

Longtermism and the Complaints of Future People

Emma J. Curran

Abstract. In this chapter, I explore a contractualist assessment of longtermism, and, in particular, the claim that we have, in most cases, a moral obligation to invest in those interventions which most improve the prospects of future people. I argue that contractualism will judge the claims that long-term interventions generate in their favour as being extremely weak, due to the uncertainty we have in controlling the very far future. As such, long-term interventions will look uncompetitive in prioritisation decisions. I then defend contractualism from a *reductio* which could be generated using my argument. In particular, I explain how contractualism can still favour short-term interventions which look seemingly analogous to unjustifiable long-term interventions, for example natural disaster preparation schemes. Finally, I highlight the prioritisation implications that contractualism might have for the longtermism movement.

Keywords. Contractualism, longtermism, risk, complaints, individuals

We are not particularly good at reliably controlling the future. Perhaps, when looking at a horizon of a few decades, or hundred years, we can make reasonable estimates about the effects of our actions. Yet, when considering the very far future – a future stretching hundreds of thousands of years ahead of us, if not more – there are deep epistemological challenges preventing us from having confidence about the impact of our present-day actions (Tarsney 2022).

Despite the uncertainty surrounding our ability to shape the very far future, *longtermists* believe we have good moral reason to invest resources into attempting to do so. Indeed, generally speaking, longtermists believe that we have *overwhelming* moral reason to try to positively shape the future. The longtermist argument is quite simple: even if anything we can do has only a very tiny chance of improving the far-future, given the massive number of people who will likely exist throughout the future, in expectation, the amount of good or value we will

bring about will be massive.¹ In fact, long-term interventions likely bring about many orders of magnitude more expected good than available short-term options. As such, according to Hilary Greaves and Will MacAskill, in many decision situations, the option that most improves the prospects of those living in the very far future will also be the option which brings about the most *good*, in expectation (2021: 3-4).² Following Greaves and MacAskill, I will call this claim “*axiological longtermism*”.

Now, if you’re the sort of agent who simply cares about bringing about the most good, then insofar as the argument sketched above is sound, you have decisive reason to divert your resources into improving the far-future. Yet, many people don’t simply care about bringing about the most good or value. Many people care about discharging their moral obligations. As such, many people don’t simply want to know if axiological longtermism is true, they want to know if this gives rise to a moral obligation to invest in long-term interventions. They care about the truth of “*deontic longtermism*”, the claim – very roughly put – that: in many decision instances, we have a moral obligation to choose the option that most improves the prospects of those living in the far future.

In this chapter, I aim to explore a contractualist assessment of the veracity of deontic longtermism, along with gesture at what form of longtermism is compatible with contractualism. However, before doing so, I must make some qualifications. When it comes to discussions of the complaints or claims of future persons, an initial, and seemingly unavoidable, hurdle is the *non-identity problem*. The non-identity problem, put simply, is that future people seemingly cannot have claims to our assistance grounded in their well-being insofar as such assistance would change the identity of future populations and they, without our assistance, would have lives worth living (Parfit 1982; 1987: ch.16). For simplicity, a detailed discussion of the non-identity problem and contractualism is absent from this chapter

1 Many think there is something problematic about allowing tiny probabilities of giant payoffs to dictate our actions. See Bostrom (2009), Balfour (2021), Wilkinson (2022), and Kosonen (ms.) for a discussion of the plausibility of ‘fanatical’ decision theories which give rise to this result.

2 See Thorstad (2021) for a discussion of the scope of cases, or decision situations, in which this claim might hold true.

A second qualification is that throughout this chapter, I will assume, when assessing long-term interventions, that the population size remains fixed across the future. Most interventions which shape the far-future will have an impact, in some way, on the size of the number of people in the future. However, to take variable population sizes into account when assessing such interventions is to enter debates about the value of bringing people into existence, and also whether or not people have claims to be brought into existence, both of which would take me far beyond the scope that this chapter permits. A complete contractualist account of the ethics of longtermism will require an account of variable population sizes. However, I will, likewise, abstain from such a discussion. I will return to these qualifications in §2, explaining why I do not think either significantly undermines the argument presented in this chapter.

I will proceed, in §1, to highlight an argument which I have developed elsewhere (Curran 2022.) The argument is to the effect that contractualism, and complaints-based theorising in general, assesses many long-term interventions as generating extremely weak claims in their favour. As such, rather than being obligatory, investing in long-term interventions will, in many cases, be impermissible on a contractualist assessment. In §2, I will discuss the significance of the qualifications with which I have presented this argument. In §3 I will respond to a significant objection to the argument of §1, which suggests that rather than being a critique of longtermism, such an argument acts as a *reductio* against longtermism. Finally, in §4, I will briefly outline the prioritisation recommendation that contractualism might have.

