

# How much life does a man need? Why Life extension is immoral

*Gregory Lewis*

## Abstract

In the near future technology may extend a human's lifespan to centuries, millenia, or more. Would this be a good thing? This paper makes a novel case *against* life extension, based not upon concerns of naturalness or possible consequences, but population ethics: 1) Increments of lifespan add more value when packaged into new lives instead of being added to existing lives already worth living; 2) If it can be said that one is worse off when one never exists versus having a life worth living, lifespan should be given with priority to bring people into existence rather than extending lives that are worth living. Given we live in a universe with finite available lifespan, and our lives are worth living, life extension should not be done. Along the way I discuss person affecting views and raise novel problems for certain species of 'soft comparativism', offer extensions to 'fair innings' and 'veil of ignorance' type arguments in different number cases, and discuss how long the morally optimal lifespan should be, coming to the surprising conclusion that this is plausibly *shorter*, not longer, than our natural lifespan.

## Introduction

There have been no reliably recorded instances of a human living more than 150 years. The majority of people, even in the affluent west, live little more than half this period. Despite remarkable increases in life expectancy, maximum lifespan appears unchanged, leading to 'rectangularization' of the survival curve: almost all of us live to old age, and almost all of us die in the interval between 70 and 100 (see Fries 1980).

This may change: perhaps improvements in biotechnology will defeat the chronic diseases that age and kill most of us (de Grey 2009), or we could freeze ourselves by cryonics to be reanimated and fixed later (Merkle 1998), or we could upload our minds into artificial computers (Martin 1971) - or something else entirely. If *any* of these are successful, then humans could live centuries, or millennia, or more.

Would this life extension be a good thing? There are well-worn concerns over whether a procedure like cryonics would allow *us* to survive (instead of someone very similar but non-identical), or whether mind uploading would produce a philosophical zombie that behaves to all appearances as if it were alive, but has no mental states (see the ‘Searlean wager’ by Agar 2011). There are other concerns about the social consequences following the availability of life extending technology, particularly if it could only be acquired by some (Pijnenburg and Leget 2007, but see Davis 2004; see generally Temkin 2008). Perhaps the main concern has been that life extension runs counter to some proper natural function of our humanity or our community, and so will make our lives much poorer (Callahan 1990, Fukuyama 2002, Kass 2004; but see Caplan 2005, Harris 2006, and many others).

I offer a different argument against life extension, independent of the previous concerns raised: instead of starting from practical concerns about feasibility or social fallout, or a particular teleology for our lives or community, I consider life extension from the point of view of population ethics and intergenerational justice. There is only so much life time that can be harvested from the universe, and if we take a greater share by extending our lives, then subsequent generations have less, and in all likelihood some in those generations will never exist at all. The goods that result from life extension are less than the goods that result without life extension, so we should not do it.

The paper proceeds as follows: first I plot out how the argument can be made against radical life extension granting a prioritarian commitment, a no-difference view on different number cases, and the fact that available lifespan is finite. I then try and motivate these assumptions, and spend the most time discussing person-affecting views, in particular comparative harm views like Arrhenius’s *soft comparativism* (Arrhenius 2009), and showing these lead to very counterintuitive results, so we should stick with an impersonalist view instead. Next, I discuss a variant of my argument that asserts that people who do not exist can be said to be *worse off* than in cases where they exist and have lives worth living - although this variant of the argument demands more, it provides elegant extensions to veil-of-ignorance and fair-innings arguments in different number cases. I then consider objections, mainly the worry that these arguments lead to a *reductio* that we should strive for as many lives of as short a length of possible, and from here discuss what makes a ‘morally optimal’ lifespan. I close by showing how the arguments can be modified to accommodate a variety of moral views, and consider the wider ramifications of this sort of argument outside the life extension case.

## The argument

For the first version of my argument, I will take the following suite of ideas (to be defended later):

First, I am going to take consequentialism, in particular welfarism. The only things of moral significance are the welfare that results from our actions, and not (for example) rights, duties, or

natural law.

Second, I take an impersonalist or ‘no-difference’ answer to the non-identity problem, common with many consequentialists: one should simply maximize value, and be indifferent to whatever changes in personal identity of who exists will occur. For example, in the famous case (Parfit 1987, p358) of a 14 year old girl choosing between:

1. Having a child now, who will have a difficult upbringing but a life that is on the whole positive.
2. Waiting until later in life to have a child, who will have a much better start in life, and consequently an even better life on the whole.

The no-difference view would judge that although it is good thing to bring a person into existence who has a life that is ‘net positive’, the girl should have waited to bring into existence the child who would have a much better life, as that would have had even greater total welfare. The fact the child born at 14 and the child born later in life would be *different people* has no bearing on these judgements.

Third, I make a distinction from Parfit’s *Impersonal Total Principle* (Parfit 1987, p387) in that instead of maximizing *aggregate* or total value, we also care about how that value is packaged between persons: one prefers improving the welfare of the worse-off than the better-off - *prioritarianism*. One could formalize the principle like this (compare Adler 2009, see also Broome 2004, 225-226):

$$V = \sum_{i=1}^N G(f_i)$$

Where  $V$  is total value of a world,  $N$  the total number of people who exist in that world (at any point in time)  $f_i$  the welfare of the  $i$ th persons life, and  $G$  is a strictly increasing, strictly concave function.<sup>1</sup> This means a given increment of welfare is worth more when added to a life with less welfare than a life with more.<sup>2</sup>

---

<sup>1</sup> Were  $G$  a linear function with fixed gradient, the equation would fit the view of total aggregate welfarism. If  $G$  were a strictly *convex* function, we would have *anti-prioritarianism* – a given increment of welfare is worth more when added to a life with more welfare than one with less. Note this equation cannot really accommodate *maximin* or *leximin* views on prioritarianism, and I will not be using versions of prioritarianism this strong in making my argument.

<sup>2</sup> Note this is a ‘whole life’ prioritarian view, where who is better or worse off (and therefore to whom welfare is better distributed to, *ceritus paribus*) depend on the welfare *across* their lives rather than their welfare at the instant a choice is being made. If we judge priority in the latter way, life extension (or at least rejuvenation) is recommended, as the elderly are generally worse off than the young due to increased disability (compare Farreley 2010). However, I find the whole life view much more plausible: it seems counter-intuitive (for example) to judge frail 120 year olds a higher priority for medical care than (previously healthy, and expected to be healthy if they survive) 10 year olds when distributing scarce life-saving

This suite of positions mean that we should aim to increase aggregate welfare (with greater weight given to changes in welfare levels among those with the lowest welfare), but have no interest in which particular people there are. Call this the *Impersonal Prioritarian Principle*.<sup>3</sup> Now the argument:

1. There is a finite upper bound of lifespan available in the universe.
2. Lifespan should be distributed according to the *Impersonal Prioritarian Principle*.
3. In cases of finite available lifespan, it is better that a given span of life is distributed in smaller parts to a greater number of people, than in larger parts to a smaller number of people who have already have lives worth living (from 1 and 2).
4. Life extension will distribute lifespan in greater parts to a smaller number of people who have already had lives worth living, instead of in smaller parts to a greater number of people.
5. Life extension is immoral. (from 3 and 4)

I will discuss another 'route' the argument could take later: if we can say that someone is worse off if they do not exist compared to when they have a life worth living, then we can adopt person-affecting restrictions and say that not being brought into existence is worse *for them*, a sort of (*possible*) *person affecting prioritarianism*. This allows attractive extensions of both fair inning and veil of ignorance arguments to different number cases, and also finds life extension to be immoral. However, let's start with a defence of the first argument.

## Some upper limit to life

There will be some hard boundaries as to how many individuals<sup>4</sup> can be alive at once. Optimal agriculture there will not provide infinite food (and expanding across the stars will only provide further finite increments). Even if technologies such as simulated reality and mind uploading can allow us to sustain orders of magnitude more people than possible biologically (see Bostrom 2003 for one estimate), our best physics suggests there will still be physical constraints. We cannot exploit our universe for infinite energy and computational resources.

---

treatment.

<sup>3</sup> A different approach to prioritarianism which retains a similar concern for how value is packaged would be a pluralist consequentialism concerned with both total value and the equality of its distribution. A similar argument against life extension can be made from this view: life span should be distributed more equally, all else equal, and having fewer people with longer lives is less equal than more people with shorter lives. I prefer the prioritarian view for 'levelling down' concerns (Parfit 1997, but see Temkin 2000, 2009), and also that measuring equality is problematic, especially so in different number cases (see Arrhenius 2009 for a discussion).

<sup>4</sup> Whether valuable individuals include post-humans, artificial intelligences, animals or whatever else has no bearing on the argument.

Thus there is a limit to the ‘carrying capacity’ for individuals, both at the present time and at any future technological level we will achieve. The disjunction of our best physical theories suggest that the life permitting conditions of the universe will cease at some point in the future (through heat death, Big Crunch, proton decay, or whatever else). So there must be an upper limit, although a very large one, as to how much lifespan can possibly be produced, even granting ‘optimal’ behaviour on humanity’s part to maximize lifespan harvested from the universe. In short, life is scarce.

