



Effective altruism despite the second-best challenge: Should indirect effects Be taken into account for policies for a better future?

Yew-Kwang Ng^{a,b,*}

^a School of Economics, Fudan University, China

^b Monash University, Australia



ARTICLE INFO

Keywords:

Effective altruism
Futures research
Indirect effects
Priority
Second best
Third best

ABSTRACT

As there are many areas of inadequate optimization (departures), and resource limitation and information costs prevent the rectification of all these departures, the pursuit of a more desirable future through either private effective altruism or governmental policies is subject to the challenge of the second-best theory (where the presence of uncorrectable distortions complicates the pursuit of desirable policies elsewhere through interdependence). This is related to the indirect effects of altruistic acts. The distinction between real and pecuniary external effects and the second and third-best theories provide insights on how to evaluate these indirect effects. Indirect effects on areas of inadequate optimization should be taken into account where possible. Despite the nihilistic implication of the second-best theory on the impossibility of piecemeal welfare policies (unless all departures from optimality are eliminated, which are almost always impossible, we cannot be certain of an overall improvement by making improvements in some specific areas), the third-best theory shows that the government or effective altruists may increase at least the expected welfare by focusing on areas of serious inadequate optimization, taking into account the indirect effects if information allows.

Futures research¹ may be concerned with both the positive question of what will be and the more prescriptive question of what are better policies and measures to promote a better future. This paper is concerned with the prescriptive question, especially with both private altruistic acts and government policies that may have important effects in the future, including on global survival. While the general idea of doing altruistic things well must be around from time immemorial, the term ‘effective altruism’ and the associated movement have been around for only about a decade. Despite this, the movement has achieved much influence. MacAskill (2015) provides a good and quite comprehensive analysis of effective altruism focusing on the direct effects of altruistic acts; this paper provides some complement to MacAskill’s analysis by discussing the indirect effects. (The precise meaning of ‘indirect effects’ will become clearer in the following discussion, and defined more precisely in Section 3. Roughly, they are the effects of acts in one area on other areas or through interactions with other areas.)

Since altruism is to make others better off, we may take the normative position of welfarism (the maximization of overall welfare as a positive function of individual welfare levels, possibly including animal welfare) for granted; justification is available in Ng (1990; 2019a, Appendix B), Pummer and MacAskill (forthcoming). Alternatively, we may so define overall welfare and turn the

* Corresponding author at: School of Economics, Fudan University, China.

E-mail address: ykng@fudan.edu.cn.

¹ On which see, e.g. Ahvenharju, Minkkinen, and Lalot (2018) and van Dorsser, Taneja, and Marchau (2018) for some recent discussions.

normative question of ‘should?’ into a scientific/positive one of: Taking account of which indirect effects will increase overall welfare so defined.

There are typically many areas with inadequate optimization (e.g. serious poverty exists in many places, excessive environmental disruption in many areas, excessive animal suffering, including from human actions); that is mainly why altruists (and governments) have scope to make improvements. However, no altruist can expect that all areas with inadequate optimization could be completely rectified soon. Only improvements in some areas can be carried out. However, the theory of second best says that, unless we rectify all departures from optimality, correcting some departures only, may actually make the overall situation worse. This raises a serious challenge to both public policy and private altruism, effective altruism in particular. An example is that, addressing the inadequate reduction of poverty by donation to charity or to the poor directly may reduce poverty there. However, if this reduction leads to much higher consumption of goods with serious untaxed environmental disruption, the overall situation may be worse. [This would not arise if environmental disruption in all activities has been adequately taxed.] This also suggests that we should not just focus on the direct effects but should also consider the indirect effects of altruistic acts or government policies.

Apart from discussing effective altruism in action, MacAskill discusses five key questions of effective altruism:

- 1 How many people benefit, and by how much?
- 2 Is this the most effective thing you can do?
- 3 Is this area neglected?
- 4 What would have happened otherwise?
- 5 What are the chances of success, and how good would success be?

These are very relevant questions for doing altruistic things effectively. To complement this analysis, this paper answers the challenge of the second-best theory. Relatedly, it provides some analysis that may help answer whether indirect effects (of altruistic acts, and possibly also of other changes/events/policies) should be taken into account. If the indirect effects are positive, they enhance the justification for the altruistic act; if negative, they may make an apparently effective altruistic act non-effective, or even counter-altruistic. Thus, this question should be very important for effective altruists. This question has not been addressed in either the welfare economics or the effective altruism literature, partly due to the infancy of the latter.² Despite the apparently nihilistic implications of the second-best theory, this paper shows that the government or effective altruists may increase at least the expected welfare by focusing on areas of serious inadequate optimization, taking into account the indirect effects if information allows. This is based on the third-best theory that takes into account also administrative and informational costs and advocates taking account of some interrelationships only, especially those on which we have information and/or are more important in their effects.

1. Should indirect effects Be taken into account? Some general discussion

It may be thought that both direct and indirect effects should all be taken into account. However, even simple traditional economic analysis suggests a more complicated picture: Some effects should and some should not be taken into account to achieve economic efficiency; these may vary under different circumstances.

Consider first the simple economic model of perfect competition (no monopolistic power), no external effects like pollution, no relevant ignorance (a consumer may not know when was the Glorious Revolution but knows which type of ice-cream she likes). Then, for pure economic efficiency (not concerned with inequality), no indirect effects need to be taken into account; the economy is already perfectly efficient (First Theorem in Welfare Economics; see [Ng, 2015](#) for an exposition). If you (as a member of this economy) make a change like shifting your consumption from commodity 1 to commodity 2 based on your own cost-benefit calculation of the direct effects on yourself, you do not have to worry about weighing the benefits and costs of any indirect effects on others. The economy remains perfectly efficient. (See the discussion of pecuniary external effects in the next paragraph.) The presence of uncorrected real (or technological) external effects like pollution warrants the taking into account of some indirect effects. However, ‘pecuniary’ external effects should still be ignored.

