

Moral demands and the far future

Andreas Mogensen

Global Priorities Institute | June 2020

GPI Working Paper 1-2020



Moral demands and the far future

ABSTRACT: I argue that moral philosophers have either misunderstood the problem of moral demandingness or at least failed to recognize important dimensions of the problem that undermine many standard assumptions. It has been assumed that utilitarianism concretely directs us to maximize welfare within a generation by transferring resources to people currently living in extreme poverty. In fact, utilitarianism seems to imply that any obligation to help people who are currently badly off is trumped by obligations to undertake actions targeted at improving the value of the long-term future. Reflecting on the demands of beneficence in respect of the value of the far future forces us to view key aspects of the problem of moral demandingness in a very different light.

I.

If our planet remains habitable for another 1 billion years and will sustainably support a population of at least 1 billion people at any time, then there could exist at least 10^{16} lives of normal duration in our future (Bostrom 2013). This is only counting human beings. Chickens alone currently have a standing population of 19.6 billion (Robinson et al. 2014). Given the number of morally statused individuals who could potentially populate the long-run future, the value at stake in choosing among actions that impact on the long-term trajectory of human civilization seems to be astronomical. What demands are placed on us as a result?

Moral theorists have typically discussed the demands of beneficence under the assumption that these represent obligations for those who are wealthy to transfer resources to their poorer contemporaries (e.g., Singer 1973; Ashford 2000; Murphy 2000; Cullity 2004). Insofar as moral philosophers have discussed obligations related to future generations, they have tended to focus on the question of what (if anything) we owe to future people as a matter of justice (e.g., Barry 1977, 1997; Rawls 1999; Meyer and Roser 2009; Steiner and Vallentyne 2009; Caney 2018). They have largely neglected the question of how the value of the future shapes our thinking about the demands of beneficence.¹ As a result, I claim, moral philosophers

¹ The issue has been discussed by some authors sympathetic to utilitarianism, including Bostrom (2003), Mulgan (2006), Beckstead (2013, 2019), Cowen (2018), and Greaves and MacAskill (2019). However, with

have either misunderstood the problem of moral demandingness and based much of the discussion of the demandingness objection to utilitarianism on false presuppositions, or else failed to recognize important dimensions of the demandingness problem that turn out to undermine many standard assumptions.²

Which of these interpretations we favour depends on the extent to which we understand the problem of moral demandingness as concerned with our obligations in the actual world.³ If we understand the demandingness problem as to do with how obligations of beneficence play out in actuality according to this or that moral theory, my arguments support the conclusion that we have misunderstood the problem of moral demandingness and based much of the discussion of the demandingness objection to utilitarianism on false presuppositions. If instead we believe that moral principles should be assessed independently of contingent facts and that candidate theories may be embarrassed by their extreme demandingness in other possible worlds, then the demandingness problem on which philosophers have focused so far need not be illusory, even if it does not correspond to how obligations of beneficence play out in actuality. Nonetheless, we will turn out to have neglected an important dimension of the problem.

In section 2, I show that the claim that astronomical value is at stake when choosing among actions that impact on the long-run future can be supported by a range of minimally plausible population axiologies, suggesting that consequentialism treats any obligation to help people who are currently badly off as decisively trumped by obligations to undertake actions directly targeted at improving the value of the long-term future. In section 3, I offer a brief discussion of previous philosophical work on the problem of moral demandingness, as well as outlining concerns about excessive demandingness that have emerged in the economic literature on optimal growth. Sections 4 through 7 highlight four different ways in which thinking about the demands of beneficence in the context of the value of the far future requires moral philosophers to revise their understanding of the problem of moral demandingness.

the exception of Mulgan, these writers do not seriously reflect on how the value of the far future intersects with the problem of moral demandingness.

² I'm grateful to an anonymous referee for suggesting this disjunctive framing of the paper's conclusions.

³ I will try to remain neutral on this question throughout this paper.

Specifically, I focus on the weight of agent-centred prerogatives, the contingency of the demandingness problem, the significance of non-compliance, and the relevance of passive demands. Section 8 concludes.

2.

Any action that makes any significant difference to how things play out over the long-run will inevitably mean that some people are born who would not otherwise have been conceived and others are prevented from existing. In claiming that the value at stake in choosing among actions that impact on the long-run future is astronomical, we are therefore forced to make value comparisons between outcomes in which the size and/or composition of the population varies. Making such comparisons is notoriously difficult (Parfit 1984; Arrhenius MS; McMahan 2013; Greaves 2017;). Fortunately, the claim that astronomical value is at stake when choosing among actions that impact on the long-run future appears to be robust across a range of minimally plausible population axiologies. Because the claim is supported by a broad range of plausible theories, we are not forced to decide which among these theories is correct.

For the sake of brevity, I restrict myself to arguing that the claim is supported by three different theories: *Total Utilitarianism*, *Average Utilitarianism*, and the *Axiological Asymmetry*. It is straightforward to extend the arguments given here to other minimally plausible population axiologies like *Critical-Level Utilitarianism* (Broome 2004; Blackorby, Bossert, and Donaldson 2005) or *Variable Value Theories* (Hurka 1983; Ng 1989), since these theories mimic the behaviour of Total or Average Utilitarianism under suitable conditions.⁴

According to Total Utilitarianism, the value of an outcome is the sum of the welfare of every individual existing in that outcome. Stated formally, if an outcome, O , contains n people

⁴ This is not to say that there is no axiology on which the claim is false. Consider the view that different number populations are incomparable (Bader 2020). Since the different possible futures we might realize in practice correspond to differently sized populations, this entails that the different futures we might realize are neither better than, nor worse than one another. On the other hand, this view may well not strike us as minimally plausible. It entails that any population of n people with blissful lives is not better than any population of m people with lives that contain only suffering just in case $m \neq n$.

and u_i denotes the (interpersonally comparable and ratio-scale measurable) welfare of person i ($i = 1, \dots, n$), then the Total Utilitarian value function, V_T , is given by

$$V_T(O) = \sum_{i=1}^n u_i$$

Because the number of people who could exist in future is so much greater than the number of individuals who currently exist, Total Utilitarianism supports the view that the value of any outcome available to us depends almost entirely on how things go in the long-run. As a result, differences in expected value between the actions available to us will be far more sensitive to differences in the probabilities of long-term outcomes than to even the most significant short-term events.

By way of illustration, consider actions that reduce the risk of human extinction. Given Total Utilitarianism, if human beings go extinct within the near future, as opposed to persisting for another 1 billion years with a population of at least 1 billion people at any point in time, very little of the loss in value will be due to the effects on people living at the time of extinction. Most of the value lost is due to the fact that trillions of people who would otherwise have enjoyed worthwhile lives are never born. Because there are so many potential people, the expected value of any action that lowers the risk of an existential catastrophe is astronomical. Given the conservative population projection we have assumed so far, Bostrom (2013) calculates that in a choice between reducing existential risk by one millionth of one percentage point and saving a hundred million human lives (without altering the risk of existential catastrophe), we ought to be indifferent between these options.

Let's now consider Average Utilitarianism. According to Average Utilitarianism, the value of an outcome is the sum of the welfare of every individual existing in that outcome divided by the size of the population:

$$V_A(O) = \frac{1}{n} \sum_{i=1}^n u_i$$

This entails that an outcome in which 100 billion people have welfare level \bar{u} is just as valuable as an outcome in which 100 trillion people experience welfare level \bar{u} . For this reason, it may seem that Average Utilitarianism conflicts with the claim that the value at stake in choosing among actions that impact on the long-run must be astronomical because there are so many morally statused beings who will exist in future. On this view, it looks as if size doesn't matter.