## The value of possible people

Some moral choices, so called ‘different number’ cases, are choices where one can not only affect the *welfare* of a particular person, but determine which people (and how many) exist. Controversy arises how we weigh up welfare between people who will exist regardless of our choice, and those who owe their existence to the choice we will make. A simple concrete example:

*To be or not to be?*

| W1   | W2         |
|------|------------|
| a=10 | a=6<br>b=6 |

We have a choice between two worlds, W1 and W2. *a* exists in either world, whilst *b* only exists in W2. We stipulate they have the levels of welfare given in the table. Which world is better? By the lights of impersonalist theories (such as total act utilitarianism) W2 is better than W1, as it has a higher aggregate welfare. ‘No difference’ views like this are uncaring as to the particular identities of those who benefit and lose out in the moral choices we make, or in which choices these people exist or not.

Many find this deeply counter-intuitive, and rather think the welfare of *a* should ‘count more’ than the welfare of *b* in cases where *a* exists in all worlds but *b* in only some: we should not ‘trade off’ *a*’s welfare to produce a ‘new person’ at a one-to-one ratio. One account that accommodates this is the (*independently existing*) *person-affecting view*: that when comparing worlds, we should weigh the welfare of those who will exist in all of the compared worlds more than those who exist in only some but not others. These views come on a spectrum: at one end the no-difference or impersonalist view described above, on the other a strict person affecting view that considers only the welfare of independently existing people and disregards the welfare of people who exist in certain worlds but not others, and intermediate views between - either holding one *lexically dominates* the other, or a mixed view where both carry some weight.

The biggest class of objections to the prior person affecting view is that it leads to non-transitive ordering of worlds. Consider this example, after Parfit’s *Mere Addition Paradox*:

*Pleasant addition paradox*

| W1   | W2          | W3         |
|------|-------------|------------|
| a=10 | a=11<br>b=1 | a=7<br>b=7 |

On a strict person-affecting view W2 is better than W1, as a's welfare improves. By the same view W3 is better than W2, as although a's welfare drops, the rise in b's welfare makes for a higher total, greater equality, and improves the welfare of the worst off. Yet W3 is worse than W1, as a is worse off.

Another common example (see Arrhenius 2009 p.295-6, compare Broome 2004 p.136) uses a similar trick of having different 'overlaps' of people existing in worlds to get intransitive pairwise comparisons.

*Intransitive cycles*

| W1         | W2         | W3         |
|------------|------------|------------|
| a=2<br>b=1 | b=2<br>c=1 | a=1<br>c=2 |

Again, on person affecting views  $W2 > W1$ ,  $W3 > W2$ , but  $W3 < W1$  (where ' $>$ ' means 'is better than', and ' $<$ ' means 'is worse than'). We can arrive at something even more unpalatable:

*Downward dutch-book*

| W1          | W2         | W3         | W4         |
|-------------|------------|------------|------------|
| a=10<br>b=5 | b=6<br>c=1 | a=1<br>c=2 | a=2<br>b=1 |

Here we progressively drop total (impersonal) utility from W1 to W4, but as the strict person-affecting view reads the pairwise comparisons as  $W1 > W2$ ,  $W2 > W3$ ,  $W3 > W4$ , we spiral around to W4 - a world strictly worse than W1. With judicious use of possible people existing in certain pairs of worlds, we can get arbitrarily close to a world with no value from any other world via a series of worlds each 'better than the last' by the lights of a strict person affecting view.

One may think that only strict person affecting views are susceptible to downward dutch books, and that a mixed view that also weighs impersonal value may escape. Suppose we compare worlds via the following measure:

$$10N + 1P$$

Where N is (impersonal) total welfare, and P person affecting welfare: in effect, we weigh total welfare, but those who independently exist count 1.1 times more than those who do not. The 'trade off' is heavily skewed to impersonal welfare - a world with one unit less of total impersonal welfare than another needs 10 units of welfare to persons existing in both worlds to compensate. Yet so long as there is a trade-off, we have the same problem:

*(Gradually) downward dutch book*

| W1                | W2               | W3               | W4               |
|-------------------|------------------|------------------|------------------|
| a = 100<br>b = 50 | b = 70<br>c = 79 | a = 49<br>c = 99 | a = 99<br>b = 48 |

Again, the pairwise comparison gives  $W1 > W2$ ,  $W2 > W3$ ,  $W3 > W4$ , and we can repeat the pattern to spiral down as close to zero as we please. So long as the 'person affecting term' is not lexically inferior to impersonal welfare, we can build a series of worlds which 'trade off' total value for increases in person affecting value, and end up distributing the chipped-away total to the persons we started with. By contrast, impersonalist views rank worlds on a cardinal scale, are insensitive to which people exist in other available worlds, and thus cannot be dutch-booked this way.<sup>5</sup>

Showing one disadvantage of person-affecting views versus impersonalist views is not enough to settle the matter in favour of impersonalism. Across the scales, impersonalist views usually lead to the repugnant conclusion, and work on impossibility theorems by Arrhenius (2010) demonstrate there is no way for impersonalist views to escape a *Very* repugnant conclusion that does not violate highly plausible intuitions. That said, I take intransitivity to be an extraordinarily heavy cost for an axiological view, and should at least give us reason to take impersonalist views seriously if we were not already inclined to accept them.

Before we move on, it is worth discussing a particular species of person-affecting (or at least person-affecting-like) view proposed by Arrhenius (2009) that does not violate transitivity. Dubbed 'soft comparabilism', this view holds we should care both about total value and *comparative harm*: if a person in a given world could have had a *higher* level of welfare in some other world available to be chosen, than that world suffers some penalty compared to the others on offer.

<sup>5</sup> If a view where person-affecting welfare is lexically *inferior* to impersonal total welfare cannot be dutch booked, but still returns intransitive orderings in the *Intransitive cycle* set of worlds given above. It also gives the same answers as the impersonalist view in all cases where impersonal welfare is not tied.

Here's my suggestion of how to formulate this view. For a set of worlds  $S$  to be chosen between, the value of a given world  $W_i$  in  $S$  is given by:

$$W_i = \sum_{j=1}^N [(p_j) + c((p_j) - (p_j)_{\max i \in S})]$$

Translating the algebra, the value of a given world,  $W_i$  depends on two things. First, the sum of welfare for all  $N$  persons ( $p_j$ ) in  $W_i$ ; second, the total *comparative harm* accrued to each person, given by comparing the value of  $p_j$  in  $W_i$  to the maximum value it could take in the set of worlds  $S$  available to be chosen, with a constant of proportionality  $c$  between the two terms. (For ease of discussion, I am going to take  $c = 1$ : for reasons similar to the *(gradually) downward dutch book*, none of the discussion turns on what value  $c$  has.)

For any set of worlds to be chosen between, this formula will yield a numerical value for each world, and so there will not be intransitivity (each of the worlds in *intransitive cycles* has a value of 1 by our formula, which accords to our intuition all are equally good). However, it *does* violate the dictum of the irrelevance of independent alternatives – that the ordering of two choices does not depend on what other choices are on offer. In soft comparabilism, the other worlds available change the comparative harm part of the equation, and thus the value of  $W_i$  depends on which other worlds are in  $S$ . A worked example:

*Dependence on irrelevant alternatives*

| W1    | W2             | W3             |
|-------|----------------|----------------|
| a = 3 | a = 3<br>b = 1 | a = 3<br>b = 5 |

If only  $W_2$  and  $W_1$  were on offer, then  $W_2 > W_1$  for reasons similar to the *Mere Addition Paradox*. However, if the choice now includes  $W_3$ , the worlds are ordered  $W_3 > W_1 > W_2$ , as now  $W_2$  has four units of comparative harm for  $b$ , and so has lower value than  $W_1$ . This seems surprising, but hardly fatal: perhaps it just is the case that our ordering of options should depend on which ‘other’ ones are available. The choices above roughly model the 14-year-old girl case ( $W_1$  = have no child,  $W_2$  = have a child young with a life worth living,  $W_3$  = have a child later, who has a much better life because of it), and to say  $W_2 > W_1$  (having a child with a life worth living is better than not having a child at all), but  $W_3 > W_1 > W_2$  (if one can have a child with a much better start in life, choosing to have a child with a poorer start in life is worse than having no child at all) commands some intuitive support.

There are other benefits and costs to soft comparativism discussed by Arrhenius; overall it does not seem to fare much worse than impersonalist views. It also secures at least a *pro tanto* reason to

favour life extension: one is comparatively harmed in worlds where one *could* have had an extended life, but didn't (see Arrhenius 2008). So the plausibility of soft comparativism hamstrings the argument against life extension.