A real external effect like the smoke of a factory (or of a cigarette smoker) that directly affects someone else (or the whole or parts of society) without compensation (either through payment or taxation) justifies taking into account the indirect effects that increase such uncorrected behavior, and hence, result in excessive levels of undesirable activities. On the other hand, a pecuniary external effect through the price system should be ignored, at least under the simple classical conditions mentioned above. Suppose that at the original equilibrium, the price of a commodity X is \$10 per unit. An increase in demand by some consumers (or from overseas) increases the equilibrium price to \$12, making existing consumers of X worse off (losing consumer surplus by an amount of \$2 times the amount consumed). This may be regarded as a pecuniary external effect or an indirect effect of the increase in consumption. However, this does not cause any inefficiency; the decrease in consumer surplus by some consumers is offset by an increase in producer surplus of the suppliers of X. Under classical conditions, the economy remains efficient both before and after the change in demand (See [Ng, 2014](#), Figure 5). Hence, such indirect effects should be ignored. This is in contrast to an increase in pollution that is

² The discussion regarding systemic/institutional changes (e.g. [Berkey, 2018](#); [Dietz, 2019](#)) may be regarded as a form of indirect effects, but the second and third best issues have certainly not been addressed. However, Gustav Alexandrie suggested to me that [Greaves’ \(2016\)](#) “distinction between ‘simple’ and ‘complex’ sources of cluelessness is in some ways similar to [the] distinction between Informational Poverty and Informational Scarcity”. Having read that paper, I find the relationship complex.

not taxed. Then, indirect effects that cause higher pollution may increase the net costs imposed and should be taken into account.

If we go beyond pure economic efficiency and take into account the issue of inequality, the situation becomes more complicated. For example, if the consumers who are made worse off by the increase in price of X from \$10 to \$12 above are very poor, their loss of say \$20,000 in consumer surplus may be larger in welfare terms than the gain by the producers of \$20,000 or more. If the government is watching both efficiency and equality and redresses this increase in inequality by costless redistribution, then no problem arises. But this may not be feasible. Thus, if a pecuniary external effect has equality implications, the indirect effect of increasing such an effect may have to be partly taken into account. However, it is more efficient for the government to address this issue through general equality promotion, than for society (including the government and altruists) to deviate from efficient measures in specific areas upon the consideration of the indirect effects on equality (Ng, 1984).

More generally, whenever the indirect effects involve areas of imperfect optimization (regarding both efficiency and equality), they have to be taken into account if significant. In fact, even for the direct effects, it is imperfect optimization that explains why altruists are concerned with those areas in particular. For example, if petrol (or electricity, water, etc.) is already optimally priced at its full social marginal costs, inclusive of the relevant environmental disruption effects (say through taxation), then even environmentally-conscious altruists do not have to be concerned with excessive consumption of petrol. They only have to consume until the private marginal benefit equals the price, as dictated by purely self-interested consumption, because the marginal disruption effect is already reflected in the price via taxation. It is only for the usual case where environmental disruption is not adequately taxed and the price of petrol does not fully reflect the full social marginal costs that deliberate under-consumption may be socially desirable. Thus, an efficient policy for the government is to impose enough taxes or price sufficiently high for such goods like petrol, electricity and water³, rather than engage in daily propaganda on radio urging people to be 'a water saving hero', as done in Singapore.⁴ In fact, such moral suasion may well prove to be very counter-productive overall, as 'moral licensing' may likely apply (Merritt, Effron, & Monin, 2010). If one may easily become a hero by just saving water, one may feel less need to behave morally elsewhere.⁵ It is much more efficient to price water and other goods sufficiently, and use moral suasion only in those areas where economic means fail or are less effective.

For another example, if inequality/poverty has already been optimally reduced to a level consistent with the economic conditions of the whole society, the need for donations to alleviate poverty may be limited. However, here, there is a significant difference between governmental poverty-reduction efforts and that of private altruists. The government is typically faced with high costs in the forms of administrative, compliance, policing, and disincentive effects (though it may also have the advantage of economies of scale). Even if the government has optimized fully by taking account of these costs, additional promotion of poverty-reduction through private charity will likely continue to increase social welfare. Not only are the various additional costs much lower, there are the additional benefits of warm glow that offset the direct costs of the donors. Though there may also be similar warm-glow effects for environmentally friendly behavior, the various additional costs are likely to be much less.

Though it is true that it is mainly the absence of full optimization that warrants the scope for private altruism for both the direct and indirect effects, a distinction is still desirable. This is so because the very fact that a certain area motivates the concern of altruists suggests that it is already an area of seriously inadequate optimization, e.g. poverty reduction or environmental protection has been very inadequate, thus motivating altruists to do something about it. On the other hand, the indirect effects of altruistic acts may occur in many different areas; whether these indirect effects are of real concern depends on whether optimization is adequate to begin with. Thus, it is more important to emphasize this point for the evaluation of indirect effects.

Some examples of likely relevant indirect effects may help. There are at least three areas of major concern by effective altruists: poverty reduction; existential risks (i.e. risks of global extinction including owing to artificial super-intelligence and environmental collapse/climate change); animal welfare (Compare: MacAskill, 2018). These are discussed briefly in Section 4 below. Additional areas may be added without affecting the conceptual logic of the discussion below significantly. Arguably, all of these areas are not yet fully optimized; we should have more equality/less poverty, more environmental protection, and impose less animal suffering. This is why they are popular areas for altruists to focus on. This makes the direct effects themselves desirable. Reducing poverty, increasing environmental protection, and reducing animal suffering efficiently are all desirable in and of themselves. (On issues on cause selection and the related systemic/institutional changes in the effective altruism community, see Gabriel, 2017; Broad, 2018, the 2017 special issue on effective altruism of *Essays in Philosophy*, and research from organizations such as the Open Philanthropy Project, 80,000 h and the Future of Humanity Institute on the topic, including OPP's ongoing list of cause area summaries: <https://www.openphilanthropy.org/research/cause-reports>.)

One possible offsetting/reinforcing source of indirect effects is the indirect effects on the other two areas. For example, we may donate money to the poor to alleviate poverty. This direct effect of poverty reduction is good. However, if these people then increase their consumption of chicken significantly, more animal suffering may result. Then, this negative indirect effect may at least partially, possibly more than fully, offset the positive direct effect, partly depending on the relative weight put on their welfare levels.