While it is true that Average Utilitarianism may count efforts to prevent the extinction of the human species as valueless if the average welfare of future people is the same as the average welfare of everyone existing up to now, it is easy to show that Average Utilitarianism still supports the claim that the value at stake when our actions have the potential to impact on the long-run future can dwarf the value of short-term improvements, so long as we keep in mind that there are other ways in which we might affect the long run.

By way of illustration, suppose that under the *status quo*, the total population across all time periods will be 1 trillion people who will experience a lifetime welfare level of 10. The last 1,000 people born among the first 100 billion are struck by an illness. If they are not helped, their lifetime welfare (and theirs alone) will be 0. You can ensure that they are cured, as a result of which they will each enjoy a lifetime welfare level of 10. Alternatively, you can perform an action that yields a probability just slightly greater than 10^{-7} that the remaining people to be born experience an average welfare level that is 1% higher. According to Average Utilitarianism, the latter action has greater expected moral value. Obviously, this is just a toy model. The more general lesson is that when the future population is very big relative to the present, what happens over the long-run plays a much bigger role in determining the average welfare of the total population of everyone who ever lives than what happens to present people.

Finally, let's consider the Axiological Asymmetry. The Axiological Asymmetry is not a population axiology in the same sense as Total or Average Utilitarianism, in that it does not tell us how to rank variable population outcomes in general.⁵ According to the Axiological Asymmetry, between two outcomes, O and O^* , differing only in that there exists some additional person in O^* but not in O , O^* is neither better nor worse than O if this person has a

⁵ The Axiological Asymmetry is, however, entailed by some population axiologies. For example, it is entailed by what Broome (1996) calls *Critical Band Utilitarianism* if the critical band is chosen to be unbounded above and bounded below by the point beneath which a life is not worth living.

life worth living, but worse than O if this person does not have a life worth living (Frick 2014; McMahan 1981; Roberts 2011). The Axiological Asymmetry reflects the widely held intuition that while there is no moral reason in favour of making happy people, there are strong moral reasons not to bring into existence people who will experience lives that are not worth living.

A view of this kind may seem an unlikely candidate to support the view that the far future has overwhelming importance. Nonetheless, it is possible to construct a plausible argument for the overwhelming importance of the long-run future given the Axiological Asymmetry, so long as we assume that the badness of additional suffering lives does not diminish. As Parfit (1984) says: "It is always bad if an extra person has to endure extreme agony. And this is always just as bad, however many others have similar lives." (406) Given this assumption, we can argue that astronomical value is at stake when it comes to the long-run future because it is reasonable to expect that there will exist astronomically many future individuals with lives in which suffering predominates and the disvalue of these lives is an increasing linear function of the total amount of suffering they contain. Beneficial trajectory changes that reduce the number of suffering individuals who populate the future of Earth-originating civilization may therefore be of the highest moral importance.⁶

Consider factory-farmed animals. Many factory-farmed animals are believed to experience lives that are not worth living. In their assessment of the welfare of US livestock, Norwood and Lusk (2011) conjecture that broiler breeder chickens, caged egg-laying hens, veal

⁶ A complication for this line of argument that I sadly do not have space to explore in full depth is the issue of 'greedy neutrality'. Broome (2004) argues that a view like the Axiological Asymmetry, which treats the mere addition of lives worth living as neutral in value, requires us to think that changes in the size of a population that involve changes in the number of lives worth living can neutralize other changes to the population that would otherwise make it better or worse, resulting in a population that is instead neither better than, nor worse than, nor exactly as good all-things-considered. For this reason, if reducing the number of lives that are not worth living results in significant changes to how many lives worth living exist in addition, this might, surprisingly, not be better overall. I am inclined to doubt that this issue is a genuine concern in respect of the concrete examples discussed here. Furthermore, if in fact neutrality turns out to be 'greedy' in this way, this arguably casts more doubt on the Axiological Asymmetry than on the view that astronomical value is at stake with respect to actions that could spare thousands or millions or billions of future individuals from lives of suffering.

calves, and all pigs except the relatively few raised in the shelter-pasture system have negative welfare scores. According to their numbers, the US annually uses around 62 million birds as broiler breeders and the annual inventory of US laying hens is 340 million, while 942,000 veal calves are slaughtered per year alongside around 115 million pigs.

Some human lives also seem sufficiently bad that they are not worth living: for example, children who are killed by infantile Tay-Sachs disease (Steinbock 1986). Fortunately, Tay-Sachs is rare, striking only 1 in 320,000 newborns in the US general population. However, for a total future population of at least 10^{16} people, an incidence of 1 in 320,000 would mean more than 30 billion cases of infantile Tay-Sachs. Reducing the incidence of Tay-Sachs by just a few hundred thousandths of a percentage point would spare more future children from a life of overwhelming suffering than there are people currently living.

While I have not canvassed every available theory, I hope I have said enough at this point to render plausible the claim that a wide range of minimally plausible population axiologies support the verdict that astronomical value is at stake when choosing among actions that impact on the long-run future. Different theories may yield different verdicts about how best to go about improving the value of the future. Total Utilitarianism assigns greater importance to reducing the risk of extinction than Average Utilitarianism, while the Axiological Asymmetry prioritizes actions that reduce the risk of astronomical future suffering.⁷

⁷ Are there concrete initiatives aimed at these goals to which we can lend our support? While in-depth assessment and/or recommendation of particular organizations is beyond the scope of this paper, I note that readers are able to switch their donations from organizations like Against Malaria Foundation or GiveDirectly to the Long-Term Future Fund set up by members of the effective altruism movement, whose stated aims are to “positively influence the long-term trajectory of civilization by making grants that address global catastrophic risks; promote, implement, or advocate for longtermist ideas; and otherwise increase the likelihood that future generations will flourish.” See <https://app.effectivealtruism.org/funds/far-future>. For concrete existential risks, readers might choose to support the John Hopkins Center for Health Security if they are worried about biological risks, the Center for Human Compatible AI or Open AI if they are worried about risks from artificial intelligence, or the Nuclear Threat Initiative if they are concerned about risks from nuclear weapons. I’m grateful to Will MacAskill for these suggestions.

Nonetheless, the theories we have surveyed agree on the overwhelming importance of posterity.⁸

For concreteness and simplicity, I will, from now on, proceed under the assumption that Total Utilitarianism is the correct population axiology, unless otherwise indicated. Of the standard theories discussed in the literature, this strikes me as the most plausible - which is not to say that its implications are always easy to swallow. However, given the results discussed in this section, much of what I say in the remainder of this paper can be generalized given suitable modifications.

3.

This section begins by reflecting briefly on the significance of the conclusions drawn in the previous section for traditional discussions of the demandingness objection to utilitarianism among moral philosophers. I suggest that much of this discussion may have limited application to the demands of beneficence we face in actuality. Economists, by contrast, have thought a great deal about the extreme demands required to maximize an intergenerational utilitarian social welfare function. I suggest that philosophers would be well-placed to help clarify the concerns debated among economists, but for the fact that we have largely neglected moral demands arising from the value of the far future. This sets the stage for my argument in sections 4 to 7, where I show that transposing existing work on moral demandingness to the intergenerational setting will be no mean feat, since many standard assumptions made in the discussion of moral demandingness are invalidated in this context.