1. An argument against deontic longtermism

Scanlonian contractualism tells us that an action is right iff it is permitted by a principle to which no one could reasonably reject. In unpacking what it means for an individual to “reasonably reject”, contractualism makes use of two types of restrictions. The first is the “personal-reasons restriction” (Scanlon 1998: 218-23; Kumar 2003a). Individuals can only legitimately reject principles on the basis of its impact on them, as opposed to “impersonal” concerns such as its impact

of the sum, or distribution, of well-being in the world. For simplicity's sake, I will say that, in a decision situation, an individual has a complaint if her well-being would be higher under one of the available options. The strength of an individual's complaint against a policy is a function of the impact the policy would have upon her wellbeing.

The second restriction contractualism contains is the "individualist restriction". In decision situations in which all options have complaints against them, agents cannot consider the aggregate complaint against any option. Rather, each complaint must be considered individually, and go through a process of pairwise comparison against the complaints with which it competes. Complaints "compete" against each other insofar as they are mutually unsatisfiable.

Presented with two options, the first to save one-hundred people from a headache, and the second to save one other person from death, contractualism tells us that the agent is not permitted to consider the weight of all one-hundred complaints of a headache together. Rather, each complaint of a headache must be compared individually against the complaint of death. The individualist restriction is designed precisely to avoid counterintuitive verdicts in which we are obligated to save a large number of people from a trivial harm – like headaches – instead of a small number of people from a significant harm – like death.³

How might contractualism assess deontic longtermism? I ask you to consider the following toy case:

Biotech Risk: Topsy wants to donate her savings to a good cause. With the money available to her, she could fund highly-successful, life-saving surgeries for ten patients, who would otherwise die (a). Topsy could also put her money into biotechnology research governance (b). Topsy knows that future biotechnology research carries with it spectacular risks, indeed *catastrophic risks*.

³ A classic example of this type that Scanlon used to motivate the individualist restriction was the "Transmitter Room" case (1998: 235).

Topsy has been advised that a country of one-hundred-million people are currently at a two-in-a-million risk of a fatal biotech-related catastrophic event happening within their lifetimes. She has also been told that her donation will reduce the risk to five-in-twenty-million.⁴

It should be fairly clear that (b) is representative of a great number of long-term interventions, and, in particular, those which the argument for axiological longtermism seems to rest upon; it brings about a lot of good in expectation – indeed, saving 175 lives – but does so by improving the prospects of a large group by a very tiny amount. (a), on the other hand, looks like, at least *some* of, the short-term interventions available to us. It improves the prospects of a small number of people by a great extent, though by doing so, brings about less good in expectation – it will only save 10 lives, compared to the 175 lives the long-term intervention expects to save.

What does contractualism tell Topsy to do in this case? This depends on which account of how to assess complaints under risk that it accompanies. There are two options discussed in the literature: the *ex-ante* interpretation and the *ex-post* interpretation. A traditional formulation of the difference between these two perspectives is with regard to the good under distribution; the *ex-ante* perspective is concerned with the distribution of individual prospects, whilst the *ex-post* perspective is concerned with the pattern of *ex-post* outcomes.

An individual's *ex-ante* complaint against a policy is simply the difference between their expected wellbeing given the policy was implemented and their expected wellbeing given the policy wasn't implemented. *Ex-post* complaints, on the other hand, focus on the likelihood that someone – not that any particular person – will incur a harm or receive a benefit. The number of *ex-post* complaints against a policy will be the number of people who could be harmed by the policy, with each complaint discounted by the difference in the improbability that the policy made to their incurring this harm. To illustrate, consider:

⁴ Adapted from a case in Curran (2022: 2).

Side Effects. There are 1,000,000 people at risk of developing a fatal disease, and we know that if untreated exactly one person will die – though it is impossible to determine who. We can treat the group with a pill which will eliminate each member’s risk of death. The pill causes nausea. The cost of nausea is judged to be greater than the benefit of having one’s risk of the disease reduced from one-in-a-million to nought, but the cost of actually having the disease is far greater than the cost of nausea.

What are the claims *ex-ante*? Well, the beneficiaries in question are each individual member of the group of 1,000,000, each of which has an interest in not having the pill; their expected wellbeing given they take the pill is lower than their expected wellbeing given they do not. So, *ex-ante*, we have 1,000,000 claims against the policy being implemented, each of which is proportional to how much worse off they expect to be made by having the pill.

Ex-post, we can compare the complaints of the patients given the intervention is administered and given it is not administered. If the intervention is not administered, it generates one complaint of certain death. If the intervention is implemented, it generates 1,000,000 complaints of nausea.

