Division of Economics, EGC
School of Humanities and Social Sciences
Nanyang Technological University
14 Nanyang Drive
Singapore 637332

Global Extinction and Animal Welfare:

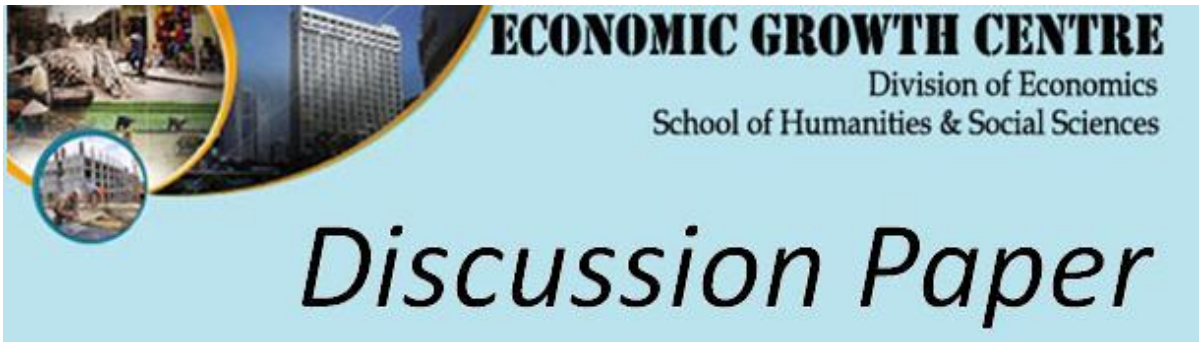
Two Priorities for Effective Altruism

Yew-Kwang Ng

September 2018

EGC Report No: 2018/08

HSS-04-86A
Tel: +65 67906073
Email: D-EGC@ntu.edu.sg



Global Extinction and Animal Welfare:

Two Priorities for Effective Altruism[1]

[This paper is based on the inaugural Atkinson Memorial Lecture, delivered on 22 February 2018, at the Global Priorities Institute & Economics Dept., Oxford University.]

Yew-Kwang Ng,

Division of Economics,

Nanyang Technological University.

ykng@ntu.edu.sg

Abstract: Effective altruism should ultimately be for the promotion of aggregate welfare. Broad altruism does not confine welfare to humans only. Thus, two priorities for broad and effective altruism may include reducing the probabilities of global extinction and the promotion of animal welfare. The former is important because if we become extinct, we lose the enormous amount of welfare into the far future. Also, we are faced with extinction probabilities that could be reduced, including through better environmental protection. Moreover, if we can avoid extinction, we will likely be able to increase our welfare enormously through such scientific and technological breakthroughs like brain stimulation and genetic engineering. Whether artificial intelligence may threaten our survival soon is also briefly discussed.

[1] I acknowledge the very helpful research assistance of Leon Sim.

An effective way to promote animal welfare is to reduce animal suffering at small or even negative costs for humans.

Keywords: Global priorities; effective altruism; animal welfare; extinction; climate change.

JEL Codes: H00 Public Economics; A10 General Economics.

Foreword

I am honoured to be invited to present this Atkinson Memorial Lecture. Professor Sir Tony Atkinson was a great economist. I have known him since 1973. It is also appropriate that, the Global Priorities Institute, which is associated with the movement of effective altruism, together with the Department of Economics, co-sponsors this lecture. Atkinson's life-long contributions in inequality/poverty were motivated by altruism after undertaking volunteer work. He has justly been described as 'one of the greatest and kindest economists'; who 'virtually single-handedly established the modern British field of inequality and poverty studies' and is known as 'the godfather of historical studies of income and wealth'. In 2016, Atkinson received the Dan David Prize for 'combatting poverty'.

I accept of course the importance of fighting poverty and reducing inequality. However, except for making three points quickly, I will not focus on issues of inequality and poverty; there are other authoritative tributes to Atkinson focusing on these (e.g. Aaberge at al. 2017, Jenkins 2017, Sandmo 2017, Harcourt 2018). Rather, I will focus on other less widely discussed global priorities related to effective altruism. This 'shift' in focus is partly due to the judgement that global poverty and inequality, though still of immense importance, have been significantly reduced in importance due to the general increase in real incomes and to the growth of formerly poor countries like China and India. However, inequality within each country has increased significantly over the last four decades.

Three Quick Points on Inequality

First, while increasing equality is desirable, we should do so efficiently (Stiglitz 2015). More than three decades ago, I argued for the adoption of efficiency supremacy in specific issues and the pursuit of equality overall (Ng, 1984a). For any degree of equality achieved, this policy involves less efficiency costs; alternatively stated, for any given amount of efficiency costs incurred, more equality may be achieved. Efficiency supremacy involves treating a dollar as a dollar irrespective of whether it goes to the rich or the poor. How could such an extreme right-wing principle be advocated by a left-wing economist? The fact was this. Ross Parish (1976) argued against the policy of treating a dollar to the poor as two dollars and a dollar to the rich as half a dollar. I tried to prove that, instead of this ratio of 4 to 1, perhaps it should be as much as 16 to 1. I ended up proving that $1 = $1 (in specific issues). This shows that my heart is left-inclined, but my head is 'right-wing'.

Secondly, in contrast to the traditional focus on the equality-efficiency/growth tradeoff (e.g. Mirrlees 1971, Okun 1975), recent research emphasizes the beneficial effects of equality on efficiency and growth (e.g. Solar 1995; Milanovic 2011*;* Stiglitz 2012; Mokyr 2014; Biswas et al. 2017). Partly, equality contributes to social harmony, which reduces crimes, softens the opposition to the introduction of new technologies, etc. and hence, promotes productivity and growth. Secondly, when there physical capital was important in former times, inequality allowed the rich to save more, leading to higher capital accumulation and growth. Now, human capital has become relatively more important, making equality beneficial to growth by allowing more people to be educated and thus attaining higher human capital.

Thirdly, it is interesting to note that most researchers (including economists and philosophers) are already more egalitarian than the general public (Weinzierl 2017). But, of course, effective altruists may adopt the view that an even higher degree of egalitarianism than that believed by researchers may be justifiable.

