Priority, Preference and Value:
A Response to Otsuka and Voorhoeve

MARTIN O'NEILL
E-MAIL: MARTIN.ONEILL@MANCHESTER.AC.UK

MANCHESTER CENTRE FOR POLITICAL THEORY
POLITICS – SCHOOL OF SOCIAL SCIENCES
UNIVERSITY OF MANCHESTER
MANCHESTER M13 9PL U.K.

In a recent article, Michael Otsuka and Alex Voorhoeve mount a significant challenge to the Priority View.[1] By means of a number of ingenious examples, they seek to show how prioritarianism is misguided, and fails to account for our considered intuitions in certain important cases. I shall argue that their challenge to the Priority View is unsuccessful. Nevertheless, it is both instructive and illuminating to consider the reasons for the failure of their challenge.

By understanding the ways in which Otsuka and Voorhoeve's arguments fail, we can come to a clearer view of a number of important issues in what Derek Parfit calls "the ethics of distribution".[2] Among these are the relationship between priority, preference and value; the relationship between prioritarianism and risk; questions of how reasonableness is (and is not) related to the moral value of outcomes; and issues regarding the ways in which well-being is (and is not) related to individual preferences. Assessing the shortcomings of Otsuka and Voorhoeve's arguments also sheds light on the perils of using empirical survey data in political philosophy, thereby addressing important questions of philosophical methodology.

---

[1] Michael Otsuka and Alex Voorhoeve, "Why it Matters that Some are Worse Off than Others: An Argument Against the Priority View," *Philosophy & Public Affairs*, (forthcoming, 2009). (All page references are to a pre-publication version of this article, dated December 1, 2008.) For the formulation of the Priority View, as targeted by Otsuka and Voorhoeve, see Derek Parfit (1991), "Equality or Priority?" delivered as the *Lindley Lecture* at the University of Kansas, (Lawrence: University of Kansas Press), and reprinted in Matthew Clayton and Andrew Williams, (eds.), (2000), *The Value of Equality*, (Basingstoke: Palgrave Macmillan), pp. 81-125.
[2] Derek Parfit, (1991), p. 82.

I. Otsuka and Voorhoeve's Rejection of the Priority View

Otsuka and Voorhoeve imagine a pair of rather complicated cases in which one or more individuals are threatened by a nasty medical condition. The core of their critique of the Priority View is motivated by considering the first of these cases, where the unpleasant fate is faced by a single individual. Given its significance for the construction of their challenge to the Priority View, I shall concentrate on this first case in my discussion.

The Risky One-Person Case[3]

There is "a young adult who is now in perfect health but who receives the distressing news that she will soon develop one of the following two mobility-affecting conditions and has a 50 percent chance of developing each."[4] The two conditions are:

Condition (a) *Slight impairment*: a condition that renders it difficult for one to walk more than 2 km.

Condition (b) *Very severe impairment*: a condition that leaves one bedridden, save for the fact that one will be able to sit in a chair and be moved around in a wheelchair for part of the day if assisted by others.

Our patient (let's call her Clara, with a nod to Johanna Spyri's *Heidi*) can choose between taking two different treatments, each of which can only be taken ahead of time (before she knows which condition she will develop), and which are mutually exclusive in their effects (i.e. she cannot take both or, if she does, then neither will be effective). The first treatment (Treatment (1)) would completely cure the condition that leads to (a) *Slight Impairment*, moving Clara instead to Condition (c) *Full Health*, but it would offer no help against Condition (b) *Very Severe Impairment*.

[3] The name for this case is my own.
[4] Otsuka and Voorhoeve, p. 1.

The second treatment (Treatment (2)) would have no effect against the condition that leads to (a) *Slight Impairment*, but would be somewhat effective against the condition that leads to (b) *Very Severe Impairment*, lessening the effect of the condition so that it leads only to:

Condition (d) *Severe impairment*: a condition in which one is no longer bedridden; rather one is able to sit up on one's own for the entire day but requires the assistance of others to move about.

(Given the descriptions of the cases, therefore, it seems unproblematic to assume that the ordering of the four conditions, from best to worst, is (c)>(a)>(d)>(b).)

Otsuka and Voorhoeve report that surveys show that "people who imagine themselves in such a predicament would typically be indifferent" between receiving Treatment (1) or Treatment (2).[5] In other words, such people are "typically" indifferent between, on the one hand, a gamble that gives a 0.5 chance of being in Condition (a) and a 0.5 chance of being in Condition (d), and, on the other hand, a gamble that gives them a 0.5 chance of being in Condition (b) and a 0.5 chance of being in Condition (c). They further claim that (following "the orthodox von Neumann-Morgenstern preference-based measure of utility"[6]) (i) the expected utility of both treatments is therefore equal, and (ii) the utility gain in moving from condition (a) to condition (c) is equal to the utility gain in moving from condition (b) to condition (d).

Let us, for the sake of argument, grant Otsuka and Voorhoeve's claims about the consequences, in terms of individual utility, of taking either treatment. Let us further grant their assumption that the survey data to which they refer is credibly informative with regard to people's informed preferences on these matters. (Although I shall turn, in Sections VI and VII, to questioning these two assumptions.) Assuming, at any rate, that our patient ("Clara") is indifferent between taking the two

[5] Otsuka and Voorhoeve, p. 2.

[6] Otsuka and Voorhoeve, pp.2-3. On the von Neumann-Morgenstern axioms for expected utility theory, see John von Neumann and Oskar Morgenstern, (1953), *Theory of Games and Economic Behaviour*, 3rd edition, (Princeton: Princeton University Press).

treatments, Otsuka and Voorhoeve imagine the choice that you would face if, as a "morally motivated stranger,"[7] you were able to provide either of the treatments to Clara.

In this situation, Otsuka and Voorhoeve claim that:

(A) It would be reasonable for you (i.e. for "a morally motivated stranger") to share Clara's indifference between these two treatments.

They further claim that, if Clara had preferred *either* Treatment (1) to Treatment (2), or *vice versa*, it would similarly be reasonable for you to act in accord with her preferences. The background claim here, therefore, is:

(B) It would be reasonable for you (i.e. for "a morally motivated stranger") in this situation "to provide a treatment that maximizes the expected increase in the utility of the recipient."[8]

Otsuka and Voorhoeve claim that Claim (B) holds in any circumstance in which there are no relevant concerns regarding the *comparative* well-being of others. This is the core of their case against the Priority view. If Claim (B) is correct, argue Otsuka and Voorhoeve, then the Priority view is false.[9]

Their argument would appear to run like this. The Priority view, as a noncomparative view about the diminishing marginal moral significance of increases in well-being, suggests that, even in a one-

[7] Otsuka and Voorhoeve, p. 3. Otsuka and Voorhoeve are careful to emphasize that the "morally motivated stranger" "is a private individual rather than a state official." (fn. 4)
[8] Otsuka and Voorhoeve, p. 4
[9] Otsuka and Voorhoeve contrast the Risky One-Person Case (i.e. Clara's Case) with (what I shall call) a "Non-Risky Many-Person Case," in which two identically sized groups each face with certainty one of the two conditions faced by Clara (i.e. the members of one identifiable group will develop Condition (a) and the members of the other group will develop Condition (b)). (See Otsuka and Voorhoeve, p. 4.) In this Non-Risky Many-Person case, Otsuka and Voorhoeve agree with the advocate of the Priority View that a "morally motivated stranger" should help the group that are facing *Very Severe Impairment* (Condition (b)) rather than aiding the group who face only *Mild Impairment* (Condition (a)). As their judgement in this case does not depart from the prioritarian view of the situation, I shall not discuss this case in any detail.

person case, there is greater moral significance to the utility gain associated with moving from Condition (b) to Condition (d) than in moving from Condition (a) to Condition (c).[10] (Given the uncontested and plausible assumption that the utility-ordering of the four states, from best to worst, is (c)>(a)>(d)>(b).) This implication of the Priority view seems to be inconsistent with Claim (B), which Otsuka and Voorhoeve treat as a fixed point on our moral reasoning, and therefore the Priority view is to be rejected.

