Soc Choice Welfare (2005) 24: 413–431
DOI: 10.1007/s00355-003-0305-z

Social Choice and Welfare

The mere addition paradox, parity and critical-level utilitarianism

Mozaffar Qizilbash

School of Economics, University of East Anglia, Norwich, NR4 7TJ, UK (e-mail: m.qizilbosh@uea.ac.uk)

Received: 26 December 2002/Accepted: 24 October 2003

Abstract. This paper uses a formal analysis of the relation of 'parity' to make sense of a well-known solution to Parfit's 'mere addition paradox'. This solution is sometimes dismissed as a recourse to 'incomparability'. In this analysis, however, the solution is consistent with comparability, as well as transitivity of 'better than'. The analysis is related to Blackorby, Bossert and Donaldson's 'incomplete critical-level generalised utilitarianism' (ICLGU). ICLGU is inspired by Parfit's work and can be related to the analysis of parity, though the distinctive 'mark' of parity suggests that the boundaries of a set of critical levels is not exact. One has to allow for vagueness to make an account based on parity plausible. These accounts are then contrasted with Broome's view which also involves vagueness.

1 Introduction

Parfit (1984, p. 443) has argued that moral theories should not lead us to accept the 'repugnant conclusion', and that they should solve the 'mere addition paradox'. One way of solving the 'mere addition paradox' which Parfit (1984, p. 431) suggests involves 'rough' rather than 'precise' comparability holding between various outcomes. The discussion of this solution

I am grateful to Ashley Piggins, Maurice Salles and Bob Sugden for very helpful comments on earlier versions. I am also very grateful to an anonymous referee for some very insightful comments. I have also benefited from comments from, and discussion on related ideas with Gustaf Arrhenius, John Broome, Robin Cubitt, James Griffin, Wlodek Rabinowicz and Bob Sugden. Finally, I would also like to thank John Broome for allowing me to read various forthcoming manuscripts. Any errors are mine.

has not progressed much, because the notion of 'rough comparability' has remained rather obscure. This notion has, nonetheless, been discussed in the recent literature on the 'incommensurability of values' (Chang 1997, 2002 a,b and Qizilbash 2002). Blackorby et al. (1996) have also developed a variation of utilitarianism – 'Incomplete Critical-Level Generalised Utilitarianism' (ICLGU) – which is inspired by Parfit's work. This paper builds on an earlier attempt to formalize the relation of 'rough equality' or 'parity' (Qizilbash 2002). It shows how Parfit's and Blackorby, Bossert and Donaldson's accounts can be understood using this formal analysis, as long as one allows for vagueness.

The paper is structured as follows; in Sect. 2, the relation of 'parity' is motivated; in Sect. 3 it is formalized; in Sect. 4, the formal analysis is related to Parfit's paradox; in Sect. 5, ICLGU is discussed; in Sect. 6, vagueness is introduced; Sect. 7 concludes; Appendix 1 contains proofs; and the 'repugnant conclusion' is discussed in Appendix 2.

2 Motivation: Parity and comparability

Chang (1997, 2002a,b), Griffin (1986, 1997 and 2000) and Parfit (1984) have all argued that sometimes we are faced with a pair of alternatives which are 'comparable' even though it is neither true that one is better than the other, nor true that they are exactly equal in value. The alternatives might be 'roughly equal in value' (Griffin), 'on a par' (Chang), 'roughly comparable' or 'in the same league' (Parfit). Griffin's discussion of 'rough equality' has come under considerable scrutiny (Broome 2000 and Qizilbash 2000, inter alia). His use of 'roughly equal in value' can be misleading. When 'roughly equal in value to' is understood as a distinct relation – so that when it holds it is definitely not true that one of the other comparative relations holds – it cannot hold between items when one is better than the other. However, in ordinary language our use of 'rough equality' allows for the possibility that options are roughly equal in value while one is better than the other. The same is true of relations such as 'in the same league as' and 'in the same class of goodness as': two items can be in the same league, or in the same class of goodness, while one is better than the other. The notion of 'rough equality' may also be problematic for those who think that the very concept of equality involves precision.[1] Chang's term 'on a par with' may be less misleading and so I use it to refer to the evaluative relation with which Griffin is concerned.[2]

[1] For example, John Broome uses 'equally as good as' to mean 'precisely as good as'. He has, nonetheless, discussed 'roughly equally as good as'. See Broome (2000).

[2] So, I use 'on a par with' and 'parity' as terms of art, just as Griffin uses 'rough equality' as a term of art. It is worth noting that Griffin himself sometimes uses 'rough equality' so that it is compatible with 'exact equality' (Qizilbash 2000, pp. 233–234). It is also worth noting that Chang thinks that her notion of 'parity' is distinct from the relations which Griffin and Parfit use (Chang 2002b, p. 661).

The examples which are invoked when 'on a par with' is used often involve alternatives which are qualitatively different. For example, Griffin (1986, pp. 80–81) uses an example involving two novelists with very diverse achievements. Reading their works realises different values – insight and amusement. When the value of the novelists is close, Griffin thinks that they are hard to rank. The problem of comparison may, he thinks, involve two possibilities. In one, 'it is hard to discriminate the differences in value that are really there' while in the other, the problem is that 'there are no fine differences really there to discriminate ... and the roughness is ... ineradicably in the values themselves' (Griffin 1986, pp. 80–81). It is the second possibility in his discussion – where 'the roughness is ... in [the relation between] the values themselves' (Griffin 1986, pp. 80–81) – which suggests that there may be a little understood relation between values. Broome has noted that Griffin's discussion focusses on the relation between values rather than alternatives (Broome 1999 and Qizilbash 2000). However, his argument clearly also relates to comparisons between alternatives. Parfit (1984, p. 431) discusses a similar example involving three candidates for a literary prize – a novelist and two poets – to motivate the notion of 'rough comparability'.