Effective Altruism

While the general idea of doing altruistic deeds well has been around from time immemorial, the term 'effective altruism' and the associated movement have been around for less than a decade. On these, readers are strongly encouraged to refer to two books by William MacAskill (2015) and Peter Singer (2015).[2] Though relatively new, the movement is worthy of support and has achieved much, including advising on more effective charities, actual donations/pledges by the activists themselves and by others, etc. Of course, a lot more has yet to be done.

Altruism should aim to increase net welfare (including the reduction of suffering). For most practical issues of effective altruism, most non-welfarists (including Kantians) need not differ much from the welfarists (including utilitarians)[3]. Nevertheless, non-welfarists should refer to my arguments for the compellingness of welfarism and utilitarianism (Ng 1975, 1990, forthcoming).

Altruism should also be broad, including the rejection of anthropocentrism, and the allowance for animal welfare to be included for consideration, at least morally speaking.[4] Altruism is complex (Wilson 2015; Robson 2017). An apparently altruistic act like a donation could be due to enlightened self-interest from the future benefits of cooperative games, future financial benefits from sales promotion, reputation benefits, etc. Though some may not wish to include these as truly altruistic, they should not be discouraged as long as positive benefits are produced. We also have the warm-glow effect (Andreoni 1990), which gives positive feelings to those doing good things. This may be called affective altruism. [Note the difference between 'affective' and 'effective' altruism.] Non-affective altruism exists when one is willing to do something that actually reduces one's own welfare,

---

[2] On a complement to MacAskill's analysis of effective altruism, particularly on the proper treatment of indirect effects, with insights from the economic theories of second and third best, see Ng (2018).

[3] As used here, welfarism is the belief that what matters ultimately or what is of intrinsic value is the welfare (defined as net happiness) levels of individuals (need not be confined to humans); utilitarianism has the further specification of unweighted or an equally weighted sum of individual welfare levels.

[4] Biologically: Altruism associated with group selection may entail hostility to outsiders (Bowles 2006, 2008). Hopefully, many of us may educate ourselves to reduce such tendencies.

even after taking into account all of the effects mentioned above. Though many economists regard non-affective altruism as impossible, I have provided an example where most people will show some degree of non-affective altruism (Ng 1999). In practice, many of these different factors may simultaneously exist for a given altruistic act. For example, even for a non-affective altruist, she will almost certainly benefit from having the warm glow of helping others, though she may be prepared to help even without the warm glow.

Global Priorities

Here, I support two priorities for the world, and for effective altruists in particular: Reducing the probabilities of global extinction; Improving animal welfare. My choices and the following discussion are influenced by the limitations of my personal knowledge. However, I am quite confident that the top priority should be the avoidance of possible global extinction, because:

A. If we become extinct, we will lose many centuries of expected welfare; 'even relatively small reductions in net existential risk have enormous expected value' (Bostrom 2013, Abstract).
B. If we could avoid extinction, our future expected welfare will likely jump quantumly, due to science and technology allowing the advent of inventions such as brain stimulation and genetic engineering.

In terms of the expected number of human lives, Sagan (1983) estimated an expected 500 trillion (human) lives or more. Bostrom (2013) gives a much larger estimate of 10 to the power of 16 [ten thousand trillion] people. Ord (forthcoming, Ch.8) paints an even more optimistic and grand flourishing of humanity (and post-human) beyond our Sun and Galaxy. If we could avoid global extinction within this and the next centuries, the likely continued advances in science and technology to levels beyond our present imagination will possibly underscore these seemingly over-optimistic predictions as serious underestimates.

While this first priority depends on the point that even a small change (in probability) of an immense magnitude is very large, the second priority of reducing animal suffering depends on the virtual certainty of our ability (given political will) to reduce animal

suffering enormously at little or even negative costs for humans. Let us consider these two priorities in turn.

Reducing the Risks of Global Extinction (Existential Risks or X-Risks)

'An existential risk is one that threatens the premature extinction of Earth-originating intelligent life or the permanent and drastic destruction of its potential for desirable future development' (Bostrom 2013). Though the second possibility does not really involve extinction, its undesirability could be as great, if the 'drastic destruction' is really permanent. In fact, an outcome worse than extinction may be possible, one that places us into great suffering permanently, described as the 'suffering risk' emphasized by Althaus & Gloor (2016). While this is possible, it is less likely than extinction, which is our main focus here.

The concern with catastrophes has been substantial in the recent decades.[5] Factors causing catastrophes may be classified into:

- Natural: Volcanic eruptions (Rampino 2008); asteroids (Chichilnisky & Eisenberger 2010); celestial collisions, cosmic rays (Dar 2008).
- Man-made: Nuclear wars (Sagan 1983, Cirincione 2008); Global-warming (IPCC 2014); Super AI (Bostrom 2014), etc.

Here, we are mainly concerned with the extreme forms involving the extinction of mankind.

Exogenous x-risks like huge celestial collisions are of lower frequencies, and the probabilities in the near to medium future are not high (Chapman & Morrison 1994, Solberg Søilen and Baback Alipour, 2011, Cotton-Barratt et al., 2016, p. 46; Global Challenges Foundation 2017, p.37; Ord, forthcoming). For example, asteroids hitting Earth causing damages on a global scale but not yet threatening survival occur about once in 120,000 years (Reinhardt et al. 2016). Thus, we may leave these largely to the future for

[5] Recent interests include: Leslie 1996; Rees 2003; Posner 2004; Bostrom & Cirkovic 2008; Tonn 2009; Pindyck & Wang 2013; Baum 2010-17; Martin & Pindyck 2015; Weitzman 2009; Bostrom 2013,2014; Bier (2018).

more effective analysis/intervention, though not ruling out the need for monitoring and research now.

Anthropogenic x-risks are more threatening but also more subject to moderation through our efforts now, making them important to focus on. Three sources of such x-risks are particularly worrying: Nuclear wars leading to nuclear winters; artificial superintelligence (ASI) taking over the world; catastrophic climate change (CCC).[6] I accept the importance of nuclear wars, but will not focus on it due to space/time limitations and a lack of expertise. I will briefly comment on ASI and CCC, before discussing an economic analysis of global extinction or x-risks. After which, I will move on to a discussion on the reduction of animal suffering.

