The Repugnant, the Sadistic, and Two "Despotic" Conclusions in Population Ethics

S. Subramanian

WORKING PAPER 01/2023
January 2023

Working Papers in Public Finance

VICTORIA UNIVERSITY OF WELLINGTON TE HERENGA WAKA

Chair in Public Finance

The Working Papers in Public Finance series is published by the Victoria Business School to disseminate initial research on public finance topics, from economists, accountants, finance, law and tax specialists, to a wider audience. Any opinions and views expressed in these papers are those of the author(s). They should not be attributed to Victoria University of Wellington or the sponsors of the Chair in Public Finance.

Further enquiries to:
The Administrator
Chair in Public Finance
Victoria University of Wellington
PO Box 600
Wellington 6041
New Zealand

Phone: +64-4-463-9656
Email: cpf-info@vuw.ac.nz

Papers in the series can be downloaded from the following website:
https://www.wgtn.ac.nz/sacl/centres-and-chairs/cpf/publications/working-papers

Title: The Repugnant, the Sadistic, and Two 'Despotic' Conclusions in Population Ethics

Author: S. Subramanian[1]

Abstract: In addressing certain questions in population ethics, Derek Parfit's 'Repugnant Conclusion' presents a well-known difficulty for classical utilitarianism. 'Critical-level utilitarian' axiologies have been proposed as a means of avoiding the Repugnant Conclusion. An objection to critical-level utilitarianism (CLU) that has been raised in the literature is the so-called 'Sadistic Conclusion' which it (CLU) may imply. In this paper it is contended that the Sadistic Conclusion may not be as serious a threat to CLU as it appears, and that the very terms in which the problem is posed carry within themselves the means of its resolution through compromise of a certain 'natural' sort. The paper also deals with two other unpleasant conclusions which could be implied by critical-level utilitarianism and critical-level generalised utilitarianism respectively. These are referred to as 'Despotism Conclusions', involving the according of undue power to the best-off and worst-off members of a society in determining the outcome of welfare comparisons across different populations.

Key Words: Repugnant Conclusion; Sadistic Conclusion; 'regret' functions; value-pluralism; Despotism of the Best-Off Conclusion; Despotism of the Worst-Off Conclusion; Principle M; maximin.

JEL Classification: D30, D31, D60, D63, I30, I31, I32, J10, J18, O15, A13.

[1] Independent Scholar; formerly: Madras Institute of Development Studies, Chennai, India. E-mail: ssubramanianecon@gmail.com

THE REPUGNANT, THE SADISTIC, AND TWO 'DESPOTIC' CONCLUSIONS IN POPULATION ETHICS

1. Introduction

Derek Parfit (1984) considers the 'awesome question' of 'how many people should there ever be?' How, in particular, would classical utilitarianism (CU) address this question? Utilitarianism is an ethical theory which judges the aggregate welfare of a society to be given by the sum of individual welfares in the society. For any given *n*-vector of individual utilities $\mathbf{u} = (u_1,...,u_i,...,u_n)$, aggregate welfare under CU is $W_{\text{CU}}(\mathbf{u}) = \sum_{i=1}^{n} u_i$. In line with this, for a society of *m* persons each of whom experiences a high quality of life as reflected in a large utility level of u^0 (so that the distribution of individual utilities is represented by the *m*-vector $\mathbf{u}^0 = (u^0,...,u^0)$), utilitarianism will dictate that the aggregate welfare of the society is $W_{\text{CU}}(\mathbf{u}^0) = mu^0$. Suppose *k* individuals are added to this society and that each person in this enhanced society experiences a utility level u^k so that the distribution of individual utilities is now represented by the (*m*+*k*)-vector $\mathbf{u}^k = (u^k,...,u^k)$, then aggregate welfare under the utilitarian calculus for this society is $W_{\text{CU}}(\mathbf{u}^k) = (m+k)u^k$. Further, the second regime will be judged to be welfare-superior to the first if $W_{\text{CU}}(\mathbf{u}^k) > W_{\text{CU}}(\mathbf{u}^0)$, or equivalently, if $u^k > \left(\frac{m}{m+k}\right)u^0$. Notice now that as *k* becomes larger and larger, the level which u^k will have to attain for $\mathbf{u}^k$ to be judged better than $\mathbf{u}^0$ becomes smaller and smaller so that, as *k* goes to infinity, the desired level of u^k goes to zero. This is the *Repugnant Conclusion*—the conclusion, yielded by classical utilitarianism, that given an initial population each of whose members enjoys a very high quality of life, one should endorse as an improvement the addition of a very large population which results in each of its members experiencing a very poor quality of life. Or, as Parfit (1984; p. 388) puts it:

> For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better though its members have lives that are barely worth living.

A possible way of avoiding the Repugnant Conclusion (hereafter RC) is through the postulation of a modified utilitarian axiology called *critical level utilitarianism* (CLU) and

one called *critical level generalized utilitarianism* (CLGU), the latter of which takes account of not just the sum total of utility but also of how equal or unequal its inter-personal distribution is. The principal proponents of CLU and CLGU have been Charles Blackorby and David Donaldson. For a miniscule sample of work in this area, the reader is referred to Blackorby and Donaldson (1984, 1991); and Blackorby, Bossert and Donaldson (1997, 2003, 2006). According to CLU, the aggregate welfare of a population is given not by the sum of individual utilities, but by first taking the difference between each individual's utility level and a certain specified critical level of utility which is supposed to represent a reasonably acceptable quality of life, and then summing these differences over all the individuals in the society.

