

On the social and personal value of existence¹

Marc Fleurbaey and Alex Voorhoeve

April 22, 2014

Abstract

If a potential person would have a good life if he were to come into existence, can we coherently regard his coming into existence as better for him than his never coming into existence? And can we regard the situation in which he never comes into existence as worse for him? In this paper, we argue that both questions should be answered affirmatively. We also explain where prominent arguments to differing conclusions go wrong. Finally, we explore the relevance of our answers to issues in population ethics.

Keywords: Population ethics; value of existence.

Introduction

It is certainly better for all of us that John Broome exists and it would have been worse for us if he had never existed. This alone would have given anyone with the power to decide whether John Broome would exist a good reason to ensure his existence. But supposing, as we hope is true, that Broome's life is also good for him, in the sense that he enjoys high lifetime well-being, is it therefore also better *for him* that he exists and would it have been worse for him if he had never existed? In this chapter, we argue that these questions should

¹ We thank Richard Bradley, Susanne Burri, and Nils Holtug for comments. Alex Voorhoeve's work was supported by the British Arts and Humanities Research Council through grant AH/J006033/1.

be answered affirmatively. We also explore the relevance of our answers to issues in population ethics.

1. The alleged absurdity of the claim that existence is better than never existing

Broome (1999, 2004) has developed an ethics of population which does not involve any comparison, from the perspective of a person's self-interest, between her existence and her never existing. Whether it is better to bring a person into existence than not to bring her into existence is only considered in terms of social welfare (or social goodness, more generally). Broome argues that such social evaluation should take the following form. If a potential individual would have a level of well-being equivalent to what we shall call the "population-value indifference level" of well-being (often referred to as the "critical level"), then her coming into existence with this level of well-being would be a matter of indifference from the perspective of social value. If she would have a level of well-being in excess of this level, then her existence would increase the value of the population, and so make the world a better place. If this individual would instead have a level of well-being below this level, then her existence would make the world worse. This approach therefore has no need for the notion of a life that is better for an individual (considering only her well-being) than her not coming into existence would be for her.

Broome offers the following argument for eschewing the notion of a life that is better for the person than her never coming into existence is for her. It cannot be better for a person to exist than never to exist, because this would mean that it would be worse for her if she never came into existence, a statement that is deemed absurd because, Broome claims, there would be no one for whom the latter would be worse. As he writes:

“[I]t cannot ever be *true* that it is better for a person that she lives than that she should never have lived at all. If it were better for a person that she lives than that she should never have lived at all, then if she had never lived at all, that would have been worse for her than if she had lived. But if she had never lived at all, there would have been no her for it to be worse for, so it could not have been worse for her” (1999, p. 168; emphasis in original).

As a consequence, in Broome’s view, the notion of “a life worth living” cannot be used in the sense of comparing a life of a particular quality to never existing. Rather, it only refers to the value of extending life (i.e., making a given life last longer, without considering the problem of the creation of a new life), and should really be called a “life worth continuing.”

In the literature, this argument is often decomposed into the following basic elements:²

Call a description of who exists and at what level(s) of well-being a “social situation.”

(1) Social situation A is better than social situation B for a person P if and only if B is worse than A for P.³

(2) If A is better than B for P, then A would be better than B for P if it obtained.⁴

(3) (a) A person who never exists has no well-being; and

² Throughout this argument, for simplicity, “better than” stands for “better than (worse than)” with all needed adjustments to the corresponding formulation, and similarly “worse than” stands for “worse than (better than).”

³ Bykvist (2007) refers to this premise as “Converse.”

⁴ Bykvist (2007) refers to this premise as “Accessibility;” Arrhenius and Rabinowicz (2010) as “Subjunctive Connection.” Bykvist (2013) refers to it as “Counterfactual Support.”

(b) Nothing can be better than anything else for a person who never exists.⁵

On the basis of these conditions, the argument runs as follows. From (1) and (2), it follows that if a particular life is better for P than never existing, then never existing would be worse for P than that life if P never came into existence. From (3), it follows that never existing could not be worse for P than anything if P never came into existence. Hence it cannot be the case that there is a life that is better for P than never existing.

2. Others' responses to the argument

Several authors who endorse the argument remark that while it rules out *comparative* judgments about whether a given life is better or worse for a person than his never existing, it permits *non-comparative* claims about the good and ill done to a person by causing her to exist with a particular level of well-being. For example, Krister Bykvist (2007, 2013) endorses the aforementioned three conditions but holds that a life may be *good* for an individual even though it is not better for her than never existing. Similarly, Elizabeth Harman (2004) argues that one can *harm* someone by creating her if her life contains particular bad features, where it is not necessary for the existence of such harm that leading the life in question is worse for her than never existing.