Artificial Superintelligence (ASI)

The concern with super AI taking over has been around long, especially at the science-fiction/movie levels. Bostrom's (2014) book marks a new height as a high-profile academic discussion. Recently (reported on 2 Feb. 2018), a news anchor robot named Erica, created by Hiroshi Ishiguro, also made world-wide headlines by being able to engage in conversation. On the other hand, we have skeptics saying that worries about ASI are "like worrying about overpopulation on Mars when we have not even set foot on the planet yet" (Andrew Ng; quoted in Shermer 2017).

An important point why ASI may be worrysome is this. When the AI gets to the stage of being able to self-improve their own competency, and the repetition of such improvements, coupled with the speed of AI may make them vastly superior to humans within a relatively short time. If such super competent beings pursue goals misaligned with ours, we might be destroyed like termites, or used up as atoms for them to make paper clips, if it is the case that they are programmed to maximize the number of paper clips produced.

I can see the point that when AI gets to the stage of being able to self-improve their competency, their degree of competency may explode quickly. However, from my limited

[6] For other sources like synthetic biology, geo-engineering, etc., readers are referred to Torres (2017) and Ord (forthcoming).

knowledge about AI, they are far from being able to do so yet. A chess-playing AI may be able to 'play' chess by choosing each move and may learn from its experience in playing. Erica may be able to engage in conversation. However, all of these are controlled by the programs, not by the AI's own intelligence. In essence, the 'intelligence' of these AIs and that of their forthcoming more advanced versions, is no different from the 'intelligence' of a hand calculator doing pre-programmed mathematical calculation.

Whether machine intelligence could reach levels posited under Bostrom's Superintelligence, including intention, motivation, etc., may depend, at least partly, or perhaps largely, on whether machine intelligence could have a mind (consciousness) of its own. Could an AI really have 'preference', intention, motivation, etc. unless it has a mind? The apparent 'goals' as in a chess playing program seems qualitatively different from the real 'goals' for an agent with a subjective mind. This touches on the 'world-knot' of the philosophy of mind which has been unsolved for thousands of years. It would be unfair to expect an AI researcher to solve it. However, the insolvability of this mind-matter problem makes the evaluation of the x-risks of ASI very difficult. However, as the problem is extremely important, even a tiny chance of relevancy should not preclude us from taking it seriously.

I understand that a 'top myth' regarding ASI is the worry that ASI might become conscious and evil (Future of Life Institute 2016). The actual worry is: ASI becomes competent, with goals misaligned with ours. However, whether this actual worry will really be worrisome depends on the substance of the 'myth'. Perhaps, 'until an AI has feelings, it's going to be unable to want to do anything at all' (Chorost 2016). It will take at least many centuries (if possible at all) for AI to have feelings and true motivation. We do not even know how these are possible within humans, how could we program them in AI soon? On the other hand, Singer's 'escalator of reasoning' will lead 'societies to greater benevolence *regardless of species origin'* (Chorost 2016), when ASI becomes more competent than humans. This would also ensure that ASI would not be harmful.

Moreover, if ASI has subjective feelings, they will then be capable of welfare, and in time, will likely be capable of enormous amounts of welfare. Then, even if ASI were to replace us, should this not be regarded as an advance? Just like how the evolution of the

era of apes into that of humans was seem as a big advance, the evolution of our era into that of ASI (or possibly a human/ASI hybrid), if possible, should be treated as another big advance. If apes were able to prevent the evolution into the human era, should it not refrain from doing so, morally speaking? Especially for altruistic apes? We should think about this deeply.

If truly altruistic, we should be worrying about the promotion of net welfare overall, irrespective of animal, human, or post-human. We should prevent ASI from actions that may destroy overall welfare. However, we should not prevent ASI that will bring about a higher level of welfare, if possible, from taking over.

Machine minds may not be possible, at least for many centuries. However, before truly subjective motivation becomes possible, it may be possible for ASI to have effective 'motivation' like that of a guided missile. Something may be mis-programmed to become very competent and do something misaligned with our interest. This may be possible, just like instruments may be misused, accidents may happen, etc. However, such problems will likely be easier handled by humans to prevent serious catastrophes. Thus, while advanced research and safeguards to reduce the problems may be prudent, whether this should be our current priority is unclear.

A relevant consideration is this. The side effects of emphasizing the x-risk of CCC (discussed below) are largely positive; there are little negative effects, if any. On the other hand, there are potentially large negative effects of 'premature/counterproductive regulation' (Karnofsky 2016) of emphasizing the x-risk of ASI. This is partly because the development of AI will most likely provide us with huge benefits for a long time yet. Perhaps it is better to focus on other priorities like CCC now and leave the ASI problem largely for the future. This does not preclude those interested in doing research on it right now.

Catastrophic Climate Change (CCC)

Though by no means certain, CCC causing global extinction is possible due to interrelated factors of non-linearity, cascading effects, positive feedbacks, multiplicative factors, critical thresholds, and tipping points (e.g. van Aalst 2006, Buldyrev et al. 2010, Hansen &

Sato 2012, IPCC 2014, Schuur et al. 2015, Barnosky & Hadly 2016, Sims & Finnoff 2016, Belaia et al. 2017, Grainger 2017, Osmond & Klausmeier 2017, Rothman 2017, Kareiva & Carranza 2018).[7]

A possibly imminent tipping point could be in the form of 'an abrupt ice sheet collapse [that] could cause a rapid sea level rise' (Baum et al. 2011). There are many avenues for positive feedback in global warming, including:

- The replacement of an ice sea by a liquid ocean surface from melting reduces the reflection and increases the absorption of sunlight, leading to faster warming;
- The drying of forests from warming increases forest fires and the release of more carbon;
- Higher ocean temperatures may lead to the release of methane trapped under the ocean floor, producing runaway global warming.

Though there are also avenues for negative feedback, the scientific consensus is for an overall net positive feedback (Roe and Baker 2007). Thus, the Global Challenges Foundation (2017, p.25) concludes, 'The world is currently completely unprepared to envisage, and even less deal with, the consequences of catastrophic climate change'.

The threat of sea-level rising from global warming is well-known, but there are also other likely and more imminent threats to the survivability of mankind and other living things. For example, Sherwood & Huber (2010) emphasize the adaptability limit to climate change due to heat stress from high environmental wet-bulb temperature. They show that 'even modest global warming could … expose large fractions of the [world] population to unprecedented heat stress' and that with substantial global warming, 'the area of land rendered uninhabitable by heat stress would dwarf that affected by rising sea level', making extinction much more likely and the relatively moderate damages estimated by most integrated assessment models unreliably low.