In analysing Griffin's discussion, Qizilbash (2002) focusses on a case involving three meals: x; x^+; and y. x is an excellent French meal, and x^+ is a very slightly better excellent French meal, while y is an excellent Italian meal. Suppose that any excellent French meal is neither better, nor worse, than any excellent Italian meal. Are they exactly as good? The relation 'exactly as good as' is, Qizilbash suggests, an *equivalence relation*. Like any equivalence relation, it is transitive: i.e. for any a, b and c, if a and b are exactly as good, and b and c are exactly as good, then a and c are exactly as good. However, the relation 'exactly as good as' does not hold between any excellent French meal and any excellent Italian meal. To see this, suppose that it did. If it did, then we would have: x is exactly as good as y and y is exactly as good as x^+. By transitivity of 'exactly as good as' it follows that x and x^+ are exactly as good. However, we know that x^+ is better than x, so that x and x^+ are not exactly as good. We are, thus, led to a contradiction. If x and x^+ are neither better than, nor worse than, nor exactly as good as y, are x and x^+ incomparable with y? Some, like Raz (1986, pp. 330–331) think that they are. Nonetheless, one is drawn to the idea that x and x^+ are just as good as y, even if they are not exactly as good as y. It is this sort of case which fits best with the second possibility in Griffin's discussion of 'rough equality': there is 'equality of value' of a sort, but the options are not 'exactly as good'. So I treat this as an example of 'parity'. When options are on a par, tiny changes in value do not tilt the balance in favour of one of the alternatives, as they do in cases of exact equality. However, some larger, or 'significant' change in value does tilt the balance. This is the distinguishing feature or 'mark' of parity. Griffin (1986, p. 81) reserved the term 'incomparability' for cases when, while one alternative or value is not better than the other, neither exact equality nor parity holds between them. It is worth noting that, when it is understood in this way, parity is a determinate relation, and does not involve vagueness.[3] Vagueness is, nonetheless, relevant to discussions of parity, as we shall see, because what makes a change in value 'significant' can be vague.

3 Defining parity

In this section, I develop a system (following Qizilbash 2002) within which parity can be defined. I use 'R' to mean 'better than or exactly as good as' and refer to it as 'narrowly at least as good as'. It is taken to be a primitive relation. I use the following symbols in the rest of the paper: '$\wedge$' for conjunction; '$\sim$' for negation; '$\vee$' for the inclusive 'or'; '$\in$', for 'is a member of'; '$\Rightarrow$' for the conditional 'if ... then'; '$\subseteq$' for set inclusion ('is included in'); '$\exists$' for the existential quantifier ('for some'); '$\forall$' for the universal quantifier ('for all'); '$\Leftrightarrow$' for 'if and only if'; and '$\neq$' for 'is not identical to'. Various properties are defined on a set of conceivable alternatives, X. Any binary relation, O, may have the following properties: *reflexivity*: $\forall x \in \mathrm{X}$, $x\mathrm{O}x$; *transitivity*: $\forall x, y, z \in \mathrm{X}$, $(x\mathrm{O}y \wedge y\mathrm{O}z) \Rightarrow x\mathrm{O}z$; *completeness*: $\forall x, y \in \mathrm{X}$, $x \neq y$, $x\mathrm{O}y \vee y\mathrm{O}x$; *symmetry*: $\forall x, y \in \mathrm{X}$, $x\mathrm{O}y \Rightarrow y\mathrm{O}x$; *asymmetry*: $\forall x, y \in \mathrm{X}$, $x\mathrm{O}y \Rightarrow \sim(y\mathrm{O}x)$; and *irreflexivity*: $\forall x \in \mathrm{X}$, $\sim(x\mathrm{O}x)$. A relation which is transitive and reflexive but not necessarily complete is a quasi-ordering. One which is symmetric, reflexive and transitive is an equivalence relation. We can define 'better than' – which is written 'B' – and 'exactly as good as', which is written 'E' in the usual way:

Definition 1. $x\mathrm{B}y \Leftrightarrow [x\mathrm{R}y \wedge \sim(y\mathrm{R}x)]$

Definition 2. $x\mathrm{E}y \Leftrightarrow (x\mathrm{R}y \wedge y\mathrm{R}x)$

Examples of parity suggest that R is not complete. In the meal example, when x and y are on a par, $\sim(x\mathrm{R}y \vee y\mathrm{R}x)$. x and y are, on some views, 'incommensurate with' each other. Broome (2000) uses the term 'incommensurate with' in this sense, and I define 'B-incommensurate' as follows:

Definition 3. x is B-incommensurate with $y \Leftrightarrow \sim(x\mathrm{R}y \vee y\mathrm{R}x)$.

If x and y are not B-incommensurate they are 'B-commensurate'. To allow for B-incommensurate cases, I assume:

Postulate 1. R is a quasi-ordering.

The properties of quasi-orderings are well-known (Sen 1979) and since R is a quasi-ordering, B and E are transitive and, furthermore, 'EB transitivity' and 'BE transitivity' hold:

---

[3] Griffin (1986, pp. 81 and 96) talks of a 'vague' ordering, but his discussion is not best understood in terms of vagueness. On this see Broome (2000) and Qizilbash (2000).

EB transitivity: $\forall x,y,z \in \mathrm{X}$: $(x\mathrm{E}y \wedge y\mathrm{B}z) \Leftrightarrow (x\mathrm{B}z)$
BE transitivity: $\forall x,y,z \in \mathrm{X}$: $(x\mathrm{B}y \wedge y\mathrm{E}z) \Leftrightarrow (x\mathrm{B}z)$.[4]

Alternatives can be B-incommensurate without being ‘incomparable’ in Griffin’s sense. If x and y are on a par, they are B-incommensurate, but, nonetheless, comparable, on Griffin’s account. So to define parity we need another primitive relation: ‘comparable with’, which is written ‘C’. I assume:

Postulate 2. C is reflexive and symmetric.

We can now define ‘on a par with’, which is written ‘P’, in terms of our primitive relations:

Definition 4. $x\mathrm{P}y \Leftrightarrow x\mathrm{C}y \wedge \sim(x\mathrm{R}y \vee y\mathrm{R}x)$.

In this system, x is incomparable with y means $\sim(x\mathrm{C}y)$. Parity and incomparability are clearly mutually exclusive. Furthermore, the two primitive relations are related as follows:

Postulate 3. $(x\mathrm{R}y \vee y\mathrm{R}x) \Rightarrow x\mathrm{C}y$

It is not assumed, however, that when options are comparable, they are necessarily B-commensurate. This is an important respect in which the system of relations presented here differs from standard formal analysis, where, in effect, comparability and B-commensurateness are treated as equivalent.

We can now define a sense of ‘at least as good as’ which is distinct from R. I shall refer to this relation as ‘broadly at least as good as’ or ‘A’, and define it as follows:

Definition 5. $x\mathrm{A}y \Leftrightarrow x\mathrm{C}y \wedge \sim(y\mathrm{B}x)$.

So x and y are broadly at least as good if and only if x and y are comparable and y is not better than x.[5]

It is worth noting that the cases of parity discussed above suggest that P and A are non-transitive. In the example of parity involving meals, $x\mathrm{P}y \wedge y\mathrm{P}x^{+} \wedge x^{+}\mathrm{B}x$. Since (from Definition 1), $x^{+}\mathrm{B}x$ implies $x^{+}\mathrm{R}x$, this contradicts $x^{+}\mathrm{P}x$ (from Definition 4), and P is non-transitive. Furthermore, when alternatives are on a par, they are ‘broadly at least as good’, and in this example, $x\mathrm{A}y \wedge y\mathrm{A}x^{+}$. However, since $x^{+}\mathrm{B}x$ implies $\sim(x\mathrm{A}\ x^{+})$ (from Definition 5), A is not transitive.