However, there are examples that reveal soft comparabilism is very implausible in different number cases. Consider this:

*Blissful addition roulette*<sup>6</sup>

| W1     | W2                                 | W3                                 | W4                                 |
|--------|------------------------------------|------------------------------------|------------------------------------|
| a = 10 | a = 11<br>b = 10<br>c = 1<br>d = 1 | a = 11<br>b = 1<br>c = 10<br>d = 1 | a = 11<br>b = 1<br>c = 1<br>d = 10 |

Were only W1 and W2 on offer, then soft comparabilism rates  $W2 > W1$ , which is intuitive. If W1, W2 and W3 on offer (or W1 and any two of W2-4), then soft comparabilism rates these worlds  $W2 = W3 > W1$ , which is again intuitive. However, when W1-4 are on offer, comparabilism rates the worlds as  $W1 > W2 = W3 = W4$ : W1 has 10 units of welfare – 1 unit of comparative harm (a could have been one unit better in W2-4), making for 9 overall; W2 (or W3, or W4) have 23 units of welfare, but 18 units of comparative harm (2 of b-d are 9 units worse off than they could be in other worlds), making for 5 units overall.

This is *very* counterintuitive, firstly that W1 is being selected over worlds that are better for the independently existing person a and more than double total welfare with only lives of positive welfare being added, but also that we have this dramatic change driven solely by introducing another *degenerate* option equivalent to the first two but in which person gains the especially good 10 welfare life.

This pathological behaviour can be generalized. Let us consider a set of worlds I, each having the same set of persons J in each, and there are as many worlds as there are people. We can describe this via a square matrix where the  $p_{j,i}$  entry is the  $j$ th persons welfare in world  $i$ . (Thus each column vector in this matrix is a world, and each row vector is a person's welfare across the set of available worlds)

$$\begin{bmatrix} p_{1,1} & \dots & p_{1,i} \\ \vdots & \ddots & \vdots \\ p_{j,1} & \dots & p_{j,i} \end{bmatrix}$$

Now consider this matrix:

---

<sup>6</sup> This example was inspired by the 'asymmetric creation' example noted by Meacham (2012, 281) in his discussion of his own take on the person affecting view that hinges on comparative harm.

$$\begin{bmatrix} 10 & 1 & \dots & 1 \\ 1 & 10 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 10 \end{bmatrix}$$

This matrix (a generalization of what happens in b-d in W2-4 above) means that as we increase the dimension of this matrix, each world in the set accrues comparative harm *faster* than it accrues total welfare. The value of a world in the set described by this matrix scales with the dimension,  $D$ , in the following way:

$$V = (10 + D) - 9(D - 1)$$

$V$  falls linearly without bound as  $D$  increases. Therefore any population, no matter how awful their total welfare is, can be preferable when on offer alongside a large enough set of worlds described by this sort of matrix, providing the awful population has no persons in common with members with this set.<sup>7</sup>

These seem pretty fatal results to soft comparabilism. Can it be avoided? Changing the balance of how we weigh comparative harms against welfare will not help: so long as there is a trade-off between comparative harm and total welfare, we can contrive examples to get similar results by adjusting the 'gap' between the especially good life and the other lives. Arrhenius notes (2009, fn. 37) there might be other alternatives for the comparator in comparative harm than the best outcome possible. One alternative that springs to mind is to use *average* welfare for  $p_j$  across all available worlds where  $p_j$  exists, and subtract this from the  $p_j$ 's welfare in a particular world to get the comparative harm.<sup>8</sup> This also has some intuitive support – we would take  $p_j$ 's complaint of comparative harm to be much stronger in the inverse case where every world but one had 10 units of welfare, but  $p_j$  was in the single world where it has 1 unit of welfare. Call this alternative *soft (average) comparativism*, in contrast to *soft (maximum) comparativism*.

Using averages means it correctly puts  $W2=W3=W4>W1$  in the *blissful addition roulette* (and extending the case to add further worlds will not result in  $W1$  being ranked above this ensemble). It also means that any set of worlds and persons such that the matrix of the set is:

---

<sup>7</sup> This is a sufficiency condition, not a necessary one: even if there are some people in the nasty world shared with the 'roulette set' of worlds, it may still be that any of the roulette set of worlds is dispreferred (by the lights of soft comparativism) to the nasty world so long as there are not too many shared members.

<sup>8</sup> This also may lead to a comparative benefit being added to a world when  $p_j$  is in a world in which its welfare is *above* the average of worlds it could have been in. For simplicity's sake let us stipulate the comparative harm term can never assume a negative value: this more closely matches our asymmetric intuitions about comparative harm being more salient than comparative benefit – and, once again, none of the subsequent discussion will depend on this assumption.

$$\begin{bmatrix} m & c & \dots & c \\ c & m & \dots & c \\ \vdots & \vdots & \ddots & \vdots \\ c & c & c & m \end{bmatrix}$$

Where  $m > c > 0$  will never result in each world in this set having negative value (and the value of each world in the set described by this sort of matrix is increasing as the dimension of the matrix increases). However, using averages leads to another paradox. Consider how the value of W1 in the matrix above scales with its dimension,  $D$ . The value of W1 using our ‘average comparative harm’ view is:

$$V = m + c(D - 1) - \left[ \left( \frac{m + c(D - 1)}{D} - c \right) (D - 1) \right]$$

The term before the square brackets is the total welfare: one helping of  $m$  and  $D-1$  helpings of  $c$ . The term in the square brackets gives the comparative harm of comparing each  $c$  to the average welfare, multiplied by each of the  $D-1$  persons below this average. What is interesting is that as  $D$  increases  $V$  tends to  $Dc$  and  $m$  falls out (if we allowed comparative *benefits*, we would get one helping of  $m$ 's comparative benefit, which would also fall out as  $D$  increases). Consider the following two sets of worlds with people described by the matrices  $\mathbf{M}$  and  $\mathbf{N}$ , such that:

$$\mathbf{M} = \begin{bmatrix} 10 & 1 & \dots & 1 \\ 1 & 10 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 10 \end{bmatrix}$$

And:

$$\mathbf{N} = \begin{bmatrix} 100 & 1 & \dots & 1 \\ 1 & 100 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 100 \end{bmatrix}$$

$\mathbf{M}$  and  $\mathbf{N}$  have no shared persons between them.  $\mathbf{M}$  has dimension 100 and  $\mathbf{N}$  dimension 99, the value of any world in the set described by  $\mathbf{M}$  is:

$$100.09 = 10 + 1(99) - \left[ \left( \frac{10 + 1(99)}{100} - 1 \right) (99) \right]$$

Whilst the value of any world in the set described by  $\mathbf{N}$  is:

$$100 = 100 + 1(98) - \left[ \left( \frac{100 + 1(98)}{99} - 1 \right) (98) \right]$$

This means that given a choice including both sets of worlds, this comparativism would rank all worlds described by  $\mathbf{M}$  above all worlds described by  $\mathbf{N}$ . Comparing the first world in each set  $\mathbf{M}$  to  $\mathbf{N}$  gives:

|    |    |
|----|----|
| M1 | N1 |
|----|----|

|   |  |
|---|--|
| $p1 = 10$<br>$p2 = 1$<br>$p3 = 1$<br>...<br>$p99 = 1$<br>$p100 = 1$ | $p1 = 100$<br>$p2 = 1$<br>$p3 = 1$<br>...<br>$p99 = 1$ |
|---|--|

If we remove  $p2$  to  $p98$  (which have the same welfare in both worlds), we end up with:

|                         |            |
|-------------------------|------------|
| M1                      | N1         |
| $p1 = 10$<br>$p100 = 1$ | $p1 = 100$ |

Average soft comparabilism ranks M1 above N1, and so violates one of the intuitions it was meant to capture: that bringing someone into existence is not preferable to a loss of welfare of a prior existing person. (Ranking  $M1 > N1$  is counter-intuitive for other reasons: N1 has 89 more points of welfare than M1, and because of how average soft comparabilism scales, we can construct examples that increase the gap in aggregate welfare to any value we please). This also looks very costly, although not as show-stopping as the difficulties with *soft (maximum) comparativism* described above.

There is a large space of possible functions that would pay attention to the *population* of possible outcomes to  $p_j$  when assessing comparative harm which do not involve the maximum or the average, and it could well be that one of these can stack up well with our intuitions without running into the problems I list above. That said, the possible functions we construct to avoid these difficulties may end up too convoluted to plausibly capture our intuitions about what we meant by comparative harm in the first place.

For present purposes we can say this: we have seen that *maximum soft comparabilism* has extremely problematic consequences. *Average soft comparabilism* has less (but still significantly) problematic results, but even if we are willing to accept these costs, this view no longer provides a secure *pro tanto* reason for life extension as it can end up preferring worlds where one possible person is brought into existence to a world where the prior existing persons welfare is greatly increased, and so analogously there will be cases where it prefers replacement to longevity. Having (hopefully) cleared the obstacles person-affecting views present to my argument, let us continue with presenting the argument itself.