While imminent extinction is very unlikely and may not come for a long time even under business as usual, the main point is that we cannot rule it out. Annan and Hargreaves

[7] The next few pages are partly based on Ng (2016a).

(2011, pp.434-5) may be right that there is 'an upper 95% probability limit for S [temperature increase] … to lie close to 4ºC, and certainly well below 6ºC'. However, probabilities of 5%, 0.5%, 0.05%, or even 0.005% of excessive warming and the resulting extinction probabilities cannot be ruled out and are unacceptable. Even if there is only a 1% probability that there is a time bomb in the airplane, you probably want to change your flight. Extinction of the whole world is more important to avoid by literally a trillion times.

An Important Distinction

Future economic (objective) benefits (consumption, income) should be discounted by the economic rate of discount r which, arguably, should equal the market rate of interest.[8] However, future subjective benefits (utility, welfare) should only be discounted at the usually much lower uncertainty discount rate δ.[9] Though this difference in discount rates is itself an important distinction (fairly well-known), it is not the distinction of this section, which is:

- Higher exogenous probabilities of extinction due to factors (e.g. celestial collision) largely beyond our control should increase our uncertainty discount rate δ and hence decrease our willingness to sacrifice present benefits for future benefits.
- The higher extinction probability due to anthropogenic factors like CCC may increase our willingness to sacrifice our present benefits for future benefits, as the higher anthropogenic extinction probability increases the scope for reducing it.

The more your house is threatened by bushfire, the less you are willing to make improvements (like interior decoration) that do not reduce the chance of fire. However, you

[8] In a simple case without complications such as taxation, the market interest rate equals the marginal productivity of capital. Thus, as a dollar now could be invested (either putting into the bank or invested in productive capital) to become (1+r) dollars next year, future dollars should be discounted at r per year.

[9] The Ramsey (1928) equation has $r = \delta + \eta g$ where g is the annual rate of growth of per-capita consumption and η is the percentage decrease in marginal utility of consumption as consumption increases by one percent. Usually, both η and g are positive and large, making r (usually in the order of 3-5%) much larger than δ (arguably much less than 0.1%).

are more willing to make improvements (external clearance, less fire-prone materials) that **do** reduce its chance of being destroyed by fire.[10]

For measures that reduce our extinction probabilities, we should not take the uncertainty discount rate on future welfare δ as given. This important point is largely ignored in most analyses. If δ is taken as given, the problem is simply one of intertemporal consumption trade-off (environmental protection now decreases our current consumption but increases the consumption levels of our grand-children as they will have to invest less in cleaning up). Then, the higher δ, the less willing we should be to sacrifice current benefits (consumption or welfare) for future ones.

Some analysts truncate the time horizon, looking only at the next one or two hundred years. Though the probability that we will survive more than 200 years should be much higher than 50%, this simplification is not too misleading, if the problem is just intertemporal consumption tradeoff, especially if the economic discount rate r is in the order of the common 3-5%. For example, a stream of value worth $10,000 in real terms from the year 201 (i.e. 200 years from now) **every year** through to infinity, discounted at 5%, has only a present value of $11.56 **in total**! This underlines the misleading nature of focusing on the consumption tradeoff for problems like global warming with long-lasting effects and catastrophic outcomes of extinction. The major problem is not that of intertemporal consumption tradeoff, but that of reducing the extinction probability. This is the important distinction.

Towards an Economic Analysis of X-Risks

Nearly 3 decades ago, I discussed 'decisions (e.g. ... environmental protection…) that may affect the probabilities of the continued survival of the human race' (Ng 1991, p.79). I noted that, if our expected welfare was infinite, we should be willing to sacrifice enormously to reduce the extinction probability marginally, making rational decisions difficult. This is similar to Weitzman's (2009, p.10) dismal theorem, which implies our willingness to

[10] However, since the realization of future benefits also become less certain, there are two opposing effects, and the net outcome may go either way, as analysed in more detail in Ng (2016).

sacrifice 100% of our current consumption[11] to avoid future catastrophic reduction in consumption, as such reductions may lead to an arbitrarily high marginal utility of consumption. However, I reached this dismal result only for cases where our expected welfare is infinite, and I argued that it is finite as our solar system and even our universe will be finite. For the relevant cases of finite expected utility/welfare, the dismal result does not apply. Obviously, if even the total expected utility of the whole world is finite, no marginal utility can be infinite, since infinite negative utilities may be safely ruled out, as suicide is a feasible option and since our capacities for both enjoyment and suffering are also finite.

Weitzman (2007) correctly mentions that in dealing with global warming, the expected growth rate should have 'a thick left tail' (p.718); but acknowledges that 'we lack a usable economic framework for dealing with these kinds of thick-tailed extreme disasters' (p.723). A method in dealing with such catastrophes including x-risks is outlined below.

In the presence of uncertainty, it is rational to maximize the value of expected welfare (Ng 1984b). For simplicity, we ignore the differences (on which see Ng 1999) between individual utility (representing their preferences) and welfare. However, traditional economics, especially with the insistence on ordinal utility, cannot analyse x-risks which requires the comparison of total utilities/welfares with marginal utilities, hence requiring cardinal utility. (Economists not comfortable with the use of cardinal utility are referred to Ng, 1997.) Three different methods may be used to compare total with marginal utilities:

- Estimate the willingness to pay to reduce risks of death by observing actual choices.
- By asking people.
- From postulated cardinal utility/happiness functions.

Economists are skeptical of the last two methods. However, recent studies in behavioural economics show that even the first method is subject to many inaccuracies due to the lack of information and imperfect rationality. Thus, we should be willing to use a combination of different methods to provide a better picture.

Results in happiness studies provide some clues. These suggest high value of total utilities (most people are rather happy) relative to marginal utilities (additional

[11] Recognising that our own survival is essential to ensure the existence of future generations, we should not go below the survival level.

consumption beyond a low level of survival and comfort does not significantly increase happiness, especially at the social level and after an adaptation period). These justify immediate and strong actions to reduce CCC, as these actions sacrifice the consumption of low marginal utilities to reduce the probabilities of losing huge amounts of expected welfare. If we do not become extinct within a century or two, our expected welfare will likely increase tremendously through scientific and technological advances (including brain stimulation, genetic engineering, and methods we may not be able to dream of; see, e.g. Ng, 2000, Appendix A; Diamandis, 2016; Ord, forthcoming). These further leaps in our expected welfare further increase the importance of reducing x-risks.