Given the above definitions, and the various postulates adopted, we can state and prove various theorems. These theorems show the structure of the relation system. While the system is non-standard, some properties *are* fairly standard. For example, we can check that E is an equivalence relation (Qizilbash 2002, p.147 and Theorem 4 in Appendix 1). Furthermore,

---

[4] Sen (1979, p. 10) refers to these as ‘IP transitivity’ and ‘PI transitivity’ respectively.
[5] Chang (2002a, p. 2) sometimes uses ‘at least as good as’ in this sense, though economists usually use ‘at least as good as’ to refer to R.

Theorem 1. B is irreflexive, asymmetric and transitive.

Proof. See Appendix 1.

However, since the system is non-standard it is worth noting the way in which some of the relations are connected. In particular, we can show:

Theorem 2. $x\mathrm{C}y \Leftrightarrow (x\mathrm{R}y \vee y\mathrm{R}x) \vee x\mathrm{P}y$

Proof. See Appendix 1.

Since Griffin uses 'comparable with' to mean 'better than, worse than, precisely equal in value to or on a par with', C captures his sense of comparability. It is also easy to check that $x\mathrm{R}y \Rightarrow x\mathrm{A}y$ (Theorem 3 in Appendix 1), though $x\mathrm{A}y$ does not imply $x\mathrm{R}y$, since (from Definition 4) when $x\mathrm{P}y$, $x\mathrm{A}y$, but $x\mathrm{C}y \wedge \sim(x\mathrm{R}y \vee y\mathrm{R}x)$. Since A is not transitive because of cases of parity, A and R have distinct properties (though it is easy to check that, like R, A is reflexive). Furthermore, $x\mathrm{A}y \vee y\mathrm{A}x \Leftrightarrow x\mathrm{C}y$ (Theorem 5 in Appendix 1), so there is a simple connection between comparability and broadly at least as good as.[6]

4 The mere addition paradox

With this system of relations in place, we can solve Parfit's 'mere addition paradox'. However, first we need to understand the notion of 'mere addition' and what Parfit calls the 'valueless level' of well-being. The valueless level is defined so that below this level '[i]f ... lives are worth living, they have ... value for the people whose lives they are. But the fact that such lives are lived does not make the outcome better.'(Parfit 1984, p. 412). On Parfit's account (Parfit 1984, p. 420) adding extra people is known as *mere addition* when extra people exist: (1) who have lives worth living; (2) who affect no one else; and (3) whose existence does not involve any social injustice. In some statements of this paradox (such as Temkin 1987, p. 141), (3) is dropped for the sake of simplicity. I follow this practice here. So *mere addition* involves the adding of extra people whose lives are worth living and who do not affect anyone else. Since these lives are worth living, and affect nobody, Parfit thinks it cannot make matters worse to add these lives.[7]

Parfit's discussion of the paradox (Parfit 1984, pp. 425–426) involves comparisons between four states of affairs. I focus on a simplified version of the paradox, which only involves three states of affairs and which is sometimes used in the literature (Temkin 1987, pp.140–141). The three states are 'a', 'a^+', 'b'.[8] Everyone in a has a level of well-being, α, which is above the

[6] It is also easy to check that $(x\mathrm{A}y \wedge y\mathrm{A}x) \Leftrightarrow (x\mathrm{E}y \vee x\mathrm{P}y)$.

[7] Parfit (1984, pp. 422–425) also constructs his examples so that any inequality which arises from mere addition does not make matters worse.

[8] I assume that all states are conceivable alternatives.

valueless level. The quality of life of everyone in b is $\beta = 0.8\alpha$.[9] There are twice as many people living in b as in a. Next consider a^+. This is the same as a except that there are some *extra people*. The number of extra people is equal to the population in a. The lives of the extra people are worth living, but their quality of life, γ, is below the valueless level. The addition of these people does not, thus, make a^+ better than a. The lives of these people also affect no-one else.[10] So a^+ arises from mere addition to a.

The paradox involves various judgements. Parfit thinks that if one judges that $b\text{B}a$, one might end up accepting the 'repugnant conclusion' (Parfit 1984, p. 430). I explain why Parfit thinks this in Appendix 2. To avoid this conclusion, one might assume that $a\text{B}b$. In fact, all one needs to assume to make Parfit's argument is $\sim(b\text{B}a)$, and this is what I assume. It is also plausible to hold that $b\text{B}a^+$, since b involves a more equitable distribution of well-being than a^+. Since a^+ is produced by mere addition to a, Parfit thinks that it is not worse than a. One might conclude from this that $a^+\text{R}a$. However, one cannot do this without contradiction. $a^+\text{R}a$ means $a^+\text{B}a$ or $a^+\text{E}a$. If the first disjunct is true – given transitivity of B (Theorem 1), and $b\text{B}a^+$ – we have $b\text{B}a$. On the other hand, if $a^+\text{E}a$, then given $b\text{B}a^+$, and BE transitivity, we have $b\text{B}a$. So, if $\sim(b\text{B}a)$ we have a contradiction. This is the essence of the paradox.

Broome (1999, p. 231) has discussed examples of the sort which Parfit uses in notation which will be more familiar to economists. So here is a variation on Parfit's example in Broome's notation. In Broome's notation (Broome 1999, p. 230 and 1996, pp. 177–181) each possible state of affairs is represented by a vector showing the distribution of well-being. Each place in the vector stands for a person who lives in at least one state of affairs being compared. The corresponding place in each vector compared stands for the same person. In a state where she does not exist, her place in a vector contains an Ω. In a place where she does exist her place contains a number which indicates her lifetime well-being. Broome takes well-being to be measurable on a cardinal scale that is comparable between people. Any positive level of well-being is good rather than bad.

Here is a version of Parfit's example in Broome's notation which involves setting $\alpha = 10$ and $\gamma = 1$, where the population in a is 2. The distribution of well-being in each of the three alternatives can now be written as follows: a: $(10, 10, \Omega, \Omega)$; a^+: $(10, 10, 1, 1)$; and b: $(8, 8, 8, 8)$. Since a^+ merely adds two extra people whose lives are good to the population in a without altering the

[9] Like Parfit I use 'p's quality of life' and 'p's level of well-being' equivalently in this paper.

[10] Parfit's actual example is a little more complicated. In it, he allows for the possibility that the extra people may be living on another planet, and other people may not even know of their existence. So the addition of these people does not, on Parfit's account, necessarily make a^+ socially unjust. Any 'natural inequality' that is introduced by the addition of extra people, furthermore, does not make a^+ worse than a (Parfit, 1984, pp. 422–425). This deals with the possibility that the addition of extra people generates natural inequality and thus makes the outcome worse. Parfit's example also includes a fourth alternative to deal with this problem.

well-being of the initial population, it is not worse than a^+.[11] Certainly, since these people have good lives, it is hard to claim that it would have been better had they not existed. However, since b involves more equality than, as well as a higher sum of well-being than, a^+, it can also be argued that $b\mathrm{B}a^+$. If $\sim(b\mathrm{B}a)$, however, contradiction follows immediately if 'a^+ is not worse than a' is taken to mean $a^+\mathrm{R}a$.