## The prioritarian case for preferring more people to life extension

From the *Impersonal Prioritarian principle* it readily follows that distributing value to new people is more valuable than distributing it to those who already have lives worth living. As Holtug (2007; p 5-6) explains:

[W]e can take the total lifetime welfare that will accrue to an individual if she comes into existence, divide it into welfare units and gradually assign less weight to these units. So the first unit has the highest value, the second a lower such value and so on.

On these assumptions, there is a sense in which prioritarianism assigns priority to possible future people over people who already exist. If we can either bestow, say, ten units of welfare on an individual by bringing her into existence or bestow ten units on an individual who already exists and has a positive welfare level, prioritarianism implies, that everything else being equal, it is better to do the former; that is, bring the first person into existence. This is where the ten units contribute most moral value...

This can be seen graphically:

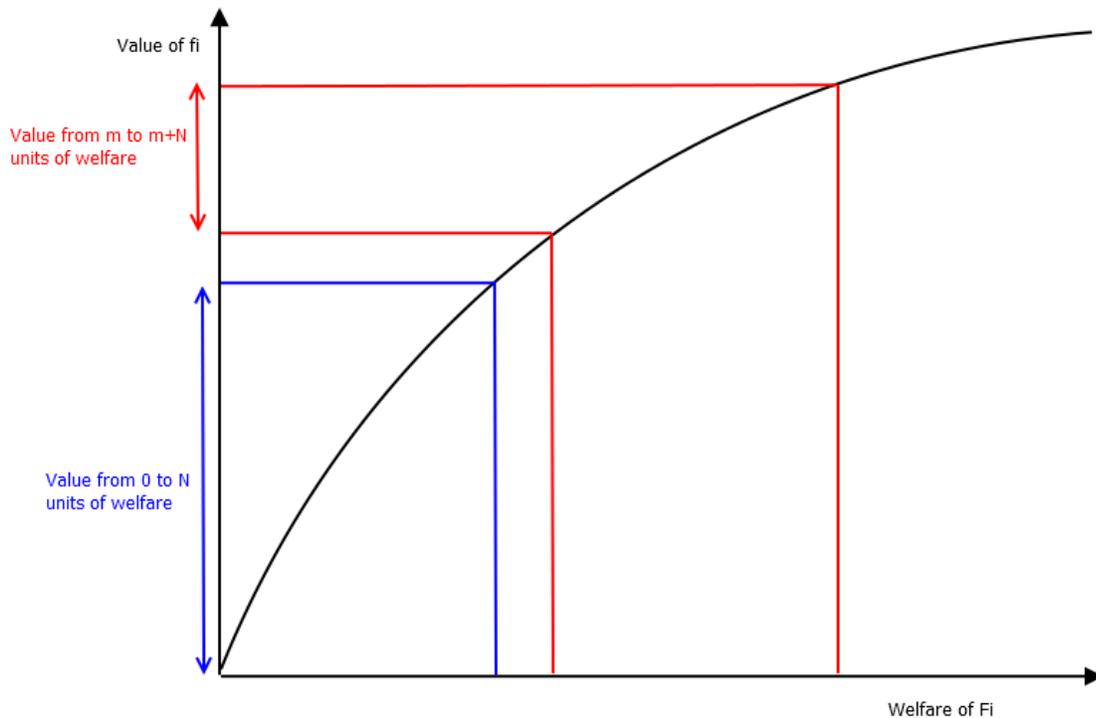


Figure 1: diminishing returns of value with welfare.

Due to the welfare-to-value function's strict concavity, value is most sensitive to changes of welfare at lower levels. So when given the choice of raising a person already above zero by 10, or creating a person with ten units of wellbeing, the latter is preferred by the *Impersonal Prioritarian principle*. Presuming it is better to have a longer life than a shorter one, then it is better to provide (for example) 80 years of life to a new person rather than to someone already 80 years of age.<sup>9</sup>

<sup>9</sup> One may suspect the *Impersonal prioritarian principle* hasn't captured the whole story, as we'd be much

Now the argument falls into place. Assuming our lives are of positive welfare, we should have a lower priority than possible persons for lifespan: by the lights of the *Impersonal prioritarian principle*, adding lifespan to possible people (and thereby making them actual people) will be more valueable than adding lifespan to our own lives. Life extension results in exactly the opposite: we take more lifespan for ourselves, and leave less to be distributed to potential successors. Therefore life extension is immoral.

## Possible persons and the harm of not existing

None of the foregoing required entering into the controversy about whether one is harmed or benefitted by being brought into existence (or *not* brought into existence), or whether people have a zero welfare level in worlds they do not exist (compare Holtug 2007, fn 6). I now enter this controversy to make a parallel argument against life extension: that life extension should not be pursued at the expense of bringing people into existence, because this *harms* them. Although it demands us to say that people are *worse off* when they do not exist compared to when they have a life worth living (and may even recommend the even stronger claim that people who do not exist have a zero level of welfare), there are attractive dividends: if we grant a (*possible*) *person-affecting prioritarianism*, we can extend and generalize both fair innings and survival lottery considerations to different number cases, and it captures the intuitions that certain actions of ours can *wrong* future generations, and that that actions that prevent possible people existing are worse *for them*.<sup>10</sup>

---

more sympathetic to distributing one year of healthy life to a twenty year old rather than bringing into existence an infant who will die at one year of age. Broome (2004, p108) approaches similar concerns ‘from the other side’ when discussing the value of packaging life between people in larger or smaller chunks:

It is surely true that our intuition normally rates continued life better than replacement. For instance, we think it better to save the life of a baby if we can, rather than let her die, even if, were she to die, her parents would replace her with another baby. There may be limits to this intuition. I am not sure we would think it better to prolong a 100-year-old person’s life for another 100 years, rather than have a new person live for 100 years.

I hope this discussion will help explain the varying intuitions we hold in ‘prolongation versus replacement’ cases. Broome also offers (p 256-259) an argument why prolongation is better than replacement, although this concern relies on his use of a positive neutral level, which I argue below (fn 11) is misguided.

<sup>10</sup> Another benefit is this can allow classically ‘impersonalist’ views to be person affecting (see Arrhenius 2009). A person affecting view which encompasses *possible* people and not only prior existing ones does not lead to intransitivity discussed above, as such a view can aggregate the value accrued to each person (irrespective of whether they are prior existing or not), and place them on a cardinal scale.

In many cases, ascribing a 'zero level' of some quality to a person that does not exist is a category error. A person who does not exist does not have zero height, zero temperature, and so on. However, I take welfare to be different from these other properties. The act of bringing a person into existence means they will assume a particular height or temperature (and not necessarily zero). By contrast, it seems plausible to say the mere act of bringing someone into existence will not give them a particular non-zero welfare level. Existence seems a *pre-requisite* for 'getting off the starting blocks' and accruing positive or negative welfare – if this never happens, one remains at the start on zero, in a similar way to, if I never open an account at a particular bank, the value of my holdings at that bank are zero (even though it is not true to say my bank account is empty).<sup>11</sup> This seems to be the view of Holtug (2001; p381) and Roberts (2003), who offer similar accounts to suggest why 'no welfare' should be 'zero welfare'.

The main problems of this view are metaphysical. Most agree that taking a person out of existence cannot harm them, as the only person who could be harmed no longer exists to be harmed. And as Arrhenius and Rabinowicz state (2010; 409) "Wellbeing presupposes being", and so assigning a welfare level, even a neutral one, is illegitimate (see also Bykvist 2006, Broome 1999 p. 168).<sup>12</sup>

I do not share these intuitions – or, at least, I think we should be able to circumnavigate them and still arrive at the result (which I find intuitive) that denying a possible person a positive-welfare life wrongs them. Perhaps we could cash things out in counterfactuals: that if Alice could (*per impossible*) reflect from her state of non-existence on the positive-welfare life she could have had, she would rightly say she was wronged (there would perhaps be some synergy here with 'deprivation' accounts of the badness of death).

A better bet would be a 'guardian angel' account: better or worse options for a person are those that a benevolent impartial observer would choose *for that person's sake* (see Arrhenius and Rabinowicz 2010). As a benevolent impartial observer would want a person to have a positive welfare life instead of non-existence *for that person's sake* non-existence is worse for that person than a positive welfare life (and vice versa).

This comparative claim is all that we need to motivate the claim that people who do not exist are

---

<sup>11</sup> This does *not* mean that a 'blank life' where each moment is neutral is of equal value to non-existence. Broome (2004, p259) argues it is plausible that a life with zero welfare at each moment is in fact a life not worth living, and so it follows that "[T]he neutral level of existence is not zero." I also find this example plausible, but a more fitting account for our intuitions is not an 'offset' given by a positive neutral level (which causes problems, see Arrhenius (2008, p.213-6) for discussion), but including the idea of *pattern goods*: certain distributions of welfare over time seem better than others, and *monotonous* distributions might be worse than 'spiky' distributions, even when total intra-temporal aggregation is equal. This 'monotony concern' (or similar) preserves the intuition a 'blank life' is worse than non-existence, but allows us to normalize our neutral level to zero.