However, like nuclear power or ASI, though the use of technology to enhance our welfare may have a high potential, it may also carry great risks, including increased x-risks. Thus, they tend to increase the probabilities that our expected welfare (through to time infinity or time of our extinction) becomes much higher, but also the probabilities that we become extinct sooner and hence result in smaller expected welfare. Therefore, they tend to re-shape our aggregate (net) welfare distribution curve from a likely inverted U-shaped into one that is U-shaped. They may thus either increase or decrease our expected welfare, depending much on whether we handle the x-risks appropriately.

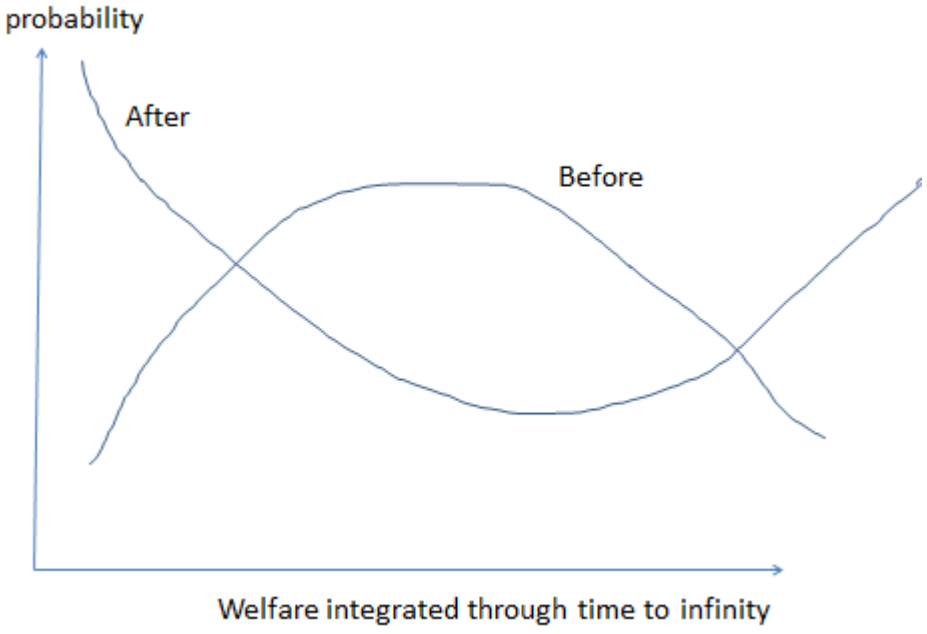


Figure 1

Using the method of maximizing expected welfare, with the comparison of marginal and total utilities, I (Ng 2016a) show:

An Important Difference

Consumption in the far future is very unimportant as r is large (around 3-5%). On the other hand, it is very important to avoid/reduce x-risks as our expected future welfare (total utility) is very high, and as the pure uncertainty discount rate δ should be very low (< 0.01% p.a.). Under very wide ranges of parameters, the benefit is many thousand times the costs. It is compelling to increase our effort in measures (such as environmental protection) that may reduce our extinction probabilities. The crux of the difference: r is large and δ is small.

However, there is an argument that δ is large at our present stage. This is based on the Fermi paradox. Since the universe is huge and our own Galaxy alone has many billions of stars and planets, it is likely that, even if a small fraction of them are suitable for life, there should be many. Thus, it is concluded that most civilizations soon become extinct after reaching our stage, as we should otherwise expect many extraterrestrial visits/conquests. 'Where is everybody?' The Great Silence means: 'Life almost never makes it beyond our current state of technological development' (Torres 2017, p.219).

If we realize the huge size of our galaxy alone (100,000 light years in diameter) and Einstein's limit of speed not exceeding that of light, there is no real paradox. Even if there were many civilizations within our galaxy, most of them would take dozens of thousands or even millions of years to reach us, even travelling at a significant fraction of the speed of light. Thus, those who could reach us within the lifespan of Homo sapiens are rather few. Moreover, some studies on possible ancient extra-terrestrial visits suggest that perhaps some of them have reached us before. Hence, if we are responsible in the next century or two (critical phase of Bostrom 2009), δ can be kept small. This is especially so after the advent of space colonization (Matheny 2007), as we would then make extinction of life on Earth not the total extinction of humans. On the one hand, this consideration increases tremendously the expected value of reducing x-risks (as δ is small), and on the other, it also makes space exploration possibly worthy, despite the high costs involved.

A Simple Way to Reduce X-Risks like CCC

A simple well-known way to protect the environment in general and to reduce x-risks like CCC in particular is to tax external costs like pollution and green-house emissions at their marginal damages imposed (Pigou 1932). While most economists agree with the rationality of this conceptually (but see below for an important exception), many are concerned with its practicability, especially regarding the difficulties of estimating the optimal tax amounts (=marginal social damages). There is a simple solution: For most cases, we should at least tax at the marginal cost of reducing the relevant pollution through abatement investment (Ng, 2004). Moreover, the amount of tax revenue so collected will be more than enough to pay for the optimal amount of abatement investment for pollution reduction, thus also solving the problem of financing for this investment.

However, an important objection to the above Pigovian taxation of external costs has been raised by an eminent economist and Nobel laureate Ronald Coase (1960). Nevertheless, this objection is based on the mistake of using an all-or-nothing comparison which ignores an asymmetry: costs of reducing pollution by polluters are negligible at the margin; costs of pollution to society are very substantial (Ng, 2007). Comparing only free-pollution with no pollution, Coase was able to show that either one option may be better than the other. This argument is only valid against the extreme environmentalist position of banning pollution completely, but not valid against Pigou's position of taxing pollution to reduce it to a socially optimal level. This reduction is certainly desirable because of the asymmetry mentioned above.

Another important point in environmental protection is that, if cap (quota) & trade are used instead of taxation, it is crucial not to give out free quotas. Free quotas (rewarding negative behavior) encourage future pollution, and might even reduce future morality.

Reducing Animal Suffering: Another Priority

This should be included as one of our priorities because big decreases in animal suffering could be achieved at small or even negative costs for humans. Some examples from a fuller discussion in Ng (2016b) are provided below.