³ Any equality implications on the real incomes of the poor could be offset by a more intensive equality promotion policy generally (Ng, 1984).

⁴ Singapore is probably a city/country that practices most efficient economic policies, such as using the most efficient Vickrey (i.e. second-price) auction in selling the COE (certificates of entitlement) for owning cars. However, for some reason, water is under-priced and we hear daily radio 'moral suasion' to be an hero of water saving.

⁵ It is true that we may also have the opposite complementary effect where morality promotion in one area increases moral behavior elsewhere, this complementary effect is unlikely to be important for the daily radio moral suasion of becoming a water saving hero. The reason for this difference is subtle but likely true.

However, this is only relevant because the chicken sector is not fully optimized. Ideally, either chicken farming should be regulated to ensure that farmed chicken do not suffer a life of misery, and/or that it should be taxed adequately to reflect the harm involved. If this adequate optimization has prevailed, then it is just like the case of the adequate pricing of petrol. We then do not have to be concerned with the indirect effect of more chicken consumption. The fact that animal welfare is an important area of our concern suggests that this optimization has not yet prevailed, and that this indirect effect should be counted.

For another example, suppose we, on the consideration of animal welfare, ban factory farming or tax it so heavily as to force it to be largely replaced by free-range farming. This direct effect in itself is good. However, if the resulting free-range farming takes up so much area formerly under forestry, we may increase the risk of global extinction so as to more than offset the gain in animal welfare. Again, this is a concern only if forestry is not preserved enough or de-forestation or environmental disruption is not taxed adequately.

Since whether an indirect (or direct) effect is of concern or not depends on whether it is in an area of inadequate optimization, the indirect effects we should pay most attention to are those that occur in other areas of the altruists' concern, since these are areas of serious inadequate optimization to begin with. We do not have to be concerned too much with indirect effects on areas beyond the altruists' direct concern, since those are likely to be areas not with serious inadequate optimization. On the other hand, if there is an important area with serious inadequate optimization, then it should arguably also be a concern of altruists, other things being equal. However, this is not the only consideration. For example, there may be an area A with serious inadequate optimization and yet altruists are not much concerned with it due to, say, the difficulties for altruists to achieve effective changes there, perhaps as a result of technical or political difficulties. Then, if a measure within an area of traditional altruists' concern happens to have indirect effects in area A, these indirect effects should be accounted for in evaluating the importance (or costs and benefits) of the relevant measure.

2. Some insights from the economic theories of second and third best

We may gain some insights from the economic theories of second and third best for the issue of what indirect effects should be taken into account. To understand these theories in a simpler setting, let us abstract away our concern with equality, and concentrate purely on efficiency to begin with.

In traditional economic theory, the first-best world is defined by the achievement of perfect efficiency in the Pareto sense (no one could be made better off without anyone being made worse off). Under classical conditions (including perfect competition and the absence of pollution), a general equilibrium (where the quantity supplied equals that demanded in each and every commodity at the prevailing set of prices for all commodities) of the market economy is perfectly efficient (First Theorem in Welfare Economics). This is first best. The first best needs not be confined to the unrealistic situation of no pollution. If all external costs like pollution are taxed (at negligible costs) at the relevant marginal social damages, the resulting situation is still Pareto efficient and first best (given the presence of technological external effects, which must be regarded as given and as parts of the technology of production, also assumed given).

The first-best situation typically involves the satisfaction of the first-best condition of the equality of the marginal social benefits of each commodity being equal to its marginal social costs ($MSB = MSC$). In the absence of external effects, these marginal social values equal the marginal private values. From this first-best world, the second-best issue arises if some of these first-best conditions cannot be satisfied and this violation is taken as an additional constraint (in addition to the technological or production possibility constraint, and the assumed given preferences or objective function). A question arises: With the additional second-best constraints, is the satisfaction of the original first-best conditions ($MSB = MSC$) in as many free (unconstrained) sectors/commodities as possible a desirable policy? Unless the first-best condition could be satisfied for all sectors, i.e. there is really no second-best constraint (such as $MSB = k$ times MSC , with $k \neq 1$ for some sectors) that exists at all, such that we go to the summit of the mountain of efficiency (first best), [Lipsey and Lancaster \(1956\)](#) show that, in general, we want to depart from the simple first-best conditions in all the free sectors. We no longer should satisfy the first-best condition for any sector (except by coincidence), not to mention doing so for as many as possible. Satisfying the first-best condition for more sectors may actually make matters worse. For the free sectors, instead of the simple first-best condition ($MSB = MSC$), we should satisfy the complicated second-best conditions that take into account the complicated interrelationships within the economy. Moreover, using the same logic (maths) of the second-best theory, we may also derive a second-and-a-half best result: Unless all the complicated second-best conditions are satisfied for all the free sectors, it is in general not desirable to satisfy as many second-best conditions as possible; satisfying more instead of less of them may make matters worse.

The intuitive rationale of the second-best theory may be explained in the following simple example. The first best requires the satisfaction of $MSB = MSC$ for all sectors. Suppose there is one (or more) commodity G, the production of which necessarily involves some heavy pollution. The first-best rule calls for taxing this commodity/pollution at the marginal external costs imposed on the society. However, suppose that this is either technically (not possible to identify the pollution, for example) or politically (the factory has strong political backing) unfeasible. If the non-taxation (and any other control) of this sector is taken as a second-best constraint, the first best is not obtainable. The second-best conditions/rules then prescribe that, for all the free sectors, we typically want to depart from the simple first-best $MSB = MSC$ (which takes into account only the direct effects of the commodities concerned, not the indirect ones through the complicated interrelationships discussed below). For example, if commodity H is highly complementary in consumption with commodity G (the one producing non-taxable pollution), other things being equal, we may want to tax H, making its $MSB > MSC$ (no longer wanting to observe the first-best rule of $MSB = MSC$), in order to discourage its consumption. This is so because the reduced consumption of H also reduces the consumption of G (G and H being highly complementary). This also reduces the production of G and hence serves to lower the undesirable pollution that we could not control directly. Thus, the second-best

conditions call for taking into account all such interrelationships (not only through complementarity and substitutability on the consumption side, but also those on the production side) to reduce pollution accordingly.