Suppose, as before, that $\mathbf{u} = (u_1,...,u_i,...,u_n)$ is an *n*-vector of individual utilities in a population of *n* persons, then aggregate welfare defined on $\mathbf{u}$ under classical utilitarianism will be given by

$$(1)\quad W_{\mathrm{CU}}(\mathbf{u}) = \sum_{i=1}^{n} u_i \,.$$

Aggregate welfare under CLU will be given by

$$(2)\quad W_{\mathrm{CLU}}(\mathbf{u}; z) = \sum_{i=1}^{n} (u_i - z),$$

where $z(>0)$ is the specified critical level of utility. Critical level *generalized* utilitarianism allows also for the ability of an impartial ethical evaluator to incorporate considerations of aversion to inequality in welfare-assessment. Such considerations can be accommodated by postulating a symmetric, increasing and concave transformation $g(\cdot)$ of individual utilities, so that under CLGU aggregate welfare would be given by the welfare function

$$(3)\quad W_{\mathrm{CLGU}}(\mathbf{u}; z) = \sum_{i=1}^{n} [g(u_i) - g(z)]..$$

A specialization of the function *g* which will be employed in this paper is the well-known 'constant elasticity of marginal utility' version employed by Atkinson (1970), which is widely invoked in the literature on the welfare basis of inequality comparisons, and involves the use of a parameter $\lambda \leq 1$ that is supposed to reflect one's degree of aversion to

inequality[2], with such aversion being a declining function of λ. The 'Atkinson-type' welfare function under CLGU is given by

$$(4)\ W^{A}_{\text{CLGU}}(\mathbf{u};z) = \left(\frac{1}{\lambda}\right)\sum_{i=1}^{n}(u_i^{\lambda} - z^{\lambda}), \lambda \leq 1, \lambda \neq 0;$$

$$= \sum_{i=1}^{n} \ln\left(\frac{u_i}{z}\right), \lambda = 0.$$

One can see now that under CLU (and CLGU), any addition to the population will increase aggregate welfare only if the additional life has a level of utility strictly greater than the critical level. In such a case, one can never end up endorsing additional lives lived at levels 'barely worth living': it is not sufficient for the additional life to have a positive level of utility—it must be large enough to exceed the critical level.

A possible objection to critical level utilitarianism is that while it may avoid the RC it may do so, under certain circumstances, only at the cost of having to accept a different unsavoury conclusion—the 'Sadistic Conclusion (SC)', due to Gustaf Arrhenius (2000, 2006). This is the conclusion that sometimes one may have to pronounce a judgement in favour of a population addition of a certain number of lives lived at negative levels of utility over an addition of a larger number of lives lived at positive, if mediocre, levels. The next section presents a version of the problem associated with the SC, and explores the nature of the ethical dilemma it poses and how one may possibly deal with it. The suggestion, elaborated on in what follows, is that the SC v RC dilemma is a choice between two ethically bad outcomes, and that having to accept the one or the other is—as in all choices between two 'bads'—not an endorsement of the ethically bad, only a judgement of which outcome is less bad under the prevailing circumstances. It is further suggested that counter-posing the SC and RC in a manner which upholds the absolute imperative of avoiding both conclusions at one and the same time in a stipulated setting where it is patently impossible to do so, represents a somewhat incoherent system of values; and insofar as the SC is a valid objection to the claimed success of CLU in avoiding the RC, the former (SC) carries within itself a 'solution' to how the 'problem' may be reasonably addressed.

[2] What is involved here is 'constant relative risk aversion'. One can also have a formulation of 'constant absolute risk aversion', where for all *i*, $g(u_i) = 1 - \exp(-\beta u_i), \beta > 0$ (see Blackorby, Bossert and Donaldson 1997).

Arguably more compelling than the SC are a pair of '*Despotic Conclusions*', a subject dealt with in Section 3. The concern here is with the possibility that while CLU and CLGU avoid the RC, they may do so at the cost of attracting two conclusions which uphold, respectively, a sort of 'dictatorship of the best-off' and a 'dictatorship of the worst-off'. The first type of despotism favours the first over the second of two types of population addition: the first addition consists of a single enormously well-off individual while the second consists of a larger number of individuals each of whom has a high level of utility (but lower than that of the single individual in the first addition). The second type of despotism is one in which of two possible additions to the population the first is judged worse—where the first addition consists of several persons all but one of whom share a high quality of life and one person has a lower level, above the critical level, while the second addition consists of a single person with a welfare level just a little higher than that of the worst-off person in the first addition.

This paper is organized as follows. Section 2 deals with the Sadistic Conclusion and Section 3 with the 'Despotism of the Best-Off (DOTBO)' and the 'Despotism of the Worst-Off (DOTWO)' Conclusions. Section 4 concludes (with an Ordinary Conclusion).

