

A STRENGTHENED IMPOSSIBILITY THEOREM IN POPULATION AXIOLOGY

A number of impossibility results in population axiology have been discovered by Gustaf Arrhenius and others, which appear to cast doubt on the prospects for finding an acceptable population axiology which avoids the Repugnant Conclusion. In particular, Arrhenius's sixth impossibility result, widely taken to be the strongest and most significant result in this area, says that the following conditions are incompatible (I have given some slightly different names):¹

- (1) Avoidance of Repugnant Addition.
- (2) Avoidance of Very Sadistic Addition
- (3) Non-Extreme Priority
- (4) Inequality Aversion

Repugnant Addition says roughly that there can be cases in which it is better to add an extremely large population of lives barely worth living, rather than a smaller population of excellent lives. Very Sadistic Addition says that for each population of very bad lives, there is a case in which it would be better to add these very bad lives to some population, rather than to add some large number of positive lives. Non-Extreme Priority and Inequality Aversion are very plausible fixed population principles which appear to be secure prior to doing population axiology. If we assume that these fixed population principles are true, one can read the sixth theorem as saying that we must choose between avoidance of Repugnant Addition and avoidance of Very Sadistic Addition: we cannot avoid both.

Impossibility results like this can be read as arguments for the Repugnant Conclusion, or variants thereof. But read in this way, Arrhenius's impossibility theorems 2-6 each suffer from a common problem. Each relies on a

Date: February 15, 2019.

¹The actual conditions used in the theorem are complicated and have been simplified for ease of presentation. The only substantial difference is that Arrhenius uses avoidance of "Very Repugnant Addition" rather than "Repugnant Addition". Since the purpose of my paper is to strengthen the theorem, taken as an argument for the Repugnant Conclusion, this difference does not matter for my purposes. The sixth theorem also uses a condition called "Egalitarian Dominance", which is an unassailable fixed population principle.

premise which either derives its support from the Mere Addition Principle, or is such that we can see how the failure of the Mere Addition Principle, together with the falsity of the Repugnant Conclusion, would cause the corresponding premise to fail. As such, the impossibility theorems extant in the literature do not really give us much more reason to believe the Repugnant Conclusion than we could already have gleaned from the so-called “Mere Addition Paradox”.

To illustrate, consider Very Sadistic Addition. It is indeed intuitive that, if we were to add an extremely bad population, the result would be worse than if we were to add any population in which each individual led a good life. But consider what would have to be true if Mere Addition were to fail robustly enough for the Mere Addition Paradox not to go through. Begin with an “A population” – a population consisting only of excellent lives. Then add an extremely large number of lives barely worth living – call the resulting population A+. By applying highly plausible fixed population principles, that there must be a “Z population” – a population consisting only of lives barely worth living – which contains the same individuals as A+, but is better. Yet if we want to robustly avoid the Repugnant Conclusion, we must say that Z is worse than A. Because Z is also better than A+, it follows by transitivity that A+ will be worse than A. Moreover: it must be the case that, given a suitable choice of A, A+ is a lot worse than A. This must be true because, if we are able to find an A’ population better than every Z population, we can also find an A population which is much better than A’. A must then be much better than every Z population; and consequently it must be much better than A+.

Given all this, by robustly rejecting Mere Addition and denying the Repugnant Conclusion, we must be committed to thinking that the addition of good lives to a population can make that population much worse under suitable circumstances. It is capable of making any A population worse than A’. All we need to do to see a case where Very Sadistic Addition fails is, given any arbitrarily bad population B, to pick a population A which is so good that, even if we add B to A, the result still remains better than A’. It must then follow that some A+ will be worse than A + B; and hence we will have a case of Very Sadistic Addition.

In order to resolve this problem with the standard impossibility results, I present a new, strengthened impossibility theorem which does not contain a premise implicitly requiring Mere Addition. Given a framework in which we can represent wellbeing levels by real numbers, I show that the following conditions are incompatible:

- (1) Avoidance of Repugnant Addition

- (2) Avoidance of the Very Sadistic Conclusion
- (3) Non-Extreme Priority (real number version)

The version of Repugnant Addition here is very slightly different to Arrhenius's, in that the order of some of the quantifiers is slightly changed. But the difference does not appear particularly philosophically significant. The Very Sadistic Conclusion says that, for any arbitrarily bad population B, there is some better population in which all lives are worth living. In contrast to Very Sadistic Addition, which implicitly relies on the Mere Addition Principle, the Very Sadistic Conclusion is clearly and unassailably false. The Non-Extreme Priority condition here says that, for any wellbeing levels $w_1 < w_2$, there exists some number n and some small quantity of wellbeing δ such that it is always better to improve the lives of n individuals in the wellbeing range w_1 to w_2 by δ each, rather than to improve the life of one individual in the range by δ .

The lesson of this impossibility result seems to me to be that we must accept Repugnant Addition: even if Repugnant Addition is not very plausible, it is clearly far more plausible than the Very Sadistic Conclusion. It also seems to me that the truth of Repugnant Addition undermines our intuitions about the falsity of the Repugnant Conclusion.