Unfortunately, not only are the administrative costs of implementing these complicated second-best conditions (accounting for all the free sectors taking into account all the interrelationships) prohibitive, the information required to define them is also not adequately available, or the costs of obtaining such information are also prohibitive. Though the second best has been described as the 'optimal feasible', it is so only assuming the non-existence of administrative and informational costs. Realistically taking these costs into account, the second best is neither optimal nor feasible. I define the second-best world as one in which some second-best constraints exist, preventing the achievement of the first best, plus the non-existence of administrative and informational costs. Under this unrealistic world, the complicated second-best conditions should be satisfied.

The real world is never second-best in this sense. Rather, it is almost always a third-best world, which is defined by the existence of some second-best constraints, plus the existence of administrative and informational costs. In this real world of the third best, what policies/rules should we follow? The theory of second best itself seems to suggest an impossibility: Either we go to the summit of first best (which is impossible in the presence of second-best constraints) or the summit of second best (which is impossible in the presence of information and administrative costs), or we do not know what is optimal. In the terminology of second best, piecemeal welfare policies are impossible or undesirable. In an atmosphere of this impossibility, a theory of third best has been provided to guide public policy (Ng, 1977; reprinted as Ng, 2017a). What policies should be followed depend much on the available amount of information and the administrative costs. For simplicity, concentrate on information.

If we have perfect information (and no administrative costs; however we are then no longer in the realistic third-best world, but in the non-existent world of second best), we follow the complicated second-best conditions. On the other extreme, we may have Informational Poverty. This is defined to be an amount of information (including zero but we may have a fair bit of non-relevant information) that does not allow us to judge which direction to diverge from the simple first-best condition. In the example above of commodity H being highly complementary with commodity G, we have more information than Informational Poverty. If we have the information that bread is complementary with butter but know nothing about the interrelationship of bread and butter with G, we may still be under Informational Poverty, despite the presence of some information. In this real world of third best, we should follow the first-best rule of $MSB = MSC$ for those sectors on which we have Informational Poverty. Though we are not sure that this is the best policy, we still maximize the expected value of our objective function. If we diverge in one direction from this first-best rule, we may gain, but we may also lose by a larger amount. The expected gain from divergence without sufficient justification is negative.⁶

For example, a commodity J may generate an external cost (e.g. pollution) of \$M per unit. Assuming a first-best world, we should tax J (or the pollution; assuming a one-to-one relationship between J and the pollution for simplicity) at \$M per unit. However, if the economy is not first-best, and there exists some second-best constraints, the theory of second best suggests that the \$M tax on J may then no longer be optimal, as we should take into account the interrelationships of J with those commodities under constraints. After accounting for these interrelationships, we may want to subsidize J instead of taxing it. The theory of third best concedes that, taking into account these interrelationships, we might want to tax J by $\$(M-N)$ and N could be larger than M, thereby turning it into a subsidy. However, we might want to tax J by $\$(M+N)$ instead, making it a much larger tax. In the absence of sufficient knowledge that suggests either $\$(M+N)$ or $\$(M-N)$ is more likely to be better, we should just stick to the first-best tax of \$M. Though we are not sure of optimality, we maximize expected value for the given amount of information. (See footnote 6 above on the rationale.)

If the amount of information is less than perfect but more than Informational Poverty, it is called Informational Scarcity. This is the case of commodity H above, in which we may want to diverge from the first-best rule of $MSB = MSC$ as mentioned. A real-world example of this third-best policy under Informational Scarcity is this. Without second-best constraints, we want to follow first-best rules all around, taxing all external costs. Thus, we should tax train/underground/MRT travel moderately and private car driving (or petrol) heavily; similarly, this should be done for all other commodities, allowing the first best to be achieved. However, suppose that, due to technical or political constraints, we cannot tax car driving which has much higher external costs of pollution and congestion per passenger than taking the train. Given that taking the train and car-driving are virtually perfect substitutes, we may then want to subsidize train rides to indirectly reduce car-driving, achieving an overall reduction in congestion and pollution. This is usually mistakenly called the second-best case for subsidizing train travel. In fact, this is a third-best policy under Informational Scarcity. The truly second-best policy has to take into account the interrelationships between all commodities, on which we do not have sufficient information. Moreover, as mentioned above, the logic of second-best theory (or the second-and-a-half best) entails that, taking into account just some interrelationships may make matters worse. It is the third-best theory that advocates proceeding with the accounting for some interrelationships only, especially those on which we have information and/or are more important. Even until just three years ago, the surviving originator of the second-best theory (Lipsey, 2017) still doubted the validity of this third-best prescription.

The insights of the theories of second and third best for effective altruism may be briefly explained. The theory of second best suggests that we should take into account the whole system, paying attention to interrelationships. This is valid and important. As mentioned above, the areas altruists (especially effective altruists) focus on are typically those that they perceive to be of inadequate optimization. In terms of the second-best theory, the first-best conditions are not satisfied in these areas. Then, this theory says that, unless we (the society or the altruists) can rectify all these departures from the first best, just making an improvement in one or more

⁶ This result is due to the general principle of increasing marginal costs or that the welfare costs of a departure from optimality typically increases at the square of the size of departure. Though this relationship may change in the presence of the second-best constraints, we do not know the direction of change under Informational Poverty, making the expected relationships unchanged; see Ng (2017a, 2017b).

areas, short of all areas, may make matters worse. This is demonstrated by the possibility mentioned above where increasing the incomes of the poor may lead to lower total welfare, if say they eat more chicken, leading to a larger increase in chicken suffering. The improvement in one area leading to an overall worse outcome will not be possible if all other areas are already fully optimized (observing the first-best condition).

The real world has many areas far from being fully optimized (excessive poverty/inequality, inadequate environmental protection, excessive animal suffering, to mention a few important ones). No altruist is optimistic enough to believe that we can achieve full optimization all around, or eliminate all departures soon. Then, the second-best theory suggests that we do not know whether and how partial improvements are possible. In its terminology, piecemeal welfare policies cannot be relied upon to improve overall social welfare. Does this mean that neither the government nor the altruists may make improvements? If we accept the second-best theory fully by its face value, yes. However, the theory of third best suggests otherwise.

