INTERGENERATIONAL DISCOUNTING AND TIME PREFERENCE

WORD COUNT: 6775

Abstract. The time-preference model of intergenerational tradeoffs posits an intergenerational time preference, in which the utility of future welfare is discounted with time. When separated from confounds with the use of thought experiments, the time-preference model predicts widespread arbitrariness and implausible or even absurd reactions to tradeoffs. In contrast, social preferences explain both diminished concern for future generations and for socially-distance contemporaries. Regardless of what is useful for economic decision making, the social-preference model is an improvement on the time-preference model for the purposes of debating the justifiability of discounting future welfare and for motivating people to care more about future people. For these purposes, it is best to assume the most charitable interpretation available that is consistent with the evidence.

1. Introduction

If you’re reading this, you probably write your own academic papers, and you probably think that authoring an academic paper is a valuable activity. If you’re like me, however, you do not find writing these papers a particularly pleasurable experience overall. Perhaps you find writing valuable because it helps to clarify your own thoughts or because it advances your own personal interests. But without a doubt, at least a small part of the answer has to do with the social context in which the activity occurs. We expect that what we write has a chance to benefit other people, in some way, and that some of these people may not be alive at the same time as we are.

When people do value activities purely for their own pleasure — i.e., when they assign things individual hedonic value — their attitudes reveal a time preference. People care not just about how pleasurable an experience will be but also where that pleasure will occur in relation to the present. Specifically, this time preference manifests a near bias: people prefer pleasurable experiences to be near rather than distant. Sometimes, this time preference is strong enough to result in less pleasure overall. This occurs when one prefers an experience that is less pleasurable but near to one that is more pleasurable but distant.1

Do time preferences exist for non-hedonic valuing that concerns future generations? For example, do we tend to care more that our research benefits near-future generations, rather than distant ones, even if the benefit is smaller? Many economists, philosophers, and governmental organizations do indeed assume that time preferences exist for such valuing in the same way that they exist for hedonic valuing. This view supports a time-preference model of intergenerational tradeoffs, in which future utility is discounted with time.

This essay combines insights from recent psychological research with thought experiments to explore the potential determinants of preferences for tradeoffs concerning the welfare of future generations. Contrary to the time-preference model, I suggest that temporal distance has little to no relevance for such tradeoffs. Instead, such preferences concern social distance.

1For philosophically accessible overviews of this phenomenon, see Loewenstein and Elster, 1992; Loewenstein et al., 2003; and, especially, Frederick et al., 2002.
In this way, preferences for the welfare of future generations are similar to preferences for the welfare of socially distant contemporaries.

Employing an accurate and charitable interpretation of preferences for intergenerational tradeoffs is important for at least three reasons. First, it helps to clarify what is at stake when philosophers and economists debate the justifiability of discounting future welfare. It matters to this debate whether discounting partly reflects time preferences or some other, less arbitrary seeming, psychological factor. Second, models create starting assumptions that lead to hypotheses tested in experimental research. More accurate models tend to generate more useful hypotheses and they also tend to better highlight potential confounds. Third, attempts to get current people to care more about the welfare of future people are unlikely to succeed without the use of a charitable model for what causes diminished concern for future welfare.

In what follows, I first introduce and survey the time-preference model of intergenerational tradeoffs. I argue that the time-preference model posits widespread arbitrariness unnecessarily and performs poorly in capturing reasonable attitudes to thought experiments. I then introduce the social-preference model and argue that it is more accurate, more useful, and more apt for debates concerning the justifiability of discounting future welfare.

2. The Time-Preference Model

Before you lie the cookies; they are very tasty. You can either have one of these cookies tomorrow or two of them next week. Which do you prefer? In isolation, having two cookies is better, but is it better enough to call for waiting an entire week? You are tempted to choose one cookie because of the temporal nearness of its benefit. The two cookies involve far more waiting.

You now face a more consequential choice. Before you lie plans for two conflicting long-term engineering projects. The first will provide a moderate benefit to the next generation of humanity. The second will provide a very large benefit to the people who live several generations from now. Is this choice like that between the cookies? While, in isolation, the larger benefit would be better, its realization involves “waiting” several generations (though, of course, only “humanity” will “wait” for it—no person will). If this decision is like that between the cookies, then you must ask yourself whether the benefit is better enough to warrant the time delay.