Despite its initial plausibility, this argument is unsuccessful. It can be rejected in one of two ways. The first is simply to reject Claim (B), while the second is to deny that accepting the Priority view need be inconsistent with accepting Claim (B). Both lines of argument are independently plausible, depending on how Claim (B) is to be understood. I shall make the case for rejecting Claim (B) in the following two sections, and then, in Section IV, discuss how, alternatively, Claim (B) can be interpreted, *pace* Otsuka and Voorhoeve, as actually being consistent with the Priority View. If either strategy is effective, then we should reject Otsuka and Voorhoeve's argument against the Priority View.

II. Priority, Preference and Value

The simplest way of rejecting Otsuka and Voorhoeve's argument is to reject their assumption that Claim (B) should be treated as a fixed point in our moral thinking, which needs to be accommodated by any plausible distributive view. Indeed, Claims (A) and (B) receives no argumentative support in Otsuka and Voorhoeve's discussion, instead of simply being treated as a fixed points of this kind. We are simply told that "it would be reasonable for you to share her indifference"[11] and that "it would be

[10] Otsuka and Voorhoeve's presentation of prioritarianism as a *noncomparative* view accords with the formulation of the view in Parfit (1991). As Parfit puts it, "on the Priority View, benefits to the worse off matter more, but that is only because these people are at a lower *absolute level*. It is irrelevant that these people are worse off *than others*." (p. 104) Parfit makes the analogy with altitude: "People at higher altitudes find it harder to breathe. Is it because they are higher up than other people? In one sense, yes. But they would find it just as hard to breathe even if there were no other people who were lower down." (p. 104). See Otsuka and Voorhoeve, p. 7.

[11] Otsuka and Voorhoeve, p. 3.

reasonable for you to provide a treatment that maximizes the expected increase in utility".[12] The nearest that these claims comes to finding support is in the empirical survey data that Otsuka and Voorhoeve cite, but the issue at stake is of course not a matter of what people *happen to think* about such cases, but a normative issue regarding what we *should* believe about such cases. It is no part of the Priority View that it accounts for, or even accords with, all of people's pre-theoretic intuitions, and it is therefore hardly an objection to the Priority view to point out that it clashes with some common intuitions, as revealed by survey data.[13]

Perhaps, then, it is true that "the Priority View as formulated by Parfit cannot account for our judgement in the one person case,"[14] but this may be because the Priority view actually gives us good reason to reject this intuitive judgement. Consider these two lemmas of the Priority View:

> (C) Benefiting people matters more the worse off those people are.
>
> (D) The goodness of some benefit enjoyed by an individual diminishes as the well-being of that individual increases.[15]

These are claims about the moral value of gains in well-being, considered from the impersonal point of view, rather than from the viewpoint of the agent whose well-being is in question. Taking the case of Clara, Claim (C) tells us that, in moving her from Condition (b) to Condition (d), we would be doing something that mattered more, morally speaking (and hence from the impersonal standpoint), than would moving Clara from Condition (a) to Condition (c). The fact that Clara is indifferent between these two changes in condition tells us only (*ex hypothesi*)[16] that, for her, the magnitude of these two gains in well-being are equal. It tells us nothing about the further question, which is the

[12] Otsuka and Voorhoeve, p. 4.
[13] I shall say more about the perils of using survey data in Section VII.
[14] Otsuka and Voorhoeve, p. 32.
[15] Claims (C) and (D) reproduce the prioritarian claims (J) and (K), as characterized in Martin O'Neill, (2008), "What Should Egalitarians Believe?" *Philosophy & Public Affairs*, 36: 2, p. 152.
[16] That is, if Otsuka and Voorhoeve are correct about the relationship between Clara's preferences and her well-being. On this, see Section VI below.

salient question at hand, of the impersonal *value* or moral significance of these two potential changes in well-being.

Similarly, Claim (D) tells us that the goodness, or value, of some benefit (such as the change in Clara's condition brought about by each of the two treatments) is inversely correlated with the level of the recipient's well-being. Thus, if Clara turns out to have been unlucky, and to have the *Very Severe Impairment* (Condition (b)), then a gain in her well-being of any given magnitude, x, will be of greater moral value than a gain of x would have been if Clara were to have been (comparatively) lucky, and had only suffered the *Mild Impairment* (i.e. Condition (a)). Thus, it is not at all the case that the Priority view here *fails* to accommodate an obvious or uncontroversial moral truth; rather, it takes a view that simply and explicitly *rejects* Otsuka and Voorhoeve's Claims (A) and (B).

The Priority View, as expressed by claims (C) and (D), is a view about the relationship between individual well-being and impersonal moral value. It specifies a strong link between the two, given that, on the Priority view, all gains in well-being are morally valuable. But it denies that these gains are straightforwardly aggregative, instead viewing their moral significance as being indexed to the level of well-being of their recipient. Hence, the Priority View explains why it would not, *pace* Otsuka and Voorhoeve, be reasonable to act so as to maximize Clara's *ex ante* expected utility, given that the impersonal significance of gains in Clara's utility will crucially depend on her *ex post* level of well-being. In other words, according to the Priority view, future improvements in Clara's well-being would matter more if she will have the *Very Severe Impairment* than they would do if she will have only the *Mild Impairment*. It is clearly not the case therefore that the Priority View has missed the uncontroversial truth of Claims (A) and (B); rather, it simply denies the truth of these claims, and does so on the basis of a clear and compelling account of the relationship between the value of gains in well-being for individual persons, and the moral value of those gains in well-being viewed from an impersonal standpoint.

It is puzzling that Otsuka and Voorhoeve treat Claims (A) and (B) as so uncontroversial when one considers what they have to say, later in their article, against the plausibility of "*ex ante* prioritarianism" (that is, the nonstandard version of the Priority View that ranges over agents' *ex ante* prospects rather than their really-achieved levels of well-being).[17] They acknowledge that, as in Parfit's standard formulation of the Priority View, "outcomes rather than prospects are carriers of moral value".[18] But, if this is the case, then it is straightforward to see why we should be more concerned to have a chance of improving the position of "Unlucky Clara" (who suffers the *Very Severe Impairment*) rather than being concerned to improve the position of "Lucky Clara" (who suffers only the *Mild Impairment*). Therefore, we have good reason *not* to take Clara's *ex ante* preferences over two sets of prospects (i.e. between [(a) or (d)] *or* [(b) or (c)]) as being fully authoritative in deciding how, morally speaking, we ought to treat her.

This is not, of course, to say that the Priority View is completely cavalier about people's preferences, or that it gives no emphasis to consideration of what is important or valuable for particular individuals, from their own perspective. Insofar as the Priority View is concerned with the maximization of weighted utility, it is concerned with whatever contributes to the utility of any particular individual, and will thus take an interest in individual's preferences and judgements of what is valuable *for them* insofar as these are reliable guides to (or determinants of) their well-being. But information about relevant individual preferences of this kind provides us only with 'inputs' that are to be weighted according to the prioritarian calculus of diminishing moral significance, rather than providing us (as in Otsuka and Voorhoeve's Claim (B)) with a fully authoritative standard for how, morally speaking, we should aim to treat other people. First-person conceptions of value or significance are given some, but only *some*, role within the prioritarian calculus; but they do not carry the day, with unimpeachable authority, as suggested by Otsuka and Voorhoeve's Claim (B).

[17] On "*ex ante* prioritarianism", see Larry G. Epstein and Uzi Segal, "Quadratic Social Welfare Functions," *Journal of Political Economy* 100 (1992): 691-712.

[18] Otsuka and Voorhoeve, p. 29. In defending the standard reading of the Priority View, Otsuka and Voorhoeve quote from a pre-publication version of Derek Parfit (forthcoming), *On What Matters* (manuscript dated September 24, 2008), §21. Parfit here characterizes the expected goodness of an act as "the goodness of these possible effects multiplied by the chance that this act would have those effects."