It is true that, increasing the incomes of the poor may make them eat more chicken and possibly lead to an overall decline in welfare. However, it is also possible that it makes them shift from chicken to beef. It is likely that while factory-farmed chicken suffer from negative welfare, cattle roaming in the field have positive net welfare (Norwood & Lusk, 2011, pp. 227–9). (There may also be different effects on the environment, ignored here for simplicity.) Thus, it is also possible that reducing poverty, apart from making improvements on that front, may have indirect beneficial effects in the area of animal welfare. The third-best theory suggests that we should take into account the relevant important effects, consider their balance, and make adjustments according to the available information. If we do not have sufficient information to suggest that reducing poverty has net positive or negative effects on other areas, we should still proceed to reduce poverty. This follows the first-best rules under Informational Poverty. The second-best theory is correct in saying that, in so doing we may actually be reducing overall welfare. However, this is counteracted by the third-best theory, which correctly says that we are at least maximizing expected overall welfare.

3. Towards a more formal analysis

Altruism is to do good unto others. We may call the good of others their welfare. Personally, I regard my welfare as my net happiness, but a more general conception of welfare may be accommodated.⁷ Altruism should also not be narrowly focused, and at least in the ideal morality perspective, should best be extended to cover the welfare of all sentients. Ideal morality may require treating the welfare of animals as on par with that of humans. Recognizing the practical difficulties of reaching this ideal level, we provide only for some concern with animal welfare, without the requirement of parity. Similarly, altruism should not be myopic, but should take full account of welfare values in the future (discounted only by the uncertainty of their realization). However, for simplicity, we will use a largely a-temporal analysis. (The relevant welfare values may be taken as the expected values of all future welfare levels.) For simplicity, we abstract away the existence of different altruists with different objective functions. We also ignore the possible divergence between the objective function of the single altruist we focus on and that of the ideal moral one (that we refer to as overall welfare above). In other words, we assume for simplicity, that the altruist is already maximizing overall welfare. Our problem of the proper treatment of indirect effects of altruistic acts already exists in this simplified framework. We may then write the objective function of an altruist as

$$W = W(W^s, W^o, W^a) \quad (1)$$

Where W is the level of overall welfare, W^s is the welfare level of the altruist herself (s for self), W^o is the welfare of all other (human) individuals (lumped together for simplicity), and W^a is the welfare level of animals. Arguably, W should just be an un-weighted sum of its three elements (utilitarianism generalized to include animal welfare). However, the generalization to just a welfarist one above (which does not rule out utilitarianism) is sufficient for our purpose here.

Focusing on variables that the relevant individuals may choose and given other exogenous variables not focused upon, we have, in general, the welfare levels of the altruist, other individuals, and animals, depending on the variables chosen by all individuals (animal behavior is taken as one of the given exogenous parameters, not as choice variables)

$$W^i = W^i(x_1^s, x_2^s, \dots, x_m^s; x_1^o, x_2^o, \dots, x_n^o); i = s, o, a. \quad (2)$$

While the welfare of an individual depends, in general, on the variables chosen by both herself and that of others, each (either s or o) can only choose variables under her own control; i.e. the altruist s chooses only $x_1^s, x_2^s, \dots, x_m^s$, and other individuals o choose only $x_1^o, x_2^o, \dots, x_n^o$, where m and n are the numbers of choice variables of s and o respectively, which may be equal or not equal to each other.

Given resources and technology, feasible values of the choice variables are subject to an overall constraint written in implicit form

$$F(x_1^s, x_2^s, \dots, x_m^s; x_1^o, x_2^o, \dots, x_n^o) = 0 \quad (3)$$

On top of this overall constraint, there may be additional individual specific constraints which we shall ignore. This simplification does not falsify the main points about indirect effects.

If all variables were chosen to maximize W in (1) subject to (3), we would already get the first best: overall welfare maximization subject to feasibility. We would not need altruism further. Typically, individuals o maximize W^o in (2) instead, not adequately taking

⁷ A defence of my welfarist and utilitarian position has also been provided elsewhere (Ng, 1975; 1990, 2019a, Appendix B).

into account their choices on the welfare levels of others (including animal welfare), leaving scope for the altruists to choose $x_1^s, x_2^s, \dots, x_m^s$ not just to maximize W^s but W instead.

Consider a certain action by s called x_p^s aimed at say, poverty reduction, like a donation to help the poor. Obviously, for W maximization, s should not just consider the effect of x_p^s on W^s and W^o , but should also consider its effect on W^a , if any. The effects on W^s (like consumption reduction effect and warm-glow effect) and on W^o (like the consumption increasing effect) may be regarded as the primary effects of the donation; any effects of x_p^s on W^a , if any, may be regarded as the secondary effects. However, they are still classified as the direct effects of x_p^s , since x_p^s appears as an element directly affecting W^a (in Eq. 2 above), though these effects may be zero or negligible. The indirect effects arise from the possible effects of x_p^s on other choice variables and the resulting changes in these latter variables have on W through their effects on the elements of W , other than the direct effects of x_p^s . In the example of a donation to increase the consumption of the poor, this increase in consumption is a direct effect. However, if apart from the general increase in consumption, there is a shift in composition, like an increase in consumption of chicken mentioned above that decreases animal welfare, this is an indirect effect. As discussed above, if there is already full optimization with respect to chicken, the increase in chicken consumption, does not cause inefficiency. If chicken farming is already regulated to ensure no overall net suffering, higher chicken consumption is not a problem. If net suffering still persists but the production or consumption of chicken is taxed fully to reflect this external cost, the higher chicken consumption still leads to lower animal welfare. However, the amount of extra taxes collected could be used to promote welfare either in the animal welfare area or elsewhere, like more environmental protection to offset the negative effect. If the original amount of tax on chicken is adequate, this offsetting will result in a zero or negligible net effect. We may summarize our discussion so far (and also that in the appendix) into the following proposition.