Side Effects should also make clear that the distinction between *ex-ante* and *ex-post* is not an ethically unimportant one. Indeed, in Side Effects, the *ex-ante* and *ex-post* perspective both capture important and seemingly contradictory considerations. On one way of looking at the world, the intervention is in no one’s interests, but from another perspective the interventions is of fundamental importance to someone’s interests – the one who would die if not treated.⁵

To further clarify the evaluation of *ex-post* complaints, consider a variant of Side Effects in which, in expectation, one person will die if not treated, but the risk is independent. In this case, as potentially all 1,000,000 individuals could

⁵ As such, the intervention in Side Effects is an example of what Stephen John calls an ‘absolute prevention paradox’: an intervention which is, in the *ex-ante* sense, in no one’s interest, yet in the *ex-post* sense, in the interests of some (2014: 30-34).

contract the disease, there are 1,000,000 complaints of death against not administering the intervention. However, each complaint is discounted at a different rate, following a binomial distribution. That is, the first is discounted by the improbability that one will die, the second is discounted by the improbability that a second would die, and so on. Once taken all together, these claims sum to approximately one claim of death, with the first claim being only a slightly discounted complaint of death.

Now, what would *ex-ante* contractualism say about Biotech Risk? Well, the complaint against (b) – the long-term analogue – is ten complaints of, near certain, death. On the other hand, (a) generates 100,000,000 complaints of a seven-in-four-millionth risk of death. Going through a process of pairwise comparison, clearly the complaint of near death will outweigh each of the 100,000,000 complaints of a tiny risk of death.

The most important upshot is that, given the fact that long-term interventions can only improve the prospects of each future person by a very tiny amount, on an *ex-ante* account, they will generate extremely weak individual claims in their favour. Indeed, much weaker than the claims generated by some available short-term interventions, and as such, we should favour such short-term interventions.

Now, there is some debate as to whether there is a genuine competition between long-term and short-term interventions; perhaps it's the case that what is beneficial in the short-term is also beneficial in the long-term. Regardless, the argument of this section demonstrates that – at least on an *ex-ante* contractualist account – if we were to engage in long-term intervention, it would be because of the strong claim generated by a correlative short-term intervention. Not because of the value brought about by the long-term intervention itself.

Next, let us consider an *ex-post* contractualist assessment of Biotech Risk. It is tempting to think interventions like (b) in Biotech Risk are easily justified *ex-post* by appealing to those people who we expect to bestow a good upon by engaging in the intervention, whomever they may be. That is, we could appeal to the *ex-post* complaints of the 175 individuals we expect to save from death (or more precisely, the 100,000,000 complaints of death discounted following a binomial

distribution, equivalent to roughly 175 complaints of death.)⁶ Contractualism permits, if not obligates, us to save the greatest number – even if it doesn't allow aggregating complaints (Kamm 1993: 101-119; Scanlon 1998: 232-33). As such, one might think contractualism would permit, if not obligate, us to engage in the long-term governance intervention, (b), in Biotech Risk.

This line of thought, however, is mistaken. The risk which the governance option is seeking to mitigate is not independent; either a biotechnological disaster occurs, killing or harming whomever it would, or it doesn't occur. And this is true of a great many of the risks longtermists are trying to mitigate, such as the risks of unaligned AI, climate change, supervolcano eruptions, or asteroid impacts. Indeed, each individual's risk of being harmed by each of these catastrophes is deeply connected. Either the supervolcano erupts, thereby killing however many people, or it doesn't erupt, meaning no one is killed by it. Likewise, either an asteroid hits or it doesn't, and AI goes rogue, or it does not.

The dependent nature of the risks facing each individual in catastrophe scenarios, including is Biotech Risk, fundamentally changes the nature of the complaints generated. In Biotech Risk, it wouldn't be true to say we *reasonably expect* to save 175 lives by engaging in (b). This is because there is no state-of-affairs in which 175 people die from the biotechnological disaster. Either no one dies, or all 100,000,000 dies. The governance intervention simply reduces the likelihood of such an eventuality. As such, *ex-post*, we don't have 175 complaints of death, but rather 100,000,000 complaints of death – as all 100,000,000 people could potential die. However, each complaint is discounted by the difference in the improbability that the intervention made to the harm actualising. That is, each is discounted by 99.999825%.

It should be clear that once we properly conceptualise the nature of the claims generated by many long-term interventions, even from the *ex-post* perspective

⁶ There is some uncertainty about how contractualism should view these 100,000,000 complaints. Given the individualist restriction, the contractualist cannot consider the aggregate sum of the 100,000,00 complaints, and therefore shouldn't – strictly speaking – consider 175 undiscounted complaints. By the law of large numbers, the 175 initial complaints should be discounted to a very small amount, which might make it appropriate to talk of 175 almost undiscounted complaints. However, they are importantly all at least slightly discounted, and as such won't be able to outweigh certain claims.

they are extremely weak. Certainly not strong enough to compete with available short-term interventions, and certainly not strong enough to “dictate” how we ought to distribute our resources.