However, many authors seem uncomfortable with the absence of any comparative evaluation of existence vis-à-vis never existing. They therefore believe that we should try to relax at least one of the three premises. One possibility is proposed by Melinda Roberts (2003, pp. 168-9), who rejects (3a) and (3b), when she assimilates never existing to a zero level of well-being, implying that in some sense a person who never exists has a level of

⁵ Bykvist (2007) refers to this as "Actualism;" Bykvist (2013) as "Well-being Implies Being."

well-being. However, we believe Roberts is wrong to reject (3a), since we do not think that it makes sense to assign a level of well-being to a never-existing person. After all, well-being implies being.

A different approach is taken by Nils Holtug (2001, p. 374) and Gustaf Arrhenius and Wlodek Rabinowicz (2010), who argue that premise (2) is too strong. What is needed for the claim that existence in social situation A with a high level of well-being is better for a person P than his never existing in B is a three-part relation between A, B, and P. This relation can exist, they claim, only if the person exists. Since, when P exists, all three relata exist, it can make sense to assert that A is better for him than B. But, contrary to (2), one cannot conclude from this that *it would have been worse for P if he had not existed*, because in that scenario, one of the elements in the relation in question would not exist. As Arrhenius and Rabinowicz put it:

“A triadic relation consisting in one state [leading a life of a particular quality] being better for a person than another state [never existing] cannot hold unless all three relata exist. Now, the states in question are abstract objects and can indeed be assumed to exist even if they do not actually obtain. However, if [as is assumed] a person is a concrete object, (...) then the relation could not hold if [the] person weren't alive, since [one of the relata] would not exist. Consequently, even if it is better for [the person] to exist than not to exist (...), [contrary to premise (2),] it doesn't follow that *it would have been worse for [the person] if [he] did not exist*, since one of the relata would then be absent” (2010, pp. 405-6).

As a consequence, they propose to weaken (2) as follows to make an exception for cases in which P doesn't exist in one of the two states being compared:

- (2)' (a) If P exists in both A and B, then A is better for P than B if and only if B would be worse for P than A, if B obtained; and
- (b) If P exists in A but not in B, then A can be better for P than B although B would not be worse for P than A, if B obtained (2010, p. 407).

3. Why it is not absurd to claim that existence is better than never existing

In contrast with the aforementioned authors, we accept (1), (2), and (3a), but propose to revise (3b) in such a way that the resulting principle is weak enough to be consistent with the claim that a good life can be better for a person than never existing.

As we have seen, Holtug (2001) and Arrhenius and Rabinowicz (2010) accept (3b) because they want to avoid assertions of the following kind about someone who did not come into existence but who would have had a good life, had he existed: “It *is* worse for this person not to exist.” They also want to avoid assertions of the following kind about someone who exists with a high level of well-being: “It *would have been* worse for her had she never come into existence.” The reason seems to be that they want their analysis to be consistent with the following principle:

No Properties of the Never-Existent: An individual who never exists cannot have any properties, not even the relational property of something being better or worse for her (Holtug 2001, p. 370).

We shall now argue that this principle should be rejected because it is too strong.

In certain circumstances, there is a clear appeal to the thesis that a never-existing person has no properties. For example, the thesis is true when the comparison is made exclusively

between social situations in which the person never exists. But the thesis seems to us to be false when the comparison involves a counterfactual situation in which the person would exist with definite, identifying characteristics.

Suppose we consider a never-existing person who could have existed in a counterfactual situation in which this person would have had a great life. Such assumptions give properties to the never-existing person we are thinking about. For one, this person has the property of “having a great life in the counterfactual situation.” More generally, this person has all the characteristics that are assigned to her in our description of her situation in the counterfactual state.