Proposition 1 : (a) Possibilities for altruism: When not all choice variables are at levels that already maximize overall welfare subject to feasibility (almost always true), there exist scopes for altruists to increase welfare further by taking certain actions, e.g. helping the poor when there is still excessive poverty/inequality.

(b) Second best: Unless all departures from overall welfare maximization are eliminated to achieve the maximum feasible overall welfare (first best), altruistic measures to make improvements in some areas may actually reduce overall welfare.

(c) Importance of taking into account indirect effects: It is therefore important to take into account not only the direct effects of altruistic acts but also the indirect effects, especially those effects on areas of serious inadequate optimization, like environmental disruption and animal suffering.

(d) Third best: Even if neither the achievement of the overall first best (eliminating departures from welfare maximization in all areas) nor the satisfaction of the second best (taking account of all interrelationships and indirect effects) are possible due to imperfect information and administrative costs, such that a definite improvement may not be ensured (second-best impossibility of piece-meal welfare policies), the government and/or altruists may still do something to increase expected welfare by focusing on areas of major inadequacy from full optimization (like poverty, environmental protection, and animal welfare), taking into account both direct and indirect effects where we have the relevant information, or can obtain without prohibitive costs. In particular, when we do not have enough information to evaluate whether the indirect effects are positive or negative, we may proceed in accordance with the direct effects alone (adopting first-best rules in a third-best world with Informational Poverty). If we have some information to estimate the indirect effects (Informational Scarcity), we make some adjustments accordingly.

4. Some priorities for effective altruism

I mention several times above that poverty, environmental protection, and animal welfare are some areas of serious inadequate optimization and hence are likely the priority focal areas for effective altruists and futures researchers. (For some other priority areas, especially those possibly involving global extinction, see [Torres, 2017](#), [Ng, 2019b](#), [Ord, 2020](#), and references mentioned at the second half of the fourth last paragraph in Section 1 above.) Here, I offer some brief discussion explaining why. Our discussion below is partly guided by the principles outlined above on how to deal with the indirect effects.

Serious poverty is an area of traditional concern, not only by altruists but also by governments. The importance of reducing serious poverty is obvious, especially in our world of relative abundance. Though global inequality and poverty have decreased significantly in the last four decades or so (mainly due to the faster growth of China, India, and other countries originally of low income levels), inequality within each country has increased significantly over the same time in most countries, making the issue of inequality important.

Secondly, while many governments are active in reducing inequality, as mentioned above, government actions here have the limitation of high administrative costs and inducing disincentive effects. Due to these extra costs, even if the government has already optimized fully given the constraints, there are still scopes left for altruists to further pursue poverty reduction and equality promotion, especially if some warm-glow effects are likely to be important here.

Thirdly, in contrast to the traditional focus on the equality-efficiency/growth tradeoff (the disincentive effects mentioned above in particular, see, 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, Chakraborty, & Ai, 2017](#)). Partly, equality contributes to social harmony, reducing crimes, softening the opposition to the introduction of new technologies, etc. and hence promotes productivity and growth. Moreover, formerly when physical capital was important, 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 gain higher human capital.

Also, ‘...economic historians such as Lindert ... have shown the complex, but on the whole favorable, effect of the Welfare State on economic performance to the point where the full economic benefits and costs may have been roughly equal, making the Welfare

State a “free lunch” (Mokyr, 2014, p.191). This suggests that more equality-improving welfare spending may be welfare improving, since equality also contributes to welfare more directly. In addition, when the importance of relative income is considered, progressivity may also be more welfare-improving; see, e.g. Bruce and Peng (2018), Yan et al. (forthcoming).

In the long run, arguably much more important than poverty reduction is the prevention (at least in reducing the probabilities) of global extinction, or existential risks (x-risks for short). If we become extinct, we lose all potential future welfare values. On the other hand, if we manage to avoid extinction, the advancement in science, technology, economy, and morality will likely lead us to flourishing lives of enormous welfare values beyond our present imagination. While there are many possible sources of x-risks, one important one is the possible catastrophic climate change (CCC), large enough as to threaten survival through cumulating, cascading, multiplicative effects with critical thresholds and tipping points (e.g. Van Aalst, 2006; Buldyrev et al., 2010; Hansen & Sato, 2012; Intergovernmental Panel on Climate Change, 2014; Schuur et al., 2015; Barnosky & Hadly, 2016; Sims & Finnoff, 2016; Belaia, Funke, & Schneider, 2017; Grainger, 2017; Osmond & Klausmeier, 2017; Rothman, 2017; Kareiva & Carranza, 2018). As this has been discussed much elsewhere, I will just briefly mention two points.

Environmental protection including the control (preferably using taxation) of greenhouse gas emissions will likely generate very high benefits even in the intermediate term; it should be emphasized even if global extinction through CCC may not be highly likely. Secondly, the emphasis on environmental protection in general and the prevention of CCC in particular is unlikely to have indirect negative effects. Hence, putting this as one of our priorities is justified. (See also Ng, 2016a, 2019b.)

Raising the reduction of animal suffering as a priority is based on two considerations. First, much of animal suffering, especially those parts that we impose upon them, could be drastically reduced at little or even negative costs on humans (Ng, 2016b for more details). Such practices like whipping horses in horse racing and cutting live eels in half (to attract the attention of customers and showing the freshness of the fish, as practiced in the wet markets of Hong Kong) serve only relative competition without important significance. Banning such practices would impose negligible if any costs on humans at the social level. Also, as most people in rich countries tend to eat too much meat for the benefits of their own health, imposing more serious farm animal protection measures (like tripling cage sizes for farm chicken or taxing suffering imposed on them) will likely also make humans better off, generating negative costs. This is similar to taxes on cigarettes that increase their prices and actually make smokers healthier and happier (Gruber & Mullainathan, 2005).

Secondly, for these parts of animal suffering imposed by us on them, we have a particularly important moral obligation to reduce if not completely eliminate them. We do not have to be Kantians to agree on this. Decreasing suffering imposed on animals even at some costs on ourselves is morally good. Promoting such measures should thus lead to an increase in general morality, allowing us humans to benefit as well. Thus, they are important even if purely from a consequentialist perspective.