So, it seems that so long as contractualists maintain the personal reasons restriction, on both the *ex-ante* and *ex-post* perspective, long-term intervention (in many cases) will generate extremely weak individual claims in their favour. And, if the contractualist also maintains their individualist restriction, these weak individual claims will unlikely be decisive when considering the various options available to us to help others.⁷ This is generalisable to all complaints-based moral theories; insofar as they limit claims to personal welfare reasons, and also do not permit the aggregation of such claims, then the claims that long-term interventions will generate will be extremely weak.

2. *Qualifications*

At this point, One might be thinking that for all I have that said, my argument may not speak to the reality of longtermism and long-term interventions as it’s failed to consider the non-identity problem and interventions which change the size of future populations. Having outlined my account of the contractualist assessment of deontic longtermism, I am now in a position to address these concerns.

As originally presented by Derek Parfit, the non-identity problem is supposed to be a problem for person-affecting accounts of morality – including contractualism. The purported problem is that as future people cannot, in many cases, generate claims on the grounds of their wellbeing, then person-affecting moral theories cannot account for the obligations we intuitively owe to future people. To avoid rejecting person-affecting moral theories, there are, broadly speaking, two routes. First, one can bite the proverbial bullet, accepting a significantly diminished realm of moral duties to future people. Second, one can

⁷ In fact, in other work I have demonstrated that even if complaint-based moral theorists weaken the individualist restriction – for example to permit limited or partial aggregation (Voorhoeve, 2014; 2017) – long-term interventions will still look extremely uncompetitive on both an *ex-ante* and *ex-post* perspective (Curran 2022).

give an account of how future people generate claims, which avoids non-identity concerns (Hare 2007; Kumar 2003b).⁸

So a contractualist assessment of deontic longtermism will likely have a background assumption of one of these positions. If we were to assume that, in most cases, future people cannot generate claims, then the argument in §1 seems somewhat redundant. On this reading, contractualism rather straightforwardly doesn't value long-term interventions, except perhaps those which will prevent people from coming into existence with lives not worth living.

Rather, the contractualist response to the non-identity problem which seems most favourable to longtermism will be to assume that there is some way of conceptualising the complaints of future-people which avoids non-identity concerns and allows them to stake claims in much the same way as you or I. Nonetheless, the argument forwarded in §1 shows that even granting this generous assumption in longtermism's favour, contractualists still have good reason to be suspicious of long-term intervention.

Let's now consider the second qualification I made, regarding the importance of variable population size. Perhaps, contractualists might recognise the importance of shaping the far-future not because of the importance of improving far-future people's wellbeing, but instead because of the importance of increasing the number of people in the future.

I am now in a position to explain why I do not believe that considerations of variable population sizes will have much impact on my argument. First, it is important to note that the relevant debate for the contractualist is not whether there is *value* in bringing people into existence, but, rather, whether people have a reasonable *claim* to being brought into existence. Whilst both are controversial, the latter is significantly more so, in part because of the obscene consequences of the view – for example plausibly obligating people to produce children if they don't have a complaint against doing so, even if they have no desire to have

⁸ This is, of course, a simplification. For one, there is a vast array of different account that could be offered of the nature of future people's claims, and each of these might have varying justifying strengths for long-term interventions.

children. Moreover, it is hard to imagine what it means for something that will never exist to have a complaint.

However, let's grant the rather strange view that people have a claim to being brought into existence. How, then, should a contractualist evaluate long-term interventions which, for example, might increase future population size? Ought we think that such interventions gain substantial support by satisfying more claims to being brought into existence? I think this is unlikely. First, it seems plausible that the same problems of uncertainty will plague interventions which aim to increase population as do those interventions considered in §1. Perhaps we can have a reasonable guess at what will improve population size in the next century, but, given the horizon of thousands of years, it's very hard to make an educated guess about the impact of any policies on population size. As such, the claims of future people to be brought into existence will look as weak *ex-ante* and *ex-post* as did the claims of future in in Biotech Risk.

3. *A reductio for contractualism*

For the committed contractualist, the argument sketched in §1 might be decisive reason to give up their longtermist activities. However, I'm sure to many readers, the argument has the distinct flavour of a *reductio*. This is especially true when we consider the further implications of these conclusions. On a contractualist account, regardless of whether you adopt an *ex-ante* or *ex-post* perspective, interventions analogous to Biotech Risk seem to be ruled out. This, in particular, might be too counterintuitive to countenance. Consider:

Natural Disaster: Above a city of 10,000,000 people there is a volcano which is at a 0.05% risk of exploding and killing the population of the city. We can reduce the risk of the population of the city dying, in the case of the volcano's eruption, to 0%, through an expensive engineering project. Elsewhere, in a faraway city, there is a patient Donald who has contracted an illness which will prove fatal unless we choose to give him an expensive treatment. Due to

our finite resources, we can only treat Donald or tend to the volcano risk, not both.