The time-preference model of intergenerational tradeoffs assumes such a symmetry between our hedonic and non-hedonic valuing. Specifically, it assumes that people are near biased in both cases. In the cookie case, a near-biased agent is pulled toward the single cookie just because of its nearness, and this pull manifests in choice behavior when the agent prefers a nearer experience over another that is at least as good in terms of non-temporal properties. Greene and Sullivan [2015, 948] define the manifestation of near bias like this:

An agent $S$ is biased toward the near with respect to pleasure iff for two exclusive future experiences, $E_1$ and $E_1$, where $E_1$ is at least as pleasurable as $E_1$ (adjusted for the subjective probabilities of the events occurring), $S$ prefers $E_1$ because it would occur nearer to the present.

The corresponding manifestation of near bias in the case of benefits to future generations can be defined like this:

An agent $S$ is biased toward the near with respect to benefits to future generations iff for two exclusive future benefits, $B_1$ and $B_2$, where $B_2$ is at least as
good as \( B_1 \) (adjusted for the subjective probabilities of the events occurring),

\( S \) prefers \( B_1 \) because it would occur nearer to the present.

Just as we find nearer pleasures preferable to distant ones, the time-preference model posits that benefits to humanity are more preferable the more temporally near they occur.

Social scientists attempt to quantify hedonic near bias with the concept of time discounting, which uses a mathematical function to capture how much an agent discounts the value of an experience based on its temporal distance. There is a crucial difference between the discounting of welfare, or utility — which John Broome [1999, 46] calls “pure” discounting — and the discounting of commodities used by economists in analyzing potential investments. The latter type of discounting need not involve pure discounting — it might only involve the fact that invested capital supports greater consumption in the future. In his seminal paper, Ramsey [1928, 553] notes that the “rate of discounting future utilities must, of course, be distinguished from the rate of discounting future sums of money. If I can borrow or lend at a rate of \( r \) I must necessarily be equally pleased with an extra £1 now and an extra £\((1+r)\) in a year’s time, since I could always exchange the one for the other.” Accordingly, when this paper discusses time discounting, it aims to discuss “pure” discounting in the sense Broome and Ramsey identify.

Furthermore, for a phenomenon to be called pure time discounting, it should concern the influence of temporal properties on our preferences. For example, in Greene and Sullivan’s definition, one must account for probabilities when identifying near bias. This is because distant experiences are typically less certain than near ones. It would make sense for even a temporally neutral agent, who does not discount with time, to tend to prefer nearer pleasures to more distant ones in situations where there is a correlation between time and probability.

There are several other factors that may influence our intertemporal preferences but do not concern nearness. An agent might prefer near pleasures (and distant pains) for reasons concerning personal identity (e.g., the agent might be skeptical that they will be the same person that feels distant pleasure and pain). Such an agent would care about identity properties and not temporal properties. Another agent might have “global” preferences about the structure of their life (e.g., they prefer an improving life to a degrading one, regardless of the total amount of pleasure in each life). Such an agent would care about the overall sequence of events, and not the nearness of any event (i.e., they would not care about an event’s location in time relative to the present).

With those caveats about hedonic time discounting in mind, let us now consider the practice of using the time discounting concept to quantify preferences for non-hedonic events. For example, a social scientist may attempt to use a discount rate to quantify preferences for events that will only be enjoyed by future generations — such as long-term economic or environmental goods. It is important to realize that preferences for intergenerational non-hedonic events and preferences for individual hedonic events might have very different determinants. Frederick [2006, 669] explains:

Strictly, the [discounted-utility] model expresses time preference, because the discount factor operates on utility itself. This time preference relevant to outcomes that occur later within one’s own life must, however, be distinguished from an intergenerational time preference, which pertains to the degree of concern one has for the welfare of future individuals. There is an important difference between discounting one’s own future utility and discounting the utility of someone else who will be alive in the future.
On Frederick’s proposed framework, we should separate factors like probabilities and opportunity costs from the concept of time preference, and when it comes to time preference there are really two concepts — individual time preference and intergenerational time preference. Since preferences for individual hedonic events and intergenerational non-hedonic events may have different determinants, it is possible that there exists an individual time preference but no intergenerational time preference (or vice versa). The existence of each type of time preference must be motivated independently.