<sup>12</sup> It should be noted that Arrhenius and Rabinowicz (2010) argue that although we cannot say non-existence is worse for a person, we can say being brought into existence is better for a person, and argue against the view expressed in Broome (1999, p.168), on similar grounds to Holtug (2001).

worse off than those who have positive welfare lives (and so, on a *(possible) person affecting prioritarianism*, the non-existent take priority). Even if it is the case that people who do not exist do not *really* have zero welfare, treating them *as if* they did is helpful: doing so does not lead to any absurd conclusions, and it allows the mathematics to be neater – for example, it allows the use of the matrix notation described above.<sup>13</sup>

## Equity, ‘fair innings’, and existence lotteries

Now having shakily supported the idea one is worse off if one does not exist compared to having a positive-welfare life, we can elegantly generalize plausible arguments and heuristics to different number cases.

### Treating the young before the old

When considering life and death, we often think there is some reason to prefer, all things being equal, the younger person to the older person when distributing scarce life-prolonging resources. There are multiple rationales for this (see Shaw 1994), but amongst them is the idea that an older person is ‘better off’ than the younger person due to them having lived more life, and so the younger person should take priority for an increment of additional lifespan (compare Veatch 2003).<sup>14</sup>

Consider this example:

*Modified house fire dilemma:* A hale and hearty 80 year old and a sickly 20 year old are both trapped in a burning building. You know that either if saved each would have ten further years of similar quality. You only have time to save one. Which do you choose?

It remains intuitive to prefer the 20 year old, even if the ‘expected payoff’ in terms of added welfare is equal between the cases.

I think there’s a similar intuition in cases where we must choose in ‘extension or replacement’ cases, where we can add lifespan to an existing person or have a new person with that lifespan instead.<sup>15</sup>

---

<sup>13</sup> Another bonus is that assigning a zero welfare level to possible persons in worlds they do not exist avoids some of the problems I levelled against *soft comparativism* earlier.

<sup>14</sup> Discussion commonly revolves around the idea of a ‘fair innings’: that although all people should have a ‘fair innings’ (commonly suggested to be around ‘three score and ten’), once they have reached this point any more life is a happy bonus, and so there is an onus in resource allocation on trying to ensure all reach a ‘fair innings’ before working to prolong the lives of those who have reached this mark (Harris 1999 p91-4, compare Williams 1997). I prefer a scheme where priority decreases smoothly as age increases, instead of a step function at the age of 70 or so: death is bad at *any* age, but those who die younger suffer a greater harm than those who die older, and thus there is a *pro tanto* reason for treating younger before the older. This is more elegant, avoids distortions, and better captures our intuitions (we’d still prefer the younger patient in the *Modified house fire dilemma* if it was a 20 versus a 40 year old, or a 75 versus a 100 year old).

<sup>15</sup> Although intuitions differ: see Broome (2004, 108) and Arrhenius (2008), who both note an intuition in

If it really is the case that one extends one's life *at the expense* of people who could have been, then extending one's life when one has already had a long and positive welfare life seems *selfish*. Our successors could have had lives similarly as good as ours, but we denied them in favour of having further life for ourselves.

We can make this clearer by generalizing the veil of ignorance heuristic to different number cases. Instead of being ignorant only to what life you will have in this future world, let us take it further and blind you to *the counterfactuals of your own existence*<sup>16</sup>, and so you are unaware about whether particular decisions in different number cases will mean you will or will not ever exist. An analogy:

*The existence lottery*: You are behind a veil of ignorance with a number of others (one for each possible person who could exist in the universe). You are informed of all facts about the universe, and are set the task to agree between yourselves how the people who will exist in this universe shall behave. Once you have done so, there is an existence lottery, and each of you are assigned to one particular possible person. If that person exists, you enjoy that life (whatever it is); if that person does not exist, nor do you.

One of the items on our agenda would surely be how lifespan should be distributed (there would naturally be a drive to maximize available lifespan, but as covered above this will remain some finite amount). There is a cause to extend lifespan and so improve the boon for the 'winners' of the existence lottery, but there is also a cause for replacement, as this means there will be more 'winners' in the first place.

Limiting the case to a more concrete example makes it clear that replacement, not extension, is what we should choose. Suppose life-extending technology becomes available, and promises to increase lifespan to 700, instead of the 70 years normally expected. If this is taken, one person (*p1*) lives for ten times longer, and 'fills their place' for 700 years. If it is refused, this person would die after seventy years, and another (*p2*) would exist to die in seventy years, and so on, resulting in ten lives of 70 years. Imagine yourself and nine others behind a veil of ignorance unaware which of *p1-10* they would be. Would they want *p1* to opt for life extension? Surely not: given some degree of risk aversity one prefers a certainty of a 70 year life over a 1/10 chance of a 700 year life (and a 9/10 chance of never existing). So, by the lights of (*possible*) *person affecting prioritarianism*, life extension is wrong too.

---

favour of longevity. I do not share this intuition, but I hope that the arguments I offer (particularly the one relying on the *impersonal prioritarian principle*) can outweigh it.

<sup>16</sup> It seems to me we this move is demanded by virtually all cases where Rawlsian thought-experiments are deployed, due to the non-identity problem: if we were to change how we were educated, or how healthcare was funded, or sundry other political changes, there would be people whose existence is contingent on what particular option was selected.

# Objections and replies

Although life extension is controversial, the arguments I have made are also controversial. Besides denying my assumptions (which I have tried to defend above, and will return to later), the two main lines of objection I anticipate are the following:

1. Either the *impersonal* or the *possible person affecting* versions of my argument lead to the repugnant conclusion that we should distribute lifetime in miniscule portions to as many people as possible (*Logan's sprint*).
2. Practical concerns: whether lifespan is really as fungible as I have implied in my examples, whether particularly exceptional individuals could warrant an extended life (even if most do not), and other 'finer print' considerations.

Let us deal with each in turn:

## Logan's sprint

My anti life-extension argument implies the following nasty reductio:

*Logan's sprint*: A distribution of life time such that two have 40 years is better than one having 80 years by the lights of prioritarianism. Better still that four have 20 years, or eight 10, or twenty billion microseconds. So we should not stop at life extension, but pursue radical life *truncation*: engineer ways of having as many people as possible live as short a time as possible and replace them with new people as quickly as we can.

This can be fleshed out to interact with either version of the argument. Following the *impersonal prioritarian principle*, the strict concavity of welfare-to-value function means that it is always better to take some fraction of welfare from one person and add it to another because the new person 'starts from zero', and so the gradient of the function is higher than at any positive value, and thus the value maximizing distribution of welfare is in infinitesimal proportions to infinitely many people. On the *(possible) person affecting prioritarianism*, if we are risk averse we should prefer the certainty of 70 years over a 1/10 chance of 700, but why not also a certainty of 35 years over a 1/2 chance of 70, and so on. The problems are similar – with a virtually limitless pool of potential people, we seem to be obliged to spread lifespan as thinly as possible, which seems absurd.

We can escape *Logan's Sprint*. For although our principles do demand we spread welfare thinly all things being equal, all things are *not* equal with respect to lifespan. We should think our early development, both *in utero* and childhood to be at least partly *investment* in the future goods of adulthood (compare to discussion in Dworkin 1993 about why it is more significant if an older fetus is lost because more is invested without return), and the goods attainable in adult life (e.g. romantic

relationships, having children, appreciation of art) have far greater value than those attainable as a child. So much so that, despite the welfare to value function being *concave*, the function of total lifetime welfare by lifespan is (at least in the early stages) even more *convex*, and so vast numbers of very short lives are not the best way of distributing lifespan. Our *possible person affecting* view with its Rawlsian gloss agrees: between a certainty of the first millisecond of life versus a one in twenty billion chance of an eighty year life (and nothing otherwise), we would choose to gamble, and behind our veil of ignorance the untold billions who would never exist would not think the lucky one who has a full life wrongs them by not distributing it in miniscule increments.

## How much life does a man need?

This reply invites further concerns. In the same way that goods in adulthood might have far more value than those attainable as a child, perhaps the goods in greatly extended lifespan likewise are far more valuable than those of normal lifespan and so we should prefer life extension after all. Moreover, we could consider less extreme trades: 70 years versus 50, or 150? Even if we find the appropriate choices behind the veil of ignorance of an existence lottery intuitively acceptable, it would be suspiciously convenient that our evolutionary history would make our natural lifespan the morally optimal one: a more likely culprit is status quo bias.<sup>17</sup>

We need to do more work on just how welfare and value changes during the course of our natural life (and beyond). Before that, some ground clearing:

First, many of our intuitions about how welfare and value change with age are predicated on particular biological facts - facts that may not obtain in a world where life extension is feasible. Most importantly, there may well be *rejuvenation* as well as extension, and we should not assume the trends of increasing disability with age will hold. If transhumanism is right, the sort of world that has the technology to extend life may have other features alien to our present human experience, such as radical intelligence enhancement. I cannot survey all possible technological hypotheticals, so for now let us discuss an *eternal youth* sort of life extension where only rejuvenation and life-extension are available.