Parfit resists this contradiction by effectively assuming that if a is neither better nor worse than a^+, nor are a and a^+ exactly as good. Since he thinks that a and a^+ are nonetheless comparable, they must, in our relation system, be on a par. Parfit's discussion also supports this. He (Parfit 1984, pp. 431–432) suggests that if a^+ is improved because the quality of life of the extra people in a^+ is made *somewhat* higher than γ, the resulting situation – '*improved a^+*'- is still not better than a. It is in this context that Parfit introduces the discussion of 'rough comparability'. The idea seems to be that this example has the same *formal* characteristics as the cases of parity discussed earlier. It has the 'mark' of parity. So one might think that if the extra people in a^+ were made *significantly* better off than they are in a^+ we may well conclude that the new situation is better. If a and a^+ are on a par, furthermore, we can solve the paradox using our relation system. In particular, we can hold the conjunction: $b\mathrm{B}a^+ \wedge a^+\mathrm{P}a \wedge \sim(b\mathrm{B}a)$. That is all that is required to solve Parfit's paradox. In fact, in this case it can be true that: $b\mathrm{A}a^+ \wedge a^+\mathrm{A}a \wedge a\mathrm{A}b$. So comparability is not necessarily violated. This is significant since some regard Parfit's discussion of this solution as a recourse to 'incomparability'.[12] Temkin (1987) has also argued that the lesson to be learnt from Parfit's paradox is that B is non-transitive. No such lesson can be learnt from this analysis.

5 Blackorby, Bossert and Donaldson's ICLGU

Taking their inspiration from Parfit, Blackorby et al. (1996) have developed ICLGU to allow for the possibility of 'incommensurability'. In explaining their position, I begin with a case with a fixed population where w_i is the well-being of person i and $i = 1, \ldots, n$, so that n is the number of people in the population.[13] It is assumed that there is a constant, positive 'critical-level' of

---

[11] This example is a little different from one which Broome himself treats as a version of the mere addition paradox (Broome 1996, p. 181).

[12] Some – like Erik Carlson (1998, p. 288) – think that this solution involves an *ad hoc* recourse to 'incomparability'. Gustaf Arrhenius (2000, pp. 81–83) discusses some reasons for 'incommensurability'. However, Parfit's claim seems to be based on the possibility that the *formal* nature of this specific example is similar to the case of the novelist and the two poets which he discusses.

[13] It is worth noting that the existence of parity can also undermine the representation of well-being in terms of a numerical index. I ignore this complication in this paper.

well-being, π. In standard critical-level generalized utilitarianism (CLGU), the value function involves transforming levels of well-being to allow for inequality aversion. Well-being levels are thus transformed using a function g(.). The value function of CLGU is:

$$W^{\text{CLGU}} = \sum_{i=1}^{n} [g(w_i) - g(\pi)] \tag{1}$$

Like Parfit, I shall put the issue of social inequality to one side in what follows.[14] I thus focus on the value function of 'critical-level utilitarianism' (CLU) which is a special case of (1):

$$W^{\text{CLU}} = \sum_{i=1}^{n} [w_i - \pi] \tag{2}$$

In developing a variation on CLGU for the variable population case, Blackorby, et al. (1986, p.139) allow for the possibility that vectors of well-being levels such as $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2 \ldots \mathbf{w}_n)$ might not be 'ranked'. These are cases where, for two vectors **w*** and **w**', it is false that **w*** is 'at least as good as' **w**', and false that **w**' is 'at least as good as' **w***. Blackorby et al. are certainly using 'at least as good as' to refer to R rather than A.[15] So when they say that two well-being vectors are not ranked, they mean that the vectors are B-incommensurate. They consider the well-being vector of one population, which I write, $\mathbf{w}^e$, with an alternative situation where someone is added to this population leaving the levels of well-being of existing people unchanged. I write this latter situation as $(\mathbf{w}^e, \mu)$ where μ is the well-being level of the additional person. They suppose that, for every population size, *n*, and every $\mathbf{w}^e \in \mathbb{R}^n$, there is a set of levels of μ, $K^n(\mathbf{w}^e)$, such that $\mathbf{w}^e$ and $(\mathbf{w}^e, \mu)$ are B-incommensurate. This is the 'critical set.' Their 'basic axiom' – the critical-set population principle – is that this set is bounded and non-empty (Blackorby et al. 1996, p. 139). If a person is added to the population above (below) all critical levels in $K^n(\mathbf{w}^e)$ then the resulting situation is better (worse) (Blackorby et al. 1996, p. 140). $K^n(\mathbf{w}^e)$ is an interval, though no assumption is made about the boundaries of the interval (Blackorby et al. 1996, pp.140–141). Blackorby et al. (1996, p. 141) add two further axioms. The first – the 'critical-set extension principle' – formalizes the notion that if each person added to the population has a level of well-being in the critical set, then the resulting well-being vector is B-incommensurate with the initial one. Finally, they assume that the set of critical levels is fixed and is *K*. These axioms are used to establish ICLGU, which has, as a special case, 'incomplete critical-level utilitarianism' (ICLU). The latter is the view that, for any population sizes *n*'

[14] To do this one might assume, for example, that the added member of the population lives on another planet.

[15] This is clear because 'at least as good as' is a 'social quasi-ordering' in the relevant part of their discussion (Blackorby et al. 1996, p. 139).

← (worse) → (better)

μ

Fig. 1. A well-being configuration

and n^* and any well-being vectors **w**' and **w***, with well-being levels in **w**' and **w*** written w'_i and w^*_i, respectively:

$$\text{(ICLU)} \quad \mathbf{w}'^{\mathrm{R}}\mathbf{w}^* \iff \sum_{i=1}^{n'} (w'_i - \pi) \geq \sum_{i=1}^{n^*} (w^*_i - \pi), \quad \forall \pi \in K$$

