

Individual Time-Bias and Social Discounting

to appear in Carla Bagnoli (ed), *Time in Action*, Routledge

Abstract

Consider two questions about appropriate attitudes to time: Within a single life, is it permissible to weight the well-being of one's near future selves more heavily than that of one's farther future selves? And as a society, is it permissible to weight the well-being of near-future people more heavily than that of farther future people? While many economists and philosophers have suggested that these two questions are independent, so that our answer to one does not tightly constraint our answer to the other, I suggest that they should be treated in parallel, so that individual time-bias is permissible if and only if social discounting is permissible.

1. Introduction

Many of us, as individuals, seem to care more about our nearer futures than our further futures. Consider a choice between getting one chocolate bar now versus getting two chocolate bars in a month's time. If you would prefer getting the one chocolate bar now, that is indicative of your being *time-biased*, and in particular *biased toward the near*, to use Parfit's (1984) helpful phrase. Similarly, consider a choice between a less painful operation tomorrow and a more painful operation a year from now. Again, if you would prefer the latter over the former, that is indicative of your being time-biased.

On a social level, many policy decisions are and should be made using cost-benefit analysis. When the costs and benefits accrue at different times, as is typically the case, we can conceptualize the procedure as first aggregating the costs and benefits *at each time*, and then aggregating those net costs or benefits *over time*. There are significant difficulties even with aggregating costs and benefits at a single time (having to do in particular with issues of

incommensurable values), but setting this problem aside, how should costs and benefits be aggregated across time? Economists typically employ a *social discount rate*, which gives costs and benefits accruing in the farther future less weight than those accruing in the near future.

In this paper, I am concerned with so-called *pure time preference* or *pure discounting*, in which it is well-being itself that is discounted. I therefore set aside other reasons why one might seem to give greater weight to the near future than the far future, having to do with uncertainty, decreasing marginal utility for resources and anticipated wealth gains, and what John Broome (1994, 139) calls ‘the fertility of technology.’¹ In what follows, I use ‘individual time-bias’ and ‘social discounting’ to refer only to the discounting of well-being itself with respect to time.

My focus in this paper is about the relation between the normative status of individual time-bias and the normative status of social discounting. Does what we say about individual time-bias constrain what we should say about social discounting, and *vice versa*? In Section 2, I show that while earlier theorists took them to be closely connected, there has emerged a near-consensus among economists and philosophers that the two issues are largely independent, so that the normative status of the one entails nothing about the normative status of the other. This is puzzling, for as I suggest in Section 3, the strongest arguments for and against the permissibility of individual time-bias carry over to the case of social discounting, and *vice versa*. And if the same arguments apply with equal strength in both the individual and the social case, then there is strong pressure to draw the same conclusion about each and thereby conclude that individual time-bias and social discounting have the same normative status. In Section 4, I consider a possible disanalogy between individual time-bias and social discounting, namely

¹ ‘Given the fertility of technology, it is often better to receive a given amount of resources sooner rather than later, since those resources can be put to work and yield greater wealth in the meantime. For instance, it is better to receive \$1,000 today rather than even \$1,000 inflation-adjusted dollars in a year, because if receiving the money now means you can invest it in stocks or start a business and thereby make substantial gains of the course of the year.

that one's treatment of one's own future selves falls only within the domain of rationality, whereas our treatment of future generations falls within the domains of morality and justice. Drawing on recent work in ethics, I reject this disanalogy and suggest that one's treatment of one's future selves can in fact be immoral or unjust. I conclude that individual time-bias and social discounting should in fact be treated in parallel and that they have the same normative status.

2. The Consensus

Some theorists have thought that whether people are, or permissibly may be, individually time-biased has implications for what the proper social discount rate is. Sidgwick (1907) endorsed a position on which both individual time-bias and social discounting are impermissible, and for the same reason, namely that they fall afoul of a principle of aggregation. Here is Rawls (1999 [1971], 259) discussing Sidgwick's view in the context of the debate over optimal societal rates of saving:

just as the good of one person is constructed by comparison and integration of the different goods at each moment as they follow one another in time, so the universal good is constructed by the comparison and integration of the good of many different individuals. The relations of the parts to the whole and to each other are analogous in each case, being founded on the aggregative principle of utility. The just savings principle for society must not, then, be affected by pure time preference, since as before the different temporal position of persons and generations does not justify treating them differently.

Other theorists have thought that a non-zero social discount rate is justified, and moreover that it is justified in part because individuals themselves are, or may permissibly be, time-biased with respect to their own well-being. Schelling (2000, 833) expresses the view as follows, though he himself rejects it:

[T]here is a near consensus among these these economists that the appropriate discount rate should be conceptualised as consisting of two components...The first is pure time preference and...deals with the impatience of consumers and reflects their inborn preferences of immediate over postponed consumption...The second component reflects the changing marginal utility of consumption with the passage of time...²

Despite these alleged connections between individual time-bias and social discounting, there is now a near-consensus that the normative statuses of the two are independent. Expressing this common sentiment, Cowen and Parfit (1992) write,

Even if this attitude [time-bias] is not irrational, it cannot justify an intergenerational discount rate. Perhaps individuals may rationally prefer smaller benefits, because they are in the nearer future. But this argument has no next step. Pure time preference within a single life does not imply pure time preference across different lives.