2. Assessing the Sadistic Conclusion

2.1 Formulation

Example 1

As in the introductory section, imagine an initial population of some *m* individuals each of whom has a high quality of life represented by the utility level u^0 which is well above the critical level z. The utility distribution in this initial situation is represented by the *m*-vector $\mathbf{u}^0 = (u^0,...,u^0)$. Let u_A and u_B be two positive numbers, with $u_B < z$; and let q and p be two positive integers such that $q > \left(\frac{u_A + z}{z - u_B}\right)p \ \ (> p)$. Now consider two possible additions A and B to the initial population. Addition A consists of p persons each with a *negative* utility level of $-u_A$, and Addition B consists of q persons each with a positive-but-less-than-critical-level utility of u_B. (These additions do not in any way alter the utility distribution $\mathbf{u}^0$ of the initial *m*-person population.) Let the corresponding utility distributions for the two additions be given, respectively, by $\mathbf{u}_A = (-u_A,...,-u_A)$ and $\mathbf{u}_B = (u_B,...,u_B)$. Then—see

Equation (2)—the aggregate welfare levels for the two population additions A and B, under Critical Level Utilitarianism, will be given, respectively, by: $W_{\text{CLU}}(\mathbf{u}_A; z) = -p(u_A + z)$, and $W_{\text{CLU}}(\mathbf{u}_B; z) = q(u_B - z)$; and it is easy to verify that $W_{\text{CLU}}(\mathbf{u}_A; z) - W_{\text{CLU}}(\mathbf{u}_B; z) = q(z - u_B) - p(u_A + z) > 0$ since $q > \left(\frac{u_A + z}{z - u_B}\right)p$ by assumption.

That is, the first population addition A, consisting of persons with negative welfare, must be considered better, under CLU, than the second addition, consisting of people with positive welfare. This is the Sadistic Conclusion.

2.2. Interpretation

How 'sadistic' the Sadistic Conclusion is, is a function of how it is stated. As stated at the end of the preceding sub-section, it does sound pretty sinister—as indeed it does in an early rendering of the Conclusion by Arrhenius (2000; p. 251): "When adding people without affecting the original people's welfare, it can be better to add people with negative welfare levels rather than positive welfare." A later version (Arrhenius 2006; pp. 9-10) is somewhat less disturbingly forceful in its impact: "For any number of lives with any negative welfare (e.g. tormented lives), there are situations in which it would be better to add these lives than some number of lives with positive welfare." This is still a less than complete description of the Conclusion since we are not told what sorts and numbers of lives are being compared with the 'tormented lives' referred to. A more informative and even less distasteful account might be of the following type: "Under certain circumstances, the addition of a given number of lives with negative welfare could be judged better than the addition of a larger number of lives lived at levels of welfare which are positive but mediocre." In particular, a statement such as "it can be better to add people with negative welfare levels rather than positive welfare" can make it seem that one is advancing the superiority of something patently bad ("people with negative welfare levels") over something patently good ("people with positive welfare levels"). 'Lives with negative welfare' are certainly bad, but there is nothing necessarily good, without further detail or qualification, about 'lives with positive welfare.'

To recover the context in which the Sadistic Conclusion is postulated, it is instructive to go back to the Repugnant Conclusion: what is repugnant about the RC is having to endorse what Parfit called "lives barely worth living". It seems reasonable to suggest that the reference to lives that are "barely worth living" must presuppose some notion of a positive critical level below which lives are barely worth living. The Sadistic Conclusion is postulated on the

failure of a critical-level welfare axiology, in certain circumstances, to endorse lives at a positive—*but sub-critical*--level. The failure to endorse such lives, however, can scarcely be a source of unqualified regret for one who finds the Repugnant Conclusion repugnant, that is to say, for one who finds repugnance in the Repugnant Conclusion precisely in having to endorse lives that are held to be "barely worth living"!

Briefly, the Sadistic Conclusion is not a paradoxical result that upholds the betterness of a Bad Thing in comparison with a Good Thing: rather, it is the altogether less disturbing result of holding one Bad Thing to be less bad than another Bad Thing. It is a bit like a person who is against capital punishment being invited to pronounce, in a presumably unavoidable situation, on whether some particular execution should be carried out by hanging or by electrocution. This general notion is persuasively advanced in Stewart Armstrong's assessment (Armstrong, 2014):

> Consider for example whether it is good to create a large permanent underclass of people with much more limited and miserable lives than all others—but whose lives are nevertheless just above some complicated line of "worth living".…[M]any systems of population ethics do feel it's a negative outcome. Then, given this underclass is a bad outcome…then we can find other bad outcomes that are *not quite as bad as this one*. Such as…a single victim, a tiny bit below the line of "worth living". So the sadistic conclusion is…simply saying that (A) creating underclasses with slightly worthwhile lives can sometimes be bad, while (B) creating a victim can sometimes be less bad…But the victim…[is] just an example of a bad outcome better than (A), only linked to (A) through some superficial similarity and rhetoric.

Apart from this, it appears that the Sadistic Conclusion is intended to convey the judgement that the addition of negative lives is *always* worse than the addition of positive lives. It is not clear why this must be the case for all possible cases involving comparisons of concentrated misery for a relatively small population, on the one hand, with diffused mediocrity for a relatively larger population, on the other. Here Armstrong's reference to continuity of welfare at the critical level is also important: if we are speaking of a very small number of 'victims' with arbitrarily small negative utility, and a very large number of the 'underclass' with arbitrarily small positive utility, then it is not necessarily shocking to judge that the addition of victims (in such a case) might be better than the addition of the underclass. More generally, whether adding victims or the underclass is better could depend on the extent of misery experienced by the victims, the extent of mediocrity experienced by the underclass, and the difference in the sizes of the two populations. It seems reasonable to suggest that there are circumstances in which adding victims could be judged to be worse than adding an

underclass, and others in which the matter is the other way around; and it may be more profitable to identify the circumstances in which the one judgement prevails over the other rather than to present the two judgements as being irreconcilably opposed to each other. The suggestion here, as in other cases of conflict between desirable values in an ethical system, is to seek the possibility of a trade-off. An illustrative example is considered in the next sub-section.