It seems sensible for a government to tend to such risks of natural disaster. In fact, it seems that governments are morally obligated to attend to such risks even if doing so means diverting some of our finite resources away from aiding individuals who are at high risk of a serious harm. However, the structure of Natural Disaster mirrors that of Biotech Risk, and, as we have seen, on both an *ex-ante* and *ex-post* perspective, contractualism compels us to choose the option to help Donald. It is tempting to think that, any minimally plausible moral theory which informs our priority setting ought to allow us to tend to natural disaster risk. As such, contractualism – and other similarly anti-aggregative complaint-based moral theories – fail to be even minimally plausible.

I agree that this is a hard conclusion for the contractualist to accept, and it certainly does not bode well for the practicability of contractualism if this is a genuine implication of the theory. To avoid such a conclusion, the contractualist needs to point to a disanalogy between Natural Disaster and Biotech Risk which will explain why we might be able to attend to natural disaster risks. I claim that such a disanalogy lies in the fact that we can reasonably expect to save a life (or prevent some harm) through mitigating natural disaster risks in a manner we can't when it comes to catastrophic risk. This opens the door to a potential *ex-post* justification of natural disaster mitigation schemes.

The first justification I want to offer operates on the level of *programmes* of interventions. That is, we can justify natural disaster interventions by evaluating larger decisions, such as general policy decisions, rather than specific interventions. Let's – for the moment – grant that each particular natural disaster intervention might look a lot like the governance intervention in Biotech Risk. As such, we will assume that we do not reasonably expect *any given* intervention to save any lives. Nonetheless, plausibly, when we engage in a programme of such interventions we really do expect to save lives at some point. This is, in part, due to the fact that natural disasters have a high antecedent

chance of actually occurring at some point and that programmes of mitigation can decisively lower the probability of fatal natural disaster events such that we expect to actually prevent the deaths of some people. This allows us to, *ex-post*, justify a programme of natural disaster risk mitigation by way of the significant claims of those people whose lives we expect to save, even if it might not be reasonable for us to expect each particular intervention in the programme to save a life.⁹

For illustrative purposes, this justification paints decisions to engage in natural disaster mitigation intervention as looking like:

Lottery: You can save one life or receive a single ticket for a lottery of N -tickets. If your ticket is drawn, 100,000 people, who otherwise would have died, will be saved.

Given an appropriately specified N ¹⁰, whilst the expected value of engaging in the lottery is larger than saving the single life, you don't reasonably expect that engaging in the lottery will have any impact on the outcome of the 100,000 people. So, *ex-post* contractualists will favour saving the one over attempting to save the 100,000 as each of their complaints of death will be discounted significantly. However, the decision to engage in a programme of natural disaster management would be analogous to:

⁹ This explanation turns around the fact that *ex-post* justifications do not decompose; on the *ex-post* perspective, the deontic status of a group of actions is different to the status of each individual sub-action. Whilst this is usually held as a troubling implication of *ex-post* theories (cf. Hare, 2016), it allows the *ex-post* theorist to avoid counter-intuitive verdicts in cases like Natural Disaster. Of course, for those already suspicious of *ex-post* contractualism due to these problems, such a solution will not be attractive. But for those who are sympathetic with *ex-post* contractualism, this provides a useful means of avoiding the reductio. Moreover, whilst many view decomposition as a problem for *ex-post* theories, there is some work suggesting that this failure in decomposition is a genuine feature of morally permissibility (see Wu (2021) and Kirkpatrick (2017).)

¹⁰ This will involve it being strictly smaller than 100,000 but larger than, for now let's arbitrarily say, 3. Each of the *ex-post* complaints of death will be discounted by $(N-1)/N$.

Grand Lottery: You can save X lives or receive one ticket for each of X lotteries of N -tickets. If your ticket for a lottery is drawn, then 100,000 people, who otherwise would have died, will be saved.¹¹

In this case, for some X and N , you can reasonably expect that by engaging in the grand lottery scheme, you will save the 100,000 people who otherwise would die. As X gets closer to N , the probability that you will save all 100,000 people will increase, meaning the *ex-post* complaints of death in favour of saving the 100,000 will be discounted to a lesser degree. Eventually these *ex-post* complaints will be discounted to a small enough extent that they will be sufficiently near certain to allow them not to be outweighed individually by the one complaint of death. If this analysis is correct, then it saves the *ex-post* theorist from troubling incorrect verdicts in cases like Natural Disaster.