Statements of the same view can be found in Cline (1992) and Greaves (2017). Unfortunately, these authors do not give further argument to back up the claim that the two issues are independent. Cowen (ms, 5) goes further. He writes that ‘Time preference within a life, however, cannot be extrapolated directly to time preference across different lives.’ His argument seems to be that individual time-bias derives from impatience, or a distaste for waiting. Within a single life, having a benefit accrue later rather than sooner does involve that individual's having to wait. But across lives, when a benefit accrues to a future person rather than a presently existing one, it is not the case that any waiting is involved. As Cowen (ibid, 6) notes, ‘The passage of time before our births does not involve waiting.’

² This quote is alluding to the so-called Ramsey formula, devised in Frank Ramsey's pioneering work on optimal rates of national saving. Ramsey himself endorse a zero rate of pure time preference, though later economists employed a positive rate of social pure time preference derived from evidence about individuals' own rates of pure time preference. Note that one could also endorse a positive rate of social pure time preference that is not derived from, or intended to reflect, individuals' own rates of pure time preference.

Along similar lines, Schelling (2000, 834) argues that ‘Any time preference pertinent to discounting the long-term benefits of greenhouse gas abatement cannot have anything to do with impatience. The alleged inborn preference for earlier rather than later consumption is exclusively concerned with the consumer's impatience with respect to his or her *own* consumption.’

As we will see, however, we need not think of individual time-bias as being rooted in impatience or a distaste for waiting, and so the fact that impatience does not extend across lives fails to show that the permissibility of individual time-bias is independent of the permissibility of social discounting.

Rawls (1971) also endorses the claim that individual time-bias and social discounting are independent, though he does so for quite different reasons. Rawls maintains that both individual time-bias and social discounting are impermissible, but he thinks (*contra* Sidgwick) that the reasons for the impermissibility are different in the two cases. Unlike Sidgwick, Rawls (1999 [1971], 259) thinks that since ‘the principles of justice are not extensions of the principles of rational choice for one person, the argument against [social discounting] must be of another kind’ from that of Sidgwick. He proceeds to give an argument against social discounting based on his own, non-utilitarian, theory of justice.³

³ Rawls holds that the participants in his original position would not consent to the application of any positive pure social discount rate. This is because they are subject to the veil of ignorance and hence will not know their temporal position with respect to other generations. They will therefore ‘not consent to any principle that weighs nearer periods more or less heavily,’ since ‘to acknowledge a principle of time preference is to authorize persons differently situated temporally to assess one another’s claims by different weights based solely on this contingency’ (1999 [1971], 260). Even given Rawls’ framework, I am unconvinced by this argument, for reasons outlined by Broome (1992, 96-8). Broome points out that exponential social discounting, described below in section 3.1, does not ‘authorize persons differently situated temporally to assess one another’s claims by different weights based solely on this contingency.’ This is because, with exponential discounting, the relative weights assigned to well-being at different times remain constant, and so persons differently situated temporally would be required to agree in the weights they use to assess one another’s claims. Moreover, Broome thinks there are positive reasons why those in the original position might accept the use of exponential social discounting, supposing history has no beginning and no end: ‘Compared with an impartial principle, exponential discounting treats each generation less favourably relative to its predecessors. But in compensation, it treats each more favourably relative to its successors. And it has the advantage, compared with

3. Parallel Arguments

Is this consensus correct? Is the normative status of individual time-bias really independent of the normative status of social discounting? One way to answer this question is to look at the arguments that have been advanced for or against the permissibility of the one, and to see whether each such argument applies with equal force for or against the permissibility of the latter. That is the tack I will pursue in this section. And while I cannot hope to survey every possible argument that has been or might be advanced in this domain, I will suggest that the strongest arguments for and against the permissibility of individual time-bias carry over with equal force to the case of social discounting, and *vice versa*. This yields good, though certainly not decisive, inductive grounds for concluding that they have the same normative status.

In some cases, the arguments that have been advanced in the context of individual time-bias have also been advanced in the context of social discounting. This is the case for arguments against the permissibility of these attitudes. But in other cases, the parallelism has not been recognized. In particular, it has not been recognized that prominent arguments in favor of the permissibility of individual time-bias carry over to the case of social discounting. Indeed, these arguments are, to my knowledge, known only in the philosophical literature and have not been discussed by economists interested in either individual time-bias or social discounting. Thus, bringing these powerful arguments in favor of individual time-bias to the attention of theorists interested in social discounting is one of the aims of this paper.

3.1 Arguments against Permissibility

There are two main arguments against the permissibility of both individual time-bias and social discounting. They are simple but powerful. The first appeals to arbitrariness. In the

an impartial principle, of putting less strain on each generation's self-control.' (97-8). Thus, it is not clear that Rawls' own argument against social discounting works, even granting the background contractualist framework.

individual case, the thought goes that it is arbitrary to care differently about your different time-slices merely due to their differing locations in time. Well-being is well-being, no matter when it occurs, and so individuals should not discount their own well-being with respect to time.

It is easy to see that the same argument applies in the case of social discounting, and indeed it has been made by numerous philosophers and economists (see Sidgwick 1907, Ramsey 1928, Pigou 1932, and Cline 1992, among others). The well-being of future people is still well-being and so should not be treated as less important than the well-being of presently existing people merely due to its temporal location.

The second argument appeals to diachronic inconsistency. If you as an individual are time-biased, then your attitudes will be diachronically inconsistent (unless your time-bias takes a specific form). Suppose that at all times, you prefer one chocolate bar right away over two a month from then. But you also prefer two chocolate bars 13 months in the future over one chocolate bar 12 months in the future. Then, your preferences will shift with the passage of time. Initially, you prefer the two chocolate bars 13 months in the future over one in 12 months time, but as the date approaches, you will come to prefer the latter over the former.