Prioritarianism treats the split between the personal and impersonal standpoints seriously, even in one-person cases. Otsuka and Voorhoeve's Claim (B), by contrast, implicitly denies that there is a normatively significant shift between these two standpoints. Some remarks of Thomas Nagel's are potentially illuminating regarding the existence of divergent standpoints. As Nagel puts it:

> There are so many people one can barely imagine it … but what happens to each of them is enormously important – as important as what happens to you. The importance of their lives to them, if we really take it in, ought to be reflected in the importance their lives are perceived to have from the impersonal standpoint, even if not all elements of those lives will be accorded an impersonal value corresponding to its personal value for the individual whose life it is...[19]

Clara's preferences are important, because they (let us assume, following Otsuka and Voorhoeve) determine her level of well-being in the various Conditions (a)-(d). They should therefore be accounted for when we decide, from the impersonal standpoint, what we should do, morally speaking. But, while we should have regard to Clara's preferences, insofar as they determine (or at least track) her level of well-being under different potential outcomes, we cannot simply "read off" the impersonal moral value of those different potential outcomes directly from Clara's preferences. Rather, if the Priority View is right about the diminishing moral significance of gains in utility, we should be more concerned to provide a utility-benefit of magnitude x to Clara if she turns out to be unlucky (and suffers the *Very Severe Impairment*) than if she turns out to be comparatively lucky (and suffers only the *Mild Impairment*).

The structure of the Priority View illustrates the important moral distinction between value *for particular individuals* and moral value considered from an impersonal standpoint. In endorsing Claims (A) and (B), Otsuka and Voorhoeve seem to be ignoring the significance of this distinction. Conversely, if we take this distinction seriously, then it is clear that we can have good reason to reject Claims (A) and (B), and thereby to abandon the core of Otsuka and Voorhoeve's case against to the Priority View.

[19] Thomas Nagel, (1991), *Equality and Partiality*, (Oxford: Oxford University Press), p. 12. See also Thomas Nagel, (1986), *The View from Nowhere*, (Oxford: Oxford University Press), Chapter IX.

III. Priority and Risk

The previous section suggested that the recommendations of the Priority View are plausible even in a one-person case, when we bear in mind the distinction between the value of some outcome for a particular individual and the impersonal moral value of that outcome. It also reveals an interesting relationship between the Priority View and the proper treatment of risk. It reveals that, if we accept the central claims of prioritarianism (e.g. as embodied in Claims (C) and (D)), then, when judging the prospect of a number of possible outcomes from the impersonal standpoint, we should be relatively risk-averse. Under conditions of risk, belief in the greater moral significance of gains to the worst-off is transformed, as it were, into a belief in the justifiability of a certain degree of risk-aversion.

Let us imagine that Clara, in being indifferent between the two treatments, was risk-neutral. She was concerned only to maximize her expected gain in utility and given that, *ex hypothesi*, the utility gains in moving (i) from Condition (a) to Condition (c), or (ii) from Condition (b) to Condition (d), were identical, Clara was indifferent between taking either of the two treatments. Now, in having risk-neutral preferences, Clara can acknowledge that, of course, ending up in Condition (c) would be much better than ending up in Condition (d), and that she would be much luckier if she has only to face Condition (a) rather than facing Condition (b). But these considerations are simply irrelevant to the choices of the rational utility-maximizer that we here take Clara to be. Given that she does not know whether she will be lucky or unlucky, and has no control over which eventuality she will face, Clara is rationally concerned only to maximize her expected gain in well-being. Given that the gain is equal under both treatments, Clara is indifferent between them.

However, even if Clara is risk-neutral in this way, the view of her prospects from the impersonal standpoint, insofar as that standpoint is characterized by acceptance of the Priority View, will be more risk-averse. Under prioritarian assumptions regarding the diminishing moral significance of gains in utility, gains to the “lucky” Clara who ends up in Condition (a) *matter less* than would gains to the

unlucky Clara who ends up in Condition (b). Therefore, it is more important to secure a gain to "unlucky" Clara than to "lucky" Clara. So, even if the purely risk-neutral strategy favoured by Clara is justifiable in terms of individual prudence, the moral view from the impersonal standpoint is that there is more value in securing the move from Condition (b) to (d) than in securing the move from Condition (a) to (c). Therefore, a "morally motivated stranger" who has reason to maximize impersonal moral value, has reason to adopt a strategy, with respect to Clara's treatment, that is more risk-averse than the strategy that Clara may have reason to choose for herself.

Now, it may be that we are assuming too much in assuming that Clara's current preferences are risk-neutral. Perhaps she is so fixated on the attractions of *Full Health* that she is prudentially "over-valuing" Condition (c), and therefore being risk-seeking or risk-loving. (Although, on Otsuka and Voorhoeve's preferred von Neumann-Morgenstern measure of utility, it is true by definition that Clara's indifference between two options shows that both options embody the same expected utility.[20]) If Clara is "over-valuing" Condition (c), then of course we have all the more reason not to treat her preferences as authoritative, and instead to maximize the expected moral value of the outcome.

However, even if Clara herself were risk-averse to some degree, such that her preferences over the two treatments reflected her particular dread of the prospect of Condition (b), and such that she therefore over-estimated the prudential badness of that condition, we still have no reason to treat her particular prudential preferences as having any final authority with regard to which treatment we ought, morally speaking, to give to her. The Priority View, after all, proposes what we might think of as a function for transforming assessments of utility into assessments of moral value or impersonal significance. The particular preferences of agents are therefore relevant, for the Priority View, only insofar as they constitute or track the real level of utility associated with different possible outcomes. The impersonally 'proper' degree of risk-aversion to be exhibited by a "morally motivated stranger",

[20] This in itself might be thought to provide a substantial problem of plausibility for the von Neumann-Morgenstern measure preference-based measure of utility. I shall discuss this and related issues further in Section VI.

with regard to choices among any set of possible outcomes, will then be given straightforwardly as a consequence of the different moral weightings of the different levels of well-being associated with each outcome. The shape of the function that links utility to value will uniquely determine the ways in which a certain degree of risk-aversion can be normatively justifiable, and will do so in a way that is structurally independent of the level of risk-aversion of the agent whose well-being is in question.

In his argument from the Original Position, Rawls suggests that, in choosing under uncertainty (i.e. a risky situation where we have no information about the probabilities of different possible outcomes), we have reason to adopt a maximin decision rule.[21] This leads to the choice of principles that maximally advantage the worst-off. The current suggestion about priority, value and risk is, so to speak, the obverse of this relationship. Rawls moves, within the argument of the Original Position, from a decision-rule that embodies an (absolutist) form of risk-aversion, to a normative commitment to giving absolute priority to the worst-off. In the case of the prioritarian approach to Clara's risky condition, we instead move from a moral commitment to a (more moderate) form of priority to the worse-off to a general decision rule that embodies some (non-absolute) degree of risk-aversion. The direction of travel is reversed, but the underlying relationship between prioritarianism and risk-aversion is structurally similar.

Acceptance of the Priority View thereby mandates a "morally motivated stranger" to display a degree of risk-aversion over the range of outcomes faced by Clara. This risk-aversion follows from the greater significance of potentially avoiding Condition (b), and of aiding the potentially "Unlucky Clara", than of providing the same degree of aid to the "Lucky Clara" who falls into Condition (a). We might even say that, on the Priority View, given that outcomes rather than prospects are the bearers of moral value, the claims of particular individuals under conditions of risk are effectively proxies or stand-ins for the claims of the possible future individuals who will endure those outcomes. Clara's moral claims, therefore, in a sense devolve fully in this instance into the possible future claims

[21] See John Rawls (1971/1999), *A Theory of Justice*, (Cambridge, Mass.: Harvard University Press), §§26-28.

of "Lucky Clara" and "Unlucky Clara". In judging the potential moral value of giving either treatment, Clara's own indifference between the two treatments is therefore, so to speak, neither here nor there.

We should therefore again conclude that accepting the Priority View, and therefore also accepting the risk-aversion that goes with it, mandates the rejection of Claims (A) and (B). Instead, a "morally motivated stranger" who was concerned to maximize the expected moral value of his actions, should be sure to give Clara the second treatment. Given that the expected gain in utility is identical for either treatment, and is thereby held constant, expected moral value is maximized through giving the treatment that allows Clara to avoid the prospect of the worst of the four possible outcomes, saving her from a 50% chance of the *Severe Impairment* (Condition (b)). Moral value is maximized by giving priority to the potential future "Unlucky Clara" rather than the potential future "Lucky Clara", whose claims are weaker precisely because she would be better-off.

IV Reasonableness and the Moral Value of Outcomes

If the argument of the previous two sections is correct, then we should reject Claim (B), and Otsuka and Voorhoeve's challenge to the Priority View fails. But, as I have mentioned, one can resist Otsuka and Voorhoeve's challenge to the Priority View even without rejecting Claim (B), providing that one resolves an ambiguity in that claim. Let us recall the content of the claim.