At this point, one might wonder if by paving the way for interventions like the one in Natural Disaster to gain justification on *ex-post grounds*, this line of thought might also justify a programme of catastrophic risk mitigation schemes. The thought here is that, if you group together a sufficient number of long-term interventions, which look like the one in Biotech Risk, then at a certain point you will have good statistical reason to believe that one of them will be successful. As such, you will also have good statistical reason to believe that you will save a great number of lives, generating significant *ex-post* complaints in support of this programme of interventions.¹²

However, I would point out that this line of justification is likely impractical for typical catastrophic risk interventions. Given the improvements to probabilities associated with long-term interventions are so tiny, the number of

11 Note, I have designed this case to look at programmes of natural disaster mitigation which look at different risks which face different people; for example a programme to mitigate the harm of flooding in multiple locations around a country. However, a similar line of justification can be offered to justify a programme of interventions which all seek to mitigate the same risk; for example, a programme of interventions all targeting the flood risk in a specific town in a country. In this case, rather than receiving one ticket for X different lotteries, the agent would receive X tickets for one lottery.

12 Whilst the likelihood of any catastrophic risk intervention being successful is unlikely to be totally independent – for example, multiple interventions are like to share similar assumptions and be based on similar models – it is also not the case that the success of catastrophic risk interventions is all or nothing either, such that all interventions work or none do (unlike the outcomes of a catastrophic risk intervention, which as discussed in §1, do have this all or nothing structure).

interventions such a programme would need to include to generate near statistical certainty of savings lives would be, at a minimum, practically preclusive. As such, at least at first glance, we seem to have a line of reasoning which allows the contractualist to avoid the counterintuitive verdicts in cases like Natural Disaster, which, nonetheless, does not commit the contractualist to permitting investment in catastrophic risk interventions.

Whilst this line of thought is promising, it does face a series of difficult problems. The first is simply that this justification can only be used when we have multiple intervention tokens; it doesn't help us justify interventions which there will only be one instantiation of, which we might nonetheless wish to engage in.¹³

Second, at least on the contractualist account we have been considering, the application of this justification is limited simply to cases in which the probability of not saving people becomes very small. Whilst we might be able to treat very slightly discounted complaint of a harm as equivalent in strength to undiscounted complaints of the same harm, there is clearly a limit to how discounted a complaint can be for us to do this. Yet, there is intuitively a gap between interventions which we do not reasonably expect to save any lives and interventions which we are almost certain will save a life, and you might still want to be able to engage in some of the interventions which fall within this gap.

Consider, for example, a programme of natural disaster risk mitigation which makes it the case that we have 70% confidence of saving lives. This seems like the sort of programme of interventions which should not be disqualified out of hand. But a complaint of death discounted by 30% is just not equivalent to an undiscounted complaint of death. So, this might be a weaker response to the problem than we had hoped; it is unable to justify a large number of intuitively permissible interventions.

I will quickly gesture at two potential responses to this problem. The first is to simply accept the limitations of this solution, whilst highlighting that at least

¹³ Thank you to an anonymous reviewer for pressing this point.

it allows the contractualist to justify those classes of interventions which would be most counterintuitive to exclude. This at least lessens the bite of the initial criticism. The second response is to highlight that the power of this justification can be extended significantly if we were to weaken one of the contractualist restrictions.

Several complaints-based theorists have weakened the individualist restriction such that the aggregation of some complaints would be permitted. These “partially-aggregative” theories allow aggregation in cases in which the lesser competing complaint is, nonetheless, of a similar moral significance to the complaint against which it competes (see, Voorhoeve, 2017). Whilst a discussion of such accounts is beyond the scope of this chapter, the contractualist could certainly use such a stance to justify *ex-post*, for example, the intervention which would give a 70% chance of saving lives from a natural disaster. In this case, the *ex-post* complaints of death, discounted by 30%, would be allowed to aggregate together to compete against the lives competing interventions could certainly save. Provided the state-of-affairs which the intervention prevented with 70% likelihood contained enough deaths, the intervention will have a large aggregate complaint in its favour.

Nonetheless, in light of these problems, I will gesture at another means for the contractualist to justify the permissibility of natural disaster intervention. This justification, unlike our last, occurs on the level of individual interventions – allowing us to avoid the first problem. Crucially there seems to be a disanalogy between natural disasters and long-run catastrophic risks which Biotech Risk has obscured – namely, the source of the low probabilities of saving a life. As noted previously, the reason we can only improve the prospects of future people by a very tiny amount, or reduce the likelihood of a catastrophic state-of-affairs by a very small amount, is because we lack the power to reliably control the future. It is not to do, principally, with the antecedent likelihood that catastrophe will occur in the very far future. The likelihood might be very high – regardless, we just can’t do that much about it.

On the other hand, the source of the low probabilities associated with the likelihood of natural disaster interventions actually preventing disaster is not so closely tied with our inability to mitigate present-day natural disasters. Of course, there are some natural phenomena which we might struggle to effectively prevent or mitigate. But there are many natural disasters for which we have management methods which definitely work. This is especially clear with mitigation schemes; take, for example, seismic designing, which reliably saves lives during earthquakes by minimising damage to buildings. Or consider a slightly more realistic version of Natural Disaster; we can confidently reduce mortality, in the case of a volcanic eruption, through city and evacuation planning.