This is potentially problematic, since it can lead to self-defeating courses of action. Suppose it is now January 1, 2019. You start off with a ticket entitling you to one chocolate bar on January 1, 2020. I offer you the option of exchanging that ticket for one entitling you to two chocolate bars on February 1, 2020, at the cost of a small fee. You accept, since you prefer the latter ticket to the former. But being time-biased, your preferences switch, and by late December 2019, you prefer the first ticket over the second. I offer you a chance to make the switch, again for a small fee, and you accept. You thereby wind up with the same ticket with which you started, minus the fees. This is a sort of diachronic exploitability familiar from the money pump argument for transitivity of preferences and diachronic Dutch Book arguments in Bayesian epistemology. Of course, the same diachronic exploitation can take place on a social

level. We can imagine a society or state that discounts the future in an analogous way that results in vulnerability to self-defeating courses of action.

On both an individual and a societal level, such vulnerability to diachronic exploitation can be avoided by discounting the future *exponentially*. Being an exponential discounter amounts to having a constant discount rate (not discounting the future at all trivially qualifies as a form of exponential discounting). As Parfit (1984, 160) puts it, a person discounts exponentially if she discounts ‘at a constant rate of n per cent per month. There will always be the *same* proportionate difference in how much this person cares about two future events.’ That is, the proportionate difference between the weight assigned to two different times depends only on how far apart in time they are, and not on how far either is from the present time. For instance, if (individual or societal) well-being today is assigned twice the weight of well-being a year from now, then exponential discounting requires also assigning well-being ten years from now twice the weight of well-being eleven years from now. In an important result, Strotz (1955-6) proves that exponential discounting will not lead to any exploitable shifts in preferences over time.

But while exponential discounting avoids diachronic exploitability, it nonetheless has problematic implications. First, exponential discounting with some positive discount rate involves not only caring more about the near future than the far future, but also caring more about the distant past than the recent past. After all, exponential discounting involves having the proportionate difference in the weights assigned to well-being at one time vs. another depend only on how far apart those points in times are, and not on their location relative to the present. In response, one might propose a modified version of exponential discounting on which one treats all past times the same and applies a constant discount rate to future times only. But this move reintroduces diachronic inconsistency, with preferences switching due to the mere passage of time. It is tempting to think that this particular sort of diachronic

inconsistency is innocuous and hence untroubling. Greaves (2017, 406) writes that ‘the *only* sort of inconsistency that can result from the discounting structure in question is the phenomenon of foreseeable regret’ and that it will not lead one to pursue self-defeating courses of action. But as we will see in Section 3.3, this tempting thought is mistaken; in certain circumstances, treating the past differently from the future can yield diachronic exploitability.

Second, exponential discounting leads to extreme differences in the weights assigned to well-being at different times, when those times are far enough apart. As Broome (2013, 150) notes, discounting at a rate of 1% per year entails that the ‘7,000 casualties of the battle of Marathon in 490 BC work out to be far, far worse than would be the slaughter of every single person alive on Earth today.’ And of course these extreme implications of exponential discounting apply not only respect to past times, but also with respect to the future.

In any event, individual time-bias or social discounting that takes a non-exponential form is vulnerable to diachronic exploitation, and even exponential discounting may be unattractive on these independent grounds.

3.2 Arguments for Permissibility: Demandingness

We now turn to arguments in the other direction. In this subsection, I consider one argument in favor of social discounting, and argue that it applies also to the case of individual time-bias. The argument is that a positive (pure) discount rate is needed in order to prevent our obligations from being overly demanding. It can be overly demanding for two reasons.

First, consider optimal savings theory (Ramsey 1928). Assuming a broadly consequentialist aim of maximizing the good, how should each generation's income and wealth be allocated between immediate consumption, on the one hand, and savings and investment, on the other? Immediate consumption improves the well-being of those doing the consuming, while savings and investment improve the well-being of those in the future. But as Arrow (1997, 1) notes, in standard economic models, ‘with zero time preference and a long horizon,

the [required] savings rates become inordinately high, possibly approaching one as the horizon goes to infinity (Koopmans 1960).’ The technical details are complex, but the basic point is simple. Suppose that some generation has the option of foregoing some consumption for the sake of investing in the future. With an infinite time horizon, ‘Each unit [of resources] sacrificed would yield a finite utility loss to the first generation, but to compensate there would be a gain, however small, to each of an infinity of generations’ (Arrow 1997, 5). Hence any sacrifice in consumption for the sake of saving and investing is on balance good, if we assume a zero discount rate. And even without an infinite time horizon, the optimal savings rate may nevertheless come out excessively high. Thus, a positive social discount rate is needed to avoid the implication that current people ought to take on extreme sacrifices for the sake of future generations.

Second, consider existential risks such as risks of the extinction of humanity from asteroids, gamma ray bursts, global pandemics, the expansion of the sun into a red giant phase, and the like (Bostrom and Ćirković 2008). The premature extinction of the human race at some time yields a massive loss of well-being, and so huge expenditures on existential risk reduction may be justified by even a small reduction in the probability of extinction (at some time). Arguably, this is grounds for thinking that we should be spending far more than we currently are on policies to mitigate these risks (Posner 2005, Bostrom 2013). But even advocates of increased efforts to mitigate existential risk may balk at the thought that we ought to spend, say, a majority of our budgets on such efforts. Posner (2005, 152), himself an advocate of increased mitigation efforts, argues:

[N]ot to discount future costs at all would be absurd...For then the present value of benefits conferred on our remote descendants would approach infinity. Measures taken today to arrest global warming would confer benefits not only in 2100 but in every subsequent year, perhaps for millions of years. The present value of \$100 billion

received every year for a million years at a discount rate of 0 percent is \$100 quadrillion, which is more than even Greenpeace wants spent on limiting emissions of greenhouse gases.