The general framework for time discounting used in this paper is in line with that of Frederick’s. As I argued above, we should be careful to distinguish hedonic (individual) time preference from other factors. When considering supposed intergenerational time preference, which is non-hedonic, we should be just as careful. For example, in the choice over engineering projects, a person might prefer the shorter project because the effects of the longer project are less certain. Such a person would care about probabilities and not temporal properties. Another person might prefer the longer project to the shorter one because they prefer a narrative in which humanity improves over time. Such a person would care about the overall sequence of benefits rather than their temporal nearness to the present.

Time preferences concern temporal nearness. These preferences must be separated from preferences for maximizing expected utility from a time-neutral perspective or creating improving sequences. The time-discounting model expresses time preference because it operates on utility — it adjusts the utility of events on the basis of temporal nearness from the perspective of the present.

For an example of the time-preference model applied to intergenerational tradeoffs, consider the US Environmental Protection Agency’s discussion of discounting in their Guidelines for Preparing Economic Analyses. It states, “Discounting reflects: (1) the amount of time between the present and the point at which these changes occur; (2) the rate at which consumption is expected to change over time in the absence of the policy; (3) the rate at which the marginal value of consumption diminishes with increased consumption; and (4) the rate at which the future utility from consumption is discounted with time” [6-2, emphasis added]. Of the things discounting reflects, according to the EPA, (1-3) are the kinds of considerations that Frederick’s framework suggests are not time preferences, whereas (4) expresses intergenerational time preference because it suggests that utility itself is adjusted on the basis of temporal nearness.

Furthermore, in the EPA guidelines, (4) is the sole suggestion of a potential psychological reason people might show diminished concern for future welfare. This is notable given that almost all philosophers, and many economists, regard time preference as arbitrary. Sidgwick [1884, 414], for example, claimed that “the time at which a man exists cannot affect the value of his happiness from a universal point of view.” Indeed, Ramsey [1928] developed the framework for applying a temporal discount rate to future utility, and even he thought it was unjustified. He identified any impulse toward pure time discounting as the result of “weakness of the imagination” [543]. Philosophers tend to think that time preference is arbitrary because it attaches normative importance to what Parfit [1984, Section 46] calls

---

2 Time preferences can also concern the distinction between past and future, as with hedonic future bias: a preference for pleasures to be future and pains to be past. Interestingly, many philosophers, including Brink [2010, 378], Hare [2013], and Dougherty [2015, 3], have claimed that people are not future biased with regard to non-hedonic events. This is striking because the time-preference model of intergenerational tradeoffs assumes that people are near-biased with regard to such events.
a “purely positional property.” Purely positional properties, like space and time, do not, in themselves, amount to relevant differences in the intrinsic desirability of events.\(^3\)

Similarly, the economist R. F. Harrod [1948, 37-40] claimed that pure time discounting is a “human infirmity.”\(^4\) More recently, in a major review of the economics of climate change commissioned by the United Kingdom government, the economist Nicholas Stern [2007, 31] treats “the welfare of future generations on a par with our own.” Stern claims that we should not discount utility on the basis of time but only uncertainty, and he takes this view to be supported by all economists with a strong interest in philosophy.\(^5\)

The EPA, like most governmental agencies, reacts to these worries by adopting a neutral view about the justifiability of intergenerational time preferences, writing: “Clearly, economics alone cannot provide definitive guidance for selecting the ‘correct’ . . . social rate of time preference” [6-12]. However, if the EPA were to accept that intergenerational time preference is not justifiable — that it can only be the result of arbitrariness — then its proposed model would posit widespread arbitrariness in the preferences of the average person.