In such cases, the low chances of saving the lives of individuals from natural disaster does not result from our lack of confidence about the efficacy of the intervention, but rather the antecedently low chance of the specific catastrophe occurring for the intervention to mitigate. The different sources of the low probabilities associated with natural disaster management and catastrophic risk mitigation is important. Namely, the probabilities associated with saving lives through catastrophic risk schemes won't grow if you extend your time frame; we don't become more confident that our action will have some large, net positive impact when we consider the whole course of history. On the other hand, this is the case for certain types of natural disaster management. Given that the probability of these interventions saving lives is currently bound by the probability of the disaster occurring, the higher the likelihood of the disaster occurring at some point in our time frame, then the higher the probability that such interventions will save lives. Therefore, by extending the time frame relevant to our intervention assessment, such natural disaster management schemes become increasingly justifiable *ex-post*. That is, individual natural disaster interventions can be *ex-post* justifiable because we really do expect them to save lives at some point in the long-run, in a way that we don't reasonably expect catastrophic risk interventions to do at some point.

This is a peculiar result, insofar as it shows the contractualism can favour some long-term interventions – at least on the *ex-post* perspective. But importantly, contractualism does not favour the characteristic long-term interventions which bring about vast amounts of goodness by tending to tiny probabilities (or risks) of enormously good (or bad) outcomes. Instead, contractualism can favour those interventions which, through the perspective that longtermism brings, decisively and reliably attend to the high probability that many significant complaints of a harm will occur at *some* point in the future.

To briefly conclude this section, it seems that contractualism can – to some extent – avoid the bite of the analogue between presently accepted activities like attending to natural disaster risk and activities it judges as generating weak claims in its favour, such as biotechnology research governance. Whilst this might disarm one of the reasons pointing to reading §1 as an argument against contractualism itself, this is limited. Given our inability to control the far-future is pervasive, contractualism seems – at least on the *ex-ante* perspective – to largely preclude attending to the wellbeing of future people. This itself might seem like too large a loss to accept. And whilst there are more available *ex-post* justifications for engaging in common short-term practices, this fails to justify individual interventions or those with short time horizons.

4. Justifiability and prioritisation

Having surveyed a range of problems facing the contractualist critique of longtermism, in this section I will highlight some of the insights contractualism might have for prioritisation within the longtermism movement.

The prospects for long-term interventions on the *ex-ante* perspective might seem dire given the discussion in §1. As a result of the low level of confidence we can have in the efficacy of many very long-term interventions, these interventions will generate extremely weak claims in their favour. One upshot of this is that the *ex-ante* contractualist who cares about helping the far-future,

will be pushed into focusing on “sure thing” interventions, which can bestow benefits on future people with a good level of confidence.

There are at two observations to accompany this point. First, given the epistemological issues we face when it comes to predicting our impact on the future, there are significant questions to be answered about the size of the set of interventions which can bestow benefits on the far-future with any degree of confidence. One fairly straightforward example would be to leave money in trust for the future. Another possible class of interventions which we might have higher confidence about are those which aim to “bring forward progress”. The conditions in which humans live has been, seemingly, improving over time. Such improvements to the human condition, some believe, are very likely to continue in the future. As such, if we can bring forward humanity’s progress – even by a small amount – then the average wellbeing of each human will be higher at a given time, than it would be otherwise. In comparison to other long-term interventions, it might be thought that we can have at least reasonable confidence in bestowing these benefits; for example increased funding for research, alongside expenditure on population-level health and literacy could provide reasonable odds of speeding “progress” along.

This leads us onto the second observation. Whilst there may be a small range of long-term interventions about which we have relatively high confidence, these interventions may still struggle to be competitive on the *ex-ante* contractualist model. This will be the case if such interventions trade-off the possibility of bestowing large benefits on beneficiaries in order to achieve larger chances of bestowing *any* benefit upon its beneficiaries. In this case, the *ex-ante* complaints in favour of such interventions are weak, not because they are discounted heavily, but rather because the benefits they bestow are meagre.

Plausibly, interventions which seek to bring forward the march of progress might look very much like this. Assuming we can bring forward progress with a higher degree of confidence – which is no small assumption – the value of doing so might be fairly trivial for each person. Let’s say we bring progress forward by ten years. Whilst in aggregate this might bring about a lot of value, for each

individual, the improvement to their wellbeing will plausibly be only slight. As such, the claims each individual will have to this intervention like also be fairly weak.

The *ex-post* contractualist enjoys more room to favour long-term interventions than her *ex-ante* cousin. Nonetheless, they will only be able to engage in long-term interventions insofar as the intervention makes a significant difference to the probability of the benefit being bestowed, or the harm being incurred. This is because the *ex-post* complaints are discounted by the difference the intervention made to the improbability of benefit or good being bestowed. As such, unless the intervention significantly changes the chances that the group of future people are facing, the *ex-post* complaints generated by the intervention will be steeply discounted.