2.3 An Analogy

A problem somewhat analogous to the RC-SC conflict is presented by the alleged conflict between the values of 'efficiency' and 'equity' in the context of income distribution comparisons. One (unhelpful) way of describing the problem is to suggest that in judging between distributions X and Y, the defender of efficiency should always choose the distribution with the higher mean income μ, while the defender of equity should always choose the less unequal distribution (as measured, say, by the Gini coefficient of inequality G). The problem for one who subscribes to the virtues of both efficiency and equity can then be portrayed as the irreconcilable conflict arising from distribution X having a higher mean than distribution Y, and distribution Y having a lower Gini than distribution X.

A conflict presented in these terms seems to point to nothing more than an underlying value system that is incoherent. One who subscribes to the values of both efficiency and equity is more reasonably seen as one who attaches some weight to both values—an egalitarian would typically attach a higher weight to equity, and a utilitarian a higher weight to efficiency—such that, in certain circumstances, she may prefer the distribution with the higher mean, and in others, the distribution with the lower Gini. Average income would typically be the guiding criterion in situations where the distributions under review do not differ too much in terms of inequality but one distribution has a much higher mean than the other; and inequality would typically be the guiding criterion in situations where the relevant distributions do not differ too much in terms of size but one distribution has a much lower Gini than the other.

These are precisely the sorts of judgements that are facilitated by a welfare measure such as Sen's (1976) 'index of real national income', given by the quantity $\mu(1-G)$. If μ_X (respectively, μ_Y) is the mean, and G_X (respectively, G_Y) is the Gini coefficient for distribution X (respectively, distribution Y), then one can conceive of situations in which $\mu_X > \mu_Y$ and $G_X > G_Y$. If it turns out that $\mu_X(1-G_X) > \mu_Y(1-G_Y)$, one would choose

distribution X over distribution Y, that is, one would appear to choose the distribution with the higher mean; and if it turns out that $\mu_X(1-G_X) < \mu_Y(1-G_Y)$, one would choose distribution Y over distribution X, that is, one would appear to favour the distribution with lower inequality. Does this mean, in the first instance, that one has chosen efficiency at the expense of equity; or, in the second case, equity at the expense of efficiency? This is an implausible construction of the more likely fact that one has merely sought a sensible compromise resolution of a problem often encountered in moral evaluations, wherein the ethical comparison of alternative states of affairs is guided by two criteria, one of which inverts the ranking by the other. Can a similar 'compromise' be effected between the demands of the Repugnant and the Sadistic Conclusions?

2.4 Application of the Analogy

The 'efficiency-equity conflict' suggests that a value-pluralist does not have to come out systematically in favour of the one or the other values of efficiency and equity, thereby having to consistently betray one value for the other, but rather that he could display his concern for both values by devising a rule which will guide him on which value to afford precedence in which circumstance, on a case-by-case basis. Similarly, and allowing that one finds neither the Sadistic nor the Repugnant Conclusion a particularly savoury conclusion, can one devise a rule which allows one, on a case-by-case basis, to choose between them in situations where there is no option but to so choose? This question is explored in what follows; and the means to the end is the adoption of some very specific and particular illustrative formulations of a general approach to the problem.

It is useful to begin by noting that the objection to CLU, in terms of its securing avoidance of the RC only at the expense of accepting the SC, is actually based on the implicit postulation of *two* critical levels of welfare (Broome, 1992a,b). Section 1 advances a unique critical level z, but for the SC to make sense, it appears that one requires two levels, a lower level $\underline{z}$ and a higher level $\bar{z}$, with some such distinction between the two as the following one: $\underline{z}$ is a (positive) level of utility above which a person is enabled to live a life that is better than "barely worth living", and $\bar{z}$ is a level at and above which a person is enabled to live a life that may be called 'decent'. A potential problem arises when considering a population addition whose members experience a welfare level that is between $\underline{z}$ and $\bar{z}$: to endorse such an addition would be to endorse lives that are less than 'decent'; and to fail to endorse

such an addition would be to fail to endorse lives that are better than "barely worth living". One may thus not comfortably either endorse nor fail to endorse.

Recognizing the dichotomy of the critical welfare levels as the source of the problem should also pave the way for a compromise aimed at eliminating the dichotomy: this is not an '*ad hoc*' response to the problem, but one which is naturally suggested by the very terms in which the problem has been presented. A possible route to reconciliation is suggested, very much in the spirit of a specialised illustration of a general approach, in what follows.