Measures that reduce farmed animal suffering enormously at little or even negative costs are certainly compelling. However, partly due to the point of the previous paragraph, 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). In the long run, it is likely counter-productive even purely from the perspective of animal welfare alone, to impose strict rules that impede scientific and technological advances.

5. Concluding remark

In the real world, there are many areas (e.g. poverty reduction, environmental protection, and animal welfare) where appropriate measures to increase overall welfare in the future have not been done sufficiently, leaving many scopes for altruists, futures researchers, and governments to make contributions. However, it is impossible to eliminate all departures from optimality soon, if ever. From the second-best theory, this imperfect elimination of all departures suggests that any particular measures in some areas, even if the direct effects are effective and positive, may actually decrease the overall welfare, through the complicated interrelationships of the indirect effects. From this valid point, the second-best theory goes to the extreme of declaring the impossibility/undesirability of piecemeal welfare policies. The third-best theory accepts the importance of taking into account the interrelationships through the indirect effects. However, due to informational and administrative costs, it is impossible to take all these into account to achieve either the first best or the second best. Instead of accepting the impossibility of ensuring an improvement and do nothing, the third-best theory shows that we may at least increase our expected welfare by following the best policies given the available information. Where we know too little to evaluate just the direction/sign of the indirect effects (Informational Poverty), we may proceed in accordance with the direct effects only (first-best rules). Where we have more information (Informational Scarcity), we make some adjustments accordingly (third-best rules).

It is also argued that, in evaluating the indirect effects, we should mainly focus on those areas with inadequate overall maximization. These are areas that are typically the main focuses of effective altruists to begin with. This provides an important simplification in evaluating the indirect effects. Indirect effects may happen in thousands of different areas. If we have to examine all the indirect effects of each act/change/event, the informational costs involved may be very large. Since indirect effects in areas without significant departures from optimization may be largely ignored, the costs of evaluating indirect effects are greatly reduced. We may thus perform effective altruism more effectively.

A reviewer of an earlier draft of this paper comments that, like other welfare analyses, this paper assumes that ‘politics and technology’ are fixed. While acknowledging the importance of politics and technology, the non-explicit analysis of them is for simplicity of focusing on the relevant issues. Where information allows, our third-best policies do not preclude considerations of the

indirect effects through changes in politics and technology.

6. Mathematical appendix

The maximization of W in (1) subject to (3), we have the Lagrangean

$$L = W(W^s, W^o, W^a) - \lambda F(x_1^s, x_2^s, \dots, x_m^s; x_1^o, x_2^o, \dots, x_n^o) \quad (4)$$

where the various W^i are as given in (2).

The first-order conditions are:

$$(\partial W / \partial W^s)(\partial W^s / \partial x_i^s) + (\partial W / \partial W^o)(\partial W^o / \partial x_i^o) + (\partial W / \partial W^a)(\partial W^a / \partial x_i^s) = \lambda (\partial F / \partial x_i^s); i = 1, \dots, m. \quad (5)$$

The altruists can only choose variables under her own control $x_1^s, x_2^s, \dots, x_m^s$, and not variables under the control of others $x_1^o, x_2^o, \dots, x_n^o$. Thus, the choice of $x_1^s, x_2^s, \dots, x_m^s$ satisfying (5) typically does not yet give us a maximum W , even if second-order conditions are assumed satisfied. This is true even if we introduce a government to choose all variables $x_1^s, x_2^s, \dots, x_m^s; x_1^o, x_2^o, \dots, x_n^o$. This is so because, among other deficiencies, the government typically does not take adequate account of animal welfare W^a , at least not in the eyes of altruists, or in accordance to (1). This means that the first best (highest feasible W) cannot be attained, and some second-best departures exist. Taking these departures as constraints in addition to the general constraint (3) gives rise to a second-best problem. These second-best constraints may be in the form of:

$$(\partial W / \partial W^s)(\partial W^s / \partial x_j^o) + (\partial W / \partial W^o)(\partial W^o / \partial x_j^o) + (\partial W / \partial W^a)(\partial W^a / \partial x_j^o) = k_j \lambda (\partial F / \partial x_j^s); k_j \neq 1 \text{ for some } j\text{'s}; \text{ the constants } k_j \text{ may differ over } j. \quad (6)$$

Maximizing W in (1) subject to (3) and (6) gives rise to a more complicated Lagrangean than (4) and to second-best first-order conditions more complicated than (5). From here, the results of both the second-best and third-best theories (Lipsey & Lancaster, 1956; Ng, 2017a,b) apply to this more complicated setting for the same reasons, resulting in Proposition 1 in the text above.

Acknowledgment

I acknowledge the helpful comments of Gustav Alexandrie and two anonymous reviewers of this journal.