One way of reducing great suffering on animals at negligible costs for humans is the prohibition of extreme cruelty to animals that really serves no good purpose. Two examples are given here. First, as reported in Bekoff (2013, pp. 230-1) on the study of David Evans

& Paul McGreevy, ‘whipping horses [in horse racing] is pointless and does not make a difference in the outcome of the race’ and ‘increased whip use was not associated with significant variation in velocity’. I would like to go even further. Even if whipping is effective in increasing the speed of the horse, it should still be banned. Suppose rigorous whipping increases speed by x%. After banning, the speed of all horses decreases by approximately x%. Since this applies to all horses in the race, it does not really matter. Problems such as unfairness arise only if some horses are allowed to be whipped and some are not. As banning should apply to all, no real problems are created, but we save much suffering on the part of the horses.

Consider another example, where fishmongers in the wet markets of Hong Kong (perhaps in other places as well, but I am not aware of) cut eels into two halves while they are still alive. This practice causes the eels to wriggle in pain, as a demonstration of their freshness, and is seen as an attractive quality to customers. I confronted the fishmongers multiple times, arguing that this would inflict pain on the eels. None refuted my claim. However, a fishmonger replied, ‘If I cannot sell the fish, I will also suffer pain!’ Assuming that the fish mongers are right and that the practice helps them to sell fish, it should still be banned. The pain of being cut into halves and wriggling in pain till death is too large to be tolerated (even after discounting some degree of possible anthropomorphism on my part). Moreover, stopping a fishmonger from employing this practice will only hurt him if others continue to do the same. If all fishmongers refrain from this cruel practice, they can still sell fish at virtually the same rates. Similar to the horse whipping example above, this is just another case of relative competition between individual producers or consumers that lead to inefficient social outcomes. We need the whole of society to take action by banning such inefficient competition that really serves no good purpose. However, this should be distinguished from competition between producers that reduce costs or increase the quality of products, since these do increase efficiency. Efficient competition (that does not involve cruelty and other external costs like pollution) should be allowed, if not encouraged. Thus, the problem is not competition as such, but the external costs imposed on other sentients.

Some simple regulations on the farming of animals may drastically reduce animal suffering at small or negative costs for humans. For example, in the factory farming of chicken, simple regulations on the minimum cage size will significantly improve chicken

welfare. The lack of action and inadequacy in the regulations are largely due to some misconceptions. Most producers believe that stricter regulations will increase their costs and reduce their profits. However, since these industries, including the factory farming of chicken, are mostly highly competitive, their long-run supernormal profits are non-existent; they can only earn average rates of return. Thus, after a temporary period of adjustment, stricter regulations will only increase prices of the final products without affecting the profitability of the producers. While consumers will have to pay higher prices, more and more consumers are willing to pay higher prices for products that involve less cruelty in the treatment of animals. Moreover, at least for the developed economies, meat consumption is excessive, and thereby unhealthy. Thus, the higher prices for meat, push consumers into consuming less, and actually make them better off. The costs on humans may well be negative! This is not just wishful thinking but is supported by research. As shown by Gruber & Mullainathan (2005), taxes on cigarettes that increase their prices actually make smokers healthier and happier. An important role for animal welfarists is to educate the public on this matter, and make society accept stricter regulations to reduce animal suffering.[12]

If we could improve the conditions of farmed animals to provide a positive net welfare in their lives, meat eating may actually be beneficial for them. Perhaps this is just my excuse for not being a vegetarian yet; Haha!

It is arguable that we should do much more; we should be prepared to incur significant costs to reduce animal suffering enormously, especially those that we impose upon them. However, strict rules on animal experiments that may impede scientific advances may well be counter-productive in the long run. Animal salvation (one of our human missions) depends on big advances in human science and technology, economy, and morality (Ng 1995). Advocates of animal welfare/liberation should be far sighted.

Concluding Remark

[12] Out of 7 top priorities for action, tackling factory farming receives the highest total number of points from Macaskill (2015, Table 2 on p.242) covering the three categories of ‘scale’, ‘neglectedness’, and ‘tractability.

Supporting a quote from Ramsey, Ord (forthcoming, middle of Ch.8) writes: 'What makes each of us special, so worthy of protection and celebration is something subtle about us, in the way that the matter which makes us up has been so delicately arranged as to allow us to think and love and create and dream.' To this poetic description, I wish to add: 'and capable of welfare, i.e. feeling good and bad'. It is the capacity to enjoy and suffer that is morally important. Let us treasure our capacity and not let the living world go extinct.

If we choose to invest heavily in environmental protection and reduce the (per-capita, annual, real) growth rate from 1.5% to 1.4%, our descendants 300 years from now will have an income nearly 65 times that of our own now. Alternatively, we may choose to proceed with business as usual (no big increases in environmental protection) in order to maintain the growth rate of 1.5%. If we are successful in achieving this rate without triggering extinction and drastic reduction in growth itself, our descendants 300 years from now will have an income 87 times that of ours now. However, they may very well not be born at all! The right choice is clear.

Mankind has faced extinction before, but the current threat is man-made and avoidable. If we manage to avoid extinction in a century or two, we face an enormously bright future with huge levels of net-positive welfare. The choice is one of committing suicide vs. going to **HEAVEN**! This is mankind's greatest crossroad!

References

AABERGE, Rolf, at al. (2017), Tony Atkinson and his legacy, Statistics Norway Research Department Discussion Paper No.863.

ALTHAUS, David, and Lukas GLOOR (2016). Reducing risks of astronomical suffering: A neglected priority. Foundational Research Institute.

ANDREONI, James (1990). "Impure altruism and donations to public goods: A theory of warm-glow giving." *Economic Journal*, 100(401): 464-477. DOI: 10.2307/2234133.

ANNAN, J.D. & HARGREAVES, J.C. (2011). On the generation and interpretation of probabilistic estimates of climate sensitivity, *Climatic Change,* 104(3-4): 423-436.

BAUM, Seth D., Jacob D. HAQQ-MISRA and Chris KARMOSKY (2011). "Climate change: Evidence of human causes and arguments for emissions reduction." *Science and Engineering Ethics*, 18(2): 393-410. DOI: 10.1007/s11948-011-9270-6.