In my view, this demandingness worry is not a good argument for a positive pure social discount rate. Insofar as we are worried about excessively demanding obligations, we would do better to reject the broadly consequentialist theory assumed as background. (It is worth noting in this regard that overdemandingness is a familiar objection to consequentialism, or at least utilitarianism, even setting aside long or infinite time horizons; see Scheffler 1982.) Perhaps some rights-based theory should be adopted in its place.⁴

Nevertheless, what I want to point out is that analogous motivations (even if they are not good ones) could also be used to support individual time-bias. As an analogue of the debate over the optimal social rate of savings, we can consider the optimal rate of savings for individuals. How much should individuals allocate to immediate consumption, and how much to saving and investment? If humans lived long enough, we would get the same sorts of results that worry Arrow in the social case, being required to make intuitively excessive sacrifices for the sake of our later selves. A positive discount rate with respect to one's own well-being would avoid this result. As an analogue of the point about existential risks, we can ask how much humans should spend (or more generally, what sacrifices they should undertake) to extend their lives or otherwise avoid death. Depending on the costs of various life-extending efforts and their probabilities of extending life by certain amounts of time, we may get the same sorts of results that worry Posner in the social case, being required to devote intuitively excessive portions of our incomes to pursuing gene therapies, telomere-protection efforts, and so on (not

⁴ See Kelleher (2017) for nuanced discussion of debates around incorporating considerations of rights and justice into the social discount rate.

to mention strict diet and exercise). Again, being time-biased would allow one to avoid these implications.

(It might be suggested that the demandingness argument for individual time-bias doesn't go through, since while society is potentially eternal, individuals' lives are necessarily finite. But in fact, society has a finite duration as well. The sun is projected to expand and turn into a red giant and sterilize the earth's biosphere in around 3.5 billion years. And even if humans manage to colonize other planets, the universe will eventually become too cold to support life (Adams 2008). And if we set aside these astrophysical considerations, it is not clear that there is any in-principle reason why an individual could not exist indefinitely.)

As with the social case, I doubt whether these demandingness considerations are good motivations for the permissibility of individual time-bias. But my main point is simply that even for this somewhat dubious motivation for social discounting, there is an analogous motivation for individual time-bias.⁵

3.3 Arguments for Permissibility: Future Bias

We turn now to some more compelling arguments. In this and the following two sections, we will look at arguments that have been advanced in support of the permissibility of individual time-bias. But they have not, to my knowledge, been taken note of in the literature

⁵ In a recent book, Scheffler (2013) explores a possible disanalogy between the individual and the social case which is relevant here. He argues that our sense of meaning and purpose in our lives depends on our belief that our society will continue, at least for the foreseeable future, but that it also depends on our belief in the finitude of our own lives. He considers a doomsday scenario in which we learn that humanity will become extinct in the near future (but after all presently existing people are dead) and suggests that in such a scenario, we would no longer value most of our projects, or at least not to the degree we currently do (and, moreover, he suggests that this loss of value would be an appropriate response to the anticipated extinction). By contrast, he suggests that if we were to become immortal, this would cause us to no longer value our projects to the same degree that we actually do (see also Williams 1973). If we do have these differing attitudes to human extinction (or the end of society as we know it) and to individual death, this suggests that the overdemandingness arguments for individual time-bias and social discounting may not be equally compelling. Overall, however, the implications for social discounting and individual time-bias of Scheffler's considerations are complex. In particular, they suggest a narrative, or at least non-aggregative, sort of value relevant to the social good and the individual good which is difficult to integrate into the standard welfare maximization framework presupposed in much discussion of these attitudes to time.

on social discounting. This is unfortunate, since once again these arguments carry over to the social case.

We have so far been dealing with time-bias understood as caring more about the near future than about the far future. This is what Parfit (1984) calls ‘bias toward the near.’ But there is another form of time-bias, which he calls ‘bias toward the future.’ You are biased toward the future if you prefer that your pleasures be in the future and your pains in the past, even if this means a somewhat worse lifetime pleasure to pain ratio. More generally, it involves weighting the well-being of your future selves more heavily than the well-being of your past selves.

There is a strong intuition that it is permissible, and perhaps even required, that one be biased toward the future (though see Dougherty (2011) and Greene and Sullivan (2015) for dissenting views; see also Hare (2015) for a survey).

If bias toward the future is indeed permissible, this puts pressure on the view that bias toward the near is impermissible. As Parfit (1984) emphasizes, many of the same reasons for thinking that bias toward the near is impermissible apply to bias toward the future as well. So there is a strong case for the conditional that if bias toward the near is impermissible, then bias toward the future is also impermissible.

First, bias toward the future, like bias toward the near, can be charged with arbitrariness. Well-being is well-being, no matter when it occurs. So it would be arbitrary to weight the well-being of various temporal selves differently due merely to their differing locations in time.

Second, bias toward the future, like (non-exponential) bias toward the near, yields diachronic inconsistency. Here is a case from Dougherty (2011) to illustrate the point. Suppose you must undergo one of two courses of surgery. The early course involves 4 hours of painful surgery on Tuesday and 1 hour of painful surgery on Thursday. The late course involves no surgery on Tuesday and 3 hours of painful surgery on Thursday. Being biased toward the future, on Monday you will prefer the late course over the early course, as it involves a lesser

amount of future pain relative to Monday. But on Wednesday you will prefer the early course over the late course, as it involves a lesser amount of future pain relative to Wednesday.