In considering additions to a population, there are potentially two sorts of regret one can entertain: first, the regret (realised, say, as a disutility) that arises from failing to tolerate lives that deserve to be tolerated by virtue of their level being better than "barely worth living"; and second, the regret that arises from tolerating lives that ought not to be tolerated by virtue of their being less than 'decent' lives. If u stands for the utility level of each member of the additional population, then the first regret—call it $R^1(u)$--can be imagined to be a declining function of u over the interval $[0, \bar{z})$, and to vanish for utility levels no lower than $\bar{z}$. For illustrative purposes, the behaviour of $R^1(\cdot)$ could be modelled as a declining, strictly convex function over a certain range of utilities, as in Equation (5), where a is a positive scalar:

$$(5)\quad R^1(u) = \frac{a}{\bar{z}^2}(\bar{z} - u)^2 \quad \forall u \in [0, \bar{z});$$

$$= 0 \quad \forall u \geq \bar{z}.$$

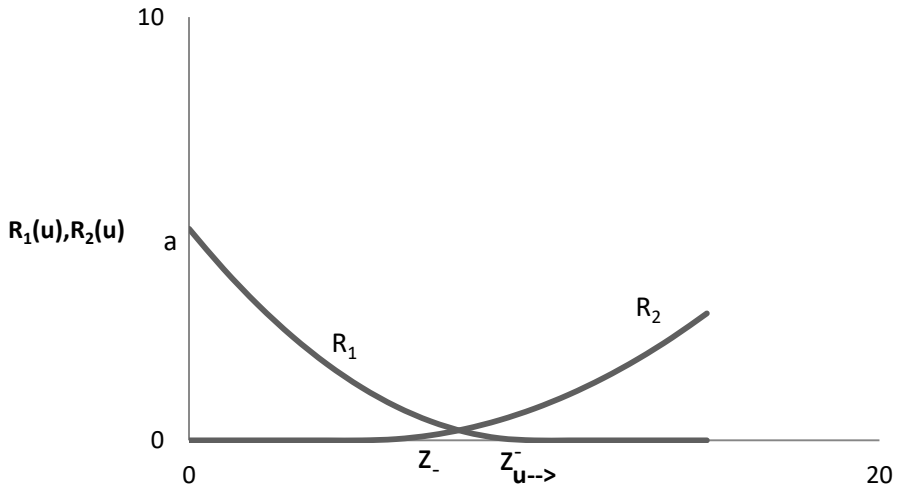
The second type of regret—call it $R^2(u)$--can be imagined to be non-existent for utility levels not exceeding $\underline{z}$, and thereafter to rise with u. Again, for illustrative purposes, the behaviour of $R^2(\cdot)$ could be modelled as an increasing, strictly convex function over a certain range of utilities, as in Equation (6), where b is a positive scalar:

$$(6)\quad R^2(u) = 0 \quad \forall u \in [0, \underline{z}];$$

$$= b(x - \underline{z})^2 \quad \forall u > \underline{z}.$$

The 'Sadistic vs Repugnant Conclusion problem' arises when the two regret curves $R^1(\cdot)$ and $R^2(\cdot)$ intersect in the positive quadrant (see Figure 1).

Figure 1: The 'Regret' Functions $R_1(u)$ and $R_2(u)$



The combined regret from failing to tolerate what ought to be tolerated and tolerating what ought not to be tolerated, is given by[3]

$$(7)\quad R(u)[\equiv R^1(u)+R^2(u)]=\frac{a}{\bar{z}^2}(\bar{z}-u)^2+b(u-\underline{z})^2 \quad \forall u\geq 0.$$

It seems reasonable to suggest that an impartial ethical evaluator would wish to choose a critical level, call it u^*, at which the combined regret from erroneous endorsement and erroneous failure of endorsement of an addition to the population—see Equation (7)—is minimized. The first order condition for a minimum, it can be verified[4], is realized at $u^*=\left(\frac{a+b\underline{z}\bar{z}}{a+b\bar{z}^2}\right)\bar{z}$. If the (absolute) slopes of the two 'regret' curves are equal ($\frac{a}{\bar{z}^2}=b$), then the regret-minimizing critical level will be pitched at the mid-point between the 'barely worth living' and 'decent' critical levels: $u^*=(\underline{z}+\bar{z})/2$; if $\frac{a}{\bar{z}^2}>b$, then u^* will lie closer to $\bar{z}$ than to $\underline{z}$, and the other way around if $\frac{a}{\bar{z}^2}<b$. The slopes of the two 'regret' curves are an expression of the relative persuasiveness of the $\underline{z}$ and $\bar{z}$ critical levels for the evaluator. At a

[3] The assumption, of course, is that the two types of regret are commensurable and lend themselves to operations like addition.
[4] It can be verified that the second-order condition for a minimum is also satisfied.

more general level, suppose $\sigma \in [0,1]$ to be a parameter reflecting the evaluator's relative preference between the $\underline{z}$ and $\bar{z}$ critical levels, so that the larger is σ, the greater is the evaluator's inclination for the smaller of the two critical levels; then the evaluator's 'compromise' choice of critical level can be identified with the weighted sum $\sigma\underline{z} + (1-\sigma)\bar{z}$.