In relating Blackorby, Bossert and Donaldson's ICLU to the discussion of parity, I use diagrammatic representations to simplify the exposition. I only consider the case where one person is added to the population, at a level of well-being μ, leaving the rest of the population with an unchanged well-being vector $\mathbf{w}^e$. I assume that all possible levels of well-being of the extra person lie in increasing order on a line. As we move to the right (left), the well-being of the extra person gets better (worse). This representative device is a close relation of one which Broome has used (Broome 2000) and I call it a 'well-being configuration'. It is shown in Fig. 1. We can represent a version of ICLU – where K has finite upper and lower bounds – in terms of this configuration. Suppose there is an interval of critical levels such that ($\mathbf{w}^e$, μ) and $\mathbf{w}^e$ are B-incommensurate. These critical levels must fall in a zone which has a lower bound of π_1 and an upper bound of π_2. It is shown in Fig. 2. If $\mu < \pi_1$, μ is in the 'worse zone' and $\mathbf{w}^e$B($\mathbf{w}^e$, μ) while if $\mu > \pi_2$, ($\mathbf{w}^e$, μ)B$\mathbf{w}^e$ and μ is in the 'better zone'. If $\pi_1 \leq \mu \leq \pi_2$, furthermore, it must be false that $\mathbf{w}^e$E ($\mathbf{w}^e$, μ). If it were true that $\mathbf{w}^e$E ($\mathbf{w}^e$, μ) and μ is slightly increased (decreased) it would be in the better (worse) zone.[16] This is clearly not the case if there is a zone or interval between the better and worse zones. The zone in between the better and worse zones is, thus, a zone of B-incommensurateness. While ICLU can involve a B-incommensurate zone of this sort, it does not tell us much about the zone.

Blackorby, Bossert and Donaldson's motivation for focussing on an interval of critical levels rather than a unique level emerged from John Broome's worry about the choice of a single 'critical level'. Such a level must be chosen, they think, so that it is '[o]n the one hand, high enough to avoid the repugnant conclusion by more than a little bit and, on the other, it should be low enough so that it does not prevent the existence of satisfactory lives' (Blackorby et al. 1997, p. 218). Since no unique level has this property, they argue in favour of a set of critical levels. They give few hints about what such an interval might look like. Nonetheless, they suggest that '[t]he interval

[16] For a similar argument see Blackorby et al. (1996, p. 140).

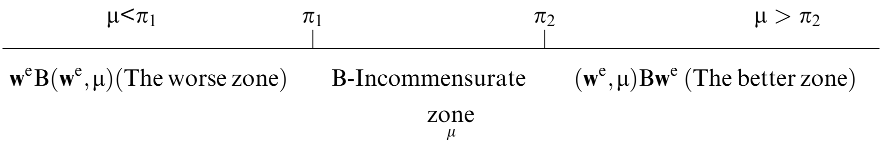

Fig. 2. B-incommensurateness with exact borderlines

might be chosen, for example, with a lifetime utility level that represents a barely satisfactory level of well-being at the bottom end and one that represents a life that is more than satisfactory but short of flourishing at the top' (Blackorby et al. 1997, p. 218).

The account of parity might help to determine the nature of the zone. Recall the 'mark of parity': if two states of affairs are on a par, a significant improvement (worsening) of one makes it better (worse) than the other, while small changes in value do not make one better than the other. If levels of μ in the zone are 'satisfactory', and it is a zone of parity, then only a significant improvement in μ would make $(\mathbf{w}^e, \mu)\mathbf{B}\mathbf{w}^e$, and only a significantly worsening of it would make $\mathbf{w}^e\mathrm{B}(\mathbf{w}^e, \mu)$. This seems to be close to Blackorby, Bossert and Donaldson's intuition in the passage just quoted. One needs, however, to be careful about what one means by a 'significant' improvement or worsening in this context. If we are to be true to Blackorby, Bossert and Donaldson's intuition, then a significant improvement is one that leaves μ outside the range of satisfactory lives so that it has become a flourishing life. At the other end, a significant worsening in μ leaves μ below any life which is even barely satisfactory, so that in the worse zone life is miserable. The idea is that if μ is a satisfactory level of well-being more than a slight improvement (worsening) in μ is required for it to become a level of well-being which constitutes a flourishing (miserable) life.

There are reasons to doubt, however, that Fig. 2 is best understood in terms of a zone of parity. Take a level of μ which is in the B-incommensurate zone, on the edge of the better zone. At this point, $(\mathbf{w}^e, \mu)$ and $\mathbf{w}^e$ are B-incommensurate, and a *tiny* increase in μ would make $(\mathbf{w}^e, \mu)\mathbf{B}\mathbf{w}^e$. If so, it is false that $(\mathbf{w}^e, \mu)\mathbf{P}\mathbf{w}^e$ at this point. A similar argument can be made for a level of μ in the B-incommensurate zone on the edge of the worse zone. So the B-incommensurate zone in Fig. 2 cannot describe a zone of parity. If the zone of B-incommensurateness is a zone of parity, it cannot have sharp borderlines.

6 Parity and vagueness

Abandoning sharp borderlines complicates matters. Thus far, we have been working with classical logic. However, what makes one level of well-being 'significantly better' or 'significantly worse' than another may be vague. One reason why one might think that there is vagueness is that 'significantly better than' can generate a Sorites paradox. If x is significantly better than y, it is plausible that something which is a little less good than x, x^{-}, is also significantly better than y. If x^{-} is significantly better than y, then by the same reasoning, so is something slightly worse than x^{-}, x^{--}. However, if we continue with this kind of reasoning, we can arrive at the conclusion that y is significantly better than itself. One can, thus, be led to a contradiction: irreflexivity of B (Theorem 1) is violated. This is a standard case of a Sorites paradox.

Exactness of the borderlines between the various zones is presupposed in Fig. 2. One might amend Fig. 2 using a specific account of vagueness – Kit Fine's supervaluationism (Fine 1975). This approach allows for various 'admissible' ways of making a borderline more precise, or admissible 'precisifications' of the borderline. If there are many admissible ways of making a borderline more precise, Fine's account only classifies statements as 'super-true' if they are true for all admissible precisifications. So if a statement is true for some admissible precisifications but not others, it is not super-true. A certain amount of 'higher-order vagueness' is allowed for because the notion of 'admissible' may itself be vague.

We can apply this approach to the population problem. One might, for example, suppose that the boundaries of the B-incommensurate zone are inexact. This is certainly plausible. If the boundaries relate to the borderlines between satisfactory lives and those which are, respectively 'flourishing' or 'miserable', it is certainly plausible that they are vague. If so, what makes a change in the B-incommensurate zone 'significantly better' (or 'significantly worse') must also be vague. Suppose next that there are vague zones between the B-incommensurate zone and the better and worse zones and that there is no 'higher-order vagueness'. In this situation, the range of admissible precisifications is exact and all vagueness is of 'first-order'. Suppose that v' is the lowest admissible precisification of the lower boundary of the B-incommensurate zone and that v'' is the highest admissible precisification of the upper boundary of the B-incommensurate zone. This situation is represented in Fig. 3.[17] In the zones between the B-incommensurate zone and the better and worse zones, there is first-order vagueness involving different admissible precisifications, which is shown by broken lines. Once this vagueness is accounted for, there is a zone of levels of μ which definitely fall in the worse zone – the levels below v'- and there is a zone of levels of μ which definitely fall in the better zone – those which are above v''. Furthermore, for all levels of μ in the B-incommensurate zone $\mathbf{w}^{e}$ and ($\mathbf{w}^{e}$, μ) are definitely on a par. Vagueness and parity are quite distinct, nonetheless.