Second, we are interested in welfare *for that person*, and disregarding (at least for now) other externalities. So for example the World Bank's own function of how we should weigh value by age (World Bank 1993) helps us little, due to its reliance on economic productivity: although there seems little reason (with rejuvenation), why productivity would have to tail off with age, in any event welfare need not rise and fall in step with productivity, and there may be a 'U'- or at least fairly level - relationship between age and happiness (*figure 3*, see e.g. Stone et al. 2010, Blanchflower and Oswald 2008, but see Frijters and Beaton 2012).

Third, we should remind ourselves we are not looking at welfare at a given time, but *whole life*

---

<sup>17</sup> A parallel objection applies to those opposed to life extension on 'unnaturalness' grounds. Barring some morally significant natural teleology, we have no reason to presume our natural state of humanity, community etc. is probably superior to whatever we end up with post life-extension.

welfare (which I will assume to be the aggregation of welfare at all moments at a life). It seems plausible that the function of *whole life* welfare by lifespan will be strictly increasing providing life at any age is better than not living, but we are interested in the behaviour of this increase: is it concave, implying there will be a point we maximize welfare by 'starting again'? Convex, implying the reverse? Some other behaviour?

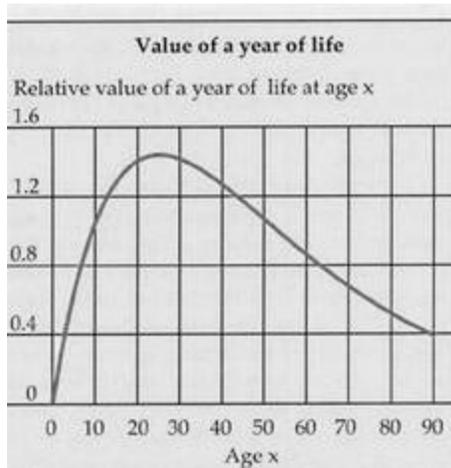


Figure 2 (from World bank 1993, Box 1.3)

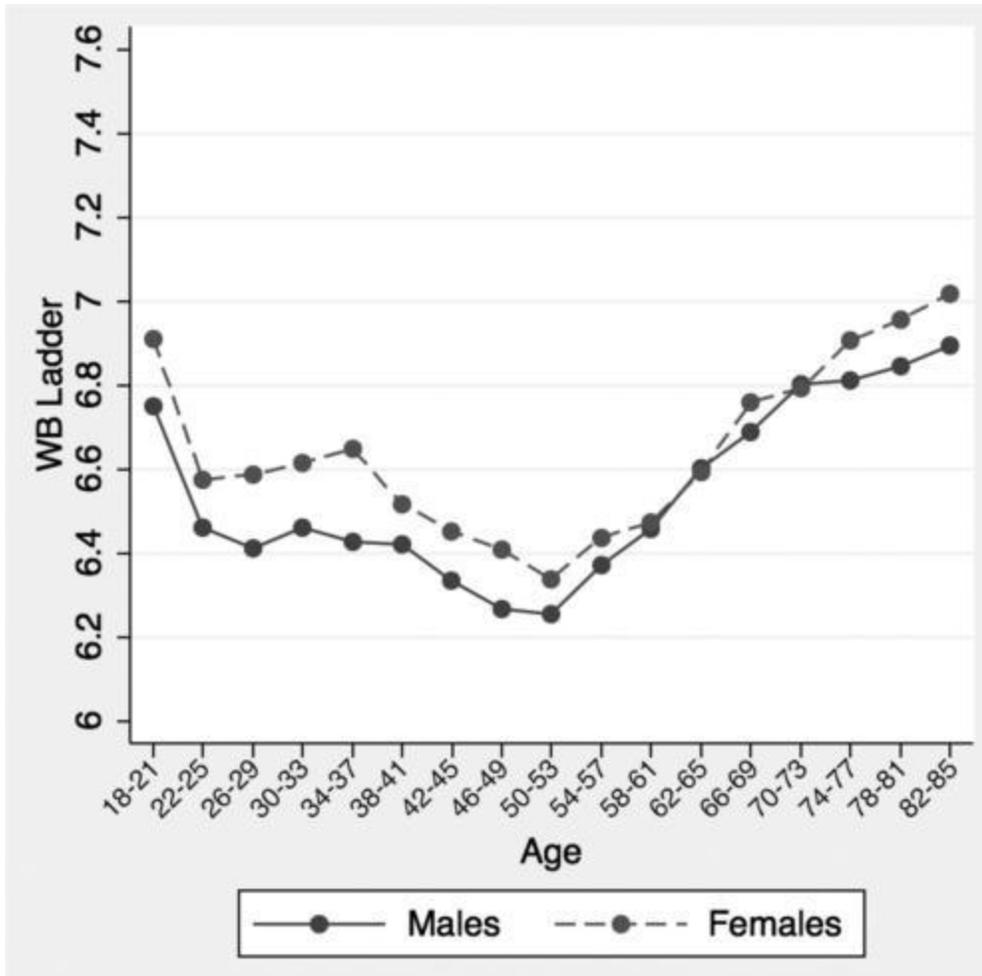


Figure 3 – Wellbeing by age (from Stone et al 2010, fig. 4)

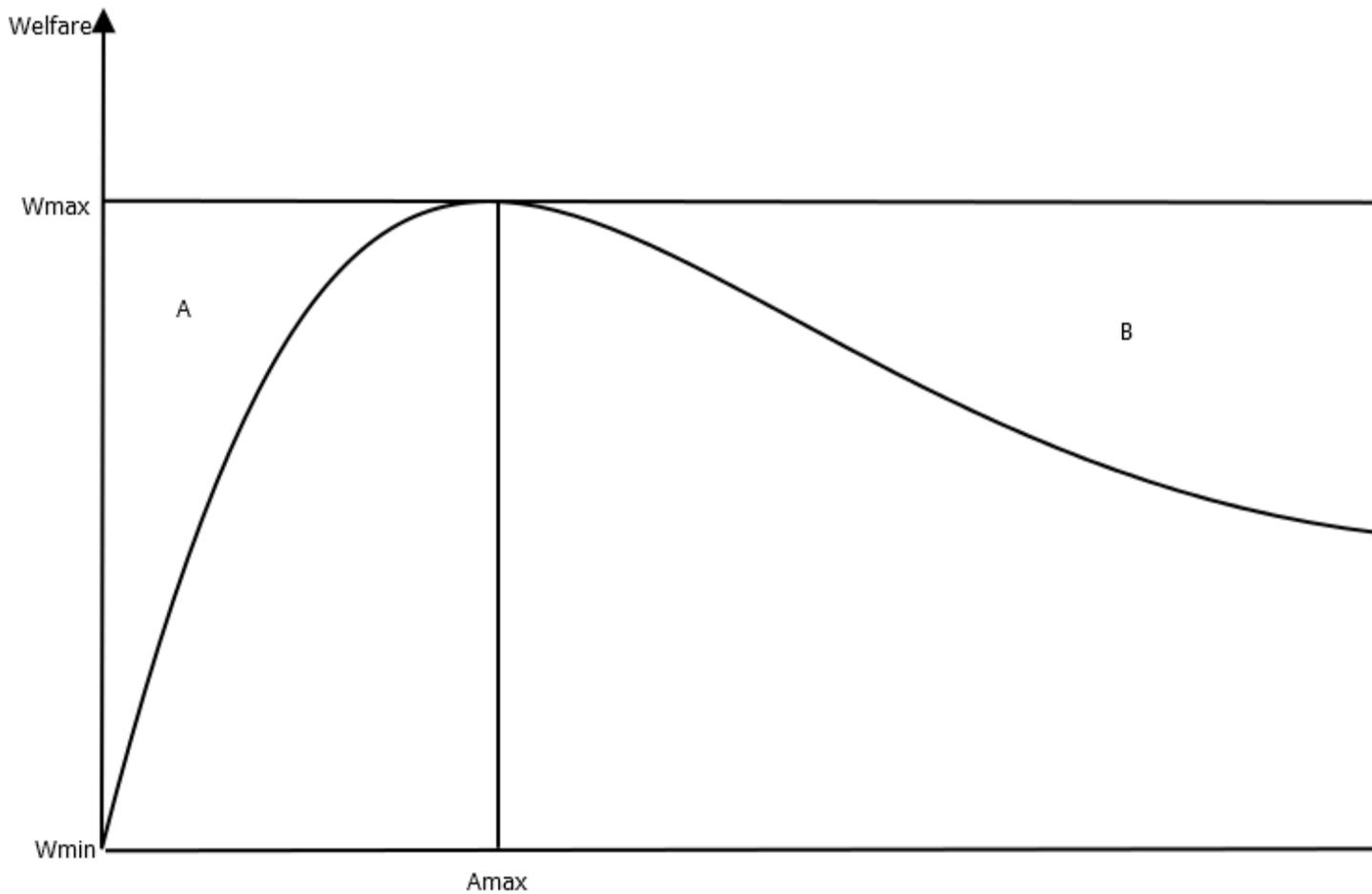
I aver that, broadly speaking, lifespan has diminishing returns: the fourth decade adds less to a life than the third, (or the fourth millennia less than the third) all else equal. One reason for thinking this is that a person will prioritize what is most valuable to them in life first, and so as time goes on they spend it doing less valuable things. Perhaps at the limit of an *immortal* life we would find things tedious and get bored (Williams 1973), although I for one am pretty sure I could keep myself interested for at least a couple of millennia.