References

- Ahvenharju, S., Minkinen, M., & Lalot, F. (2018). The five dimensions of futures consciousness. *Futures*, 104, 1–13.
- Barnosky, A. D., & Hadly, E. A. (2016). *Tipping point for planet earth: How close are we to the edge?* Thomas Dunne Books.
- Belaia, M., Funke, M., & Schneider, U. (2017). *Integrated assessment of climate tipping points. Doctoral dissertation.* Universität Hamburg.
- Berkey, B. (2018). The institutional critique of effective altruism. *Utilitas*, 30(2), 143–171.
- Biswas, S., Chakraborty, I., & Ai, R. (2017). Income inequality, tax policy, and economic growth. *The Economic Journal*, 127(601), 688–727.
- Broad, G. M. (2018). Effective animal advocacy: Effective altruism, the social economy, and the animal protection movement. *Agriculture and Human Values*, 35(4), 777–789.
- Bruce, D., & Peng, L. (2018). Optimal taxation in the presence of income-dependent relative income effects. *Social Choice and Welfare*, 1–23. <https://doi.org/10.1007/s00355-018-1118-4>.
- Buldyrev, S. V., et al. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291), 1025–1028. <https://doi.org/10.1038/nature08932>.
- Dietz, A. (2019). Effective altruism and collective obligations. *Utilitas*, 31(1), 106–115.
- Van Dorsser, Cornelis, W. E. W., Taneja, P., & Marchau, V. A. W. J. (2018). Improving the link between the futures field and policymaking. *Futures*, 104, 75–84.
- Gabriel, I. (2017). Effective altruism and its critics. *Journal of Applied Philosophy*, 34(4), 457–473.
- Grainger, A. (2017). The prospect of global environmental relativities after an Anthropocene tipping point. *Forest Policy and Economics*, 79, 36–49. <https://doi.org/10.1016/j.forpol.2017.01.008>.
- Greaves, H. (2016). XIV—Cluelessness. *Proceedings of the Aristotelian Society*, 116, (pp. 311–339).
- Gruber, J., & Mullainathan, S. (2005). Do cigarette taxes make smokers happier? *The BE Journal of Economic Analysis & Policy*, 5(1), 1–43.
- Hansen, J. E., & Sato, M. (2012). *Paleoclimate implications for human-made climate change. Climate change.* Vienna: Springer 21–47.
- Intergovernmental Panel on Climate Change (2014). *Climate change.* Intergovernmental Panel on Climate Change.
- Kareiva, P., & Carranza, V. (2018). Existential risk due to ecosystem collapse: Nature strikes back. *Futures*.
- Lipsey, R. G. (2017). Generality versus context specificity: First, second and third best in theory and policy. *Pacific Economic Review*, 22(2), 167–177.
- Lipsey, R. G., & Lancaster, K. (1956). The general theory of second best. *The Review of Economic Studies*, 24, 11–32.
- MacAskill, W. (2015). *Doing good better: How effective altruism can help you make a difference.* London: Guardian.
- MacAskill, W. (2018). *Understanding effective altruism and its challenges. The Palgrave Handbook of Philosophy and Public Policy.* Cham: Palgrave Macmillan 441–453.
- Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing when being good frees us to be bad. *Social and Personality Psychology Compass*, 4(5), 344–357.
- Milanovic, B. (2011). More or less. *Finance & Development*, 48(3), 6–11.
- Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38, 175–208.
- Mokyr, J. (2014). A flourishing economist. *Journal of Economic Literature*, 52(1), 189–196.
- Ng, Y.-K. (1975). Bentham or Bergson? Finite sensibility, utility functions and social welfare functions. *Review of Economic Studies*, 42(4), 545–569.
- Ng, Y.-K. (1984). Quasi-Pareto social improvements. *American Economic Review*, 74(5), 1033–1050.
- Ng, Y.-K. (1990). Welfarism and utilitarianism: A rehabilitation. *Utilitas*, 2(2), 171–193.
- Ng, Y.-K. (1995). Towards welfare biology: Evolutionary economics of animal consciousness and suffering. *Biology and Philosophy*, 10, 255–285.
- Ng, Y.-K. (2014). Why is finance important? Some thoughts on post-crisis economics. *Singapore Economic Review*, 59(5), 1450037.
- Ng, Y.-K. (2015). "Welfare economics", In: James D. Wright (editor-in-chief), *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition, 2015, Vol 25. Oxford: Elsevier. pp. 497–503.
- Ng, Y.-K. (2016a). The importance of global extinction in climate change policy. *Global Policy*, 7(3), 315–322. <https://doi.org/10.1111/1758-5899.12318>.
- Ng, Y.-K. (2016b). How welfare biology and commonsense may help to reduce animal suffering. *Animal Sentience*, 007.

- Ng, Y.-K. (2017a). "Towards a theory of third best", (reprinted from Public Finance 1977). *Pacific Economic Review*, 22(2), (2017)).
- Ng, Y.-K. (2017b). Theory of third best: How to interpret and apply. *Pacific Economic Review*, 22(2), 178–188.
- Ng, Y.-K. (2019a). *Marteks and Morals*. Cambridge University Press.
- Ng, Y.-K. (2019b). "KEYNOTE: Global extinction and animal welfare: Two priorities for effective altruism" (Atkinson Memorial Lecture at Oxford University 2018). *Global Policy*, 10(2), 258–266. <https://doi.org/10.1111/1758-5899.12647>.
- Norwood, F. B., & Lusk, J. (2011). *Compassion, by the pounds: The economics of farm animal welfare*. New York: Oxford University Press.
- Okun, Arthur M. (1975/2015). *Equality and Efficiency: The Big Tradeoff*. Washington: Brookings Institution.
- Ord, T. (2020). *The Precipice: Existential Risk and the Future of Humanity*. London: Bloomsbury.
- Osmond, M. M., & Klausmeier, C. A. (2017). An evolutionary tipping point in a changing environment. *Evolution*, 71(12), 2930–2941. <https://doi.org/10.1111/evo.13374>.
- Pummer, T. & W. MacAskill (forthcoming). "Effective Altruism". In Hugh LaFollette (ed.), *International Encyclopedia of Ethics*. Available here: <https://philpapers.org/archive/PUMEA>.
- Rothman, D. H. (2017). Thresholds of catastrophe in the Earth system. *Science Advances*, 3(9), 1–12. <https://doi.org/10.1126/sciadv.1700906>.
- Schuur, E. A. G., et al. (2015). Climate change and the permafrost carbon feedback. *Nature*, 520(7546), 171–179. <https://doi.org/10.1038/nature14338>.
- Sims, C., & Finnoff, D. (2016). Opposing irreversibilities and tipping point uncertainty. *Journal of the Association of Environmental and Resource Economists*, 3(4), 985–1022.
- Solar, P. (1995). Poor relief and English economic development before the industrial revolution. *The Economic History Review*, 48(1), 1–22.
- Stiglitz, J. J. (2012). *The price of inequality: How today's divided society endangers our future*. Norton.
- Torres, P. (2017). *Morality, foresight & human flourishing: An introduction to existential risks*. Durham, North Carolina: Pitchstone.
- Van Aalst, M. K. (2006). The impacts of climate change on the risk of natural disasters. *Disasters*, 30(1), 5–18. <https://doi.org/10.1111/j.1467-9523.2006.00303.x>.
- Yan, Eric, Qu Feng & Yew-Kwang Ng (forthcoming). Do we need Ramsey taxation? our existing taxes are largely corrective, *Economic Modelling*.