As stated, I begin by reflecting on the significance of the conclusions drawn in the previous section for the demandingness objection to utilitarianism. The demandingness of utilitarianism is typically illustrated via the obligation to support charitable causes that help badly off people living right now: typically, the many millions who live in extreme poverty in developing countries (see, *inter alia*, Kagan 1997: 154-5; Mulgan 2001: 3-4; Driver 2006: 56-60). However, if the argument of section 2 is correct, utilitarianism may be thought to entail that

⁸ One other issue that I have not addressed is whether we can be sufficiently confident in our ability to impact the long-run future for far-future value to achieve overwhelming practical importance, even granting that the relevant probabilities need only be quite small. Satisfactory treatment of this issue is beyond the scope of this paper, but see Tarsney (2019) for defence of the affirmative answer.

any obligation to help people who are currently badly off is decisively trumped by obligations to undertake actions directly targeted at improving the value of the long-term future (Bostrom 2003; Beckstead 2013, 2019; Cowen 2018; Greaves and MacAskill 2019).

This assumes that helping people who are currently badly off is not the best way of improving the long-term future. However, it would be a striking coincidence if short-termist and long-termist priorities coincided in this way (Greaves and MacAskill 2019). Cowen (2018: 89-91) even suggests that when the value of the long-term future is kept in mind, an argument can be made that utilitarianism supports redistribution from people who are currently poor to people who are currently rich. This is because the rich earn higher returns on accumulated wealth (Piketty 2014) and Cowen believes that sustainably higher economic growth over the long-run may yield sufficient benefits in aggregate to outweigh significant one-time costs once the full extent of the future is kept in mind. Obviously, many doubts can be raised about this line of argument,⁹ but its aim is not to convince us, so much as to put us on the backfoot. Cowen writes: “Direct, short-term redistribution to today’s poor is no longer the default option for a moral theory that emphasizes individual well-being. Instead, in many cases utilitarianism has to work to avoid the conclusion of redistributing more resources to the wealthy.” (90) We therefore have grounds to believe that most moral philosophers have misunderstood - perhaps quite radically - what utilitarianism concretely demands of us.

By contrast, the demands placed on the current generation in maximizing an intergenerational utilitarian social welfare function are a well-worn subject of debate among economists. One particular focus of debate is the savings rate required for optimal growth in the modelling framework pioneered by Ramsey (1928). Ramsey himself notes that the rate of savings implied by his analysis “is greatly in excess of that which anyone would normally suggest” (548).

To illustrate the problem, I follow Dasgupta’s (2008) presentation. Assume an indefinite sequence of generations in which each generation is of the same size and perfectly homogeneous, so that generational welfare at time t is summarized by the utility of consumption of a representative agent, $u(c_t)$. In stipulating that the sequence of generations is

⁹ For example, it may be argued that inequality inhibits economic growth: see Cingano (2014), Ostry et al. (2014).

indefinite, we mean simply that it has no guaranteed endpoint. We assume there is a constant exogenous per period probability of extinction given by $\delta > 0$.¹⁰ From the perspective of the current (zeroth) generation, the expected value of the indefinite consumption stream $\{c_t\}$ is given by

$$W_{\{c_t\}} = \sum_{t=0}^{\infty} \frac{u(c_t)}{(1 + \delta)^t}$$

As is standard, we assume the utility function has the so-called *isoelastic* form, with $u(c_t) = c_t^{1-\eta} / (1 - \eta)$ for $\eta > 0$ and $\eta \neq 1$ and $\ln(c_t)$ for $\eta = 1$. Here, η is the elasticity of the marginal utility of consumption. If we follow Stern (2006) and set η equal to 1, this means that the same proportional increase in consumption is equally valuable regardless of a person's ex ante consumption level.

We consider a very simple pure capital model. At the beginning of each period, some portion of the inherited capital stock is consumed and the portion remaining earns a rate of return, given by a positive constant, r . The optimal saving-output ratio, s^* , can then be shown to be approximated by the expression $(r - \delta) / \eta r$. (The equality is exact in continuous time.) Therefore, if $\eta = 1$ and δ is very small in comparison to r nearly the entirety of total output must be saved. For example, if $r = 4\%$ and $\delta = 0.1\%$, then $s^* \approx 97.5\%$.

The model is obviously a gross simplification of economic reality, but even highly simplified models can provide valuable insights into the complex realities they represent. In his classic discussion of economic methodology, Friedman (1953) goes so far as to assert: "the more significant the theory, the more unrealistic the assumptions" (153). DeLong (2006) and Stern (2008) object that the model under discussion here nonetheless misleads us, since substantial reductions in the optimal saving-output ratio occur if we relax the model's assumption of zero exogenous productivity growth, which they regard as artificial. On the other hand, Dasgupta

¹⁰ For simplicity, we assume this to be the only source of uncertainty in the model. If the extinction event obtains, the value in each subsequent period drops to zero.

(2008: 153) defends the use of this assumption, noting that ‘capital’ in the model is intended to cover all physical, human, or natural capital. Dasgupta (2008: 152-3) similarly appeals to the inclusiveness of the capital measure to justify eschewing diminishing returns to the factors of production, although he concedes that he makes “outrageous assumptions about the aggregation of capital” (153 n 18).

Setting aside object-level questions about the irrationalism of the model, two further points should be kept in mind in reflecting on what can be learned from unrealistic growth models such as the above. Firstly, since the optimal savings rate was derived to be close to 100% in the simplified model, it could fall quite a long way in a more realistic setting and nonetheless strike us as excessive. Thus, in highlighting the demandingness of an undiscounted intergenerational utilitarian social welfare function, Zuber and Asheim (2012) highlight the attempt by Mirrlees (1967) to bring Ramsey’s model closer to economic reality, objecting that by Mirrlees’s estimates, “present generations should save up to 50% of their net income for the sake of future generations.” (1573) Secondly, regardless of their value as representations of the actual economy, simple economic models like the one analysed here describe possible worlds. If we believe that moral theories can be tested by their implications in even relatively distant possible worlds, the implications of our ethical assumptions for the world of the simple growth model presented above are not at all irrelevant in reflecting on what moral principles should be thought to govern how much to save in actuality - even if the model is a poor model of economic reality.¹¹

One stock reaction among economists to the kind of result derived above is to treat it as an argument for *pure time discounting*, whereby we are to down-weight the utility of future people over and above the adjustment already incorporated in our model to take account of the probability that some exogenous extinction event wipes out the human race (Arrow 1996; Nordhaus 2007). Under this approach, our concern for the welfare of future people diminishes at a constant rate merely because of the passage of time, and we incorporate this positive rate of pure time preference into the value of δ alongside the exogenous risk of extinction.

¹¹ Compare Dasgupta (2008: 155 n 25): “‘non’-artificial models, such as those used by Stern in his computer runs, don’t reveal which parameter is doing what work in generating his findings. How is one to test the robustness of ethical assumptions if not by putting them to work in stark, artificial models?”

By suitably increasing the value of the δ parameter in this way, we can avoid the excessively high rate of saving that we derived earlier. However, many philosophers and economists argue that discounting future utility merely due to distance from the present is morally unacceptable (see, *inter alia*, Sidgwick 1874; Ramsey 1928; Parfit 1984; Rawls 1999; Stern 2006; Caney 2014).¹² Even if we are comfortable with the idea that present people count for more than future people, the rate of pure time preference required to avoid demands for excessive accumulation on behalf of the present generation requires assigning absurdly little weight to the welfare of people who live sufficiently far in future. For example, if the constant rate of pure time preference is just 1% per annum then the expected value of reducing the risk of fatality for one currently existing person by one in a million is about the same as preventing the deaths of 4×10^{15} otherwise similar people living 5,000 years from now.