(B) It would be reasonable for you (i.e. for "a morally motivated stranger") in this situation "to provide a treatment that maximizes the expected increase in the utility of the recipient."[22]

[22] Otsuka and Voorhoeve, p. 4

As stated, this is a claim about *reasonableness*.[23] It is not a claim about moral value. What it is *reasonable* for an individual to do, in a particular set of circumstances, will depend on a wide variety of facts about relevant personal relationships and the expectations of others, as well as various facts about the social and institutional background in which their action is to take place. These kinds of personal, social or institutional background conditions are stripped away when we consider what it would be reasonable for a "morally motivated stranger" to do, and we accordingly have only the sketchiest idea of what would definitively count as a *reasonable* course of action for an agent operating under such a sparse and under-specified set of conditions. In itself, this sketchiness might suggest that Otsuka and Voorhoeve's direct appeal to our intuitions about *reasonableness* in such cases might not present a promising way of proceeding. For it is difficult to say what counts as the *reasonable* thing to do when we know so little about the structure of the situation, and where the situation has been described in such a way as to bracket many of the most salient features of the moral and political landscape in which real agents operate. As Thomas Nagel put it, "neither ethics nor political theory have as their aim to provide advice to a powerful and benevolent outsider capable of affecting the welfare of human beings."[24] And accordingly there are potentially potent problems in appealing to our moral intuitions in cases where were are told to imagine ourselves as just exactly this kind of "powerful and benevolent" intervener.

Nevertheless, we can move beyond this problem of the underspecification of situation, and hence also the underspecification of the content of what would be *reasonable* in this situation, by entertaining some different hypotheses regarding what would count as the reasonable course of action for a morally motivated stranger of this kind. Here is one suggestion.

---

[23] On the idea of the reasonable, and its contrast with the idea of the rational, see T. M. Scanlon, (1998), *What We Owe to Each Other*, (Cambridge, Mass.: Harvard University Press), Chapter 1, esp. pp. 32-33, and Chapter 5, esp. pp. 191-97. See also John Rawls, (2001), *Justice as Fairness: A Restatement*, (Cambridge, Mass.: Harvard University Press), pp. 6-7.

[24] Nagel, (1991), p. 14.

Claim (E): The uniquely reasonable course of action for a morally motivated stranger to undertake is that course of action that maximizes the expected moral value of the consequences of his action.

If Claim (E) is true, then Claim (B) is false. Expected moral value is, as we have seen, maximized by giving the second treatment rather than the first treatment. But we can also find alternative ways of understanding the content of what it would be reasonable for a morally motivated stranger to do with regard to Clara's treatment. So, here is an alternative suggestion.

Claim (F): The reasonable course of action for a morally motivated stranger to undertake will be given by consideration of the justifiability of that action to the agents whom it may affect, where this justification is understood to depend on accommodating the significance of a broad range of values.

If Claim (F) is true, then Claim (B) may also be true. For, if Claim (F) is true, then it may well *not* be the case that the only reasonable course of action is that course of action that maximizes expected moral value. If we instead think that the most reasonable course of action will be the action that we can best justify to Clara herself, then we may think that we should subsume the promotion of the expected moral value of the (Clara-affecting) outcome, in favour of giving greater concern to communicating our respect of Clara herself as an autonomous agent, and thereby by giving greater weight to the authority of her own preferences.

Given that, *ex hypothesi*, Clara will be the only person affected by her condition, and by the choice of treatment that she will undergo, we may think that there are good reasons to let her determine her own choice of treatment, even if this will be at the cost of failing to maximize overall expected moral value. We may, for example, think it important that the particular risks faced by Clara reflect her own choices and tastes (and hence that there would be what Scanlon calls "representative value" in leaving the choice of treatment in Clara's hands), thereby ensuring that what happens to Clara is at least to

some degree a matter of her own choosing, rather than an outcome enforced from outside. We may also think that it would be demeaningly paternalistic or disrespectful for us to choose Clara's treatment in any way that did not faithfully track her own preferences (and hence that there would be what Scanlon calls "symbolic value" in leaving the choice in Clara's hands).[25]

It is important to note that following this strategy in interpreting the demands of *reasonableness* in this situation leads us only to an indirect form of support for Claim (B). For it would be true that we have reason to maximize the expected utility of the recipient (i.e. Clara, in this case) not because it would be the most valuable course of action, *per se*, but because taking the course of action that maximized Clara's expected utility might itself also be the best way of respecting her preferences, and thereby of acknowledging the way in which the most reasonable course of action will be the one in which we give special weight to the value of Clara's own choices. This indirect support for Claim (B) holds only insofar as we assume the relationship between preferences and utility that Otsuka and Voorhoeve themselves assume when they invoke the von Neumann-Morgenstern conception of utility. That is, on the von Neumann-Morgenstern conception of utility, individual preferences constitute (or, at the very least, unerringly track) expected utility.[26] Therefore, if we act so as to maximize expected utility, we do so by respecting the authority of the recipient's own preferences. But, given the representative and symbolic value of making sure that Clara's preferences are respected, our *reason* for maximizing her utility is derivative from the independent value of considerations that are conceptually separable from the promotion of utility itself. Thus, it may be reasonable to do what Claim (B) recommends, insofar as it is reasonable to provide a course of treatment that maximizes the expected increase in the utility of the recipient, but it is only reasonable insofar as the posited (von Neumann-Morgenstern) relationship between preference and utility holds, and it is reasonable *not* because it is always reasonable to maximize expected utility, but because the

[25] On representative and symbolic value, see T. M. Scanlon, (1998), pp. 251-3. See also T. M. Scanlon, (1988), "The Significance of Choice," in *The Tanner Lectures on Human Values*, volume 8, edited by Sterling McMurrin (Salt Lake City: University of Utah Press), pp. 149-216.
[26] I shall discuss these issues in more detail in Section VI.

demands of justifiability mean that it is often reasonable to respect the authority of the choices and preferences of other people on matters that affect their lives.

Most significantly, it should be emphasized that this (indirect) acceptance of Claim (B) is not remotely inconsistent with the endorsement of the Priority View. If we reject Claim (E) and instead accept Claim (F), and thereby decouple our view about *reasonableness* from our conception of the moral value of outcomes, then we can endorse Claim (B) while simultaneously endorsing the Priority View, and its consequences as captured by Claim (C) and Claim (D). This is because, if we recall Claims (C) and (D), we see that the Priority View is a view about the moral value of outcomes, and not a view about all-things-considered reasonableness. Therefore, if we understand all-things-considered reasonableness to involve something other than, and separable from, the promotion of overall expected moral value, then we can simultaneously hold, without any danger of inconsistency, both that it is *reasonable* to track Clara's preferences over her choice of treatment, but also that tracking her preferences in this way will nevertheless fail to maximize the expected moral value of the outcome. Tracking Clara's preferences will fail to maximize expected moral value given the consequence of the Priority View that the potential gain in utility in moving from Condition (b) to Condition (d) is more valuable than the potential gain in utility in moving from Condition (a) to Condition (c). Thus, if we accept Claim (F), we can endorse all of Claims (A), (B), (C) and (D) together consistently, and thereby Otsuka's and Voorhoeve's challenge to the Priority View will again be seen to fail, albeit in a different way to that discussed in the previous two sections.