This is just a straightforward generalization of common discussions of counterfactual states. When we talk about what would have happened if Pope Francis had married, there is no metaphysical conundrum about the identity and characteristics of Jorge Maria Bergoglio in this counterfactual state. Similarly, when we talk about what would have happened if Vincent van Gogh would have had a fourth sister (with particular identifying characteristics) instead of his brother Theo, we imagine a person who is different than an existing person without making this hypothetical person a mysterious entity. By a very minor extension, we can talk meaningfully about hypothetical persons who never exist and who are not variations or replacements of existing persons, but whom we can nevertheless richly describe. We can talk about the additional children that Nelson would have had with Emma Hamilton if he had not died at Trafalgar (describing them in a manner that would uniquely identify them), and imagine the different lives they could have had depending on various counterfactual assumptions. There is no greater mystery in establishing the truth conditions for the statement that it is worse for Nelson’s hypothetical son not to have been born

(supposing he would have had a great life) than in establishing the truth conditions for the statement that it is worse for the Pope not to have married (supposing he would have had a good marriage).

In ordinary discourse, a person is not just a concrete entity who exists in the actual world. Instead, it is the *set* of possible descriptions associated with the same identity in all the counterfactual states we care to describe. This set may include states in which the person never exists. The literature seems to have failed to see the difference between persons as concrete objects and persons as they figure in ordinary discourse. A concrete object has properties only in the states in which it exists, and one cannot say that a person would be taller in a state in which she does not exist than in the current state in which she does. But when we discuss possible people, the topic of our debate is not only concrete persons. Instead, it is persons as they are considered by evaluators of possible worlds, evaluators who assess these worlds from the perspective of these persons' interests. For an observer who compares two different possible worlds, in one of which a person exists and in one of which she never exists, it makes sense to compare these worlds for the sake of this person. For a *discourse about* a person and how different states can be ranked for her sake, it does not even matter whether the person is purely fictional. The only requirement for meaningful comparative evaluation of states from the perspective of a person's interests (i.e., as better or worse for her) is that the description of every state in which she exists includes information on her level of well-being in that state and the description of every state in which she never exists mentions the fact that she does not come into existence.⁶

⁶ Arrhenius and Rabinowicz (2013), very briefly (in a footnote), and Bykvist (2013), more extensively, consider the possibility that merely possible persons have a weak form of existence. They rightly reject the idea that

In sum, suppose that in situation A, person P would have a good life, and that in B, he would never exist. Accepting (1) and (2), it makes sense to hold not merely that if A were to obtain, then A would be better for P than B, but also that, if B were the actual situation, then his never existing would be worse for him than A. If B were actual, then the fact that this person does not come into existence in the current situation is compatible with the fact that in the counterfactual situation there would be a person with characteristics that are sufficiently definite to make possible the latter comparison.

Note that, as mentioned, we endorse (3a). Since a person who never exists has no well-being, the statement “it is worse for her never to exist than to exist” cannot mean that in the former case, she has a lower level of well-being than in the latter. But one can sensibly hold that a particular life can be better for a person than never existing without assigning a level of well-being to never existing. It is sufficient that there is a level of well-being, when existing, that is deemed *equivalent* to never existing. (Call this the “personal-value indifference level” of well-being; it is often referred to as the “neutral level.”) Then, we submit, enjoying a greater well-being than this level implies that a person’s life is better for her than never existing. It also implies that her never existing would be worse for her than existence.

Our proposal is, therefore, to replace (3) with the following, weaker condition.

(3') (a) A non-existing person has no level of well-being; and

such objects can have something like a level of well-being. However, they do not seriously consider the possibility that if a merely possible person has well-defined well-being levels in counterfactual states, this is sufficient to make the comparison between the person’s lives in these states and his never existing meaningful in states in which he does not exist.

(b') Nothing can be better than anything else for a person who exists in none of the (actual and counterfactual) states under consideration.

The combination of (1), (2), and (3') is compatible with the idea that existence at a high level of well-being is better for P than non-existence and that this is true whether or not he actually exists. It is also compatible with holding that existence at a very low level of well-being is worse for P than non-existence and that this is true whether or not he exists.

4. A difficulty for other views

As we shall now explain, this lack of dependence of the truth of such judgments on the actual state of the world gives our proposal an advantage over the proposal advanced by Holtug and by Arrhenius and Rabinowicz. Consider a not-yet-existing, potential person P who would have an awful life if he were to come into existence and whose existence depends on our decision. On the view we propose, it would be worse for P to come into existence and it would be better for him if he never came into existence. This is just to say that, if we were acting solely for his sake, we would not create him. By contrast, the view proposed by Holtug and by Arrhenius and Rabinowicz holds that if P were to come into existence, then his existence would be worse for him and that if he did *not* come into existence, then his never existing would be neither better than, nor worse than, nor equally good as existence for him. These judgments alone do not determine how one should rank not-yet-existing, contingent person P's existence and his never existing when considering only P's interests. To do so, their view requires further assumptions.