Rather than trying to decrease the value of s^* by beefing up δ , we can try to avoid imposing excessive sacrifices on the current generation by increasing our aversion to consumption inequality (Asheim and Buchholz 2003; Dasgupta 2008; Broome 2012). Because future generations are expected to be richer on average than the current generation, greater aversion to consumption inequality reins in the extent to which the current generation ought to save in order to augment future consumption. Within the Ramsey model, aversion to consumption inequality is determined by η . Thus, Dasgupta (2008) suggests that in order to avoid demands for excessive accumulation, we should reject $\eta = 1$ in favour of a value for η in the range $[2, 4]$. But, as Dasgupta himself notes, increasing the value of η in this way seems to have the effect of making the requirements on people who are presently well-off to benefit people who are presently badly-off even more extreme than they are generally thought to be under utilitarianism. Thus, if the utility of the present generation is the sum of the utility of each individual and the utility of each individual is given by a common isoelastic utility function with $\eta = 2$, a person whose annual consumption is \$36,000 is required to reduce her

¹² Drupp et al. (2018) surveyed two hundred economists who had been (co-)authors of at least one publication in the field of social discounting in a leading economics journal and found that 38% favour the Ramsey/Stern view on which $\delta \leq 0.1\%$.

consumption by up to 50% in order to prevent a decline of 1% in the annual consumption of a person whose status quo consumption per annum is \$360.¹⁵

As Parfit (1984: 484-5) suggests, a more reasonable response to these problems is simply to reject the assumption that we are obligated to maximize the value of a utilitarian social welfare function. We should instead impose some kind of limit on the sacrifices we can be asked to make for the sake of promoting the good by benefiting future generations. By analogy, philosophers who have thought about the traditional demandingness objection to utilitarianism do not suppose the objection can be satisfactorily answered by revising our conception of what welfare consists in, nor even by revising our account of what constitutes the impartial good, so long as we remain wedded to the maximizing act-consequentialist theory of right action (see Scheffler 1993: 41-2). Instead, we need to find an alternative criterion of right action that avoids the unpalatable implications that follow when the maximizing act-consequentialist theory of right action is married to any minimally plausible axiology.

If we reject the assumption that maximally promoting the good is morally required whenever it is morally permissible, is there any other plausible theoretical account of the demands of beneficence that we can put in its place? That has been a central question in normative ethics for the past fifty years (see, e.g., Singer 1972, 1993; Slote 1985; Kagan 1989; Shiffrin 1991; Scheffler 1992, 1993; Murphy 1993, 2000; Unger 1996; Ashford 2000, 2003; Cullity 2004; Arneson 2004, 2009; Norcross 2006). But discussion of how (if at all) to limit the demands of beneficence has largely been conducted against the background of the assumption that the demands of beneficence involve obligations for those who are wealthy to transfer resources to contemporary people who are very badly off. Reflecting instead on the demands of beneficence in relation to the value of the far future forces us to view key aspects of the problem of moral demandingness in a new light, as I will argue in sections 4 through 7.

¹⁵ An additional complication arises from the fact that it is merely *on average* that future people are expected to be richer than current generations. For discussion, see Schelling (1995), Rendall (2019), and especially, Fleurbaey and Zuber (2013), who argue that disaggregating intragenerational consumption supports the use of a negative discount rate for assessing climate policies since “the discount rate should be negative when the poorest contributors to the policy are richer than the poorest beneficiaries” and “[i]t is plausible that many climate policies satisfy this condition.” (586)

This section has noted that the conclusions drawn in section 2 suggest that moral philosophers have misunderstood what utilitarianism concretely demands of us, in light of which much of the discussion of moral demandingness may be thought to have limited application to the demands of beneficence that we face in actuality. Economists have thought a great deal about the demandingness associated with maximizing intergenerational social welfare, but have been unable to arrive at a satisfactory position on this issue. Parfit has suggested that the lessons philosophers have learned in reflecting on the demandingness problem may be of help in reflecting on the issues debated among economists. While I agree with this sentiment, the next three sections will show that switching our focus to the intergenerational context problematizes or invalidates many standard assumptions made by moral philosophers in debates on moral demandingness.

4.

This section considers what is arguably the most straightforward solution to demandingness of utilitarian moral theories and argues that concerns already expressed in the literature regarding the satisfactoriness of this solution ramp up when we take into account the value of the far future.

The most natural solution to the demandingness of utilitarianism is, I take it, that proposed by Scheffler (1993): adopt a revised theory incorporating an *agent-centred prerogative* that allows the agent to weight her own welfare more heavily than others', as opposed to being required to treat her own welfare as having no more significance for her than it has from the point of view of the Universe.¹⁴ Assuming that the theory we are revising starts out incorporating a Total Utilitarian axiology alongside the act-consequentialist criterion of right

¹⁴ For simplicity, my discussion here interprets the prerogative purely as a rule for weighting utilities. Strictly speaking, under Scheffler's theory "each agent would have the prerogative to devote energy and attention to his projects and commitments out of proportion to their weight in the impersonal calculus." (44) Although Scheffler goes on to say that the prerogative "allow[s] each agent to assign a certain proportionately greater weight to his own interests than to the interests of other people" (20), the identification of a person's projects and commitments with her welfare is generally questionable. The points I make in the ensuing discussion can easily be transposed to a more expansive conception of an agent's interests, *mutatis mutandis*.

action, the simplest revision of the theory incorporating an agent-centred prerogative would tell us that an act is morally required for agent i only if it maximizes the following value function:

$$V_{ACP}(O) = u_i + \frac{1}{k} \sum_{i \neq j} u_j$$

where k is some constant greater than 1. By choosing a value for k that is suitably great, we can limit the sacrifices each person is asked to make on behalf of others.

A key concern attached to Scheffler's theory is whether it is possible to choose a value for k that suitably limits the demands of beneficence without putting extreme weight on the agent's own interests (Murphy 2000: 64-5; Arneson 2004: 44-5; Noggle 2009). When we take into account our ability to impact the astronomical value at stake over the long-run future, I claim that the problem becomes overwhelming. The sheer size of the future is liable to swamp even values of k that strike us as obscene.

To illustrate, suppose there could be 10^{16} people in our future but for a range of existential risks, such as nuclear war or bioengineered pandemics. Recall that in a choice between reducing existential risk by just slightly more than one millionth of one percentage point and saving a hundred million human lives (without altering existential risk), act-consequentialism in conjunction with Total Utilitarianism entails that we ought to choose the former. It follows that the postulate of an agent-centred prerogative will acquit you of an obligation to sacrifice your own life in order to reduce the risk of existential catastrophe by ever so slightly more than one millionth of one percentage point only if you value your own life at more than one hundred million times that of a stranger.

To make matters worse, this result depends on adopting a reasonably conservative projection of the potential future population. Giving a mere one percent credence to less conservative estimates that take into account the potential for (post-) humanity to spread to the stars and for future minds to be implemented in computational hardware, Bostrom (2013) calculates the expected value of reducing existential risk by as little as one billionth of one billionth of one percentage point to be one hundred billion times the value of a billion human lives. To acquit yourself of an obligation to stand ready to sacrifice your life in order to achieve

a barely perceptible reduction in the risk of existential catastrophe, you would therefore need to assign astronomically greater agent-relative importance to your own welfare.

It may be objected that we cannot determine reasonable bounds for the value of k in the abstract. Whether it is unreasonable to value one's own life at k times that of a perfect stranger should be assessed not by reflecting on whether the chosen value for k strikes us as a really big number, but by considering its concrete implications. If values below some threshold, k^* , require people to be willing to sacrifice their lives for very small reductions in the risk of extinction, this may be thought to indicate that values for k at least as great as k^* are not in fact unreasonable.