Reverting to Example 1, suppose $\underline{z} < u_B < \bar{z}$. Confronted by Example 1, the ethical evaluator may be imagined to have one of two options. Option 1 would be for him to say: 'since $u_B \geq \underline{z}$ and I would like to avoid the Sadistic Conclusion I should prefer the population addition $\mathbf{u}_B$ to the addition $\mathbf{u}_A$; but then since $u_B < \bar{z}$ and I would also like to avoid the Repugnant Conclusion, I should prefer the addition $\mathbf{u}_A$ to the addition $\mathbf{u}_B$. But I cannot have both $\mathbf{u}_A$ preferred to $\mathbf{u}_B$ and $\mathbf{u}_B$ preferred to $\mathbf{u}_A$, only, it happens that I do have both preferences and I wish to uphold both of them even if I can't, and given that I do and I can't I insist on having it both ways and making myself miserable in the knowledge that I can't…' Or, recognizing that it is a matter of making up her mind, the evaluator could resort to Option 2, which would consist in her saying: 'if $u_B < u^*$, I'll express a preference for $\mathbf{u}_A$ over $\mathbf{u}_B$, that is, I'll act as if I preferred to avoid the Repugnant Conclusion; and if $u_B > u^*$, I'll express a preference for $\mathbf{u}_B$ over $\mathbf{u}_A$, that is, I'll act as if I preferred to avoid the Sadistic Conclusion. Being a value-pluralist (as in the context of the 'efficiency-equity conflict'), I'll seek a compromise whereby I'll be sometimes seen as coming out on the side of the one value and sometimes on the side of the other, depending upon the precise circumstances which describe the case under review.' It does appear that Option 2 is the more sensible one to exercise, one which permits a sort of natural 'dissolution' of the apparent paradox of the Sadistic vs Repugnant Conclusion.

The next section considers what could be a less tractable problem for CLU and CLGU.

3. Two Despotic Conclusions

Critical level utilitarian theories were conceived to address Parfit's Repugnant Conclusion. However, there could be other unpalatable conclusions implied by Classical Utilitarianism which are not satisfactorily disposed of by critical level axiologies. The 'Despotism Conclusions' described in what follows are a case in point. In this section, problems

involving negative and very low positive levels of welfare, such as would not, in the first instance, be entertained as desirable population additions, are not considered: Tortured Souls and Mediocre Lives are kept out of the picture.

Starting with the initial utility distribution $\mathbf{u}^0$ considered in Section 3, imagine two possible additions C and D to this population. Addition D consists of a large population of p persons (say ten billion), each of whom has a reasonably high quality of life $\underline{u}$ which is above the critical level z. Addition C consists of a single person with a massively large level of welfare $\bar{u}$ which exceeds $p(\underline{u}-z)+z$. It is immediately clear that under CLU, aggregate welfare is greater for the population addition C than for the addition D: the single person in C must, by virtue of his monstrously high quality of life, be afforded priority over the ten billion persons in D who each enjoy a comfortable even if not opulent level of welfare. This is the 'Despotism of the Best-Off (DOTBO)' Conclusion:

The DOTBO Conclusion. Of two population additions to an initial population whose welfare levels are unaffected by the additions, the first addition must be preferred to the second if the first consists of a single person with an extremely high quality of life, even if the second consists of several individuals who share a more than decent, but not extravagant, quality of life.

It does appear to be unacceptable that a single extraordinarily lavish life should be permitted to block the alternative of several lives which, while not similarly fabulous, are nevertheless well worth living. Yet, this is an implication of Critical Level Utilitarianism. Perhaps Critical Level *Generalized* Utilitarianism might be expected to perform better?

Before addressing this question directly, it is instructive to consider an aspect of CLGU that is reflected in a conclusion called the Critical Level Repugnant Conclusion (CLRC) (Broome 1992a, b) which Blackorby et al. (1997) hold—rightly, it can be argued—not to display anything particularly repugnant about it. As they put it (Blackorby et al. 1997; p.10):

> Although members of the CLU and CLGU families with positive critical levels avoid the repugnant conclusion, some critics have claimed in response that CLGU suffers from a 'critical-level repugnant conclusion'. A principle implies the critical-level repugnant conclusion if any state in which each member of the population experiences a utility level above the critical level is ranked as worse than a state in which a sufficiently large population has a utility level that is above the critical level but arbitrarily close to it. We do not find this to be ethically unattractive, as long as the critical level is chosen in a reasonable way. As Sen argues, the critical utility

> level…should be high enough so that a 'scenario in which more people enjoy a utility level…[above the critical level]…must be seen as a better outcome' (Sen [1991, p. 19]). This view is consistent with (but not the same as) Griffin's view of the repugnant conclusion. His argument suggests that the critical level be set at that point where people have the 'capacity to appreciate beauty, to form deep loving relationships, to accomplish something with their lives beyond just staying alive' (Griffin [1986, p. 340]). This suggests that the ethical judgements needed to choose a critical level require a fairly complex theory of the good. In addition, a critical level expresses not only a minimal level of well-being necessary to make the creation of people socially desirable, but also the kinds of actions, experiences, and states of mind that we believe to be necessary for a good and valuable life.