One might think that the diachronic inconsistency resulting from bias toward the future is innocuous since it is practically inert; you cannot affect whether some event occurs in the past versus the future. So this diachronic inconsistency, unlike the diachronic inconsistency associated with bias toward the near, will not leave you vulnerable to diachronic exploitation, or performing self-defeating courses of action. But Dougherty (2011) shows that this is not quite right. If you are biased toward the future and also risk averse in a certain way, then there are cases in which you will in fact be led to perform self-defeating courses of action.⁶

Moreover, while it is true that you cannot affect whether some experience of pleasure or pain occurs in the past versus the future, this may not be true of well-being more generally. Many non-hedonic theories of well-being entail that it is possible to affect the well-being of past selves.⁷ Consider desire-satisfactionism, according to which the satisfaction of your desires adds to your well-being, and their frustration subtracts from your well-being. Crucially, it is not *believing* some desire to be satisfied that adds to well-being, but rather its in fact being satisfied. Your past selves likely had desires about the future, and hence you may be able to affect whether their desires were in fact satisfied. If your younger self desired that you eventually travel to Africa, you can satisfy that desire by booking a trip.

The preceding suggests that if bias toward the near is impermissible, then bias toward the future is also impermissible. Parfit (1984) himself holds that both are impermissible. But one could instead apply *modus tollens* and hold that bias toward the future is permissible, and so bias toward the near is permissible as well. Thus, the arbitrariness and diachronic inconsistency arguments against bias toward the near must not be sound.

⁶ But see Greene and Sullivan (2015), who argue that the diachronic inconsistency is in fact due to the kind of risk aversion involved, rather than to the future bias as such.

⁷ Cf. Broome (2004, 46-7).

In any event, this argument—from the permissibility of future bias to the permissibility of near bias—applies to the social case as well. There is at least a strong intuition that policy-makers can permissibly assign less weight to the well-being of past people, even if they are in a position to causally affect their well-being by satisfying their future-directed desires. If such societal bias toward the future is permissible, this puts pressure on the widely held view that societal bias toward the near (i.e. pure social discounting) is impermissible. For the same sorts of reasons standardly appealed to in arguing for the impermissibility of societal bias toward the near (namely, arbitrariness and diachronic inconsistency) would condemn societal bias toward the future as well. So if societal bias toward the future is permissible, then societal bias toward the near should be permissible as well, absent some relevant disanalogy between the two. Thus, in both the individual and the social case, the intuition that bias toward the future is permissible supports the claim that bias toward the near is permissible as well.

3.4 Arguments for Permissibility: Bias in One's Favour

It is natural to think that it is permissible to favour one's own interests over others' interests, at least to some degree. But Parfit (1984) argues that this claim is in tension with the claim that individual time-bias is impermissible. After all, if it is permissible to be biased in one's own favour, why should it not also be permissible to be biased in one's own *current and near-term* favour?

Parfit (1984) calls this the *appeal to full relativity*. There are two natural views: that both person neutrality and temporal neutrality are required, and that neither is. But the intermediate position, that temporal-neutrality but not person-neutrality is required, is unstable and can be attacked from both sides.

Here it is worth noting that person-neutrality can be motivated by the same two considerations that motivate temporal neutrality. First, bias in one's own favour is arguably arbitrary. If well-being is well-being, no matter *when* it accrues, similarly well-being is well-

being, no matter *to whom* it accrues. Second, bias in one's own favour yields an analogue of diachronic inconsistency. In particular, it yields cases of interpersonal inconsistency, in which each person prefers performing her member of some set of actions, despite both wanting that they not collectively perform the set of actions as a whole. The Prisoner's Dilemma is one such case. Each prisoner prefers to defect, no matter what the other one does, but they both prefer that they both cooperate rather than both defect. Indeed, cases of diachronic inconsistency are structurally just like Prisoners Dilemmas, with your different time-slices or temporal selves as the prisoners (Hedden 2015a).⁸

Whether or not Parfit's Appeal to Full Relativity shows that bias toward the near is permissible, what is important for our purposes is that an analogue can be used to support the permissibility of social discounting. Pigou (1932) held that a government ought not only to promote the interests of current citizens but also to safeguard the interests of future citizens as well. He advocated a (pure) social discount of zero.

Marglin (1963) suggests (though it is unclear whether he ultimately endorses) an opposing view, on which government ought to ignore the interests of future citizens, except to the extent that those interests are taken into account by the preferences of current citizens:

I want the government's social welfare function to reflect only the preferences of present individuals. Whatever else democratic theory may or may not imply, I consider

⁸ Parfit (1984, 187) makes a related point about predictable regret, imagining the following accusation levelled against an agent who is biased toward the near:

You do not *now* regret your bias towards the near. But you *will*. When you pay the price—when you suffer the pain that you postponed at the cost of making it worse—you will wish that you did not care more about your nearer future. You will regret that you have this bias. It is irrational to do what you know that you will regret.