However, we have more to go on than the sheer size of the numbers involved. Note, for example, that for any given value of k , plausible assumptions about choice under uncertainty¹⁵ suggest that a person cannot be required to accept a risk of death of n/k in order to certainly save n lives that will otherwise certainly be lost. Thus, suppose we insist that a person cannot be required to sacrifice her life for the sake of reducing existential risk by 10^{-18} even given a 1% probability for (post-) humanity to spread to the stars and for future minds to be implemented in computational hardware, and we therefore set k to be greater than 10^{20} . It would follow that a person could not be required to take on a risk of death much less than that of an ordinary drive to the supermarket in order to save billions of lives with certainty. This seems incredible.

The conclusions we have reached here should be unsurprising, given what was said about pure time preference in section 3. There I noted that the rate of pure time preference required to avoid demands for excessive accumulation on behalf of the present generation in the growth model we considered leads us to assign absurdly little weight to the welfare of people who live sufficiently far in future. While the agent-centred prerogative does not involve pure time discounting, it does involve the use of a utility discount factor whose value changes, albeit discontinuously over people, as opposed to declining gradually over time. The problems

¹⁵ The assumption is that under conditions where her impact on her own welfare is uncertain but her impact on others is not, agent i is morally required to perform some act only if it maximizes $\mathbb{E}(u_i) + \frac{1}{k} \sum u_j$. Stated otherwise, the agent is not morally required to be risk-seeking in respect of her own welfare when facing gambles of this kind, valuing the contribution of a gamble involving her own welfare in excess of its expected value to her.

facing the agent-centred prerogative are thus recognisable as analogues of the troublesome implications of pure time discounting highlighted in section 3: namely, that in order to avoid demands for excessive sacrifice when the size of the long-term future is taken into account, we seem required to drive down the utility discount factor that we apply to the welfare of others to implausibly low values.

This section has considered the idea that the problem of moral demandingness can be satisfactorily addressed by letting agents weight their own interests out of proportion to their impartial significance. I have argued that the concern that excessive demands can be avoided via this approach only if the agent is allowed to give her own interests excessive weight is put into overdrive when we consider the demands of beneficence in relation to the value of the far future. Thus, taking account of the value of the long-term future significantly dampens the plausibility of the idea that the problem of moral demandingness can be satisfactorily solved in this way.

5.

This section turns our attention to an issue that has already cropped up at points in our discussion: namely, whether the demandingness problem essentially concerns how obligations of beneficence play out in actuality according to this or that moral theory. There is a *prima facie* plausible line of argument suggesting that the demandingness problem for utilitarianism is defanged if we reject the idea that what happens in the actual world should be our central concern when thinking about demandingness. However, this line of argument turns out to be much less plausible once we reflect on ways in which we might improve the value of the far future.

The reason why it matters whether we conceive of the demandingness problem as concerned with how a theory's obligations of beneficence play out in actuality is that, plausibly, it is contingent whether any given moral theory is highly demanding (Railton 1986: 160-2; Ashford 2000: 435-9; Murphy 2000: 11, 26-31). It presumably could have been the case that utilitarianism asked very little of us. If other theories turn out to make only modest demands of us, this too is contingent. Deontological theories would impose very serious costs on us if securing our basic needs required violating deontological constraints. Utilitarianism might be less demanding on us in these circumstances.

Suppose, then, that we believe that fundamental moral principles should be assessed independently of contingent facts. We may then infer that the plausibility of a foundational moral theory cannot depend on whether the world just so happens to be of a kind that makes the theory highly demanding. Thus, it could not be a serious objection to utilitarianism that it is more demanding than other theories given the way the world happens to be. It could be claimed that utilitarianism is extremely demanding in *more* possible worlds than other moral theories, but it is not clear how to interpret or justify this claim.

A salient alternative is to interpret the demandingness objection as insisting that a theory should not be extremely demanding in worlds that are, in some sense, *morally normal*. While extreme circumstances call for extreme measures, within worlds in which circumstances are not extreme (as viewed from the moral point of view) we require that a candidate moral theory does not impose severe burdens on agents who comply with its demands. Thus, in explaining his own moderate conception of morality, Scheffler (1992) states his conviction that “under favorable conditions, morality permits people to do as they please within certain broad limits” (100). This, of course, shifts the discussion to the question of how to characterise morally normal worlds. It invites the obvious rejoinder that the demandingness of utilitarianism in the actual world is to be blamed not on the theory, but on the fact that the circumstances are unfavourable.

This is exactly the view taken by Ashford (2000) in her defense of utilitarianism against Williams (1973). According to Ashford, “the source of the extreme demandingness of morality is that the current state of the world is a constant emergency situation; there are continually persons whose vital interests are threatened and, given modern communications, the relatively well-off are continually able to help them.” (430) If not for the constant state of emergency represented by widespread extreme poverty, Ashford argues, there would be no incompatibility between the demands of impartial beneficence and Williamsian integrity. Similarly, Railton (1986) blames the demandingness of consequentialism on “how bad the state of the world is” (161), noting that the theory would not be nearly so disruptive of our personal projects and commitments if “wealth were more equitably distributed” and/or “political systems were less repressive and more responsive to the needs of their citizens” (161).

Once we recognize the ways in which utilitarian beneficence may ask us to subordinate our own projects to the value of the far future, it becomes clear that Ashford and Railton are

mistaken to suppose that the demandingness of utilitarianism depends on the existence of intragenerational extremes of wealth and poverty. Consider the growth model analysed in section 3. Within this model, we assumed wealth to be exactly equally distributed within each generation. Nonetheless, we were able to derive results concerning the required rate of saving that most people regard as excessively demanding. The world represented by the model is not one in which there exists a constant state of emergency. The ability of the agents in the model to generate very significant benefits for distant others arose not from extreme disparities in consumption within a generation, but from the productivity of capital and the resulting possibilities afforded by economic growth across generations. The optimal rate of saving was not derived by assuming short-lived, historically unique circumstances, and the high rate of saving derived within the model is in fact time invariant.

In reflecting on whether utilitarianism is extremely demanding in morally normal worlds, we must therefore avoid assuming that utilitarianism is highly demanding of us only because of the extremes of poverty and wealth that currently exist. We need to address whether conditions faithfully modelled by the growth model analysed in section 3 are to count as 'morally normal' or 'favourable'. Because the essential features of the model that drive the derivation of a very high rate of optimal saving seem so innocuous - optimistic, even - I presume that we are under much greater pressure to answer 'yes' than when we reflect on worlds marred by intragenerational extremes of poverty and wealth.

We should conclude that an otherwise appealing response to the demandingness problem that seems available to utilitarians who are attracted to the thought that what really matters is the demandingness of a moral theory in morally normal worlds is put into doubt once we recognize the ways in which individuals may in principle contribute to improving the value of the far future.

6.

This section considers responses to the demandingness problem that emphasize the idea that a person is obligated to do no more than her fair share of the collective effort. Whereas theories of this kind curb the demands of beneficence faced by individual people in relation to global poverty, I show that extant theories of this kind are toothless in the face of demands of beneficence that relate to the value of the far future.

The kind of theories we are to consider lead off from the insight that a demand to maximize intragenerational aggregate utility would not represent nearly so heavy a burden on those who heed its call if everyone could be expected to comply with this demand, even granting that there are millions of people in extreme poverty. If everyone who could help did their bit, then each of us might be required to make only modest sacrifices. Some philosophers argue that a key failing of utilitarianism in specifying the demands of beneficence is precisely its failure to moderate its demands in the face of others' non-compliance (Cohen 1981; Murphy 1993, 2000). When others who could help refuse to do so, utilitarianism requires us to pick up the slack. A more plausible conception of beneficence is supposed to index what is required of us under conditions of imperfect compliance to our fair share of the total effort defined under conditions of perfect compliance.