There is much that is appealing in the above reservation with the notion that the CLRC is ethically unacceptable. Having said that, the general tendency of the view as it has been articulated in the quoted passage suggests, and convincingly at that, something a good deal stronger than merely the following: that there is nothing wrong for a '*sufficiently large*' population of lives above the critical level to be preferred to a smaller population of lives, even if the lives in the smaller population all experience a much higher level of welfare. In particular, why is it not enough that the population with persons having the lower level of welfare be simply larger (without having to be '*sufficiently*' larger) than the population with persons having the higher welfare level—given that the welfare levels in both populations are above the critical level? Indeed, this is precisely what Sen's sentiment, as quoted by Blackorby et al., seems to suggest: a "scenario in which more people enjoy a utility level…[above the critical level]…must be seen as a better outcome". If we are comparing populations which consist only of individuals with worthwhile lives, that is to say lives which are above the critical level (a critical level that is determined along the lines endorsed by Blackorby et al. in their quoted passage), then surely the appropriate ranking rule to adopt would be to favour that population which has the largest number of persons: this is the rule that gives expression to the sentiment that what matters in the end is to endorse the largest possible additions of lives worth living. A weak formulation of this notion, one which also accommodates the Pareto Principle in fixed population comparisons, is what may be called the Principle of Maximization of Worthwhile Lives. The nomenclature is no doubt unwieldy, so it can be simply rechristened the Maximum Numbers Principle, or just Principle M for short:

Maximum Numbers Principle (Principle M). Consider two possible additions to an initial population whose welfare levels are unaffected by the additions: the first addition consists of q individuals each of whom has a welfare level of u_1 that is above the critical level z, and

the second addition consists of p individuals each of whom has a welfare level of u_2 that is also above the critical level z but below the level u_1. Then, by Principle M, the first addition should be preferred to the second if and only if $q \geq p$, otherwise the second addition should be preferred to the first.

Where does Critical Level Generalized Utilitarianism stand with respect to Principle M? Consider Example 2 below.

Example 2.

There is an initial population of *m* persons whose utility distribution is given by the *m*-vector $\mathbf{u}^0$ described in Example 1. Consider two additions E and F to this initial population. Addition E consists of q individuals each with a high utility level of $\beta_1 z$, where z is the critical level and $\beta_1 > 1$; and addition F consists of p individuals each with a utility level a little above the critical level but well below the level shared by the members of addition E: each individual in F has a welfare level of $\beta_2 z$, with $1 < \beta_2 << \beta_1$. The population addition E is represented by the *q*-vector of utilities $\mathbf{u}_E = (\beta_1 z, \beta_1 z, ..., \beta_1 z)$, and the addition F by the *p*-vector of utilities $\mathbf{u}_F = (\beta_2 z, \beta_2 z, ..., \beta_2 z)$. For specificity, suppose that $\beta_1 = 10{,}001$ and $\beta_2 = 1.1$. If CLGU is captured by an Atkinson-type welfare function, then—making use of Equation (4) and performing some routine manipulations—it can be verified that

$W^A_{\text{CLGU}}(\mathbf{u}_E; z) > / = / < W^A_{\mathbf{CLGU}}(\mathbf{u}_F; z)$ according as

$$\gamma \equiv \frac{q}{p} > / = / < \gamma^* \equiv \frac{\beta_2^\lambda - 1}{\beta_1^\lambda - 1} \quad \text{for } \lambda \leq 1, \lambda \neq 0; \text{ and } \gamma \equiv \frac{q}{p} > / = / < \gamma^* \equiv \ln\left(\frac{\beta_2}{\beta_1}\right) \text{ for } \lambda = 0.$$

For the assumed values of $\beta_1 = 10{,}001$ and $\beta_2 = 1.1$ in Example 2, Table 1 furnishes the values of γ^* for a range of values of λ, from $\lambda = 1$ downward toward $\lambda \to -\infty$.

Notice that, as would be expected, the cut-off level of γ (γ^*) above which $\mathbf{u}_E$ will be pronounced better than $\mathbf{u}_F$ according to CLGU (when it is represented by an Atkinson-type social welfare function) increases as aversion to inequality increases, that is, as λ declines, until, in the limit, as λ goes to $-\infty$, γ^* goes to unity—which, as it happens, is the minimum value at which Principle M certifies that $\mathbf{u}_E$ is better than $\mathbf{u}_F$. That is, CLGU agrees with

Principle M only for λ tending to $-\infty$: this is the Rawlsian, or maximin, criterion by which population additions are judged solely according to the welfare level of the worst-off person in each addition.

Table 1: How γ^* varies with λ in Example 2

λ	γ^*	λ	γ^*
1	0.00001	:	:
0.9	0.00002	-1	0.0909
0.8	0.00005	-2	0.1736
0.7	0.00011	-3	0.2487
0.6	0.00024	:	:
0.5	0.00049	-5	0.3791
0.4	0.00100	:	:
0.3	0.00195	-10	0.6150
0.2	0.00362	:	:
0.1	0.00634	-20	0.8514
0	0.01004	:	:
-0.1	0.01576	-50	0.9915
-0.2	0.02244	:	:
-0.3	0.03009	$\rightarrow -\infty$	1

Thus, for CLGU[5] to endorse Principle M, the maximin criterion would have to be embraced—entailing judgements of the type that a population addition consisting of a single individual should be preferred to an addition with several persons when the welfare level of the single person in the first addition is just a little higher than that of the worst-off person in the second addition, even if all other individuals in the second addition share a much higher level of welfare. This amounts to endorsing a sort of 'Despotism of the Worst-Off (DOTWO) Conclusion':

The DOTWO Conclusion. Of two population additions to an initial population whose welfare levels are unaffected by the additions, the first addition, which has several members, must be pronounced worse than the second addition which has a single member, if all but one of the members in the first addition share a very high level of welfare while the worst-off individual has a smaller, but decent quality of life, which is however just a little bit lower than the welfare level of the only person in the second population addition.