But he points out that an analogous objection can be targeted at one who is biased in his own favour: he may regret that in the past he had his bias towards the near. But this does not show that he must regret having this bias now. A similar claim applies to those who are self-interested. When a self-interested man pays the price imposed on him by the self-interested acts of others, he regrets the fact that these other people are self-interested. He regrets their bias in their own favour. But this does not lead him to regret this bias in himself.

it axiomatic that a democratic government reflects only the preferences of the individuals who are presently members of the body politic. (p. 97)

One might object to Marglin's suggestion on various grounds. But for present purposes, the important thing is that we have here an analogue of Parfit's Appeal to Full Relativity. Suppose we think that a government may permissibly favour the interests of its own citizens over those of foreigners. This view is in tension with the thought that a government must not favour the interests of current citizens over those of future citizens. As with the individual case, there are two natural views: that a government must be neutral both with respect to its own citizens versus foreigners and also with respect to its own current citizens versus future citizens, and alternatively that a government may permissibly favour its own citizens over foreigners and also may permissibly favour its own current (and perhaps near-future) citizens over future (or far-future) citizens. The intermediate position can be attacked from both sides, by defenders of full neutrality and by defenders of pure social discounting. Insofar as one is sympathetic to the position that governments may permissibly favour their own citizens, one should therefore likewise be sympathetic to the permissibility of social discounting.

3.5 Arguments for Permissibility: Similarity

We turn now to the final argument. As with the previous two, it comes from Parfit in the context of individual time-bias. Parfit (1984) defends reductionism about personal identity over time, the view that facts about whether an earlier person (or time-slice) and a later person (or time-slice) are identical (or part of the same temporally extended person) are not metaphysically deep facts, nor are they what matter to us in thinking about our own survival. Looking closely at Parfit's arguments for reductionism would take us too far afield, but a major role is played by puzzle cases in which the facts about identity over time are murky; that is, cases in which some event occurs, such that it is unclear and controversial whether the person who is around after the event is or is not the same as the person who was around beforehand.

Cases include teletransportation, operations which alter many but not all of one's physical and psychological characteristics (and where these changes seem to near some threshold such that it is a borderline case of whether or not the post-operative person is identical to the preceding person), and cases involving fission, such as where one enters a teletransporter and two, instead of one, perfect duplicates are created (so that it cannot be that both are identical to the pre-fission person).

Parfit thinks that what really matters to us are psychological connections with later time-slices. He proposes that:

The value to me of my relation to a resulting person depends both (1) on my degree of [psychological] connectedness [i.e. similarity in relevant respects] to this person, and (2) on the value, in my view, of this person's physical and psychological features. (p. 299)

He argues that this view justifies 'a new kind of discount rate' (314) which correlates with, but is not quite identical to, a discount rate with respect to time:

My concern for my future may correspond to the degree of connectedness between me now and myself in the future. Connectedness is one of the two relations that give me reasons to be specially concerned about my own future. It can be rational to care less, when one of the grounds for caring will hold to a lesser degree. Since connectedness is nearly always weaker over longer periods, I can rationally care less about my further future.

Again setting aside whether this view is right, my point is that it carries over to the social case. Indeed, Parfit explicates his reductionism about person identity with an analogous reductionism about the identity of nations over time:

A nation is in many ways unlike a person. Despite these differences, the identity of persons over time is, in its fundamental features, like the identity of nations over time. Both consist in nothing more than the holding over time of various connections, some of which are matters of degree. It is true that in my old age it will be just as much me. But this truth may be fairly compared with the truth that (say) modern Austria is still *just as much* Austria. A descendant of the Habsburg Emperors could justifiably call this truth trivial.

Insofar as Parfit's reductionism about personal identity, and concomitant emphasis on a notion of psychological connectedness that comes in degrees, can justify the permissibility of something very much like individual time-bias, similar views about nationhood can justify the permissibility of something very much like social discounting. In fact, Schelling (2000, 834) suggests (without endorsing) such a view:⁹

Actually, time may serve as a measure of 'distance.' The people who are going to be living in 2150 may be considered 'further away' than the people who will be living in 2050...In redistributing income via transfer payments—providing foreign aid, contributing to charity, and so forth—people are expected to differentiate, and *do* differentiate, among recipient peoples according to several kinds of distance or proximity. One is geographical... Another is political: East Coast Americans are more interested in the people of Los Angeles than in the people of Quebec. Yet another is cultural: Some people are closer in language, religion, and other kinds of heritage...Deciding whether one cares more about the people who will be alive in 2150 than the people who will be alive in 2050 is a little like deciding whether one cares more about people in one continent than in another, or about English-speaking people

⁹ Parfit (1984, 485) also considers such a view, but notes, correctly, that does not involve discounting with respect to time as such, but rather with respect to something that correlates with time.

more than people who speak other languages, or about those with whom one shares history and culture more than those who do not.

Insofar as we think it permissible to discount one's own future well-being according to degrees of psychological connectedness (which correlates with time), we should think it likewise permissible for a society to discount future generations' well-being according to degrees of political or cultural similarity (which correlates with time). Once again, the individual and social cases are parallel.

4 Rational vs. Moral Permissibility

We have seen that a number of arguments—in my view, the strongest ones in the literature—for and against the permissibility of individual time-bias have analogues in the case of social discounting, and *vice versa*. This provides inductive grounds for thinking that individual time-bias and social discounting have the same normative status. Being inductive, however, it is not a knock-down argument. Why might individual time-bias and social discounting nonetheless differ in their normative status? First, it could be that there is some other sound argument for or against individual time-bias that doesn't carry over to the case of social discounting, or *vice versa*.

