separability has a derived status, and one might well suspect that the same should be true for the priority view. But investigating this further would involve looking closely at how strong separability emerges from the Harsanyi premises, and that would mean looking closely at risk.

For what it is worth, I believe that there is an interesting way of trying to understand both the antiutilitarian and antiegalitarian components of the priority view in terms which essentially involve risk which is more promising than any risk-free approach. But whether the priority view does turn out to be an important new position on the ethics of distribution seems to me to be very unclear. This needs separate discussion.¹¹

Acknowledgments I am extremely grateful to three anonymous referees for extensive comments. Work on this article was partially supported by an AHRC Research Leave grant, and is gratefully acknowledged.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Appendix: A Proof of Proposition 1

Assuming *RF1* through *RF5*, we wish to show the left to right direction of the biconditional in Proposition 1. So assume that the betterness relation is anonymous, the risk-free principle of personal good is true, and that benefits have absolute and declining marginal value. We wish to show that for some strictly increasing and strictly concave w , $P(w)$ is a goodness measure. The case $n = 1$ is trivial so assume $n > 1$. Let the interval I be as per the definition of a rich domain.

By *RF1* let G^* be a goodness measure, unique up to positive affine transformation. By anonymity we can write $G^*(h)$ as $G(x_1, \dots, x_n)$ where $[x_1, \dots, x_n]$ is the individual goodness profile of h . Since benefits are absolute we have that for all $x_1, \dots, x_n, w_1, y_1, \dots, y_n \in I$

$$G(x_1, x_2, \dots, x_n) - G(w_1, x_2, \dots, x_n) = G(x_1, y_2, \dots, y_n) - G(w_1, y_2, \dots, y_n)$$

This implies that $G(x_1, x_2, \dots, x_n) - G(w_1, x_2, \dots, x_n)$ is independent of x_2, \dots, x_n so we can write it as $H_1(x_1, w_1)$. Hence $G(x_1, x_2, \dots, x_n) = H_1(x_1, w_1)$

¹¹ I attempt it in McCarthy (2011).

+ $G(w_1, x_2, \dots, x_n)$. Fix w_1 to obtain functions $h_1(x_1) =_{df} H_1(x_1, w_1)$ and $g_2(x_2, \dots, x_n) =_{df} G(w_1, x_2, \dots, x_n)$ so

$$(8) \quad G(x_1, x_2, \dots, x_n) = h_1(x_1) + g_2(x_2, \dots, x_n)$$

Since benefits are absolute we also have

$$\begin{aligned} G(x_1, x_2, x_3, \dots, x_n) - G(x_1, w_2, x_3, \dots, x_n) \\ = G(y_1, x_2, y_3, \dots, y_n) - G(y_1, w_2, y_3, \dots, y_n) \end{aligned}$$

By (8) it follows that

$$g_2(x_2, x_3, \dots, x_n) - g_2(w_2, x_3, \dots, x_n) = g_2(x_2, y_3, \dots, y_n) - g_2(w_2, y_3, \dots, y_n)$$

A similar argument then shows that there are functions $h_2(x_2)$ and $g_3(x_3, \dots, x_n)$ such that

$$g_2(x_2, \dots, x_n) = h_2(x_2) + g_3(x_3, \dots, x_n)$$

Hence

$$G(x_1, x_2, \dots, x_n) = h_1(x_1) + h_2(x_2) + g_3(x_3, \dots, x_n)$$

Proceeding in this way we obtain functions h_3, \dots, h_n such that

$$G(x_1, x_2, \dots, x_n) = h_1(x_1) + h_2(x_2) + \dots + h_n(x_n)$$

Anonymity then implies that the h_i 's are identical. Write them as w . The risk-free principle of personal good implies that w is strictly increasing. And since benefits have declining marginal value it follows that for all $y > x$, $w(\frac{x+y}{2}) - w(x) > w(y) - w(\frac{x+y}{2})$, hence w is strictly concave. Thus for some strictly increasing and strictly concave function w , G is identical to $P(w)$ as required. \square