When forming hypotheses, it is bad to privilege interpretations on which people are arbitrary unless there are good reasons to think otherwise. Are there good reasons to think otherwise in this case? If there were good experimental data proving that people have an intergenerational time preference, even though such a preference is arbitrary, then that would one thing. However, the experimental data does not prove this, and may even call the idea into question.

Even in the individual hedonic case, there is reason to doubt that people have time preferences. In their reviews of the hedonic time preference literature, Frederick et al. [2002] and Soman et al. [2005] conclude that no experiments manage to isolate time preference from plausible confounds. Such confounds include the effects of uncertainty, inflation (for monetary tradeoffs), and anticipatory utility (or anticipatory “disutility,” as in the case of negative waiting). Frederick et al. [2002, 384] write, “It is possible that the responses . . . can be entirely explained in terms of these confounds.”

Frederick et al. [2002, 377] also find “tremendous variability” in estimates of people’s average discount rate. In their review, they show that existing experiments estimate an annual discount rate between -6% and infinity, with no clustering of moderate results. Furthermore, there is no evidence of progress: the variability is not lessening over time. These factors — inability to separate time preference from confounds, inconsistent results, and no improvement in consistency of results — may indicate that time preference may not be a strong determinant of preferences over individual hedonic tradeoffs. Frederick et al. [2002, 377] conclude: “To truly understand intertemporal choices, one must recognize the influence of many considerations besides pure time preference.”\(^6\) Similarly, Soman et al. [2005, Sections

\(^3\)Lowry and Peterson [2011] call this the “standard argument” while Sullivan [2018] calls it the “non-arbitrariness argument.”

\(^4\)See also Parfit, 1984, 480–6; Pigou, 1932, 29–30; and Solow, 1974, 9. As Broome [2016, 10] writes, the rejection of pure discounting is supported by the popular view that morality must be impartial.

\(^5\)The discussion here draws from a review of Stern’s work in de Lazari-Radek and Singer, 2014, 361.

\(^6\)A more recent review by Story et al. [2014] comes to the same conclusion. In addition to the wide divergence in estimates of discount rates, they focus on the lack of correlation between discount rates in differing choice domains and the fact that studies find significant amounts of “negative discounting” for bad events (i.e., people who prefer hasten the arrival of bad events). Story et al. [2014, 2–3] write, “The low correlations between health and monetary discounting as well as the prevalence of negative and zero time
1.2 and 2.2] claim that as the study of time preferences has progressed it has become increasingly clear that researchers need to abandon the “elegant simplicity” of the economic discounted-utility model in favor of psychological approaches that understand intertemporal choices as the joint product of conflicting motives.\footnote{As Loewenstein [1987, 666–7] and Frederick et al. [2002, 352–3] show, this was the predominant view of economists before the introduction of the discounted-utility model in 1937. Frederick et al. write, “When the [discounted-utility] model eventually became entrenched as the dominant theoretical framework for modelling intertemporal choice, it was due largely to its simplicity and its resemblance to the familiar compound interest formula, and not as a result of empirical research demonstrating its validity.”}

In the case of intergenerational time preferences, there are more factors to consider. Imagine an experiment showing that people tend to prefer the shorter engineering project (in the case introduced at the beginning of this section). To know whether people are reacting to nearness in time, we would need to isolate the influence of time from that of other factors. In addition to the factors that are also relevant to hedonic tradeoffs, such as probabilities, there are factors that seem to have a large effect on non-hedonic tradeoffs, such as social distance. Do we prefer the shorter engineering project because the people it benefits are closer \textit{temporally}, or because they are closer \textit{socially}?

Since it is bad to privilege interpretations on which people are arbitrary without good reason, and such reason is lacking in this case, we should seriously consider the possibility that there are alternatives to the time-preference model that do not involve arbitrariness (or at least involve less obvious arbitrariness), and do a better job of capturing common attitudes toward intergenerational tradeoffs. This is one reason to doubt the usefulness of the time-preference model. Given this reason for doubt, it is worthwhile to reflect on the relative plausibility of the time-preference model with thought experiments, which I turn to in the next section.