Second, it could be that one or more of the arguments I have surveyed is stronger in the one case than in the other. For instance, it could be that it is more clearly impermissible for a country to favour its own citizens over foreigners than it is for an individual to be biased in her own favour. If so, this would mean that Parfit's appeal to full relativity yields a stronger argument for the permissibility of individual time-bias than its analogue does for the permissibility of social discounting. Or it could be that it is more clearly impermissible for a country to proportion concern for future generations according to the degree to which they are culturally or politically similar to the present generation than it is for an individual to proportion concern for future time-slices according to their degrees of psychological connectedness, in

which case we would have an argument for the permissibility of individual time-bias that is stronger than its analogue for the case of social discounting. For my part, however, I view the arguments surveyed previously to be equally forceful in both the individual and the social cases.

Third, it might be that different flavours of normativity apply in the individual and the social cases. I will focus on this possibility in the remainder of this section. So far, I have talked in terms of permissibility *simpliciter*. But it is important to distinguish between rational permissibility and moral permissibility. And many theorists hold that morality only applies to our treatment of other people. One cannot treat oneself immorally (or unjustly). Mill writes in Ch. 4 of *On Liberty* that ‘self-regarding faults’ are ‘not properly immoralities and, to whatever pitch they may be carried, do not constitute wickedness’ except when ‘they involve a breach of duty to others, for whose sake the individual is bound to have care for himself.’ Rawls (1999 [1971], 260) endorses this thought with respect to attitudes to time:

In the case of the individual, pure time preference is irrational: it means that he is not viewing all moments as equally parts of one life. In the case of society, pure time preference is unjust: it means (in the more common instance when the future is discounted) that the living take advantage of their position in time to favour their own interests.

Indeed, some theorists have argued that the very idea of duties to the self is incoherent. For instance, Singer (1959) argued that whenever you have a duty to someone to perform a given act, that person can release you from that duty. But this means that if you have some duty to yourself to do something, you can release yourself from that duty. But a duty from which you can release yourself at will is no duty at all. Hence there can be no duties to the self.

But recently, some ethicists have pushed back against Singer's argument for the incoherence of duties to the self and against the broader view that morality does not apply within a single life.

Hills (2003), for instance, argues that everyone has a duty to promote one's own well-being. She appeals to three premises: first, that everyone has a duty to promote others' well-being; second, that this duty is at least sometimes unwaivable; and third, that reasons for action are universal. Thus, 'If you have an unwaivable duty to promote the well-being of others, then, since reasons are universal, you must have a counterpart duty to promote your own well-being' (136). She concedes that duties to the self may not be enforceable via blame and punishment but rejects the underlying 'juridical' model of duties which treats moral duties like legal duties.

Schofield (2015) endorses Darwall's (2006) view that what distinguishes moral reasons from other kinds of reasons is that they are second-personal. This might initially seem to entail the impossibility of moral reasons to treat oneself in a certain way. But Schofield argues that what is distinctive about relating to someone second-personally is that the two persons have different standpoints or perspectives. It is not crucial that such a second-personal relationship involve two metaphysically distinct relata. And, Schofield notes, a temporally extended individual occupies distinct standpoints or perspectives at different times, and so an agent at one time can relate second-personally to herself at another time. And he argues that in some cases, as when one's younger self smokes and thereby causes serious harm to her later self, or when one's younger self seriously constrains the autonomy of her later self, that later self may justifiably feel resentment—a paradigmatic second-personal attitude—toward her earlier self.

This is of course not the last word on the matter. But I hope to have said enough to cast some doubt on the idea that one cannot treat oneself immorally or unjustly, and so also on the idea that individual time-bias cannot be morally impermissible.

Putting my own cards on the table, I think that individual time-bias and social discounting are both morally impermissible and rationally permissible. I think that the arguments surveyed in the previous section in favour of the permissibility of these attitudes are unconvincing when it is moral permissibility that is at issue. This is because I think it is morally impermissible to weight the interests of future members of society by their degree of cultural or political similarity to current society, and I think that it is likewise morally impermissible for an individual to weight the interests of her future selves by their degree of psychological connectedness. And, in response to the Parfitian appeal to full relativity, I think it is morally impermissible for an individual to weight her own interests more heavily than the interests of others, and it is likewise morally impermissible for a country to weight the interests of its own citizens more heavily than the interests of foreigners. And while I feel the pull of the argument from the permissibility of bias toward the future, I think we should employ *modus tollens* and conclude that bias toward the future is morally impermissible on both a societal and an individual level, while avoiding counterintuitive conclusions about our duties to past selves and past people by adopting a view of well-being (such as hedonism) on which it is impossible to affect past levels of well-being.¹⁰

By contrast, I think that rationality requires only that preferences be coherent and avoid the most egregious forms of arbitrariness like Parfit's Future Tuesday Indifference (Parfit 1984). Hence it is rationally permissible for an individual or society to care more about the future than the past, more about those psychologically or culturally connected to it, more about itself than others, and hence also more about its near future than about its far future. And I have argued elsewhere (Hedden 2015b) against the claim that vulnerability to diachronic inconsistency is *ipso facto* irrational.

¹⁰ As for the argument from overdemandingness, I am happy to concede that morality is quite demanding. But if you are concerned about morality being overdemanding, this would motivate not the adoption of a rate of pure time preference, but rather some non-consequentialist moral theory that permitted agent-centered prerogatives.

5. Conclusion

There is a near-consensus among economists and philosophers that individual time-bias and social discounting are largely independent issues. I agree that descriptive facts about how individuals in fact discount their own well-being with respect to time do not entail anything about whether and how society should discount the well-being of future generations. But this does not mean that the normative status of individual time-bias is independent of the normative status of social discounting. And I have suggested that a close look at the arguments that have been levelled for and against the permissibility of each kind of attitude suggests that they should have the same normative status. For the arguments for and against individual time-bias yield analogous arguments for and against social discounting, and vice versa. Moreover, I find the arguments equally compelling in the one case as in the other. One might worry, however, that while individual time-bias may be irrational, only social discounting can be immoral or unjust, for morality only applies to our treatment of other people, and not to our treatment of ourselves. But I side with a number of ethicists who have recently rejected this common view and argued that morality and justice can apply intrapersonally as well as interpersonally. Thus, far from being independent issues that can be debated and evaluated separately, individual and societal attitudes to time are in fact intimately connected.