Arrhenius and Rabinowicz appeal to the figure of potential, contingent P's "guardian angel," who must make decisions only for P's sake. Of this angel, they write:

“[I]f [P’s] guardian angel has a choice between bringing her charge into existence with negative welfare [and] not bringing [him] into existence at all, she would choose the latter. Moreover, if the guardian angel had the choice between bringing her charge into existence with a positive welfare [and] not bringing [him] into existence, she would choose the former. (...) [I]t seems reasonable to say that in the choice between bringing P into existence with negative welfare [and] not bringing [him] into existence at all, one ought to prefer the latter for P’s sake. Likewise, in the choice between bringing P into existence with positive welfare [and] not bringing [him] into existence at all, one ought to prefer the former for P’s sake” (2010, pp. 410-11).

They also propose that we take the expression “one ought to prefer never existing to existing at a low level of well-being for P’s sake” simply to mean “never existing is better for P than existing with a low level of well-being” (2010, p. 411).

Jointly, these proposals have unappealing implications. For they imply the following two statements about P’s situation:

- (i) When P does not yet exist, but might be brought into existence, the situation in which he exists is worse for him than the situation in which he does not exist, and the latter is better for him.
- (ii) If P is not brought into existence (and he can no longer be brought into existence), then his never-existing is no longer better for him than his existence.

This shift in the circumstances under which the statement “P’s never existing is better for him than his existence at a very low level of well-being” is held to be true is very peculiar. If never being brought into existence is better for P *now* (when he does not yet exist, but can

be brought into existence), then how can it fail to *remain* better for him once an irrevocable decision has been made not to bring him into existence?

Of course, an analogous problem arises for Arrhenius and Rabinowicz' view if the not-yet-existing, contingent P would have a wonderful life. On their view, it would then be true that:

(iii) When P does not yet exist, but might be brought into existence, the situation in which he exists is better for him than the situation in which he does not exist and the latter is worse for him; and

(iv) If P is irrevocably not brought into existence, then never existing is no longer worse for him than existence.

Again, such a shift seems to us very implausible.

Holtug (2001, p. 375ff) takes a different approach, but nonetheless seems to face the same problem. According to Holtug, we "extrinsically harm" a person by failing to bring him into existence rather than causing him to exist if and only if (a) bringing him into existence would have been intrinsically better for him; and (b) bringing him into existence would not intrinsically harm him. As a consequence, Holtug holds that if not-yet-existing, contingent P's life would be a good one, then it would harm P not to bring him into existence and it would benefit P to bring him into existence. He adds that, on his view, "it is difficult to resist the claim that, everything else being equal, we ought to cause [P] to exist" (p. 384). The reason that Holtug finds this claim difficult to resist is, we surmise, that, on his view, if one takes only the interests of potential, contingent P into account and so chooses for P's sake, one ought to prefer P's existence to his never existing. If one then makes the further, natural assumption that "preferring P's existence to his never existing for P's sake" is just to

say that one takes P's existence to be better for him than his never existing, then Holtug's view leads to the same conclusion as Arrhenius and Rabinowicz' view. When P does not yet exist and his existence depends on our choice, then the state in which P is brought into existence with a high level of well-being is judged to be better for him, but if we choose not to bring him into existence, and this choice is irreversible, then it is no longer considered worse for P that he was not brought into existence.

In sum, we have argued that views that reject (2) in favour of (2') face a dilemma. They must either refuse to rank potential, contingent P's existence against P's never existing for P's sake, or give up the natural assumption that ranking them for his sake is just to rank them as better or worse for him. The view we have proposed does not face this dilemma.

5. Implications for population ethics

Why is it so important whether one can compare existence to never existing from the point of view of the affected person's interests? The stakes are often traced to the following principle.

Person-Affecting Principle (PAP): A social situation cannot be better than another if it is not better for someone.

Suppose this principle is correct. Also suppose (contrary to our arguments) that one cannot compare existence to never existing from the perspective of a person's interests. Then one faces the unpleasant prospect of being unable to compare social situations with totally disjoint populations. This is so even when social situation A is obviously better than B, because, say, everyone who exists in A (call them the x-people) has a very high level of well-being and everyone who exists in B (the y-people) has an awful life.