3. \textbf{Thought Experiments Show the Time-Preference Model is Implausible}

A primary consideration against the time-preference model of tradeoffs concerning future generations is a simple and fundamental truth about near bias: it concerns \textit{experienced time} and not \textit{objective time}. Consider a variant of a standard test for hedonic near bias.\footnote{Modeled after experiments described in Ainslie, 2001.} Imagine you have a choice between a nice meal this evening, which is worth about $50 to you, or a better meal one year from now, which is worth about $100 to you. Many people choose the immediate reward in this situation, and researchers take this to be evidence for near bias. However, consider the following variant:

\begin{quote}
Tomorrow morning you will be put into a consciousness-free stasis for a year. You have a choice between the $50 meal this evening and the $100 meal after you exit the stasis. You know that it would feel as if the $100 meal occurs tomorrow.
\end{quote}

There is, of course, license to worry about the stasis lowering the probability that one will get to enjoy any meal — and merely stating that entering stasis does not increase one’s chance of death or disablement is not enough to assuage this worry (as Frederick et al., 2002, 384 argue) — but it does seem that the effect of time on one’s preferences is not the same as it would be without stasis. For a more mundane example of the stasis phenomenon, consider a child who preference for health and aversive outcomes demonstrate that discounting cannot be considered a universal mechanism by which all delayed events are evaluated."
requests an earlier bedtime on Christmas Eve. The objective time before Christmas morning is always the same; only experienced time changes. Both examples suggest that models of hedonic time discounting that reference objective time are only satisfactory in cases in which there is a correlation between objective time and experienced time.\footnote{Intuitive arguments for the rational permissibility of future bias also typically reference experienced time. We feel, intuitively, that things are better when painful experiences are “over and done with” [Heathwood, 2008, 57], we say things like “thank goodness that’s over” [Prior, 1959], and we are “greatly relieved” that a painful experience has already occurred [Parfit, 1984, 165]. In line with these arguments, a common prediction in the philosophical literature is that there is no future bias when subjective experiences are absent. For example, Hare [2013] discusses cases in which future bias seems to be absent when people do not experience events directly. In one example, Hare imagines that his wife either will be or has been unfaithful to him, and he is uncertain whether the event is in the immediate past or immediate future. In another, he imagines that a hated team either will win or has won the Super Bowl, and he is again uncertain of the temporal properties of the event. In each case, it seems that future-biased preferences are nonexistent, which is in stark contrast to standard cases involving pure hedonic bads. Greene et al. [2020] tested this prediction and found it to be correct.}

That near bias is intimately connected to our experience of the passage of time has been noted before. Hedonic near bias, some have claimed, is concerned with the experience of waiting, and not by the passage of objective time. Schelling (2000, 234) writes, “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.” To be patient, one must endure \textit{experienced} time. If one spends the time unconscious, then patience is not a factor.

When one is not alive, one does not experience the passage of time. Past the point of death, there is only objective time by which to measure tradeoffs, and thus no correlation between experienced and objective time. If the time-discounting concept is only satisfactory when there is a correlation between experienced and objective time, then the concept is not satisfactory for intergenerational tradeoffs. Intergenerational tradeoffs do not involve experienced time. This would explain the lingering absurdity in the time-preference model of the engineering projects tradeoff (introduced at the start of Section 2). The reason trading between engineering projects is so unlike trading between cookie experiences is that the latter involves waiting — i.e., the cookie case involves an agent experiencing the temporal distance between the events — while the former does not. The absurdity has been noted by Cowen and Parfit [1992, 155]: “Abstinence from consumption does not involve waiting when consumption is postponed across generations. Such abstention cannot be a meaningful bad for people who are not yet born; in the meantime, nobody is left waiting.”

The absurdity is most apparent when the concept is applied to very long time frames. Cowen and Parfit [1992, 147] point out that a discount rate of just five percent would make a death next year count for more than a billion deaths in four hundred years. Perhaps this phenomenon is even more outrageous when we look backwards. Consider Cowen and Parfit [1992, 145] question: “Imagine finding out that you, having just reached your twenty-first birthday, must soon die of cancer because one evening Cleopatra wanted an extra helping of dessert. How could this be justified?” The inexorable progression of temporal distance assures these absurd results, given any discount rate, when the time scales become large.