Another reason for diminishing returns is that, from the perspective of *whole lifetime* value, we tend to be risk averse and time discount: given the certainty of another ten years versus a half chance at another twenty, we usually prefer the former. There might be irrational reasons driving these preferences (time discounting, scope insensitivity, excessive risk adversity, various other confounds)<sup>18</sup>, but the underlying intuition seems robust even when accepting these concerns.

The other major effect on the welfare/lifespan curve is our mental development. We rapidly

<sup>18</sup> Bennett Foddy made this point to me.

appreciate in welfare over early development as we become able to appreciate goods like family, friends, lovers, art, and things like that. At a sketch, the average welfare function over time would look something like this:



Welfare by age rapidly increases, reaches a maximum, and then steadily declines, plausibly to some limit above zero. If we assume whole lifetime value is the time-integral of this function, then it is clear by inspection that whole-welfare by lifespan is *convex* up to  $A_{max}$  and *concave* thereafter. If this is about right, then so long as area B is greater than area A, or:

$$\int_0^{A_{max}} (W_{max} - f(t)) dt > \int_{A_{max}}^{\infty} (W_{max} - f(t)) dt$$

Aggregate welfare is not maximised by allowing a life to continue *ad infinitum*: one eventually loses more potential welfare by letting a life continue indefinitely as opposed to replacement, 'resetting' the diminishing the returns for the investment cost during early development. So even consequentialist views solely concerned with aggregate welfare (and not how it is packaged) should not get behind unlimited longevity.

To get to the question of what sort of lifespan really is morally optimal, we would need a much closer analysis: the exact shape of the curve between age and welfare, and the balance between

investment costs and diminishing returns; the degree of prioritarianism we should hold, and how concave should the curve between total welfare and value be; all the externalities we have ignored about different dependency ratios, productivity, the costs of the life-extending method, wider societal fall out, and many others. However, we have principled reasons to avoid *Logan's Sprint*, and have reason to think the optimal lifespan is not extremely short, nor extremely long.

Although one should distrust Rawlsian thought experiments in this circumstance due to all the cognitive biases likely to be getting in the way, they are perhaps the best guide we have. On the various existence lotteries I contemplate, it seems the optimal lifespan is *shorter*, not *longer*, than the 'three score and ten' or so most of us can expect: although I might be willing to gamble on a 1/3rd chance of a sixty year life versus a certainty of twenty years, I find myself less willing to trade a half chance at eighty years versus a certainty of forty years. Most of all, I find myself willing to trade a three-quarters chance at eighty years for a certainty of getting sixty; and an average lifespan of 50-60 years with a third more people getting to enjoy such a life looks a better deal than our present circumstances.

## **(Im)practicalities**

Last (and least) are some practical considerations that weigh against the argument made: It has been naively assumed we could instantly replace people who died with their successors, and that prolonging our lifespan inevitably takes up life time that could have been used for another life. Both are implausible: births are not synchronized to deaths, and it is plausible there would be cases where one would deprive no one else of life time they extended their life by 100 years. However, it is pretty clear that these effects are unavoidable when the numbers of people (or the length of life extension) grows: if life extension was widely practiced, then it seems plausible there would be fewer future persons, and so the fact there may be no discernible effect in the limit case of a single person matters little (compare Glover 1975, Kagan 2011).

It may be the case that although in general we should not have extended lives, perhaps certain individuals should. Perhaps they so greatly improve aggregate welfare with their extended lives it is worth the anti-prioritarian costs: perhaps they are brilliant innovators or explorers or leaders who will 'unlock' vast amounts of life time which would otherwise be lost, or perhaps they are artists or thinkers who vastly improve the quality of that lifetime. This is plausible, but such exceptions have little impact on the general principle.

Finally, it may be the case that life extension, even if a bad thing, may be connected with other good things (such as treatments for the diseases of age) such that the 'package deal' is on balance good (see Gems 2011). If so, then so be it.<sup>19</sup>

---

<sup>19</sup> One 'third way' open to thoroughgoing consequentialists is (*Logan's Run* style) mandated suicide/killing at a certain age, thereby getting around any unwanted life-extending side effects.

# Conclusions, Consequences, and a Normative ‘sensitivity analysis’

For foregoing has hopefully presented a prioritarian case against life extension, and defended it on its own terms against objections. To finish, let us see what other consequences the argument (or the assumptions behind it) suggests for other moral questions, and whether the case made against life extension has any pull for those who reject the assumptions I make.

## Over-living and under-population

Many find the idea of life extension disquieting, and the sort of thing they would not want, and so the main conclusion of my argument (that life extension should generally not be done) is not counter-intuitive. Our reasoning so far gives us a very strong concern for the welfare of future generations, and this is also broadly matches up to our intuitions: most believe we should be doing more about (for example) climate change than we currently do.

What people *will* find counter-intuitive is the idea that our notion of a ‘fair innings’ is actually unfairly long, and that ideally our lives should be ending around the close of what we call ‘middle age’. This has health policy implications: we should prioritize the young over the old in healthcare allocation even more strongly than a ‘fair innings’ argument would suggest. We should invert our current beliefs about the purpose of medical care: save when the patient is young, we should be aiming for *palliation* instead of *longevity*, and indeed we should consider longevity an unwanted side-effect of medicine, now directed primarily to improving quality of life.

If we are already living too long, changing our patterns of medical care might not be enough: the most generous estimates give the life expectancy change due to medical care as on around five years on average (Bunker 2001). So not only should we redirect our medical care away from longevity, there are cause to pursue policies that *shorten* lives: perhaps public initiatives that use take some of the health budget to promote life shortening bacchanalia (smoking, alcohol, fatty foods) would be most acceptable.

Second, the common belief that the world is already overpopulated may well be mistaken: prioritarianism will likely oblige us to have more people whose lives are worth living over improving the lives of those who already exist. This also suggests a *super-repugnant* conclusion: that for any population with very high average welfare, there is a much larger population with *lower* total welfare that should be preferred by the lights of prioritarianism (see Holtug 2007 p. 6-8). The ‘optimal population’, much like the ‘optimal lifespan’, will turn on a number of empirical considerations about how far above the neutral level we are, the trade off between adding people and adding welfare to people, how prioritarian should we be, etc. But again we can intuit that our

current western lifestyles are probably going 'too far'. Dividing the UK's (purchasing power parity adjusted) GDP in two and having twice the population having a standard of living around the level of Argentinians, Puerto Ricans or Latvians seems a good deal to me. So we have *pro tanto* reasons for pro-natalism in the affluent west, although other concerns (global poverty, existential risk) will likely dominate it.

## Normative sensitivity

The argument I have used touches on a lot of controversial topics: what normative theory we should use (and what species of consequentialism is best), the value we should assign to possible persons and empirical points about how welfare changes with age. Although I cannot go through every plausible combination of views on these recondite issues, I hope some brief remarks will show this argument (or an argument like it) can cater to those with differing assumptions, and the anti-extension camp can be a fairly broad church.

*What if we are not consequentialists?* There may be other concerns around life and death which are not about the consequences. Perhaps death is a fundamental human rights violation that we should endeavour to prevent at almost any cost, or perhaps there is a duty on the old not to burden the young, or perhaps there is some natural order to the world that mandates we do not extend our lives. Non-consequentialist concerns can go either way, but presuming one's normative view pays *some* heed to consequences it would need very strong concerns in favour of life extension to outweigh the consequentialist cost.

*Is welfare what is valuable?* Even if we accept consequentialism, the consequences we should be interested might not be welfare, however broadly construed: perhaps what is valuable is the satisfaction of preferences, or other items on an objective list. It seems likely that we can cash out the *Impersonal prioritarian principle* or the *possible person affecting prioritarianism* in terms of preferences or other plausible candidates on our objective list (indeed, in the latter case a preference account might be more plausible). Providing one accrues more of whatever is valuable as one ages, the general thrust of this paper seems to hold.

*Should we be prioritarian?* Much of the paper relies on prioritarianism to provide the 'packaging' concern. There are those inclined to deny we should have any concerns about how value is packaged, and just aim to maximize it. Even these people should not be too enthusiastic about life extension given that aggregate value may not be maximized by extending life indefinitely if lifespan does have diminishing marginal returns. There are other accounts which would have a similar packaging concern: egalitarianism, putting justice on our objective list, or maximin/leximin. The argument can be recast using any of these, although these other accounts run into troubles along the way (maximin or leximin cannot escape *Logan's sprint* like I do, equality in different number cases is perplexing, etc.)