Accepting that existence can be better for a person than never existing seems a way out of this conundrum. If one adopts Holtug's and Arrhenius and Rabinowicz' proposal, then the PAP is compatible with the claim that A is better than B, because it is better for the *x*-people to exist with a high level of well-being rather than never to exist. However, as Arrhenius and Rabinowicz (2013, p. 15) note, when paired with their proposal for ranking existence vis-à-vis never existing, the PAP is still too restrictive. Consider the choice between A, which contains only *x*-people who lead wonderful lives, and C, in which the *x*-people are as well off as in A, but which also contains *y*-people who have awful lives. Since, on Holtug's and on Arrhenius and Rabinowicz' view, A is not better for anyone (it is not better for the *x*-people because they are equally well off in C, and it is not better for the *y*-people, since they do not exist in A), it follows from the PAP that A is not better than C.

By contrast, on our proposal for ranking existence vis-à-vis never existing, these two difficulties for the PAP do not arise: A can be said to be better than B because it is better for the *x*-people that they exist than that they never exist, and A can be said to be better than C, because it is better for the *y*-people that they are not brought into existence to lead awful lives. Nonetheless, in combination with our view, the PAP might still be regarded as too strong, because it rules out several seemingly reasonable views in population ethics, including Broome's theory supplemented with a notion of a personal-value indifference level that is below the population-value indifference level. To illustrate: suppose that A consists entirely of well off *x*-people, while D consists of these *x*-people at the same level of well-being plus *y*-people at a relatively low level of well-being which is barely above the personal-value indifference level. Since, on our view, D is better for the *y*-people and worse for no one, the PAP holds that A cannot be better than D. But if, as is typically supposed, the

population-value indifference level is significantly higher than the personal-value indifference level, then A is better than D. Paired with our view, the PAP would therefore rule out such a variant of Broome's view.

In sum, when paired with our view, the PAP is too restrictive, because it rules out reasonable views like a variant of Broome's view. But one would also like to say more than the PAP says.

Our diagnosis of some of the PAP's shortcomings is this: insofar as the motivation for the PAP is that it is well-being that ultimately matters, this motivation is ill-served by a principle that focuses on the identity of the well-being bearers. As a starting point for a different basic principle, we therefore propose the obviously correct claim that a social situation with only one person is better than another social situation with only one person if well-being is greater in the former.

We propose the following way of extending this to situations with a fixed number of people. It is enough to find a one-to-one mapping from one social situation to the other so that each individual in the former is better off than her counterpart in the latter. This is the Suppes-Sen Principle, which combines the Pareto Principle for a given population size with anonymity.

Suppes-Sen Principle (SSP): A social situation A is better than B if they have the same population size and the level of well-being is better at every anonymized position (rank in the distribution of well-being), from the worst-off to the best-off position.⁷

⁷ This is equivalent to Parfit's (1984, p. 360) Principle Q. While we find this principle attractive, we should note that, as Parfit mentions, it rules out some so-called "claim-based" views of population ethics. On these views,

Can we extend this principle to comparisons of social situations with different numbers of people, without ruling out reasonable views in population ethics? This is indeed possible as follows:

Variable-Population SSP: A social situation A is better than B if there are two situations A' and B' such that:

- a. A' is as good as A and B' is as good as B;
- b. A' and B' have the same population size; and
- c. The level of well-being in A' is higher at every position than in B'.

This principle does not commit us to anything particular about the value of increasing population size. It is, for instance, satisfied by total utilitarianism,⁸ average utilitarianism,⁹

people do not have a claim to be brought into existence, but they do have a claim to have their quality of life improved if they exist. Now consider a choice between state E, in which Charles comes into existence with 60 utils, Dan with 80, and Edward never exists, and state F, in which Charles has 79 utils, Dan never exists, and Edward exists with 60 utils. (Assume that a life with more than 0 utils is better for a person than never existing.) Then no one has a claim against F, whereas Charles has a strong claim against E. On some claim-based views, Charles' claim should override impersonal goodness. Parfit (2010, section 78) argues that such views are therefore problematic; Otsuka (2014) offers a defence of such views. Indeed, there seems to be a tension between the Suppes-Sen principle and claim-based views even in fixed-identity cases (see Voorhoeve [2014] for an example).

⁸ A' (B') can have a single individual whose utility is the total utility in A (B).