Consider another thought experiment involving the potential source of value with which we started: philosophical inquiry. This is a potential source of value well-suited to considerations of future generations since people often pursue philosophical inquiry without the illusion that a specific goal will be reached within their lifetime. Moreover, the value one places in studying
philosophy often seems to depend on the idea that one is engaging with a tradition that will continue after one is dead. (If one does not feel that way about philosophical inquiry, then other potential sources of non-hedonic value that meet these criteria can be substituted in without losing the point.)

Consider the following case:

_Hiatus_: You learn that the practice of philosophical inquiry will go on hiatus for a thousand-year period shortly after your death. After this thousand-year period has elapsed, people will take up philosophy again, and they will start by reading the work published before the hiatus.

Does the prospect of _Hiatus_ threaten the value you place in philosophical inquiry? Would increasing the length of the hiatus cause it to lose even more value? If _Hiatus_ significantly lessened the value of philosophical inquiry, then that would be evidence for a time preference. Moreover, if time preference affects the perceived value of philosophical inquiry, then we should expect its value to decrease as the length of the hiatus increases. As the hiatus increases, at some point future philosophical inquiry should not matter to you at all.

None of this seems to be true, and for good reason. If philosophical inquiry matters to us partly because of its engagement with a tradition that will continue after we die, then it should still matter to us even if this tradition will go on hiatus for a period after our deaths. In other words, philosophical inquiry should continue to matter to us regardless of which future generations we engage with, if we hold constant the qualities of the engagement. To think otherwise would be to draw an arbitrary distinction based solely on the temporal locations of the people that continue our traditions.

One might prefer, less arbitrarily, that the future generations that continue our traditions are like us, and one might expect the next generation to be more like us than subsequent ones. This, however, would not be a time preference. Such a person would care about similarity and not nearness. Finally, there are strong intuitions that point in the opposite direction of near bias — favoring distance over nearness. It is a great achievement that the work of some ancient Greek philosophers engages generations existing thousands of years after their death, and in some cases — that of Aristotle, for example — the ideas were indeed discovered after a significant hiatus. Contrary to the time-preference model, the vast temporal distance seems to enhance, rather than threaten, the value of Aristotle’s work. Even if temporal distance did not enhance the perceived value of Aristotle’s contributions, it certainly wouldn’t hurt it. It is not hard to imagine that Aristotle would care more about the High Middle Ages, which he so profoundly influenced, than he would about the Early Middle Ages, even though the Early Middle Ages were nearer to him temporally. What seems to matter in this case are relations between people that are not temporal in nature, and the importance of these relations is not discounted with time. Since the time-preference model is not proven by experimental results and seems implausible in thought experiments,

---

10Cf. Nozick, 1981, 584: “You want there to be some time after which you continue to leave a mark, but this time needn’t be precisely at your death. Artists as well as those who anticipate resurrection are quite willing to contemplate and tolerate a gap.”

11Parfit [1984, 485] makes a corresponding point regarding the view that degree of kinship justifies discounting: “We believe that our government ought to be especially concerned about the interests of its own citizens. It would be natural to claim that it ought to be especially concerned about the future children of its citizens, and, to a lesser degree, about their grandchildren. Such claims might support a new kind of Discount Rate. We would be discounting here, not for time itself, but for degrees of kinship.”
we should seriously consider alternatives. Depending on one’s reactions to these thought experiments, the time-preference model seems wildly inaccurate in the worst case, or, in the best case, a crude approximation of the things that matter to intergenerational tradeoffs. Even if the model is an approximation, there is an opportunity to find something better: there may be more useful models available that can more effectively guide our thinking about intergenerational tradeoffs.

4. The Social-Preference Model

Simon [1995, 368] suggested that people’s tendencies to allocate resources can be broken up into three dimensions: 1) resources for current consumption by the same person, 2) resources for future consumption by the same person, and 3) resources for current or future consumption by other individuals. Perhaps we should, indeed, model (1) and (2) as a function of temporal distance (as in hedonic time discounting). However, Simon claimed that (3) should be modeled as a function of “perceived social distance.” This idea is now called “social discounting” in the economics and psychology literature.