References

- Adams, Fred. 2008. 'Long-Term Astrophysical Processes.' In Bostrom and Ćirković, eds. 2008, pp. 33-47.
- Arrow, Kenneth. 1997. 'Discounting, Morality, and Gaming.' Manuscript, Stanford University. Available: https://www.researchgate.net/publication/23740962_Discounting_Morality_and_Gaming
- Bostrom, Nick, and Milan Ćirković, eds. 2008. *Global Catastrophic Risks*. Oxford: Oxford University Press.
- Bostrom, Nick. 2013. 'Existential Risk Prevention as Global Priority.' *Global Policy* 4(1): 15-31.
- Broome, John. 1992. *Counting the Costs of Global Warming*. Cambridge: The White Horse Press.
- Broome, John. 2004. *Weighing Lives*. Oxford: Oxford University Press.
- Broome, John. 2013. *Climate Matters*. New York: W.W. Norton and Co.
- Cline, William. 1992. *The Economics of Global Warming*. Institute for International Economics.
- Cowen, Tyler. ms. 'What is the Correct Intergenerational Discount Rate?' George Mason University. Available: <https://www.gmu.edu/centers/publicchoice/faculty/%20pages/Tyler/DISCOUNT.pdf>
- Cowen, Tyler, and Derek Parfit. 1992. 'Against the Social Discount Rate.' In Laslett and Fishkin 1992, pp. 144-61.
- Dougherty, Tom. 2011. 'On Whether to Prefer Pain to Pass.' *Ethics* 121 (3): 521-37.
- Dougherty, Tom. 2015. 'Future-Bias and Practical Reason.' *Philosophers' Imprint* 15 (30): 1-16.

- Greaves, Hilary. 2017. 'Discounting for Public Policy: A Survey.' *Economics and Philosophy* (3):391-439.
- Greene, Preston and Meghan Sullivan. 2015. 'Against Time Bias.' *Ethics* 125 (4): 947-970.
- Hager, Mark. 1989. 'Bodies Politic: The Progressive History of Organization 'Real Entity' Theory.' *University of Pittsburgh Law Review* 50: 575-654.
- Hare, Caspar. 2008. 'A Puzzle about Other-Directed Time-Bias.' *Australasian Journal of Philosophy* 86: 269-77.
- Hare, Caspar. 2015. 'Time: The Emotional Asymmetry.' In Bardon and Dyke 2015, pp. 507-20.
- Hedden, Brian. 2015a. *Reasons without Persons: Rationality, Identity, and Time*. Oxford: Oxford University Press.
- Hedden, Brian. 2015b. 'Options and Diachronic Tragedy.' *Philosophy and Phenomenological Research* 90 (2): 423-51.
- Hills, Alison. 2003. 'Duties and Duties to the Self.' *American Philosophical Quarterly* 40(2): 131-42.
- Kelleher, J. Paul. 2017. 'Pure Time Preference in Intertemporal Welfare Economics.' *Economics and Philosophy* 33, 441-73.
- Laslett, Peter, and James Fishkin (eds). 1992. *Justice Between Age Groups and Generations*. New Haven: Yale University Press.
- List, Christian, and Philip Pettit. 2011. *Group Agency*. Oxford: Oxford University Press.
- Marglin, Stephen. 1963. 'The Social Rate of Discount and the Optimal Rate of Investment.' *The Quarterly Journal of Economics* 77: 95-111.
- Mill, John Stuart. 1859. *On Liberty*.
- Nordhaus, William. 2007. 'A Review of the *Stern Review on the Economics of Climate Change*.' *Journal of Economic Literature* 45: 686-702.

- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Pigou, Arthur. 1932. *The Economics of Welfare*, 4th Edition. London: Macmillan.
- Posner, Richard. 2005. *Catastrophe: Risk and Response*. New York: Oxford University Press.
- Ramsey, Frank. 1928. 'A Mathematical Theory of Saving.' *The Economic Journal* 38 (152): 543-59.
- Rawls, John. 1999 [1971]. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Scheffler, Samuel. 1982. *The Rejection of Consequentialism*. New York: Oxford University Press.
- Scheffler, Samuel. 2013. *Death and the Afterlife*. New York: Oxford University Press.
- Schelling, Thomas. 2000. 'Intergenerational and International Discounting.' *Risk Analysis* 20(6): 833-7.
- Schofield, Paul. 2015. 'On the Existence of Duties to the Self (and Their Significance for Moral Philosophy).' *Philosophy and Phenomenological Research* 90(3): 505-28.
- Sidgwick, Henry. 1907. *The Methods of Ethics*, 7th Edition. Reprinted by Hackett Publishing Company, 1981.
- Singer, Marcus. 1958. 'On Duties to the Self.' *Ethics* 69(3): 202-5.
- Stern, Nicholas. 2007. 'The economics of climate change: the Stern review.' Cambridge, UK: Cambridge University Press.
- Williams, Bernard. 1973. 'The Makropulos Case: Reflections on the Tedium of Immortality.' In *Problems of the Self*. Cambridge: Cambridge University Press, pp. 82-100.