In developing this idea, both Cohen (1981) and Murphy (1993) foreground global poverty as a source of moral demands. If we look instead to the value of the long-run future, everything seems to turn on its head. Consider the growth model analysed in section 3. Here, extreme demands are derived while assuming perfect compliance. As Ramsey (1928) noted in setting up the model originally, we are "to assume that the community will always be governed by the same motives as regards accumulation" (544). This allows current savings to keep on paying returns indefinitely, yielding extraordinary benefits over the long run. If some near-future generation would be expected to consume everything we had saved, there would be no similarly stringent demand for accumulation. The picture is much the same if we focus instead on the mitigation of existential risk. If future people will be reckless in courting existential catastrophe, then the reasons for us to reduce the risk of extinction within our own time are significantly weaker, since a long and flourishing future for humanity is much less likely to obtain than it would be if future people could be counted on to be more cautious in safeguarding humanity's potential.

Therefore, when we consider the value of the far future as a source of moral demands, the extreme demandingness of utilitarianism apparently cannot be blamed on the failure of the theory to moderate its demands in the face of others' non-compliance. A theory of beneficence that indexes what is required of us under conditions of imperfect compliance to our fair share of the total effort defined under conditions of perfect compliance would be no less demanding. Or so it seems.

Our discussion so far has only taken account of the compliance of future people, however. What about past people? Consider the issue of optimal saving. We clearly do not observe perfect compliance with the utilitarian principle of accumulation over the course of human history. Had previous generations complied, we would presumably have been better off now. Suppose that any viable principle of beneficence must satisfy a ‘compliance condition’ such that “the demands on a complying person should not exceed what they would be under full compliance with the principle” (Murphy 2000: 7). Then, we may argue, we are not required to save nearly as great a percentage of total output as we would have been required to save under conditions of full adherence to the utilitarian principle of accumulation, because in doing so we would render ourselves worse off than we would have been had all previous generations adhered to that principle. In a similar way, we might conjecture that a demand on the current generation to optimally promote existential security would not be nearly so onerous if not for the recklessness of previous generations who pushed ahead with dangerous technologies like nuclear weapons and bequeathed to us a suite of global institutions that might again allow our technological capabilities to outrun our collective wisdom, be it in the form of poorly regulated synthetic biology or mismanaged applications of advanced machine intelligence (Ord 2020). In this way, taking account of imperfect compliance on behalf of previous generations may be thought to mitigate the demands on the current generation to sacrifice for the sake of the future.

While plausible on its face, in adjusting the demands of beneficence to take account of the non-compliance of past people, this line of reasoning turns out to conflict with the theoretical justifications for imposing a ‘compliance condition’ on principles of beneficence that have been offered by both Cohen (1981) and Murphy (2000).

When Murphy spells out his ‘compliance condition’ fully, it takes on an explicitly forward-looking character: “a person’s maximum level of required sacrifice is that which will reduce her expected well-being to the level it would be, all other aspects of her situation remaining the same, if there were to be full compliance *from that point on.*” (117) (My emphasis.) This is no accident. The forward-looking character of the constraint follows from its underlying theoretical justification (Murphy 2000: 114-5). The idea that motivates the imposition of the ‘compliance condition’ is that when I know that you will not perform your duty, I am not to treat you like a force of nature whose failure to behave optimally obligates me to respond appropriately, but as an autonomous moral agent whose responsibilities are still hers to fulfil

even when she refuses. Thus, even when I know you will not perform your duty, I am not required to fulfil it instead, since it is your duty, and not mine. This presumes, of course, that the obligations at issue are ones that you (or I) could as yet fulfil. Irrevocable failures to fulfil past obligations are therefore not covered by the 'compliance condition'.

The justification proposed by Cohen (1981) for thinking that the demands of beneficence should not rise in response to others' non-compliance also appears to support a purely forward-looking application of the constraint, although he does not say so explicitly. Cohen relies on the methodological principle that we have reason for thinking that a moral rule is unsound if the consequences of its widespread acceptance would be bad. He argues that widespread acceptance of a moral principle that requires people's contributions to rise in the face of others' non-compliance will have bad effects, because it would incentivize people not to contribute, under the belief that others will cover for them. When it comes to past people who are dead and gone, it presumably makes no sense to worry that we would be incentivizing them not to do their part by adopting moral principles that require us to do more because they did less. Cohen's argument therefore also seems to presuppose that we are dealing with responsibilities that people still should and could take on, to the exclusion of irrevocably lapsed past duties.

This clearly does not settle the matter. Alternative principled justifications may emerge. Nonetheless, we should conclude that reflecting on the value of the far future requires us to rethink the nature and significance of a 'compliance condition' on demands of beneficence. Without the emergence of a new theoretical justification for adjusting the 'compliance condition' to be backward-looking as well as forward-looking, it seems plausible that the significance of non-compliance is the inverse of what it has generally been thought to be when we reflect on the demands of beneficence, with these demands becoming more and more burdensome the nearer we approach conditions of perfect intergenerational compliance (starting from the present time).

This section has considered responses to the demandingness problem that treat a person as obligated to do no more than her fair share of the collective effort. I have shown that extant theories of this kind are toothless in the face of demands of beneficence that stem from the value of the long-term future.

7.

In this penultimate section, I want us to focus on the role played by *passive effects* in assessing the demandingness of moral theories: i.e., benefits and costs conferred on individuals, not as a result of their compliance with a moral theory's demands, but as a result of others' compliance. It has been argued that taking account of passive effects (more than) levels the playing field with respect to the demandingness of utilitarianism vis-à-vis other moral theories. While this may be true when we consider demands to alleviate the suffering of people in extreme poverty, it does not hold true when we reflect on the value of the far future.

As Murphy (2000: 47-61) and Sobel (2007) have emphasized, discussions of moral demandingness tend to neglect passive effects, focusing almost exclusively on the active demands of a theory: i.e., benefits and costs conferred on individuals as a result of their compliance with the theory's demands. If not for this, they claim, utilitarianism would not seem so demanding when compared to other theories. Other theories permit us to spend money on frivolous luxuries while some people's basic needs go unmet: people starve to death, children are killed by malaria, and so on. These other theories thus ask people living in poverty to accept very heavy burdens. The active demands put on wealthy Westerners by utilitarianism are not nearly so harsh. From this, Sobel infers that the demandingness objection to utilitarianism must "presuppose that people have greater claim against aiding others than they have for claiming aid from others" (8). Sobel argues that this presupposition begs the question against consequentialists, who reject the doing/allowing distinction. He concludes that the demandingness objection is ineffective as a complaint against utilitarian moral theories.

Sobel's argument is on shakier ground when we take on board the idea that agents who comply with utilitarian demands of beneficence will focus their efforts on improving the value of the long-term future. That is because the Non-Identity Problem makes it hard to see other theories as being more demanding on the intended beneficiaries of actions that comply with the demands of utilitarianism: namely, future people. (Arguably, it makes it hard to speak of such people as 'beneficiaries'.) If we do succeed in positively affecting the long-term future of Earth-originating civilization, our actions will almost certainly change the size and/or composition of the future population. If we do not perform these actions, the outcome may be worse, but there may be no one for whom it is worse. Future people may have a lower quality of life, but those same people would not have had a higher quality of life had we chosen otherwise.