References

- Broome, J. (1989). What's the good of equality?, In J.D. Hey (Ed.) *Current issues in microeconomics*. (pp. 236–262). New York: Macmillan.
- Broome, J. (1990–1991). Fairness, *Proceedings of the Aristotelian Society* 91, 87–102.
- Broome, J. (1991). *Weighing goods*. Oxford: Blackwell.
- Broome, J. (2004). Equality versus priority: a useful distinction, In D. Wikler & C. Murray (Eds.), *Fairness and goodness in health*. Geneva: World Health Organization.
- Brown, C. (2007). Prioritarianism for variable populations, *Philosophical Studies* 134, 325–361.
- Fleurbaey, M. (2004). Equality versus priority: how relevant is the distinction?, In D. Wikler & C. Murray. (Eds.) *Fairness and goodness in health*. Geneva: World Health Organization.
- Gibbard, A. (2008). *Reconciling our aims*. Oxford: Oxford University Press.
- Hammond, P. (1991). Interpersonal comparisons of utility: why and how they are and should be made, In J. Elster & J. Roemer (Eds.) *Interpersonal comparisons of well-being* (pp. 200–254). Cambridge: Cambridge University Press.
- Harsanyi, J. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *Journal of Political Economy* 63, 309–21.
- Hurley, S. (1989). *Natural reasons* Oxford: Oxford University Press.
- Jackson, F. (2000). *From metaphysics to ethics: A defense of conceptual analysis*. Oxford: Oxford University Press.

- Jensen, K.K. (2003) What is the difference between (moderate) egalitarianism and prioritarianism?, *Economics and Philosophy* 19, 89–109.
- Krantz, D., Luce, R.D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement. Vol. 1.* New York, Academic Press.
- Lewis, C.I. (1946). *The analysis of knowledge and valuation.* La Salle: Open Court Publishing.
- McCarthy, D. (2008). Utilitarianism and prioritarianism II, *Economics and Philosophy* 24, 1–33.
- McCarthy, D. (2011). The priority view, manuscript. Pokfulam: University of Hong Kong, Department of Philosophy.
- McCarthy, D. (forthcoming). Distributive equality, *Mind*.
- McKerlie, D. (2001). Dimensions of equality, *Utilitas* 13, 263–288.
- Myerson, R. (1981). Utilitarianism, egalitarianism, and the timing effect in social choice problems *Econometrica* 49, 883–897.
- Nagel, T. (1970) *The possibility of altruism.* Princeton: Princeton University Press.
- Narens, L. (2007). *Introduction to the theories of measurement and meaningfulness and the use of symmetry in science* London: Lawrence Erlbaum Associates.
- Parfit, D. (1984). *Reasons and persons* Oxford: Oxford University Press.
- Parfit, D. (2000). Equality or priority?, In M. Clayton & A. Williams (Eds.) *The ideal of equality.* (pp. 347–386). New York, Macmillan.
- Parfit, D. (2011). *On what matters.* vol. 1 New York: Oxford University Press.
- Rabinowicz, W. (2001). Prioritarianism and uncertainty: on the interpersonal addition theorem and the priority view, In D. Egonsson, J. Josefsson, B. Petersson & T. Rønnow-Rasmussen (Eds.) *Exploring practical philosophy: From action to values.* Ashgate: Aldershot.
- Rabinowicz, W. (2002). Prioritarianism for prospects, *Utilitas* 14, 2–21.
- Rawls J. (1971). *A theory of justice.* Cambridge: Harvard University Press.
- Resnik, M. (1987). *Choices: An introduction to decision theory.* Minneapolis, MN: University of Minnesota Press.
- Roberts, F. (2009). *Measurement theory* Cambridge: Cambridge University Press.
- Segal, U. (2006). Fair bias, *Economics and Philosophy* 22, 213–29.
- Sen, A. (1973) *On economic inequality.* Oxford: Clarendon Press.
- Temkin, L. 1993. *Inequality.* Oxford: Oxford University Press.
- Tungodden, B. (2003). The value of equality, *Economics and Philosophy* 19, 1–44.
- Wakker, P. (1989). *Additive representations of preferences.* Dordrecht: Kluwer.