This account cannot be right if the B-incommensurate zone is a zone of parity. Suppose μ is in the B-incommensurate zone, on the edge of the vague zone which falls between the B-incommensurate zone and the better (worse) zone. A slight increase (decrease) in μ would put μ in the vague zone. If the broken lines relate to admissible precisifications of the borderline between the

[17] This picture is closely related to one presented by Chang (2002a, p. 168).

$\mu < v'$ v' v'' $\mu > v''$

$\mathbf{w}^e B(\mathbf{w}^e, \mu)$ (The worse zone) B-Incommensurate zone μ $(\mathbf{w}^c, \mu) B \mathbf{w}^c$ (The better zone)

Fig. 3. B-incommensurateness with vague borderlines

zones, there must be some admissible precisification according to which a point just to the right (left) of the B-incommensurate zone makes ($\mathbf{w}^e$, μ) better than (worse than) $\mathbf{w}^e$. If, for all levels of μ in the B-incommensurate zone $\mathbf{w}^e P(\mathbf{w}^e, \mu)$, this cannot be right: slightly increasing (or decreasing) μ cannot make the difference between states being on a par and one being better (worse) than the other, on *any* admissible way of making the borderlines more precise.

We must thus reject the view that we can allow for vague zones between the B-incommensurate zone while only allowing for first-order vagueness. The obvious way of responding to this problem is to allow for higher-order vagueness, so that the edges of the vague zones in Fig. 3 are also imprecise. So at the edge of the zone of first-order vagueness bordering the better (worse) zone more than a slight increase (decrease) in μ is needed for μ to fall into the better (worse) zone. Similarly, at the edge of the B-incommensurate zone, it requires more than a slight increase (decrease) in μ for μ to fall into the zone of first-order vagueness. If we reinterpret Fig. 3 to allow for such higher-order vagueness, so that the vague zones involve both higher-order and first-order vagueness that might restore the plausibility of the idea of a zone of parity with vague borderlines. Furthermore, there is no reason to think that there is any further vagueness once such vagueness is taken into account. In Fine's account, higher-order vagueness only applies to the range of admissible precisifications, and does not spread any further.

There is another way to object to the argument in favour of interpreting Fig. 3 in terms of first-order vagueness. It suggests that when thinking about the vagueness of 'significantly better than' in making that argument the key comparison is between ($\mathbf{w}^e$, μ) and $\mathbf{w}^e$ for some specific value of μ. It is a question of what is significantly better or worse than that value of μ. Imprecision about this generates a pair of vague zones for any value of μ in the zone, rather than a unique pair of vague zones on either side of a zone of B-incommensurateness.[18] In particular, a pair of vague zones generated by one level of μ in the zone of B-incommensurateness may not be the same as a pair of zones generated by some other value of μ in the same zone.

Here it is worth returning to the way in which the notion of a 'significant improvement' ('significant worsening') was related to Blackorby, Bossert and Donaldson's intuition earlier. The idea is that an increase (decrease) in the

[18] I am very grateful to an anonymous referee for pointing this out to me.

level of μ is 'significant' if it is large enough to make for flourishing (to make for a miserable life). So what makes a change 'significant' will depend on where μ falls in the B-incommensurate zone. The vague zones relate, respectively, to the borderline between a life which is at the top end of the range of satisfactory lives and one which is flourishing, and the borderline between one which is at the bottom end of the range and one which is miserable. The vague zones are, thus, fixed, and do not depend on any particular value of μ.

Broome (2001, p. 14) has argued against the picture described in Fig. 3. There are two arguments which are relevant. One only applies if the B-incommensurate zone is interpreted as a zone of parity. Broome argues against Griffin's notion of 'rough equality'. If the zone of parity is a zone of 'rough equality' – in the ordinary sense – it cannot be *very wide*. Everything in the zone must be very close in value. Broome thinks that the zone that falls between the better and worse zones *might* actually be very wide.[19] For this reason, he does not think it is a zone of rough equality. However, even if the zone of parity is narrow,[20] the zone between the better and worse zones can be very wide if the zones of vagueness (incorporating higher-order vagueness) are wide enough.

Broome's other argument depends on a principle which he adopts – the 'collapsing principle' (see Broome 1997, 2001, 2004). This principle is controversial, is rejected by Chang (2002a), and I do not discuss it here.[21] Broome thinks that even if one does not accept this principle, one might find his view attractive (Broome 2001, pp. 12–13). So I focus on Broome's own view. Broome thinks that there is a zone of levels of well-being between the better and worse zones and that this is a zone of vagueness. There is, on his view, just one level of well being above which $(\mathbf{w}^e, \mu)\mathrm{B}\mathbf{w}^e$ and below which $\mathbf{w}^e\mathrm{B}(\mathbf{w}^e, \mu)$. This is a unique level π such that $(\mathbf{w}^e, \mu)\mathrm{E}\mathbf{w}^e$ when $\mu = \pi$.[22] However, there

[19] While Broome has not made this argument in print, he often makes this argument in seminar discussions and it is related to a view he takes in his discussion of Griffin's 'rough equality' view. See Broome (2000, pp. 29–30).

[20] It is certainly plausible that a zone of parity will be narrow in one sense. If the zone is 'very wide', then a level of μ at the top end of the zone might be much higher than one at the bottom end. If the level of μ at the top end of the zone is much higher than one at the bottom end, the quality of life at the top end is likely to be significantly better than at the bottom end. If so, the B-incommensurate zone cannot be a zone of parity. The zone of parity must be 'narrow' in the sense that it cannot be 'very wide'. Qizilbash (2002, pp.153–155) makes much the same point. This does suggest that if Blackorby, Bossert and Donaldson's position is interpreted in terms of parity, the range of satisfactory lives in the B-incommensurate zone cannot be 'very wide'.

[21] Broome's discussion (Broome 1997) also involves degrees of truth, which I am not assuming here.

[22] Broome (2001, 2004) actually describes it as the 'neutral level', though he relates his position to Blackorby, Bossert and Donaldson's, and it is harmless to interpret his position in terms of an unique critical level which is vague.

$\mu < \pi$' π' π'' $\mu > \pi$''

$\mathbf{w}^e B(\mathbf{w}^e, \mu)$(The worse zone) Vague zone $(\mathbf{w}^e, \mu) B \mathbf{w}^e$ (The better zone)

μ

Fig. 4. One vague zone

is vagueness about π. Broome uses supervaluationism to describe this situation. Suppose that the level could be made more precise in various ways and that π' and π'' are the lowest and highest admissible precisifications. Then there is a single vague zone with an upper and lower bound, as shown in Fig. 4. This view certainly has the attraction that it dispenses with problems relating to B-incommensurateness while still preserving a zone between the better and worse zones. A similar picture is presented by Broome (2001, p. 13).