[5] More accurately, a CLGU that is captured within the framework of the widely employed Atkinsonian formula.

Briefly, if Critical Level Utilitarianism implies a 'Despotism of the Best-Off Conclusion', a plausible version of Critical Level Generalised Utilitarianism implies a 'Despotism of the Worst-Off Conclusion'.

4. Conclusion

This paper has dealt with two issues in population ethics. The first issue relates to the significance of the Sadistic Conclusion as a convincing objection to the efficacy of Critical Level Utilitarianism in addressing and avoiding Parfit's Repugnant Conclusion. The suggestion is that the Sadistic Conclusion is not an embarrassment to CLU, and that it carries within itself the seeds of a certain sort of natural dissolution of the problem it raises.

The second issue relates to the possibility that there are unpalatable conclusions, different from the Repugnant Conclusion, which critical level axiologies still have to contend with. One such conclusion is that of the 'Despotism of the Best-Off (DOTBO)', which upholds a certain unacceptable privileging of very small populations with enormously high welfare levels over larger populations with more modest but nevertheless decent qualities of life: CLU is unable to avoid this conclusion.

A natural principle to invoke—referred to in this paper as the Principle of Maximization of the Numbers of Worthwhile Lives (Principle M)—takes the form of a rule which avoids the discriminatory favouring of opulent minorities by favouring larger populations in all comparisons involving populations whose members experience levels of welfare in excess of the critical level. It turns out that a plausible version of CLGU (involving aggregation through an Atkinson-type social welfare function) is compatible with Principle M only if one subscribes to the maximin criterion of always favouring that outcome in which the worst-off person is better off. This would imply a 'despotism conclusion' at the other end of a spectrum which has the DOTBO Conclusion at one end—a 'Despotism of the Worst-Off (DOTWO) Conclusion' which entails an unacceptable privileging of very small populations in which the worst-off person may be just a bit better off than the worst-off person in a much larger population all other members of which enjoy a high quality of life.

Briefly, it is contended in this paper that while the ability of critical level utilitarian axiologies in avoiding the Repugnant Conclusion is not seriously threatened by outcomes such as the Sadistic Conclusion, these axiologies are nevertheless susceptible to certain other

unacceptable outcomes, different from the Repugnant Conclusion, which involve allowing unduly large power to the best-off or the worst-off in determining what sorts of population additions to favour.

References

Armstrong, S. (2014). 'Embracing the "Sadistic Conclusion",' Blog on *Practical Ethics*, published 13 February 2014. Available at: http://blog.practicalethics.ox.ac.uk/2014/02/embracing-the-sadistic-conclusion/

Arrhenius, G. (2000). 'An Impossibility Theorem for Welfarist Axiologies,' *Economics and Philosophy*, 16(2): 247-266.

Arrhenius, G. (2006). 'The Repugnant Conclusion,' Stanford Encyclopaedia of Philosophy. Available at: https://www.academia.edu/26885796/The_Repugnant_Conclusion_in_Stanford_Encyclopedia_of_Philosophy

Atkinson, A.B. (1970) 'On the Measurement of Inequality,' *Journal of Economic Theory*, 2(3): 244-263.

Blackorby, C. and D. Donaldson (1984). 'Social Criteria for Evaluating Population Change,' *Journal of Public Economics*, 25(1-2): 13-33.

Blackorby, C. and D. Donaldson (1991). 'Normative Population Policy: A Comment,' *Social Choice and Welfare*, 8(3): 261-267.

Blackorby, C., W. Bossert and D. Donaldson (1997). 'Critical-Level Utilitarianism and the Population Ethics Dilemma,' *Economics and Philosophy*, 13(2): 197-230.

Blackorby, C., W. Bossert and D. Donaldson (2003). 'The Axiomatic Approach to Population Ethics.' *Politics, Philosophy and Economics*, 2(3): 342-381.

Blackorby, C., Bossert, W., Donaldson, D. (2006). 'Population Ethics and the Value of Life,' In: McGillivray, M. (ed.) *Inequality, Poverty and Well-being*: *Studies in Development Economics and Policy*. Palgrave Macmillan, London.

Broome, J. (1992a). 'The Value of Living,' *Recherches Economiques de Louvain*, 58(2): 125-142.

Broome, J. (1992a). 'The Value of Living,' *Recherches Economiques de Louvain*, 58(2): 167-171.

Griffin, J. (1986). *Well-Being: Its Meaning, Measurement, and Moral Importance*. Clarendon: Oxford.

Sen, A. (1976). 'Real National Income,' *Review of Economic Studies*, 43(1): 19-39.

Sen, A. (1991): 'Welfare Economics and Population Ethics,' presented at the Nobel Jubilee Symposium on Population, Development and Welfare, December 5-7, 1991, Harvard University (mimeo.).

Parfit, D. (1994). *Reasons and Persons*. Clarendon: Oxford.

About the Author

S. Subramanian is an Independent Scholar; formerly, Madras Institute of Development Studies.
Email: ssubramanianecon@gmail.com

VICTORIA UNIVERSITY OF WELLINGTON TE HERENGA WAKA
Chair in Public Finance
Working Papers in Public Finance