Now, Otsuka and Voorhoeve do consider an objection to their challenge to the Priority View that is somewhat related to the points made in this section, when they consider whether it is the value of *autonomy* that explains our reluctance to go against Clara's wishes in the Risky One-Person Case. Acknowledging the distinct value of autonomy in this way could be consistent with endorsing the Priority View as a view about the moral value of outcomes. Otsuka and Voorhoeve's response to this line of objection from the value of autonomy is unconvincing. They suggest that "[Clara's] autonomy … would not be threatened if our morally motivated stranger acted contrary to [her] wishes by giving

her the treatment that maximizes expected priority-weighted utility instead of the treatment she prefers."[27] They hold that there is no threat to autonomy here because "the morally motivated stranger" "rightfully possesses each of the treatments" and is thereby entitled to provide whichever he chooses to Clara, and moreover he thereby "would not interfere with [Clara's] exercise of any choices she has a right to make."[28]

In responding in this way, Otsuka and Voorhoeve rely on an implausibly narrow conception of individual autonomy, whereby nobody's autonomy is ever threatened by anyone else acting in such a way that does not violate their rights. Indeed, it is difficult to see how this could be a conception of *autonomy* at all, where autonomy is standardly understood to involve a positive conception of an agent's self-rule or self-direction, as against simply being a variety of negative liberty, understood as non-interference.[29] Moreover, Otsuka's and Voorhoeve's tacit conception of what counts as non-interference here is itself controversial, given their assumption, which has been notably and convincingly criticized by writers such as G. A. Cohen and Jeremy Waldron, that no instance of the rightful exercise of individual property rights could ever count as an instance of interference.[30]

One can at any rate find many cases where it would be a strain to say that the rightful exercise of one's property rights did not undermine (or at least fail to respect) the autonomy of others. If I am the rightful owner of a vast library of books, and you have nothing to read, it seems perverse to say that I

[27] Otsuka and Voorhoeve, p. 18.
[28] Otsuka and Voorhoeve, p. 18.
[29] On the distinction between negative liberty and positive conceptions of liberty involving ideas of autonomy, see Isaiah Berlin, "Two Concepts of Liberty," in David Miller, ed., (2006), *The Liberty Reader*, (Edinburgh: Edinburgh University Press). On the Kantian idea of autonomy, see Immanuel Kant, *Groundwork of the Metaphysics of Morals* (1785), translated by Mary Gregor, (Cambridge: Cambridge University Press, 1998), Third Section, 52-66 (Ak. 4: 446-463). See also Christine Korsgaard, "Morality as Freedom," in her *Creating the Kingdom of Ends*, (Cambridge: Cambridge University Press, 1996), and Onora O'Neill, "Reason and Autonomy in Kant's *Grundlegung* III" in her *Constructions of Reason: Explorations of Kant's Practical Philosophy*, (Cambridge: Cambridge University Press, 1989).
[30] See G. A. Cohen, "Freedom and Money," (Buenos Aires, Argentina: Universidad Torcuato di Tella, 2001), published online at http://www.utdt.edu/Upload/_115634753114776100.pdf. See also G. A. Cohen, "Capitalism, Freedom and the Proletariat," in *The Idea of Freedom: Essays in Honour of Isaiah Berlin*, ed. Alan Ryan (Oxford: Oxford University Press, 1979); G. A. Cohen, "Appendix: On Money and Liberty," in *Equality*, ed. Jane Franklin (London: Institute of Public Policy Research, 1997); Jeremy Waldron, 'Homelessness and the Issue of Freedom', *University of California Law Review* 39 (1991), and reprinted in his *Liberal Rights* (CUP, 1993), pp. 309-338; Jeremy Waldron, "Mr. Morgan's Yacht," in *The Egalitarian Conscience: Essays in Honour of G. A. Cohen*, ed. Christine Sypnowich (Oxford: Oxford University Press, 2006).

would not be displaying a greater concern with the value of your autonomy if I lent you whichever book you chose, rather than instead lending you a volume of my own choice. Perhaps, as the rightful owner of the library, I would not thereby have wronged you or violated your rights if I decided to dictate your supply of reading matter instead of letting you choose, but this is a different issue to the question of your autonomy. If it is your autonomy that concerns me then, even if I am happy to offer advice, I should give weight to your choices rather than providing you with whatever I happen to decide upon.

At any rate, setting aside this issue concerning the characterization of autonomy, it is certainly the case that Otsuka and Voorhoeve's "rights-based" response will be unlikely to succeed against a strategy that instead points to the significance of the positive *representative* and *symbolic* value of choice in the justifiability of our actions to Clara, and hence to the determination (in accord with Claim (F)) of what it would be *reasonable* to do, all-things-considered. After all, a "morally motivated stranger", who acknowledges that acting reasonably towards Clara may involve giving special weight to her preferences, would be strangely motivated indeed if he ignored those preferences, justifying his behaviour on the grounds that, in doing so, he had used his own property rightfully and without violating any right of Clara's. That would provide rather cold comfort to Clara, and hardly seems like a promising line of justification, or a reasonable response to the challenge of accommodating the special value of choice (or, indeed, of autonomy) within a conception of all-things-considered reasonableness. In short, the most plausible conclusion to draw here is that, keeping Claim (F) in view, a morally motivated stranger would be very likely to have good reason to act *in accord* with the recommendations of Claim (B), albeit not because he has any good independent reason to maximize expected utility *per se*. Moreover, in accepting that it would be reasonable to act in accord with the recommendations of Claim (B), our benevolent stranger need not for one moment reject the truth of the Priority View, or act in any way that assumes the falsity of prioritarianism.

V. Acknowledging Egalitarian Considerations without Rejecting the Priority View

The main weight of Otsuka and Voorhoeve's argument against the Priority View is carried by their examination of Clara's predicament in the Risky One Person Case. As we have seen in the foregoing sections, they can be shown to have failed (in two distinct ways) to demonstrate that Clara's case presents a real problem for the Priority View. We may well have sufficient reason to reject Claim (B) outright; and, even if we endorse Claim (B), we can understand it as being fully consistent with the endorsement of the prioritarian Claims (C) and (D).

Nevertheless, Clara's case does not exhaust Otsuka and Voorhoeve's critique of prioritarianism. They offer a further pair of cases in which, they contend, the Priority View again fails to explain a divergence in our considered moral judgement between one-person and many-person cases. The cases are these:

> **The One- and Two- Person "Natural Course of Events" Cases:**[31]
>
> (i) Imagine that you are a morally motivated stranger who learns that unless you intervene in an unpredictable natural course of events, a person will either, with 50 percent probability, receive a gain in utility, or instead, with 50 percent probability, suffer a smaller loss in utility. If you intervene, this person will face neither the prospect of the gain nor the risk of loss. Suppose that you opt for nonintervention on grounds that the expected gain to the person of the uninterrupted unfolding of this course of events is just great enough relative to the expected loss to justify a gamble rather than the risk-free option. (ii) Imagine that you are a morally motivated stranger who learns that unless you intervene in a natural course of events, there is a 50 percent chance that the following will happen: of *two* people who are equally well off, the first will receive a gain in utility and the second will suffer a smaller loss in utility, where these gains and losses are the same as in the one-person case.[32]

Otsuka and Voorhoeve go on to claim:

> It should immediately be apparent that this second case involving two people differs in morally important respects from the first case, in which only one person's interests are at stake. In the second case, there is no single person for whom the prospect of a greater gain is the desirable flipside of exposure to the risk of a lesser loss and for whom the prospect of such gain might be worth the exposure to such risk. Rather, if you do not intervene, there are two people: a first person who would face just a prospect of a gain and who would, if this gain materializes, be better off than the second person, and a

[31] The names for these cases are my own.
[32] Otsuka and Voorhoeve, p. 10.

> second person who would face just a risk of a loss, and who would, if this loss materializes, become worse off than the first person. It follows that rather than simply deciding whether the potential gain outweighs the potential loss to the same person, you must now decide whether the potential gain to the first person outweighs the potential loss to the second person, who would, if this loss materializes, be worse off than the first person. These differences between the one-person and the two-person case imbue the potential loss to a person with greater negative moral significance in the two-person case. You should therefore intervene in a two-person case to prevent the second person from facing the risk of loss, thereby also eliminating the first person's prospect of gain, even though this prospective gain is, by hypothesis, just large enough relative to the potential loss to justify refraining from intervention and letting the chips fall where they may in analogous one-person cases. The Priority View, as we have seen, cannot countenance any such shift. Given that the separateness of persons renders such a shift appropriate, it follows that the Priority View is not adequately responsive to this morally significant fact.[33]

As in Clara's case, the advocate of the Priority View has sufficient resources to meet this renewed challenge. Indeed, just as with Clara's case, there are two independent lines of response open here to the prioritarian, one of which we might characterize as more concessive, and the other as more staunchly recalcitrant.