One appealing upshot of the *ex-post* contractualist assessment is that it, effectively, precludes agents from engaging in interventions which aim to minimise risk which is already very low. This is because, provided the antecedent risk is small enough, even if the intervention reduces the risk to zero, it will have still only changed the probability by a small amount. Thus the *ex-post* claims generated by the intervention will be very severely discounted. Whilst axiological or consequentialist approaches to longtermism might blush at the mention of how their theories justify interventions which look much like Pascal's Mugging (Bostrom 2003), contractualist longtermists can avoid such embarrassment. By a similar token, *ex-post* contractualism views those interventions which increase an already very high likelihood of an extremely good outcome as extremely uncompetitive.

5. Conclusions

In this chapter, I have given an account of the contractualist assessment of longtermism. I have outlined an argument demonstrating that contractualism, on both the *ex-ante* and *ex-post* perspectives, will view long-term interventions as uncompetitive in the context of available alternative interventions. I have then

considered the view that such a conclusion speaks more to the falsity of contractualism than that of longtermism. In particular, I have considered the claim that by precluding long-term interventions which seek to mitigate catastrophic risk, contractualism arrives at a deeply counterintuitive verdict concerning a number of, seemingly permissible, present-day behaviours. I have attempted to distinguish between long-term interventions targeting catastrophic risk and short-term interventions managing natural disaster risk on two fronts. First, I argued that, at least on *ex-post* grounds, natural disaster management can be justified on the level of programmes of interventions, as opposed to individual interventions. Whilst this justification is open to present-day natural disaster management, it does not seem available for long-term interventions. Following a survey of potential issues for this account, I outlined a second distinction between the two interventions concerning the source of their low probabilities associated with benefitting any given person. Having now defended, to some extent, contractualism from this paper's *reductio*, in §4 I consider what the implications contractualism might have for prioritisation within the longtermism movement.

6. References

- Balfour, D. (2020). 'Pascal's Mugger Strikes Again', in *Utilitas*, 33/1: 118-24
- Bostrom, N. (2009). 'Pascal's Mugging', in *Analysis*, 69/3: 443-445
- Curran, E.J. (2022). 'Longtermism, Aggregation, and Catastrophic Risk',
retrieved from: <https://globalprioritiesinstitute.org/longtermism-aggregation-and-catastrophic-risk-emma-j-curran/>
- Greaves, H., and MacAskill, W. (2021). 'The case for strong longtermism', *GPI Working Paper*, retrieved from <https://globalprioritiesinstitute.org/wp-content/uploads/The-Case-for-Strong-Longtermism-GPI-Working-Paper-June-2021-2-2.pdf>
- Hare, C. (2007). 'Voices from Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?', in *Ethics*, 117: 498-523
- (2016). 'Should We Wish Well to All?', in *The Philosophical Review*, 124/4: 451-472

- John, S. (2014). ‘Risk, Contractualism and Rose’s “Prevention Paradox”’, in *Social Theory and Practice*, 40/1: 28-50
- Kamm, F. M. (1993). *Morality, Mortality*, 1 (Oxford University Press)
- Kirkpatrick, J. R. (2017). ‘Permissibility and the Aggregation of Risks’ in *Utilitas*, 30/1:107-119
- Kosonen, P. (ms). ‘Tiny Probabilities and the Value of the Far Future’, retrieved from: <https://petrakosonen.files.wordpress.com/2022/07/chapter-6-tiny-probabilities-and-the-value-of-the-far-future.pdf>
- Kumar, R. (2003a). ‘Reasonable reasons in contractualist moral argument’, in *Ethics*, 114/1: 6-37
- (2003b). ‘Who Can Be Wronged?’ in *Philosophy & Public Affairs*, 31/2: 99-118
- Scanlon, T. M. (1998). *What We Owe to Each Other* (Belknap Press of Harvard University Press)
- Tarsney, C. (2022). ‘The Epistemic Challenge to Longtermism’, *GPI Working Paper*, retrieved from: <https://globalprioritiesinstitute.org/christian-tarsney-the-epistemic-challenge-to-longtermism/>
- Thorstad, D. (2021). ‘The Scope of Longtermism’, *GPI Working Paper*, retrieved from: <https://globalprioritiesinstitute.org/the-scope-of-longtermism-david-thorstad-global-priorities-institute-university-of-oxford/>
- Voorhoeve, A. (2014). ‘How Should We Aggregate Competing Claims?’ in *Ethics*, 125/1: 64-87
- (2017). ‘Why One Should Count Only Claims with Which One Can Sympathize’, in *Public Health Ethics*, 10/2: 148-156
- Wilkinson, H. (2022). ‘In defence of fanaticism’, in *Ethics*, 132/2: 445-77
- Wu. P. (2021). ‘Aggregation and Reductio’, in *Ethics*, 132/2: 508-525