One way of objecting to this picture goes as follows. What notion are we making precise when we select a level of well-being associated with π? Our intuition cannot be very dissimilar to Blackorby, Bossert and Donaldson's. Consider then the borderlines of the vague zone in Fig. 4. When we look at one end of the vague zone – close to the better zone – we are looking for a level of well-being which is 'just short of flourishing'. 'Just short of flourishing' is a vague predicate. One might then think that the vague zone in Fig. 4 is generated by the vagueness of this predicate. On the other hand, at the bottom end of the vague zone we are looking for a level of well-being at which life is 'barely satisfactory'. This is also a vague predicate. We can again define a range of levels of well-being which are generated by the vagueness of this predicate. The picture in Fig. 4 suggests that the ranges of vagueness generated by 'just short of flourishing' and 'barely satisfactory' are one and the same: there is only one zone of vagueness as one moves from left to right. However, it is plausible that there is at least one level of well-being which is definitely better than *barely* satisfactory, while not definitely being just short of flourishing. So the two zones of vagueness need not be identical. If so, the picture in Fig. 4 is wrong. The correct picture, on this line of reasoning, should involve two non-identical vague zones – one for each predicate.

Alternatively, one might argue in defence of the picture in Fig. 4 that there is only one predicate involved. There is a unique level of well-being which is 'satisfactory'. This is the unique critical level π. There is vagueness about what level of well-being this is, because 'satisfactory' is a vague predicate. Vagueness about this generates a unique zone of vagueness. This position is weak because it is implausible that only one level of well-being classifies as 'satisfactory'. These arguments are clearly not 'knock-down' arguments against Broome's position. They relate Broome's position to the intuition in Blackorby, Bossert and Donaldson's discussion of their own view. That may not be a fair test. These arguments merely signal worries about the intuition that there is a single vague zone between the better and worse zones. Nonetheless, Broome's position certainly has the advantage of parsimony.

7 Conclusions

Parfit's mere addition paradox can be solved using two primitive relations in terms of which parity is here defined. Blackorby, Bossert and Donaldson's ICLU also allows for B-incommensurateness, and can, plausibly, be understood in terms of parity. However, the 'mark' of parity suggests that the zone of B-incommensurateness has vague borderlines. ICLU has to be qualified to allow for this, if it is understood in terms of parity. The resulting position involves two vague zones, allows for higher-order vagueness and contrasts with Broome's recent discussion, which involves an unique vague zone.

Appendix 1

In this appendix, certain theorems which are stated in the paper are proven.

Theorem 1. B is irreflexive, asymmetric and transitive.

Proof. Suppose not. First, suppose B is not irreflexive. So $\exists x \in \mathrm{X}$, $x\mathrm{B}x$. Then, from definition 1, $x\mathrm{R}x \wedge \sim(x\mathrm{R}x)$ and we have a contradiction: B is irreflexive. Next suppose that it is not asymmetric, then $\exists x,y \in \mathrm{X}$: $x\mathrm{B}y \wedge y\mathrm{B}x$. Then (from Definition 1) $[x\mathrm{R}y \wedge \sim(y\mathrm{R}x)] \wedge [y\mathrm{R}x \wedge \sim(x\mathrm{R}y)]$. From which it follows that $x\mathrm{R}y \wedge \sim(x\mathrm{R}y)$ which is a contradiction, and $y\mathrm{R}x \wedge \sim(y\mathrm{R}x)$ which is also a contradiction. So B is asymmetric. Finally, suppose that B is not transitive. From postulate 1, R is a quasi-ordering, and B is transitive. Again we have a contradiction and B is transitive. ■

Theorem 2. $x\mathrm{C}y \Rightarrow (x\mathrm{R}y \vee y\mathrm{R}x) \vee x\mathrm{P}y$

Proof. First we need to show that $x\mathrm{C}y \Leftrightarrow (x\mathrm{R}y \vee y\mathrm{R}x) \vee x\mathrm{P}y$. Suppose not. Then $\exists x,y \in \mathrm{X}$: $(x\mathrm{C}y) \wedge \sim[(x\mathrm{R}y \vee y\mathrm{R}x) \vee (x\mathrm{P}y)]$. The second conjunct implies $\sim(x\mathrm{R}y \vee y\mathrm{R}x) \wedge \sim(x\mathrm{P}y)$. If $x\mathrm{C}y \wedge \sim(x\mathrm{R}y \vee y\mathrm{R}x)$ then $x\mathrm{P}y$, which contradicts $\sim(x\mathrm{P}y)$. If, on the other hand, $\sim(x\mathrm{P}y)$, then (from Definition 4) $\sim[x\mathrm{C}y \wedge \sim(x\mathrm{R}y \vee y\mathrm{R}x)]$. This in turn implies either $\sim(x\mathrm{C}y)$ which contradicts $x\mathrm{C}y$, or $\sim\sim(x\mathrm{R}y \vee y\mathrm{R}x)$, which contradicts $\sim(x\mathrm{R}y \vee y\mathrm{R}x)$. Next we need to show that $(x\mathrm{R}y \vee y\mathrm{R}x) \vee x\mathrm{P}y \Rightarrow x\mathrm{C}y$. Suppose not. Then $\exists x,y \in \mathrm{X}$:$(x\mathrm{R}y \vee y\mathrm{R}x) \vee (x\mathrm{P}y) \wedge \sim(x\mathrm{C}y)$. If $x\mathrm{R}y \vee y\mathrm{R}x$ then from postulate 3, $x\mathrm{C}y$ so that we have a contradiction. On the other hand, if $x\mathrm{P}y$ then (from Definition 4) again $x\mathrm{C}y$, and we have a contradiction. ■

Theorem 3. $x\mathrm{R}y \Rightarrow x\mathrm{A}y$

Proof. Suppose not. Then $\exists x,y \in \mathrm{X}$: $x\mathrm{R}y \wedge \sim(x\mathrm{A}y)$. $\sim(x\mathrm{A}y)$ means – from definition 5 – that $\sim[x\mathrm{C}y \wedge \sim(y\mathrm{B}x)]$, which implies $\sim(x\mathrm{C}y) \vee \sim\sim(y\mathrm{B}x)$. If $\sim(x\mathrm{C}y) \wedge x\mathrm{R}y$, we have a contradiction, because from postulate 3, $x\mathrm{R}y$ implies $x\mathrm{C}y$. On the other hand, $\sim\sim(y\mathrm{B}x)$ implies $y\mathrm{B}x$, which means $y\mathrm{R}x \wedge \sim(x\mathrm{R}y)$. However, we have $x\mathrm{R}y$, so again we have a contradiction. ■

Theorem 4. E is an equivalence relation.