The more concessive of the two responses is very simple. The advocate of the Priority View can simply make clear what follows (and what does not follow) from endorsing the prioritarian Claims (C) and (D), being careful to characterize his view in its most plausible form. The relevant contrast here is between what one might describe as *Pluralist* and *Fundamentalist* versions of the Priority View. The "fundamentalist" prioritarian holds that Claims (C) and (D) are, so to speak, the whole truth about the ethics of distribution, and denies that any essentially comparative or egalitarian considerations should have a role in our thinking. The *pluralist* prioritarian, by contrast, allows that there may be a variety of comparative (and, particularly, egalitarian) considerations that should play a role in our moral thinking in many-person distributive cases, but which are nevertheless of course irrelevant in one-person cases. There need be no inconsistency whatsoever in holding a complex distributive view that allows a role for deontic egalitarian considerations of fairness, or indeed for a wide variety of non-intrinsic egalitarian considerations relating to the relationship between distribution and the character of interpersonal relations, while nevertheless endorsing the Priority

[33] Otsuka and Voorhoeve, pp. 10-11. Otsuka and Voorhoeve describe this paragraph as containing "the crucial argumentative move" of their article.

View, and with it Claims (C) and (D).[34] Just as the most plausible versions of egalitarianism are pluralist, in allowing a role for non-egalitarian considerations, so too the most plausible versions of prioritarianism will be pluralist rather than absolutist or fundamentalist. With regard to egalitarianism and prioritarianism, "to see the two kinds of views as straightforward adversaries is to mischaracterize the conceptual terrain, and to miss some of the most significant and plausible of the available distributive views."[35] Therefore, in acknowledging that there are special moral considerations that appear in the two-person case, one need not thereby reject the core of the Priority View.

Otsuka and Voorhoeve go too far when they say that "the Priority View cannot countenance any such shift" between one-person and two-person cases.[36] A plausible pluralist prioritarianism makes no objection to the claim that there may be salient comparative or egalitarian considerations of this kind, that apply only in many-person cases. It is just that the elaboration of such considerations is not itself a core part of the Priority View. It is similarly excessive to claim that "[t]he Priority View is mistaken because, in ruling out such essentially comparative considerations, it ignores the moral significance of the separateness of persons."[37] A pluralist Priority View, which allows space for the integration of comparative distributive considerations, rules out nothing of the sort. Thus, rather than saying that the Priority View is mistaken, it would be more accurate to describe it as being *incomplete*. And this incompleteness is surely itself no fatal failing, given that it would be fanciful to think that any "single principle" distributive view, whether egalitarian or prioritarian, could capture the full truth about the ethics of distribution. Plausible egalitarian views are themselves similarly incomplete, and need supplementing by the central axiological insight of prioritarianism – that is, the insight of the diminishing marginal moral significance of greater benefits.[38] [39] Given this, arguments in favour of

[34] On the variety of "Non-Intrinsic egalitarian" considerations, see Martin O'Neill, (2008), "What Should Egalitarians Believe?" *Philosophy & Public Affairs*, 36.2: 119-156, esp. pp. 121-134. On the possibility of a complex distributive view that combines prioritarianism with the acknowledgement of egalitarian considerations, see O'Neill, (2008), ibid., esp. § IV "Equality or Priority, or Equality *and* Priority?", pp. 152-55.
[35] O'Neill, (2008), p. 153.
[36] Otsuka and Verhooeve, p. 11.
[37] Otsuka and Voorhoeve, p. 16.
[38] See O'Neill, (2008), pp. 152-55.

the salience of egalitarian considerations, as given by Otsuka and Voorhoeve, are not themselves arguments *against* the truth of the Priority View (as centrally captured by Claims (C) and (D)).

The more staunch or recalcitrant prioritarian response to Otsuka and Voorhoeve's second set of examples takes a very different line. Instead of emphasizing the pluralist nature of the plausible versions of prioritarianism, it instead points to the ways in which, on the Priority View, even one-person cases, when they involve risk, can involve the sorts of trade-offs that Otsuka and Voorhoeve see as being characteristic only of many-person distributive cases. This line of response effectively proceeds by undermining the exaggerated contrast that Otsuka and Voorhoeve draw between the normative characterization of one-person and many-person cases, showing how we should still view one-person cases under risk as cases where there are salient issues regarding trade-offs between better and worse-off individuals (albeit that, in such situations, the individuals in question are potential rather than actual). In effect, the prioritarian response here is directly to challenge Otsuka and Voorhoeve's assumption that there is a deep difference in kind between interpersonal tradeoffs and intrapersonal tradeoffs under risk.

As mentioned above in Section III, according to the Priority View, given that outcomes rather than prospects are the bearers of moral value, the claims of particular individuals under conditions of risk are effectively proxies or stand-ins for the claims of the possible future individuals (that is, their different potential future selves) who will endure those outcomes. Thus, in deciding in the first of Otsuka and Voorhoeve pair of "Natural Course of Events" cases on a strategy of nonintervention, it is insufficient for the justification of one's decision that one has simply managed to maximize the *ex ante* expected utility of the individual in question. (Just as in Clara's case, where the recommendations of the Priority View departed from the imperative to maximize Clara's *ex ante* expected utility.) For, on the Priority View, one has to give greater weight to the interests of the possible "unlucky" individual who would endure the outcome that involves the loss in well-being, as

[39] One may therefore conclude that Otsuka and Voorhoeve are guilty of a certain overstatement when, at the conclusion of their article, they claim that: "Egalitarian or otherwise comparative views … can account for our judgements in all cases…" (p. 32). On why this claim may be mistaken, see O'Neill, (2008), p. 155.

opposed to being neutral between the interests of the individual in question under either the "lucky" or "unlucky" outcomes. As argued in Section III, prioritarianism carries with it a commitment to a particular degree of risk-aversion in cases such as this.

Otsuka and Voorhoeve seem to miss the full significance of the fact that, on the Priority View, it is outcomes rather than prospects that bear moral value. It is certainly true that, in the "Two-Person Natural Course of Events" case "there is no single person for whom the prospect of a greater gain is the desirable flipside of exposure to the risk of a lesser loss and for whom the prospect of such a gain might be worth the exposure to such risk."[40] But, even in the "One-Person Natural Course of Events" case, we can similarly say that there is no possible outcome for our single individual in which the occurrence of his potential loss would be compensated by the occurrence of a compensating gain. Thus, given that the badness of a utility loss of size y would be an impersonally bad outcome, at the bar of prioritarianism, of greater magnitude than the goodness of a utility gain of the same size, we should only recommend a strategy of nonintervention in the One-Person Natural Course of Events case when the size of the potential gain is substantially greater than the size of the potential loss. Here, we find ourselves on similar conceptual territory to that of the Two-Person case. For, even with the Two-Person case, one would assume that, when the gain to one individual becomes sufficiently large in comparison to the loss to the other individual, it would be impersonally better to recommend intervention rather than nonintervention.[41] This is not, of course, to say that there are no salient moral differences between one-person and two-person cases, or that intrapersonal tradeoffs under risk are identical in structure to interpersonal tradeoffs. But it is to say that, once one acknowledges, as Otsuka and Voorhoeve themselves acknowledge, that outcomes rather than prospects are the bearers of value in prioritarian moral thinking, the difference between intrapersonal tradeoffs under risk and interpersonal tradeoffs comes to seem much more like a difference in degree rather than a difference in kind.

[40] Otsuka and Voorhoeve, pp. 10-11.
[41] I assume that Otsuka and Voorhoeve would not dispute this claim. Otherwise, their view would face the charge of itself being an implausibly "absolutist" form of egalitarianism. Integration of the prioritarian Claims (C) and (D) would be one way for a *pluralist* version of egalitarianism to avoid this problem. (See O'Neill, (2008), § IV, esp. pp. 154-5.

Overall, the correct prioritarian response to Otsuka and Voorhoeve's challenge should be a mixture of the concessive and the recalcitrant strategies outlined above. It is certainly true that the most plausible versions of prioritarianism will be pluralist rather than absolutist or fundamentalist, and hence the Priority View should not be understood as making the mistake of *denying* the significance of comparative or egalitarian considerations. Nevertheless, although there *are* important moral differences between (risky) one-person cases and many-person cases, it should also be stressed that the moral difference between the two sorts of cases is not so great as Otsuka and Voorhoeve suppose. While it is undeniable that "a shift occurs in the moral importance of benefits and burdens when we move from a case involving intrapersonal tradeoffs to a case involving interpersonal tradeoffs,"[42] this shift is most certainly *not* a shift from a context in which we should be straightforward maximizers of *ex ante* expected utility to a context where distinctively distributive considerations come into play. For, as the prioritarian treatment of risk shows, distinctively distributive considerations are at the very core of our thinking about the right way to act, even in cases that involve only the outcomes that might be faced by one single individual.