BARNOSKY, A. D., and E. A. HADLY (2016). *Tipping Point for Planet Earth: How Close are We to the Edge?* Thomas Dunne Books.

BEKOFF, Marc (2013). *Why Dogs Hump and Bees Get Depressed*. New World Library.

BELAIA, M., FUNKE, M., & SCHNEIDER, U. (2017). *Integrated Assessment of Climate Tipping Points* (Doctoral dissertation, Universität Hamburg Hamburg).

BIER, Vicki (Ed.), 2018. *Risk in Extreme Environments: Preparing, Avoiding, Mitigating, and Managing*. Routledge, 174-184.

BISWAS, S., CHAKRABORTY, I., & AI, R. (2017). "Income inequality, tax policy, and economic growth." *Economic Journal*, *127*(601), 688-727.

BOSTROM, Nick (2009). "The future of humanity." *New Waves in Philosophy of Technology*, edited by Olsen J.K.B. Palgrave Macmillan, pp.186-215.

BOSTROM, Nick (2013). "Existential risk prevention as a global priority." *Global Policy*, 4(1): 15–31. DOI: 10.1111/1758-5899.12002.

BOSTROM, Nick (2014). *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press.

BOSTROM, N. & CIRKOVIC, M.M., eds. (2008). *Global Catastrophic Risks*, Oxford: Oxford University Press.

BOSTROM, N., and E. YUDKOWSKY (2014). "The ethics of artificial intelligence." *The Cambridge Handbook of Artificial Intelligence*, edited by Keith FRANKISH and William M. RAMSEY, Cambridge University Press, 316-334.

BULDYREV, Sergey V., et al. (2010). "Catastrophic cascade of failures in interdependent networks." *Nature*, 464(7291): 1025-1028. DOI: 10.1038/nature08932.

CHOROST, Michael (2016). "Let artificial intelligence evolve." *Slate*, 18 April 2016, http://slate.me/1SgHsUJ.

COTTON-BARRATT, Owen, et al. (2016). *Global Catastrophic Risks 2016.* Global Challenges Foundation: Global Priorities Project.

CHAPMAN, C. R., and D. MORRISON (1994). "Impacts on the Earth by asteroids and comets: assessing the hazard." *Nature,* 367(6458): 33-40. DOI: 10.1038/367033a0.

CHICHILNISKY, Graciela, and P. EISENBERGER (2010). “Asteroids: Assessing Catastrophic Risks.” *Journal of Probability and Statistics*, 2010: 1-15. DOI: 10.1155/2010/954750.

CIRINCIONE, Joseph (2008). *Bomb Scare: The History and Future of Nuclear Weapons*. Columbia University Press.

COASE, Ronald H. (1960). 'The problem of social cost', *Journal of Law and Economics,* 3:1-44.

DAR, Arnon (2008). "Influence of Supernovae, gamma-ray bursts, solar flares, and cosmic rays on the terrestrial environment." *Global Catastrophic Risks* (2008): 238.

DIAMANDIS, Peter H. (2016). “The Road to Abundance—Innovation, Disruption, and Opportunity: 2016 IRI Medal Address.” *Research-Technology Management,* 59(6): 20-24. DOI: 10.1080/08956308.2016.1232135.

Future of Life Institute (016). “The Top Myths About Advanced AI.”, https://futureoflife.org/background/aimyths/.

Global Challenges Foundation (2017). *Global Catastrophic Risks 2017*. Global Challenges Foundation.

GRAINGER, Alan (2017). “The prospect of global environmental relativities after an Anthropocene tipping point.” *Forest Policy and Economics*, 79: 36-49. DOI: 10.1016/j.forpol.2017.01.008.

GRUBER, J., and S. MULLAINATHAN (2005). “Do cigarette taxes make smokers happier?” *The B.E. Journal of Economic Analysis and Policy*, 5(1): 1-43.

HANSEN, James E., and M. SATO (2012). “Paleoclimate implications for human-made climate change.” *Climate Change*, Springer Vienna, 21-47.

HARCOURT, Geoffrey C. (2018). Tony Atkinson (4 September 1944 to 1 January 2017): A personal tribute. Economic and Labour Relations Review. 2018 Feb 13:1035304618756824.

Intergovernmental Panel on Climate Change (2014). *Climate Change 2014.* Intergovernmental Panel on Climate Change.

JENKINS, Stephen P. (2017). “Anthony B. Atkinson (1944–).” *The Palgrave Companion to Cambridge Economics*, Palgrave Macmillan, 1151-1174.

KAREIVA, Peter, and Valerie CARRANZA (2018). “Existential risk due to ecosystem collapse: Nature strikes back.” *Futures*.

KARNOFSKY, Holden (2016). “Potential risks from advanced artificial intelligence: The Philanthropic Opportunity.” *Open Philanthropy Project*, 6 May 2016, https://www.openphilanthropy.org/blog/potential-risks-advanced-artificial-intelligence-philanthropic-opportunity.

LESLIE, John & Freeman J. DYSON (1996). *The End of the World: The Science and Ethics of Human Extinction.* Routledge.

MACASKILL, William (2015). *Doing Good Better*, London: Guardian.

MARTIN, Ian WR, and Robert S. PINDYCK (2015). "Averting catastrophes: the strange economics of Scylla and Charybdis." *American Economic Review* 105(10): 2947-85.

MATHENY, Jason G. (2007). “Reducing the risk of human extinction.” *Risk Analysis,* 27(5): 1335-1344. DOI: 10.1111/j.1539-6924.2007.00960.x.

MIRRLEES, James A. (1971). “An exploration in the theory of optimum income taxation.” *Review of Economic Studies*, 38(2): 175-208. DOI: 10.2307/2296779

MILANOVIC, B. (2011). “More or Less.” *Finance and Development*, 48(3): 7-11.

MOKYR, J. (2014). “A flourishing economist.” *Journal of Economic Literature*, 52(1): 189-196. DOI: 10.1257/jel.52.1.189

NG, Yew-Kwang (1984a). “Quasi-Pareto social improvements.” *American Economic Review*, 74(5): 1033-1050.

NG, Yew-Kwang (1984b). “Expected subjective utility: Is the Neumann-Morgenstern utility the same as the neoclassical’s?” *Social Choice and Welfare*, 1(3): 177-186.

NG, Yew-Kwang (1990). “Welfarism and utilitarianism: A rehabilitation.” *Utilitas,* 2(2): 171-193.