It seems mis-placed to speak of a theory that permits us to bring about these suboptimal future outcomes as imposing heavy burdens on the people who exist in those outcomes, unless they have lives that are not worth living.

Therefore, once we keep in mind that utilitarianism orients the demands of beneficence toward improving the long-term future, the claim that the heavy active demands of the theory must be weighed alongside the heavier passive demands of other theories begins to seem dubious. It is hard to escape the impression that utilitarianism is extremely demanding relative to other theories not only in terms of its active demands, but also in terms of its passive demands. Other theories at least permit us to help people living right now who are very badly off. Utilitarianism imposes very heavy burdens on such people by requiring those who could help them to direct their energies elsewhere.

My discussion so far has neglected one of the ways in which we can improve the value of the future: namely, by working toward beneficial trajectory changes that result in a smaller number of future lives falling below the threshold at which a life is not worth living. When a less demanding theory permits us to allow a suboptimal outcome to come about in which there are more people whose lives are not worth living, it does not seem mis-placed to speak of the less demanding theory as imposing heavy burdens on some of the people who exist in those outcomes: namely, those whose lives are so bad that they are not worth living.

Now, if utilitarianism is combined with the Axiological Asymmetry, then insofar as the future has overwhelming importance, its importance reduces to that of preventing the existence of astronomically many lives in which suffering predominates. On the other hand, some of the activities recommended as among the optimal means of improving the value of the long-run future given a Total Utilitarian axiology increase the expected number of people with lives that are not worth living. Consider extinction risk mitigation. Insofar as we increase the probability that the future population will be astronomically greater than it could otherwise have been, we presumably greatly increase the total number of people throughout all time whom we expect to experience lives that are not worth living. Even if the vast majority of people who live in future have lives that are very good, there will be some small percentage of the population whose lives are net negative due to sheer bad luck. If humanity persists for trillions of years, there will be extraordinary numbers of such unlucky individuals. Insofar as actions that bring into existence individuals with lives not worth living count as imposing heavy

burdens on those individuals, actions that counteract threats to the continued survival of our species may count as imposing heavy burdens on very large groups of people.

Therefore, insofar as we presuppose a Total Utilitarian axiology, Sobel's criticism of the demandingness objection falls flat. The demandingness objection to utilitarianism need not presuppose that a theory's active demands matter more than its passive demands, since utilitarianism turns out to be more demanding than other theories whether we focus on the former or the latter, so long as we keep in mind what utilitarianism in fact concretely demands of us. No questions need be begged against the theory.

I conclude that whereas highlighting the passive demands of a theory may represent a forceful dialectical move when we reflect on the demands of beneficence in relation to extreme poverty and ignore the value of the far future, it does not succeed when we cease to ignore the value of the far future. It may even be said to backfire.

8.

I expect there is much more to be said on this issue. My aim has not been to provide an exhaustive catalogue of key contrasts and reversals, but to highlight enough to make clear that we have either misunderstood the problem of moral demandingness by neglecting the value of the long-run future, or at least failed to recognize important dimensions of the demandingness problem that come into view when we keep the value of the long-run future in mind. A key error has been to assume that utilitarianism directs us to maximize welfare within a generation by transferring resources to people currently living in extreme poverty. In fact, utilitarianism seems to imply that any obligation to help people who are currently badly off is trumped by obligations to undertake actions directly targeted at improving the value of the long-term future.

Once we take on-board this conclusion, key aspects of the problem of moral demandingness take on a very different character. Simply allowing the agent to weight her own welfare more heavily than the welfare of other people avoids the imposition of extreme demands only if we introduce weights that are obscene. The demandingness of utilitarianism can no longer be so easily blamed on unfavourable circumstances. The significance of imperfect compliance could well be the opposite of what we thought. Taking account of passive effects in estimating the demandingness of a theory no longer favours utilitarianism; the opposite may

be true. If we hope to gain a concrete understanding of the demands of beneficence and their moral significance for how we conduct our lives, it is clear we need to think again.

Setting aside the fact that additional complications are likely to emerge upon further reflection, does utilitarianism emerge as more or less plausible once the demandingness problem is reconceptualized in the ways I have discussed in sections 4 through 7? Clearly, there is no overall trend: the points raised in sections 4 and 6 tell against the plausibility of rival theories of beneficence; the points raised in section 5 and 7 raise significant doubts about some responses that have been made to the demandingness objection on behalf of utilitarianism. Holistic conclusions must depend on what weight we give to the different issues covered here. For my own part, I have up to now been especially struck by the rejoinder that other moral theories impose extreme passive demands on people in dire poverty, and the loss of this response to the demandingness objection on behalf of utilitarianism inclines me to believe that the theory is less plausible, all things considered. But I know that others do not share this attitude, and will judge the matter differently.

Bibliography

- Arneson, Richard (2004) Moral limits on the demands of beneficence? In Chatterjee, ed. *The ethics of assistance: morality, affluence, and the distant needy*, 33-58. Cambridge: Cambridge University Press.
- Arneson, Richard (2009) What do we owe to distant needy strangers? In Schaler, ed. *Peter Singer under fire*, 267-293. Chicago and La Salle, IL: Open Court.
- Arrhenius, Gustaf (MS) *Future generations: a challenge to moral theory*. Unpublished manuscript.
- Arrow, Kenneth (1996) Discounting, morality, and gaming. Working Papers 97004, Stanford University, Department of Economics.
- Asheim, Geir and Buchholz, Wolfgang (2003) The malleability of undiscounted utilitarianism as a criterion of intergenerational justice. *Economica* 70, 405-22.
- Ashford, Elizabeth (2000) Utilitarianism, integrity, and partiality. *Journal of Philosophy* 97, 421-439.
- Ashford, Elizabeth (2003) The demandingness of Scanlon's contractualism. *Ethics* 113, 273-302.
- Bader, Ralf (2020) Person-affecting utilitarianism. In Arrhenius, Bykvist, Campbell, and Finneron-Burns, eds. *Oxford handbook of population ethics*. Oxford: Oxford University Press.
- Barry, Brian (1977) Justice between generations. In Hacker and Raz, eds. *Law, morality, and society: essays in honour of H.L.A. Hart*, 268-284. Oxford: Clarendon Press.
- Barry, Brian (1997) Sustainability and intergenerational justice. *Theoria* 45, 43-65.