VI. Preferences, Utility and Well Being

The ways in which Otsuka and Voorhoeve's arguments fail to undermine the plausibility of the Priority View have now been demonstrated. In this and the next section, I want to address two related issues about the relationship between preferences, utility and well-being, both of which are raised by Otsuka and Voorhoeve's arguments. This section will briefly address some problems related to their use of the von Neumann-Morgenstern preference-based approach to utility, while the next section will examine their use of the empirical individual preferences that are revealed in survey data.

By virtue of "assuming the soundness of the orthodox von Neumann-Morgenstern preference-based measure of utility"[43] Otsuka and Voorhoeve treat individual preferences (as revealed, for example, by

[42] Otsuka and Voorhoeve, p. 12.
[43] Otsuka and Voorhoeve, pp. 2-3.

the answers that people give in surveys) as directly revealing individuals' expected utilities. Furthermore, they assume, in setting up the position of the "morally motivated stranger" that "preferences … [are] an accurate measure of her [i.e. Clara's] utility, where "utility" is understood to refer to how well her life is really going (or would go)."[44] In other words, Otsuka and Voorhoeve explicitly here endorse a preference-based measure of well-being, such that an individual's well-being (in the general sense of how well his or her life is going) can be directly measured by examining the degree to which their preferences are satisfied.[45] Such a conception of well-being is problematic, and involves assumptions that are far from being either innocent or plausible.

Consider the case of Clara. Otsuka and Voorhoeve assume that the morally motivated stranger, if he is concerned to make Clara's life go as well as possible, should aim to maximize the expected satisfaction of her preferences. But what if some of Clara's preferences were for the welfare of others, or for the completion of some project (saving Venice from the sea, perhaps)[46] that she took to be her duty, but the pursuit of which she knew would bring her pain and suffering? For example, perhaps Clara would, if thinking purely about her own well-being, definitely prefer Treatment (2) to Treatment (1), on the basis that, considering her own pleasure rather than what she could do to aid others or to pursue some impersonal good, her highest priority would be to avoid the terrible suffering associated with falling into Condition (d). Nevertheless, given the great potential advantages in mobility that would be offered by having a 50% chance of attaining full health (Condition (c)), and given the advantages that mobility brings for one's ability to help others or for pursuing certain difficult projects, Clara skews her preferences towards Treatment (1), to the degree that she becomes indifferent between the two treatments. If Clara's preferences had this structure, then it would be very odd to say that a "morally motivated stranger", who was concerned to promote Clara's well-being,

---

[44] Otsuka and Voorhoeve, p. 3.

[45] In the typology offered by Derek Parfit of different views of well-being, or of "what makes someone's life go best", Otsuka and Voorhoeve are hereby assuming the truth of some version of the "Desire-Fulfilment Theory" of well-being. See Derek Parfit, (1984), *Reasons and Persons*, (Oxford: Oxford University Press), Appendix I, pp. 493-502.

[46] The example is from Parfit, (1984), § 59, esp. pp. 151-3.

should be indifferent between giving either of the two treatments. Rather, he should give her Treatment (2).

When Clara is settling her preferences between a life of great personal pleasure and a life of selfless self-sacrifice, she is settling upon the life that she thinks best to live, or most valuable. But this is something importantly separate from settling upon the preferences according to which her life would go best *for her*. If the two ideas were not separable, then the very idea of self-sacrifice would make no sense, as acting in accord with one's preferences would always and everywhere be equivalent to acting in a way that promotes one's own well-being, no matter how selfless those preferences might be. By folding all preferences, including preferences for things that are quite separate from, or even opposed to, one's own pleasure or flourishing, into a single "utility-function" that is taken to be a measure of overall well-being, the von Neumann-Morgenstern preference-satisfaction approach, when followed in the way we see in Otsuka and Voorhoeve's article, leads us to an unhelpfully broad and uninformative account of well-being.[47] It rules out even the possibility that our preferences might embody trade-offs between our own well-being and other goods (whether the good of other people, or some impersonal good), because the measure of our well-being itself is always referred back to the structure of those preferences themselves.[48]

The truth is that it is a misuse of the von Neumann-Morgenstern axiomatic account of rational choice to translate it willy-nilly into a model of well-being. For what axiomatic rational choice theories of this kind give us is not an account of well-being at all, but a much broader account of all-things-considered choiceworthiness. As Scanlon rightly describes things, formal theories of rational choice, such as those specified by the von Neumann-Morgenstern axioms, "are offered not as accounts of well-being (of "what makes someone's life go better") but rather of what a person has reason to do or

[47] This problem is structurally identical to the difficulties face by the von Neumann-Morgenstern preference-based account of well-being in dealing with second-order preferences with regard to the risk of falling into different levels of well-being, as discussed in Section III and footnote 21 above.

[48] As Scanlon puts it, "the question of choiceworthiness is not the same as the question of well-being, since it makes sense to say that a person had good reason to choose a certain plan of life even though it involved a lower level of well-being – was worse from the point of view of the person who lived it – than some available alternative." (Scanlon, (1998), p. 131)

to choose all things considered, and the grounds on which these choices are to be assessed are explicitly intended to include preferences for things other than the person's own well-being."[49] Thus, we simply *cannot* justifiably move, by use of von Neumann-Morgenstern, from Clara's preferences to an account of what would make her life go best, or what would promote her well-being; rather, we can move *only* from her preferences to an account of what it would be formally rational for her to choose. And, crucially, there is no reason to think that a "morally motivated stranger" need have reason to make choices that track whatever it would be formally rational for Clara to choose, as opposed to making choices that would promote Clara's well-being, more narrowly defined.

As well as these special problems of other-regarding or impersonal preferences, there are other familiar problems with the identification of well-being with preference-satisfaction or desire-fulfilment. As Scanlon further emphasizes, identifying well-being with desire-satisfaction seems to get the order of normative explanation exactly backwards: typically, we do not consider that some outcome promotes our well-being *because* we desire it, but rather we desire a particular outcome because we take ourselves to have reason to think that this outcome is conducive to our well-being.[50] Otsuka and Voorhoeve attempt to circumvent such problems by denying that their use of the von Neumann-Morgenstern expected utility measure of well-being actually commits them to the *identification* of well-being with desire-satisfaction. As they put it, "we are not spelling out a proposal regarding what utility *is* … Rather we are presenting an account of how to measure the magnitude of one's expected utility. One might believe that two options have the same expected utility for a person when she is indifferent between these options without also believing that utility is identical with, or even that it consists of, preference satisfaction. One might maintain that utility is, or consists of, something other than preference satisfaction, while also maintaining that the specified idealized preferences unerringly track the magnitude of this other thing."[51] However, Otsuka and Voorhoeve here identify what we might call a merely logical possibility, and certainly a possibility without much

[49] T. M. Scanlon, (1998), p. 116.
[50] See T. M. Scanlon, (1998), Ch. 3, esp. p. 114.
[51] Otsuka and Voorhoeve, p. 3, fn. 3.

plausibility or normative significance. For it would be nothing short of an endless miracle if an individual's preferences unerringly "tracked" her underlying well-being, without themselves constituting or being identified with her well-being. Such a suggestion would only point to a significant possibility if we believed that all preferences are necessarily preferences for one's own well-being, but the examples of preferences regarding the fate of Venice or the flourishing of others, as mentioned above, clearly show that we can have preferences whose objects are not our own well-being, and we can even have preferences whose objects are, in some ways, inconsistent with the promotion of our own well-being (as when working to save Venice involves a life of self-sacrifice). Given the possibility of preferences for objects other than our own well-being, there can be no reason to expect the endless miracle that would be embodied by the possibility that our preferences always "tracked" some underlying measure of well-being, itself constituted by an unknown substratum made up of something other than the satisfaction of those preferences.[52] The merely logical possibility of such a chain of miracles is far from sufficient to rescue Otsuka and Voorhoeve's approach to preference, utility and well-being.