⁹ A' (B') can have a single individual whose utility is the average utility in A (B).

and “critical-level,” or (as we would call it) “population-value indifference-level utilitarianism.”¹⁰

One may note that for fixed populations, the PAP is weaker than the Pareto Principle because it never requires a comparison to be made in a certain way (it only precludes certain relations). A similar variant can be formulated for the Variable Population SSP:

Weak Variable Population SSP: Social situation B cannot be better than A if there are two situations A' and B' such that:

- a. A' is as good as A and B' is as good as B;
- b. A' and B' have the same population size; and
- c. The level of well-being in A' is higher at every position than in B'.

Finally, we note that the Suppes-Sen Principle involves interpersonal comparisons. Can one formulate a principle which, like Pareto and the PAP, involves only *intrapersonal* comparisons but which applies to different populations and does not imply the notorious Mere Addition Principle? (This principle holds that raising the well-being of incumbents and adding new people just barely above the personal well-being indifference level always improves the social situation.) This seems impossible. In particular, it is impossible to sensibly compare disjoint populations without making interpersonal comparisons.

¹⁰ A' (B') can have a single individual whose utility above the population-value indifference level is the total utility in A (B) minus the population-value indifference level multiplied by the population size of A (B).

6. Challenges to our view

The conjunction of our view on the value of existence and the Variable-Population SSP does not, as far as we can see, have counterintuitive implications for population ethics. However, the idea that existence at a high level of well-being is better for a person than never existing may be thought to have such implications when paired with other apparently attractive principles. First, if one accepts our view and the Pareto Principle, then the Mere Addition Principle follows: if adding an individual does not have negative effects on others, then it is better to add an individual to the population so long as his level of well-being is in excess of the personal-value indifference level. If one also assumes inequality aversion, then one cannot avoid the infamous Repugnant Conclusion, which holds that a sufficiently populous social situation in which everyone's quality of life is just barely in excess of the personal-value indifference level is better than a social situation with a smaller, uniformly well-off population.¹¹

¹¹ Once adding persons just above the personal-value indifference level is deemed acceptable, then redistributing between a very large number of such additional people and the incumbents can create a very large population with an equalized level of well-being that is arbitrarily close to the personal-value indifference level.

Note that a similar problem arises for Arrhenius and Rabinowicz' view when combined with a version of the Pareto Principle which they propose, viz.:

Subjunctive Weak Pareto (SWP): If state A would be better than state B for everyone who would exist if A were to obtain, and for everyone who would exist if B were to obtain, then A is better than B.

As Arrhenius and Rabinowicz (2013) note, their view and SWP together imply a variant of the Mere Addition Principle, and therefore, assuming inequality aversion, entail the Repugnant Conclusion.

Second, as we discuss elsewhere in detail (Fleurbaey and Voorhoeve 2014), the view that existence with a high level of well-being is better than non-existence may have unpalatable implications when it is paired with a common view about fairness. To illustrate, suppose that Fiona and Georgina both exist. Matters are so fixed that precisely one of them will, due to an untreated illness, have a life at the personal-value indifference level, which we can set at 0 utils. The other will get a treatment which will give her a good life. If Fiona is treated, she will have a lifetime well-being of 70 utils; if Georgina is treated, she will have 69 utils. You must choose the probability p that the treatment will go to Fiona, with $0 \leq p \leq 1$. (Georgina's chance of receiving the treatment is therefore $(1-p)$.) In this case, a plausible view of fairness requires that you give them both equal chances of treatment.¹²

Next suppose that neither Fiona nor Georgina exist and that precisely one of them will come into existence. If Fiona exists, her well-being will be 70 utils; if Georgina exists, her well-being will be 69 utils. You must choose the probability p that Fiona will exist. If existence is better than non-existence and if fairness applies to these benefits as it would to benefits to already existing individuals, then fairness requires you to equalize their chances of existence. However, it is counterintuitive that you would be acting unfairly to Georgina in this case if you minimized her chance of existence and maximized Fiona's chance of existence.¹³

¹² See, for example, Broome (1990).

¹³ Bykvist (2013) expresses similar concerns about the idea of valuing never existing as equivalent from the perspective of the person's interests to living at well-being level of zero. If a never-existing, but at one time possible, person were treated just like an existing person with a well-being of zero, then this would imply that there is a staggering and heretofore unrecognized amount of inequality.

It seems to us that these challenges reveal that the good that one could do for a contingent person by creating her may not always have the same moral force as the good one could do for an already existing person (or a person whose future existence is determined independently of our choices). Naturally, this difference in moral force requires explanation, which we cannot offer here.¹⁴ But if there is indeed such a difference, then one should not simply extend the Pareto Principle to variable-population cases; nor should one apply standard ideas about the fairness of equal chances to a benefit for already existing people to cases in which the benefit is coming into existence with a high level of well-being.