NG, Yew-Kwang (1991). Should we be very cautious or extremely cautious on measures that may involve our destruction? On the finiteness of our expected welfare. *Social Choice and Welfare*, 8(1): 79-88.

NG, Yew-Kwang (1995). Towards welfare biology: Evolutionary economics of animal consciousness and suffering. *Biology and Philosophy*, 10: 255-285.

NG, Yew-Kwang (1997), ‘A case for happiness, cardinal utility, and interpersonal comparability’, *Economic Journal,* 107(445): 1848-1858.

NG, Yew-Kwang (1999), 'Utility, informed preference, or happiness?', *Socail Choice and Welfare,* 16(2): 197-216.

NG, Yew-Kwang (2000). *Efficiency, Equality and Public Policy: With A Case for Higher Public Spending.* Macmillan.

NG, Yew-Kwang (2007). Eternal Coase and external costs: A case for bilateral taxation and amenity rights. *European Journal of Political Economy,* 2007, 23: 641-59.

NG, Yew-Kwang (2016a). "The importance of global extinction in climate change policy." *Global Policy*, 7(3): 315-322. DOI: 10.1111/1758-5899.12318.

NG, Yew-Kwang (2016b). How welfare biology and commonsense may help to reduce animal suffering. *Animal Sentience,* 2016, 007.

NG, Yew-Kwang (2018). A welfare-economic analysis of effective altruism: Should indirect effects be taken into account? Manuscript.

OKUN, Arthur M. (1975/2015). *Equality and Efficiency: The Big Tradeoff*. Brookings Institution.

ORD, Toby (forthcoming). *Our Legacy: Existential Risk and the Future of Humanity*. Bloomsbury.

OSMOND, Matthew M., and Christopher A. KLAUSMEIER (2017). An evolutionary tipping point in a changing environment. *Evolution*, 71(12): 2930-2941. DOI: 10.1111/evo.13374.

PARISH, Ross M. (1976). "The scope of benefit-cost analysis." *Economic Record*, 52: 302-314.

PIGOU, Arthur C. (1932), *The Economics of Welfare,* London: Macmillan.

PINDYCK, R.S. & WANG, N. (2013). The economic and policy consequences of catastrophes, *American Economic Journal: Economic Policy, 5(4): 306–339*

RAMPINO, Michael R. (2008). "Super-volcanism and other geophysical processes of catastrophic import", *Global Catastrophic Risks*, *1*, 203.

RAMSEY, Frank P. (1928) "A mathematical theory of saving." E*conomic Journal,* 38(152): 543-559.

REES, Martin (2003). *Our Final Century: Will the Human Race Survive the Twenty-First Century?* William Heinemann.

REINHARDT, J. C., et al. (2016). "Asteroid Risk Assessment: A Probabilistic Approach." *Risk Analysis,* 36(2): 244-261. DOI: 10.1111/risa.12453.

ROBSON, A. J. (2017). "Group Selection: A Review Essay on Does Altruism Exist? by David Sloan Wilson." *Journal of Economic Literature*, 55(4): 1570-82. DOI: 10.1257/jel.20151363.

ROE, Gerald H., and BAKER (2007). "Why is climate sensitivity so unpredictable?" *Science*, 318(5850): 629-632. DOI: 10.1126/science.1144735.

ROTHMAN, Daniel H. (2017). "Thresholds of catastrophe in the Earth system." *Science Advances*, 3(9): 1-12. DOI: 10.1126/sciadv.1700906

SANDMO, Agnar (2017). "Tony Atkinson 1944-2017: A lifetime commitment to the study of inequality." *European Journal of the History of Economic Thought,* 24(3): 612-623. DOI: 10.1080/09672567.2017.1307628.

SAGAN, C (1983). "Nuclear war and climatic catastrophe: some policy implications." *Foreign Affairs,* 62(2): 257-292. DOI: 10.2307/20041818.

SCHUUR, Edward A. G., et al. (2015). "Climate change and the permafrost carbon feedback." *Nature,* 520(7546): 171-179. DOI: 10.1038/nature14338

SIMS, Charles, and David FINNOFF (2016). "Opposing irreversibilities and tipping point uncertainty." *Journal of the Association of Environmental and Resource Economists*, 3(4): 985-1022.

SHERMER, Michael (2017). "Artificial Intelligence is not a threat – yet." *Scientific American*, 1 March 2017, https://www.scientificamerican.com/article/artificial-intelligence-is-not-a-threat-mdash-yet/.

SHERWOOD, Steven C., and Matthew HUBER (2010). "An adaptability limit to climate change due to heat stress." *PNAS*, 107(21): 9552-9555. DOI: 10.1073/pnas.0913352107.

SINGER, Peter (2015). *The Most Good You Can Do*. Yale University Press.

SOLAR, Peter (1995). "Poor relief and English economic development before the industrial revolution." *Economic History Review*, 48(1): 1-22. DOI: 10.2307/2597868

SOLBERG Søilen, K., & Baback Alipour, P. (2011). "Semantics, Classifications and Evidence in a Model for Global Catastrophic Risks." *Knowledge Organization*, 38(5): 438-454.

STIGLITZ, Joseph J. (2012). *The Price of Inequality: How Today´s Divided Society Endangers our Future.* Norton.

STIGLITZ, Joseph J. (2015). *How to Fix Inequality*. Boston Globe.

TORRES, Phil (2017). *Morality, Foresight & Human Flourishing: In Introduction to Existential Risks*, Durham, North Carolina: Pitchstone.

VAN AALST, Maarten K. (2006). "The impacts of climate change on the risk of natural disasters." *Disasters*, 30(1): 5-18. DOI: 10.1111/j.1467-9523.2006.00303.x.

WEINZIERL, Matthew (2017). "Popular acceptance of inequality due to innate brute luck and support for classical benefit-based taxation." *Journal of Public Economics*, 155: 54-63. DOI: 10.1016/j.jpubeco.2017.09.007

WEITZMAN, Martin L. 2007. "A Review of the Stern Review on the Economics of Climate Change." *Journal of Economic Literature,* 45(3): 703-724.

WEITZMAN, Martin L. (2009). "On modeling and interpreting the economics of catastrophic climate change." *The Review of Economics and Statistics*, 91(1): 1-19.