VII. Methodology in Political Philosophy and the Perils of Survey Data:

I shall end with some brief remarks about methodology in political philosophy, and especially regarding the special problems of relying on survey data. One of the most distinctive aspects of Otsuka and Voorhoeve's discussion is the degree to which it discusses results gained from health economics, with regards to people's preferences between various possible trade-offs or gambles. What is more striking is that, at each stage, Otsuka and Voorhoeve are concerned to endorse the normative view that would best accord with people's (apparent) empirical preferences on these

[52] Perhaps *some* of the foregoing problems for the use of "the orthodox von Neumann-Morgenstern preference-based measure of utility" might be circumvented if the individuals' preferences in question could be restricted to "self-interested" preferences, thereby avoiding preferences that range over the well-being of others or, say, the fate of Venice. Such a restriction presents at least three significant problems. Firstly, it does not seem correct to say that no "other-regarding" preferences (say, for the flourishing of our children or of our football teams) contribute to our well-being. Secondly, following from this, and crucially for the whole enterprise, the boundaries of what should count as "self-interested" as opposed to "other-regarding" preferences are impossibly vague and difficult to police (and, as Scanlon points out, this boundary is anyway of no practical interest from the agent's own point of view (see Scanlon, (1998), pp. 126-133)). Thirdly, there is no reason to think that empirical preferences as revealed by surveys are themselves only "self-interested" rather than unrestricted, which brings us on to Section VII.

matters. For example, Otsuka and Voorhoeve agree with these surveys that, in Clara's position, we should be indifferent between taking Treatment (1) and Treatment (2). They also agree with empirical surveys showing that people prefer to help the worse-off group in the Non-Risky Many Person Case, seemingly offering the empirical evidence of corroborative survey data as a means of supporting their view.[53]

Otsuka and Voorhoeve's treatment of survey data as being morally authoritative reaches its greatest height when they comment on the fact that surveys indicate that surveyed individuals are typically indifferent between treating 1,500 individuals with the Mild Impairment (Condition (b)) as opposed to a single individual with the Very Severe Impairment (Condition (a)). Here, Otsuka and Voorhoeve claim that, "[i]f the aforementioned estimates of the typical priority people give to the less well off are correct, a treatment for the Very Severe impairment that brought a person up to the condition of the Severe impairment will have an expected moral value of 1,500 times the expected moral value of the treatment for the Slight impairment."[54] This claim seems to embody the (rather surprising) background assumption that surveys of individual preferences over distributive cases not only reliably track the correct distributive view in some general sense, but that societal preferences across distributive outcomes give a *precise* measure of "expected moral value". If a social survey says that members of the public prefer one case of outcome x to 1,500 cases of outcome y then, Otsuka and Voorhoeve seem to suggest, x is fifteen hundred times more morally valuable than y.

There are many reasons why we should want to avoid placing so much moral authority on the results of empirical surveys of people's preferences. Here are eight such reasons.

1. If surveys are set up to find people's raw *preferences* between different distributive outcomes, then they may be asking the wrong question, especially in many-person rather than one-person cases. Even if our concern is to grant some moral authority to people's "empirical" or "really-

[53] See Otsuka and Voorhoeve, pp. 4-5.
[54] Otsuka and Voorhoeve, p. 8, fn. 15.

existing" moral judgements (as opposed to what we independently judge to be the truth about distributive ethics),[55] this is not the same as granting some special authority to their raw preferences over different outcomes. Such preferences may themselves be a mixture of moral and other preferences, including various prejudices and whims.

2. When presented with a complicated example, such as Clara's "Risky One-Person Case", with its central gamble over four possible outcomes (i.e. Conditions (a), (b), (c) and (d)) many people, when surveyed, will simply fail fully to grasp the structure of the situation. A report of indifference between two lotteries may simply represent a lack of understanding of what each lottery involves, and the structure of the gamble between them. One must be very careful that survey data does not simply embody confusion over the situation that is being presented.

3. Even when survey results do not end up embodying downright confusion, as in (2), they may nevertheless embody a certain lack of imagination. One assumes that not every respondent to the dilemma of the Risky One-Person Case was really able to imagine themselves into the situations associated with each of the four Conditions (a)-(d).

4. Survey results that do not embody confusion (as in (2)) or lack of imagination (as in (3)) may nevertheless embody over-quick or casual deliberation. Someone faced with the *real* choice between Treatments (1) and (2) would presumably deliberate with a great deal more care and attention than someone who is merely acting as an experimental subject for a health economist.

5. Many people will, when surveyed with regard to a situation like Clara's, "fight the example" in some way. Rather than failing to understand the situation (as in 2) they may simply choose to

[55] The distinction between these two goals is well-captured by the title of one of the articles in health economics cited by Otsuka and Voorhoeve: Nord et al's "Incorporating Societal Concerns for Fairness". Note that the article is *not* entitled "Incorporating Fairness…". That which is to be incorporated into health policy here is public perceptions of fairness, not fairness itself. (By the same token, though, that which is to be incorporated into health policy here is *also* not mere aggregative *preferences*.) See Erik Nord, Jose-Louis Pinto Prades, Jeff Richardson, Paul Menzel and Peter Ubel, "Incorporating Societal Concerns for Fairness in Numerical Valuations of Health Programmes," *Health Economics* (1999): 25–39.

resist it, imagining that, even if they were so unlucky as to fall into Condition (b), some other cure would surely soon come around the corner. Aggregative survey results can never communicate the number of respondents whose decision is based on an illicit mental readjustment of the case in question in this way.

6. Other people will have quasi-principled reasons for being reluctant to express clear preferences in such cases, on the grounds that they would be "tempting fate" or "playing God". Thus, reports of indifference between two options may actually be disguised reports of non-standard considerations of these kinds.

7. Some people may simply be very bad judges of how their own well-being would be affected by living under different sets of conditions. One might imagine oneself Stoically coping with *Very Severe Impairment* out of mere self-deception, whereas in reality one would fall to pieces if facing anything worse than the *Severe Impairment*. This sort of lack of information about our own potential future reactions to different conditions can itself reduce the informativeness of our *ex ante* preferences over possible gambles.

8. One should be somewhat concerned with claims that tell us that, for example, "people who imagine themselves in such a predicament would *typically* be indifferent"[56] between two options. This presumably does not mean that *everyone* is indifferent between the options in this way. Perhaps a majority are indifferent; or perhaps indifference is the median position. Nevertheless, it seems perverse to grant a special normative authority to the mode or median preference in a society, rather than acknowledging the real diversity of viewpoints or preferences within that society, which could be obscured by reports of "typical" sets of preferences.

[56] Otsuka and Voorhoeve, p. 2. [*My italics*]

One could no doubt think of other particular problems generated by ceding normative authority to the results of empirical surveys. But the central problems with such a methodological approach are more general. Crucially, we simply have no reason to think that the truth about distributive ethics is already known, let alone that it is already embodied in the empirical views or preferences of our fellow citizens. Thus, it need be no more than a superficial objection to some distributive view that it does not find support from surveys of such views or preferences. Views in what Parfit calls the "ethics of distribution" should stand or fall at the bar of reason, through the method of reflective equilibrium, not at the bar of public popularity. To think otherwise is to discount the very possibility of moral progress, for it gives a normative fundamentality to preferences themselves, whatever the quality of their foundation. Moreover, if ethics is simply about fitting theory to the stable fixed-points of public preference, then one should surely wonder what possible gain there could be in adding the gilding of philosophical theorizing to these already authoritative fixed points.

VIII. Conclusion

Despite the misgivings of the foregoing section, Otsuka and Voorhoeve are to be lauded for bringing political philosophy into closer contact with health economics. Nevertheless, one must be very careful not to cede too much in the way of normative authority either to von Neumann-Morgenstern measures of individual utility-functions, or to society's preferences over multi-person distributive cases. In a number of places, Otsuka and Voorhoeve do seem to go too far in granting an unwarranted normative authority to empirical preferences, in both of these ways.

Otsuka and Voorhoeve are also to be congratulated for subjecting the Priority View to a particularly ingenious and thorough investigation. It is clear that the structure and implications of prioritarianism, as well as its limits, can be shown in much finer detail when held up to the unforgiving glare of their examination. Nevertheless, the priority view as formulated by Parfit can either rebuff or accommodate all aspects of Otsuka and Voorhoeve's challenge.