Proof. To be an equivalence relation E must be symmetric, transitive and reflexive. Is E reflexive? Suppose it is not. Then $\exists x \in \mathrm{X}$: $\sim(x\mathrm{E}x)$. In which case, from Definition 2, we have $\sim(x\mathrm{R}x)$. But from postulate 1, R is reflexive and thus $x\mathrm{R}x$. So we have a contradiction: E is reflexive. Is E symmetric? Suppose not. Then $\exists x,y \in \mathrm{X}$: $x\mathrm{E}y \wedge \sim(y\mathrm{E}x)$. But then, from Definition 2, we have $x\mathrm{R}y \wedge y\mathrm{R}x \wedge \sim(y\mathrm{R}x \wedge x\mathrm{R}y)$, which yields $(y\mathrm{R}x \wedge x\mathrm{R}y) \wedge \sim(y\mathrm{R}x \wedge x\mathrm{R}y)$ which is a contradiction. So E is symmetric. Is E transitive? Suppose it is non-transitive. From postulate 1, R is a quasi-ordering, and E is transitive. So again we have a contradiction and E is transitive, and an equivalence relation. ■

Theorem 5. $(x\mathrm{A}y \vee y\mathrm{A}x) \Leftrightarrow x\mathrm{C}y$.

Proof. Suppose not. First suppose: $\exists x,y$: $(x\mathrm{A}y \vee y\mathrm{A}x) \wedge \sim(x\mathrm{C}y)$. Consider $x\mathrm{A}y \vee y\mathrm{A}x$. If the first disjunct is true then (from Definition 5) $x\mathrm{C}y$. However, we know that $\sim(x\mathrm{C}y)$. So we have a contradiction. Alternatively, $y\mathrm{A}x$. This implies $y\mathrm{C}x$ which, given symmetry of C (from postulate 2), implies $x\mathrm{C}y$. Again we have a contradiction. Suppose, on the other hand, that $\exists x,y$: $x\mathrm{C}y \wedge \sim(x\mathrm{A}y \vee y\mathrm{A}x)$.$\sim(x\mathrm{A}y \vee y\mathrm{A}x)$ implies $\sim(x\mathrm{A}y) \wedge \sim(y\mathrm{A}x)$. The first conjunct means $\sim[x\mathrm{C}y \wedge \sim(y\mathrm{B}x)]$. However, $x\mathrm{C}y$, so it must be that $\sim\sim(y\mathrm{B}x)$, which implies $y\mathrm{B}x$. However, $x\mathrm{C}y$ implies $y\mathrm{C}x$ (given postulate 2), and if $y\mathrm{B}x$ then $\sim(x\mathrm{B}y)$ (given B is asymmetric from Theorem 1). So we can conclude that: $y\mathrm{C}x \wedge \sim(x\mathrm{B}y)$. However, we have $\sim(y\mathrm{A}x)$ which means $\sim[y\mathrm{C}x \wedge \sim(x\mathrm{B}y)]$. So we have a contradiction. ■

Appendix 2. The repugnant conclusion

In the text, it is argued that Parfit claims that if we accept $b\mathrm{B}a$ we might accept the repugnant conclusion. For this reason, one might claim that $a\mathrm{B}b$, and it was assumed that $\sim(b\mathrm{B}a)$. In state a the quality of life is α, the quality of life in b is $\beta = 0.8\ \alpha$, and the quality of life in c (d etc.) is $0.8\ \beta$ ($0.8^2\ \beta$ etc.). There is a finite sequence of states of affairs $a, b, c, \ldots, z$. At the end of this sequence, we have a state of affairs z, involving a life which is barely worth living. Suppose that 10 billion people are alive in a, and that in b (c etc.) twice as many people are alive as in a (b etc.) Suppose also that the people living in a have a very high quality of life.

Parfit's repugnant conclusion (1984, p. 388) is the following: for any possible population of at least 10 billion people, all with a very high quality of life, there must be a much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living. If we suppose that bBa, because it is better to double the size of a population, if the people in the larger population are living at 0.8 the level of well-being which is enjoyed by people in the smaller population, the same reasoning implies cBb, dBc, etc. ... and zBy. Then, by transitivity of B, zBa. This implies the repugnant conclusion.

References

Arrhenius G (2000) Future generations. A challenge for moral theory. Uppsala University, Uppsala
Blackorby C, Bossert W, Donaldson D (1996) Quasi-orderings and population ethics. Soc Choice Welfare 13: 129–151
Blackorby C, Bossert W, Donaldson D (1997) Critical-level utilitarianism and the population-ethics dilemma. Econ Philos 13:197–230
Broome J (1996) The welfare economics of population. Oxford Econ Papers 48:177–193
Broome J (1997) 'Is incommensurability vagueness?' In: Chang R (ed) Incommensurability, incomparability and practical reason. Harvard University Press, Cambridge (Mass)
Broome J (1999) Ethics out of economics. Cambridge University Press, Cambridge U.K.
Broome J (2000) Incommensurable values. In: Crisp R, Hooker B, (eds) Well-being and morality. Essays in Honour of James Griffin. Clarendon Press, Oxford
Broome J (2001) Greedy neutrality of value. In: W Rabinowicz (ed) Value and choice. vol. 2. Some common themes in decision theory and moral philosophy. Lund Philos Reports 1: 7–16
Broome J (2004) Weighing lives. Oxford University Press, Oxford
Carlson E (1998) Mere addition and two trilemmas of population ethics. Econ Philos 14: 283–306
Chang R (1997) Introduction. In Chang R, (ed) Incommensurability, incomparability and practical reason. Harvard University Press, Cambridge (Mass)
Chang R (2002a) Making comparisons count. Routledge, New York and London
Chang R (2002b) The possibility of parity. Ethics 112: 659–688
Fine K (1975) Vagueness, truth and logic. Synthese 30: 265–300
Griffin JP (1986) Well-being: Its meaning, measurement and moral importance. Clarendon Press, Oxford
Griffin JP (1997) Incommensurability: What's the problem? In: Chang R, (ed) Incommensurability, incomparability and practical reason. Harvard University Press, Cambridge (Mass)
Griffin JP (2000) Replies. In: Crisp R, Hooker B, (eds) Well-being and morality: Essays in honour of James Griffin. Oxford University Press, Oxford
Parfit D (1984) Reasons and persons. Clarendon Press, Oxford
Qizilbash M (2000) Comparability of values, rough equality and vagueness: Griffin and Broome on incommensurability of values. Utilitas 12: 223–240
Qizilbash M (2002) Rationality, comparability and maximization. Econ Philos 18: 141–156

Sen AK (1979) Collective choice and social welfare. North-Holland, Amsterdam
Temkin L (1987) Intransitivity and the mere addition paradox. Philos Pub Affairs 16: 138–187