7. A challenge to welfare-based approaches to the value of existence

So far, we have followed the literature in focusing only on the well-being (understood as what is in the person's self-interest or what is of prudential value for him) that is enjoyed in the envisioned life in order to determine the personal or social value of a person's existence. But this is a highly limiting assumption, because it ignores people's views on how values besides well-being determine the value of their existence.

Consider a man whose life consists mostly of struggles and suffering and who has few pleasures and achievements. Suppose a neighbour says to him: "Your well-being is so low that your life is not worth living." This man can reasonably feel insulted by this remark. He may sensibly regard his existence as of great value, even though he would much prefer being spared his trials. He may, for example, have acted well towards others, intelligently pursued noble aims (even though he did not achieve these aims), and responded to

¹⁴ Some have argued that the explanation is that a person cannot be wronged by an action if she would never exist if that action would be performed (Roberts 1998; Vallentyne 2000).

adversity with fortitude. And he may, quite sensibly, believe that this makes his life a valuable one, albeit one with a very low level of well-being. (Or, if he is a religious man, he may believe that his existence is valuable because it is part of God's plan.) Symmetrically, a villain who repents of his wrong-doing at the end of his days may have a dim view of the value of his existence even though standard measures of lifetime well-being would put him well above the personal-value indifference level and even though, as chance would have it, the consequences of his crimes for others' well-being were minor.

What these examples suggest is that well-being is not the only consideration that matters in the evaluation of a person's life relative to his never existing. Whether a life is worth living or not *from a person's reasonable¹⁵ comprehensive moral perspective* is a deeper question than the question whether his well-being exceeds a particular level. Comparing existence to never existing from the viewpoint of the personal value of individual well-being may, at least on some people's reasonable comprehensive views, be just as beside the point as comparing Matisse to Cezanne by the size of their paintings.¹⁶

In sum, there is a mismatch between the evaluation of existence in terms of individuals' comprehensive moral views and the welfare-based evaluation of existence. Individuals can answer the existential question on the basis of different philosophical and religious values, which cannot be exhaustively accounted for by a suitable notion of well-being.

¹⁵ We here employ the term "reasonable" in Rawls's (2005) sense.

¹⁶ A similar problem occurs for the notion of a "life worth continuing" when well-being is defined in such a way that it may run against the individual's own judgment (based on her comprehensive moral view) about whether her life is worth continuing.

These observations might suggest the following argument in favour of an approach that, like Broome's, avoids appealing to the notion of a life worth living.

Social welfare evaluation relies on a population-value indifference level that is defined in terms of well-being only. From the perspective of social welfare evaluation, a judgment that a person's well-being is below the population-value indifference level and therefore lowers social welfare does not imply that his existence is bad for him, or bad in terms of a more comprehensive set of values. It just means that social welfare is improved only when new members are above his level. Being told that one's life is not worth living is very different from being told that it is bad for the social distribution of welfare. It is therefore more legitimate to evaluate the contribution to social welfare of additions to the population in terms of individual welfare only than to judge the value of an individual's existence (versus his never existing) in terms of his welfare only.

However, obviously, this defence of Broome's approach is fragile and observations similar to the preceding ones could be used to raise an objection against his theory. Why make social evaluation so distant from the assessments that people will make on the basis of their reasonable comprehensive moral views? And, why shouldn't one feel insulted when the criterion for social evaluation assesses one's existence negatively even though one deems one's own existence valuable in spite of one's low level of well-being?

In sum, both an exclusively welfare-based assessment of whether an individual's life is worth living and an exclusively welfare-based assessment of whether an individual's existence improves the distribution of well-being are problematic.

One may object to this conclusion in two ways. First, one may reject the idea that individuals' evaluations of their existence are generally very different from well-being evaluations. This is an empirical counter-argument which may be correct, but which does not address the philosophical problem.¹⁷ Moreover, even if only a handful of individuals reject an exclusively well-being based approach to the value of existence, it remains an open question how to take their perspective into consideration.

Another possible response is to broaden the measure of the value of an individual's life that one uses in assessing whether an individual's existence adds personal or social value. The idea would be to move beyond well-being narrowly construed as self-interested or prudential value. On a broadly defined preference-and-value approach, the value of life could be determined by whether people think their life is worth living, taking full account of their reasonable comprehensive moral views. One would then develop a notion of individual and social value that would incorporate the diversity of values of the members of the population.