- Beckstead, Nick (2013) *On the overwhelming importance of shaping the far future*. PhD thesis: Department of Philosophy, Rutgers University. <<https://rucore.libraries.rutgers.edu/rutgers-lib/40469/PDF/1/play/>>
- Beckstead, Nick (2019) A brief argument for the overwhelming importance of shaping the far future. In Greaves and Pummer, eds. *Effective altruism: philosophical perspectives*, 80-98. Oxford: Oxford University Press.
- Blackorby, Charles, Walter Bossert, and David Donaldson (2005) *Population issues in social choice theory, welfare economics, and ethics*. Cambridge: Cambridge University Press.
- Bostrom, Nick (2003) Astronomical waste: the opportunity cost of delayed technological development. *Utilitas* 15, 308-14.
- Bostrom, Nick (2013) Existential risk as a global priority. *Global Policy* 4, 15-31.
- Broome, John (1996) The welfare economics of population. *Oxford Economic Papers* 48, 177-93.
- Broome, John (2004) *Weighing lives*. Oxford: Oxford University Press.
- Broome, John (2012) *Climate matters: ethics in a warming world*. London: W. W. Norton & Company.
- Caney, Simon (2014) Climate change, intergenerational equity and the social discount rate. *Politics, Philosophy & Economics* 13, 320-42.
- Caney, Simon (2018) Justice and future generations. *Annual Review of Political Science* 21, 475-493
- Cingano, Federico (2014) Trends in income inequality and its impact on economic growth. OECD Social, Employment and Migration Working Papers, No. 163, OECD Publishing, Paris.
- Cohen, Jonathan L. (1981) Who is starving whom? *Theoria* 2, 65-81.
- Cowen, Tyler (2018) *Stubborn attachments: a vision for a society of free, prosperous, and responsible individuals*. San Francisco, CA: Stripe Press.
- Cullity, Garrett (2004) *The moral demands of affluence*. Oxford: Oxford University Press.
- Dasgupta, Partha (2008) Discounting climate change. *Journal of Risk and Uncertainty* 37, 141-69.
- DeLong, Brad (2006) Partha Dasgupta makes a mistake. <https://delong.typepad.com/sdj/2006/11/partha_dasgaptu.html> Last accessed 26.09.2019
- Driver, Julia (2006) *Ethics: the fundamentals*. Oxford: Blackwell.
- Drupp, Moritz, Mark Freeman, Ben Groom, and Frikk Nejse (2018) Discounting disentangled. *American Economic Journal: Economic Policy* 10, 109-34.
- Fleurbaey, Marc and Zuber, Stephane (2013) Climate policies deserve a negative discount rate. *Chicago Journal of International Law* 13, 565-95.
- Frick, Johann (2014) 'Making people happy, not making happy people': a defence of the asymmetry intuition in population ethics. Doctoral dissertation, Harvard University.
- Friedman, Milton (1953) The methodology of positive economics. In Hausman, ed. *The philosophy of economics*, 145-78. Cambridge: Cambridge University Press.

- Greaves, Hilary (2017) Population axiology. *Philosophy Compass* 12
- Greaves, Hilary and MacAskill, William (2019) The case for strong longtermism. Global Priorities Institute Working Paper < https://globalprioritiesinstitute.org/wp-content/uploads/2019/Greaves_MacAskill_The_Case_for_Strong_Longtermism.pdf>
- Hurka, Thomas (1983) Value and population size. *Ethics* 93, 496-507.
- Kagan, Shelly (1989) *The limits of morality*. Oxford: Oxford University Press.
- Kagan, Shelly (1997) *Normative ethics*. Abingdon: Routledge.
- McMahan, Jeff (1981) Problems of population theory. *Ethics* 92, 96-127.
- McMahan, Jeff (2013) Causing people to exist and saving lives. *The Journal of Ethics* 17, 5-35
- Meyer, Lukas and Roser, Dominic (2009) Enough for the future. In Gosseries and Meyer eds., *Intergenerational justice*, 219-48. Oxford: Oxford University Press.
- Mirrlees, James (1967) Optimum growth when technology is changing. *Review of Economic Studies* 34, 95-124.
- Mulgan, Tim (2001) *The demands of consequentialism*. Oxford: Oxford University Press.
- Mulgan, Tim (2006) *Future people*. Oxford: Oxford University Press.
- Murphy, Liam (1993) The demands of beneficence. *Philosophy and Public Affairs* 22, 267-292.
- Murphy, Liam (2000) *Moral demands in nonideal theory*. Oxford: Oxford University Press.
- Ng, Yew-Kwang (1989) What should we do about future generations? Impossibility of Parfit's Theory X. *Economics and Philosophy* 5, 235-253.
- Noggle, Robert (2009) Give till it hurts? Beneficence, imperfect duties, and a moderate response to the aid question. *Journal of Social Philosophy* 40, 1 – 16.
- Norcross, Alastair (2006) The scalar approach to utilitarianism. In West, ed. *The Blackwell Guide to Mill's Utilitarianism*, 217-32. Oxford: Blackwell.
- Nordhaus, William (2007) A review of the the Stern Review on the economics of climate change. *Journal of Economic Literature* 45, 686-702
- Norwood. F. Bailey and Lusk, Jason (2011) *Compassion by the pound: the economics of farm animal welfare*. Oxford: Oxford University Press.
- Ord, Toby (2020) *The precipice: existential risk and the future of humanity*. London: Bloomsbury.
- Ostry, Jonathan, Andrew Berg, and Charalambos Tsangarides (2014) Redistribution, inequality, and growth. IMF Staff Discussion Note <<https://www.imf.org/external/pubs/ft/sdn/2014/sdn1402.pdf>>
- Parfit, Derek (1984) *Reasons and persons*. Oxford: Oxford University Press.
- Piketty, Thomas (2014) *Capital in the 21st century*. Cambridge, MA: Harvard University Press.
- Railton, Peter (1984) Alienation, consequentialism, and the demands of morality. *Philosophy and Public Affairs* 13, 134-71.
- Ramsey, F. P. (1928) A mathematical theory of saving. *The Economic Journal* 38, 543-59.

- Rawls, John (1999) *A theory of justice, revised edition*. Cambridge, MA: Harvard University Press.
- Rendall, Matthew (2019) Discounting, climate change, and the ecological fallacy. *Ethics* 129, 441-63.
- Roberts, Melinda (2011) An asymmetry in the ethics of procreation. *Philosophy Compass* 6, 765-776.
- Robinson, Timothy, G. R. William Wint, Giulia Conchedda, Thomas Van Boeckel, Valentini Ercoli, Elisa Palamara, Giuseppina Cinardi, Laura D'Aietti, Simon Hay, and Marius Gilbert (2014) Mapping the global distribution of livestock. *PLoS One* 9(5): e96084
- Scheffler, Samuel (1992) *Human morality*. Oxford: Oxford University Press.
- Scheffler, Samuel (1994) *The rejection of consequentialism: a philosophical investigation of the considerations underlying rival moral conceptions, revised edition*. Oxford: Oxford University Press.
- Schelling, Thomas (1995) Intergenerational discounting. *Energy Policy* 23, 395-401.
- Shiffrin, Seanna (1991) Moral autonomy and agent-centred options. *Analysis* 51, 244-54.
- Sidgwick, Henry (1874) *The Methods of Ethics*. London: MacMillan.
- Singer, Peter (1973) Famine, affluence, and morality. *Philosophy & Public Affairs* 1, 229-243.
- Singer, Peter (1993) *Practical ethics, 2nd ed.* Cambridge: Cambridge University Press.
- Slote, Michael (1985) *Common-sense morality and consequentialism*. London: Routledge and Kegan Paul.
- Sobel, David (2007) The impotence of the demandingness objection. *Philosophers' Imprint* 7: 1-17.
- Steinbock, Bonnie (1986) The logical case for 'wrongful life.' *Hastings Centre Report* 16, 15-20.
- Steiner, Hillel and Vallentyne, Peter (2009) Libertarian theories of intergenerational justice. In Gosseries and Meyer eds., *Intergenerational justice*. Oxford: Oxford University Press.
- Stern, Nicholas (2006) *The economics of climate change*. Cambridge: Cambridge University Press.
- Stern, Nicholas (2008) The economics of climate change. *American Economic Review* 98, 1-37.
- Tarsney, Christian (2019) The epistemic challenge to longtermism. Global Priorities Institute Working Paper <https://globalprioritiesinstitute.org/wp-content/uploads/2019/Tarsney_Epistemic_Challenge_to_Longtermism.pdf>
- Unger, Peter (1996) *Living high and letting die: our illusion of innocence*. Oxford: Oxford University Press.
- Zuber, Stéphane and Asheim, Geir (2012) Justifying social discounting: the rank-discounted utilitarian approach. *Journal of Economic Theory* 147, 1572-1601.