*Should we value possible people?* If we concern ourselves only with prior existing people, then life extension seems a very good thing: it enables us to improve the lot of these people dramatically.

However, the arguments given above suggest such *person affecting views* have heavy costs to bear regarding intransitivity, and accounts that avoid this (such as *soft comparativism*) have their own problems.

*The harm of never existing.* If one is not worse off not existing than having a positive-welfare life, than the (*possible*) *person affecting prioritarianism* variant of my argument cannot fly. Yet even if my reasons for accepting this are not enough, the *impersonal prioritarian principle* does not require this claim, and so that argument is still aloft.

*Are there really diminishing returns with increased lifespan?* A lot of work in avoiding *Logan's Sprint* and coming to the conclusion that the optimal lifespan is less than we think is done by the case that there are diminishing returns. If they are flat or *increasing*, then the optimal lifespan is much longer than this (and possibly indefinite). So if one does not find the *survival lottery* cases persuasive, then this is a show-stopper for my argument.<sup>20</sup>

## Conclusion

We have seen that the case against life extension is not restricted to empirical concerns about implementation and abuses, nor to particular teleologies about our humanity, but to thoroughgoing consequentialism. Although it demands controversial answers to a number of normative and empirical questions, these are plausible answers, and the overall argument can be modified to accommodate those with different positions on many of the controversies: a broad swathe of consequentialist views suggest life extension is immoral, and so we should be much less enthusiastic about it than we were before.<sup>21</sup>

## References

- Adler, M.D., 2009. Future Generations: A Prioritarian View. SSRN eLibrary.
- Agar, N., 2011. Ray Kurzweil and Uploading: Just Say No! *Journal of Evolution and Technology* 22, 23–36.
- Arrhenius, G., 2003. The person-affecting restriction, comparativism, and the moral status of potential people. *Ethical Perspect* 10, 185–195.
- Arrhenius, G., 2008. Life Extension Versus Replacement. *Journal of Applied Philosophy* 25,

---

<sup>20</sup> A related view would be a denial of the continuity of personal identity, which makes the sense of saying value accrues to people and has diminishing returns tenuous. I find it plausible that we could preserve a sense of diminishing returns even if 'people' are just a succession of experience moments, although such an account will demand more empirically. If that fails, then my life extension argument will have little pull on these people - although, that being said, the motivation *for* life extension will have little pull on these people too: without continuity of personal identity, why care about ensuring a certain procession of experience moments continues, versus starting a different one?

<sup>21</sup> Thanks to Julian Savulescu, Bennett Foddy, Eitan Fisher, Michael Peyton-Jones, Emma Casey, Michelle Hutchinson, Toby Ord, Will MacAskill and the Cambridge 'Doctors of Philosophy' group for discussion and criticism of the paper and ideas within it.

211–227.

- Arrhenius, G., 2009a. Can the Person Affecting Restriction Solve the Problems in Population Ethics?, in: Roberts, M.A., Wasserman, D.T., Weisstub, D.N. (Eds.), *Harming Future Persons, International Library of Ethics, Law, and the New Medicine*. Springer Netherlands, pp. 289–314.
- Arrhenius, G., 2009b. Egalitarianism and Population Change, in: Gosseries, A., Meyer, L.H. (Eds.), *Intergenerational Justice*. Oxford University Press, Oxford, pp. 324–348.
- Arrhenius, G., Rabinowicz, W., 2010. Better to Be Than Not to Be?, in: *The Benefit of Broad Horizons: Intellectual and Institutional Preconditions for a Global Social Science: Festschrift for Bjorn Wittrock on the Occasion of His 65th Birthday*. Brill.
- Blanchflower, D.G., Oswald, A.J., 2008. Is well-being U-shaped over the life cycle? *Soc Sci Med* 66, 1733–1749.
- Bostrom, N., 2003. Are We Living in a Computer Simulation? *The Philosophical Quarterly* 53, 243–255.
- Broome, J., 2004. *Weighing Lives*. Clarendon Press, Oxford.
- Bunker, J.P., 2001. The role of medical care in contributing to health improvements within societies. *Int. J. Epidemiol.* 30, 1260–1263.
- Bykvist, K., 2006. The Benefits of Coming into Existence. *Philosophical Studies* 135, 335–362.
- Callahan, D., 1990. *What kind of life?* Simon and Schuster, New York.
- Caplan, A.L., 2005. Death as an unnatural process. *EMBO reports* 6, S72–S75.
- Davis, J.K., 2004. Collective suttee: is it unjust to develop life extension if it will not be possible to provide it to everyone? *Ann. N. Y. Acad. Sci.* 1019, 535–541.
- de Grey, A.D.N.J., 2009. The singularity and the Methuselarity: similarities and differences. *Stud Health Technol Inform* 149, 195–202.
- Dworkin, R., 1994. *Life's dominion: an argument about abortion, euthanasia, and individual freedom*. Vintage Books.
- Farrelly, C., 2010. Equality and the duty to retard human ageing. *Bioethics* 24, 384–394.
- Fries, J.F., 1980. Aging, natural death, and the compression of morbidity. *N. Engl. J. Med.* 303, 130–135.
- Frijters, P., Beatton, T., 2012. The mystery of the U-shaped relationship between happiness and age. *Journal of Economic Behavior & Organization* 82, 525–542.
- Fukuyama, F., 2002. *Our Posthuman Future*. Farrar, Straus and Giroux, New York.
- Gems, D., 2011. Tragedy and delight: the ethics of decelerated ageing. *Phil. Trans. R. Soc. B* 366, 108–112.
- Glover, J., 1975. It Makes no Difference Whether or Not I Do It. *Proceedings of the Aristotelian Society, Supplementary Volumes* 49, 171–190.
- Harris, J., 1990. *The Value of Life: An Introduction to Medical Ethics*. Routledge.
- Harris, J., 2004. Immortal Ethics. *Annals of the New York Academy of Sciences* 1019, 527–534.
- Holtug, N., 2001. On the Value of Coming Into Existence. *Journal of Ethics* 5, 361–384.

- Holtug, N., 2007. On giving priority to possible future people, in: Rønnow-Rasmussen, T., Petersson, B., Josefsson, J., Egonsson, D. (Eds.), *Homage à Wlodek*. Philosophical Papers Dedicated to Wlodek Rabinowicz. Available at: <http://www.fil.lu.se/hommageawlodek>
- Kagan, S., 2011. Do I Make a Difference? *Philosophy & Public Affairs* 39, 105–141.
- Kass, L.R., 2004. *Life, Liberty and the Defense of Dignity: The Challenge for Bioethics*. Encounter Books.
- Martin, G.M., 1971. Brief proposal on immortality: an interim solution. *Perspect. Biol. Med.* 14, 339.
- Meacham, C.J.G., 2012. Person-Affecting Views and Saturating Counterpart Relations. *Philosophical Studies* 158, 257–287.
- Merkle, R.C., 1992. The technical feasibility of cryonics. *Medical Hypotheses* 39, 6–16.
- Parfit, D., 1986. *Reasons and Persons*. Oxford University Press.
- Parfit, D., 1997. Equality and Priority. *Ratio* 10, 202–221.
- Pijnenburg, M.A.M., Leget, C., 2007. Who wants to live forever? Three arguments against extending the human lifespan. *J Med Ethics* 33, 585–587.
- Rawls, J., 1971. *A Theory of Justice*. Harvard University Press.
- Roberts, M.A., 2003. Can it ever be better never to have existed at all? Person-based consequentialism and a new repugnant conclusion. *J Appl Philos* 20, 159–185.
- Shaw, A.B., 1994. In defence of ageism. *J Med Ethics* 20, 188–191, 194.
- Stone, A.A., Schwartz, J.E., Broderick, J.E., Deaton, A., 2010. A snapshot of the age distribution of psychological well-being in the United States. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9985–9990.
- Temkin, L.S., 2000. Equality, Priority, and the Levelling Down Objection, in: Clayton, M., Williams, A. (Eds.), *The Ideal of Equality*. London/New York, pp. 126–161.
- Temkin, L.S., 2003. EQUALITY, PRIORITY OR WHAT? *Economics and Philosophy* 19, 61–87.
- Temkin, L.S., 2008. Is Living Longer Living Better? *Journal of Applied Philosophy* 25, 193–210.
- Veatch, R.M., 1993. How Age Should Matter: Justice as the Basis for Limiting Care to the Elderly, in: Winslow, G.R., Walters, J.W. (Eds.), *Facing Limits: Ethics and Health Care for the Elderly*. Westview Press, Boulder, Colorado, pp. 211–229.
- Williams, A., 1997. Intergenerational equity: an exploration of the “fair innings” argument. *Health Econ* 6, 117–132.
- World Bank, 1993. *World development report 1993: investing in health*. Oxford University Press, New York.
- n.d. .
-