There are two difficulties with this solution. First, unless we adopt the above-mentioned idea that benefits that come from being brought into existence have a different moral force than benefits to existing people, relying on people's views on the value of existence would suggest that it is good to add people to society when they think their life is worth living in this broader sense (and no incumbent is affected). A version of the Mere Addition Principle will therefore obtain, as will a variant of the Repugnant Conclusion (with a large population

¹⁷ The argument also doesn't address the objection that an individual who evaluates his existence solely in terms of his own well-being and who has a decent level of well-being just below the population-value indifference level may still find his existence judged to be a social bad, even if it is bad for no one.

full of individuals who think their life is barely worth living). This problem would become even more pressing if there were people whose well-being (narrowly construed) is very low but who nevertheless believe strongly in the value of their existence. One may then obtain a variant of the Repugnant Conclusion in which a sufficiently large number of people who all have an arbitrarily low level of well-being (narrowly construed) but who all believe their existence is valuable has to be declared better than a smaller, well-off population.

The second difficulty is that this approach suggests that one should use a lower population-value indifference level of well-being (narrowly construed) for creating lives among populations who strongly believe in the value of their existence than for creating lives among populations who have no such beliefs. It would be strange to evaluate social situations with such a mechanical treatment of diverse philosophical views about the value of existence! In performing social evaluations, could it be right to assign greater value to the creation of individuals in certain sects whose members believe that their own existence fulfils God's plan than to the creation of equally well-off individuals in other segments of society with a less grandiose view on the value of their existence? The practice of leaving it to genitors to decide whether to have children typically leads to larger families wherever the value of existence is deemed greater. But should the social evaluation criterion take these views on board and really adopt different population-value indifference levels of well-being for different sects?

There is a clear dilemma here for liberals. Either one takes account of the value of existence as assessed by the members of the population. Then the social evaluation of variable population choices will, unappealingly, depend on sectarian views. Or only ordinary well-

being considerations are taken on board, as in Broome's approach, and the project is potentially divorced from what really matters to some people.

References

- Arrhenius, G. and W. Rabinowicz 2010, "Better to be than not to be?" in H. Joas (ed.), *The Benefit of Broad Horizons: Intellectual and Institutional Preconditions for a Global Social Science*. Brill, pp. 399-421.
- Arrhenius, G. and W. Rabinowicz 2013, "The value of existence," ms. (extended version of Arrhenius and Rabinowicz, 2010).
- Broome, J. 1990, "Fairness," *Proceedings of the Aristotelian Society, New Series* 91: 87-101.
- Broome, J. 1999, *Ethics out of Economics*, Cambridge: Cambridge University Press.
- Broome, J. 2004, *Weighing Lives*, Oxford: Oxford University Press.
- Bykvist, K. 2007, "The benefits of coming into existence," *Philosophical Studies* 135: 335-362.
- Bykvist, K. 2013, "Being and well-being," paper delivered at workshop Well-Being and Preferences, Collèges d'Etudes Mondiales, Paris.
- Fleurbaey, M. and A. Voorhoeve 2014, "Priority or Equality for Possible People?" in progress.
- Harman, E. 2004, "Can we harm and benefit in creating?" *Philosophical Perspectives* 18: 89-113.
- Holtug, N. 2001, "On the value of coming into existence," *The Journal of Ethics* 5: 361-384.

Otsuka, M. 2014, "How it makes a moral difference that one is worse off than one might have been," ms.

Parfit, D. 1984, *Reasons and Persons*, Oxford: Clarendon Press.

Parfit, D. 2010, *On What Matters, Vol. I*, Oxford University Press.

Rawls, J. 2005, *Political Liberalism, 2nd edition*, New York: Columbia University Press.

Roberts, M. 1998, *Child versus Child-Maker: Future Persons and Present Duties in Ethics and the Law*. Lanham, MD: Rowman & Littlefield.

Roberts, M. 2003, "Can it ever be better never to have existed at all? Person-based consequentialism and a new Repugnant Conclusion," *Journal of Applied Philosophy* 20: 159-185.

Vallentyne, P. 2000. "Review of Child versus Child-maker: Future Persons and Present Duties in Ethics and the Law by Melinda Roberts," *Noûs* 34: 634-647.

Voorhoeve, A. 2014 "How Should We Aggregate Competing Claims?" *Ethics* (forthcoming).