A social discount function measures the value of a reward allocated to another person at a given social distance. Studies of social discounting proceed similarly to studies of time discounting. Participants are asked to imagine a ranking of the 100 people “closest to them” and to consider making various tradeoffs between a reward for another person and a reward for themselves. For example, a participant might be asked whether she would prefer $25 for herself or $100 for another person of a given closeness rank. The amounts are then varied to determine an indifference point for each rank.12

Social discounting and time discounting have some corresponding properties. Trope and Liberman [2010] found that temporal and social distance are both dimensions of psychological distance and have an influence on decision making. Jones and Rachlin [2006, 2009] discovered that human social discounting, like human time discounting, is hyperbolic.13 Indeed, Jones and Rachlin [2009, 62] conclude that similar hyperbolic equations can be used to describe both time and social discounting.

Part of what matters to perceived social closeness is kinship,14 but this is just one of several factors. Empirical research into the exact nature of social distance is still evolving. Nevertheless, there is a genuine phenomenon here, which has a systematic impact on our valuing, and which concerns the amount of social connection we feel with other people.

The corresponding manifestation of social bias in the case of benefits to future generations can be defined like this:

An agent $S$ is socially biased with respect to benefits to future generations iff for two exclusive future benefits, $B_1$ and $B_2$, where $B_2$ is at least as good as $B_1$ (adjusted for the subjective probabilities of the events occurring), $S$ prefers $B_1$ because it concerns people of greater perceived social closeness.

Whereas the time-preference model posits that benefits to humanity influence us to a lesser degree as the people involved become more distant temporally, the social-preference model, more plausibly, suggests that benefits to humanity influence us to a lesser degree as people

---

12See, e.g., Jones and Rachlin [2006, 2009]; Yi et al. [2011], or Ma et al. [2015].
13For more on hyperbolic time discounting, see Ainslie, 2001.
14As per the thesis discussed by Parfit (see footnote 9).
become more distant socially. In other words, it posits that what matters for intergenerational tradeoffs is not the temporal location of other people, but the degree to which we feel socially connected to them. Indeed, this is a more justified response: to base one’s values on temporal location would be to draw an arbitrary distinction.\footnote{Some consequentialists claim that social bias is similarly arbitrary. I remain neutral on this; instead, I am merely claiming that social bias is less obviously arbitrary than near bias. In other words, increased concern for a person’s welfare on the basis of social connection is less obviously arbitrary than increased concern on the basis of temporal location.}

The social-preference model performs much better in thought experiments: it gives plausible predictions in situations where the time-preference model’s predictions are implausible or even absurd. To start, consider again Cowen and Parfit’s Cleopatra case (introduced in Section 3). Cleopatra might favor herself over those with whom she feels distant socially, but her social distance to someone living in 1520 is about the same as it is to someone living in 2020. Therefore, Cleopatra presumably would have the same attitude toward the dessert/death tradeoff whether the person who would die lives in 1520 or 2020. The progression of time, in itself, does not warrant further discounting. The time-preference model, in contrast, embraces the absurdity. For any discount rate, there will be a timescale on which future deaths become outweighed by present desserts.

The stark contrast in plausibility between the time- and social-preference models is also apparent in \textit{Hiatus}. In \textit{Hiatus}, we imagine that philosophical inquiry stops entirely after our deaths but then picks up again where it left off. Since social connection falls rapidly after our deaths and then remains relatively constant, it should matter little whether the hiatus lasts 1000 or 10000 years, as long as inquiry picks up where it left off. In other words, while we may prefer that there be no hiatus so that those who engage with our work have preexisting social closeness, given a hiatus of any significance the key features of our connection will be that our readers are doing philosophy and pursuing related questions. These, after all, are the factors that connect Aristotle to philosophers alive today. Aristotle was able to forge a connection to people existing thousands of years in the future. We are in awe of his accomplishment, and expect that he would be in awe of it too, precisely because it has crossed such a vast temporal distance to connect with other people. The time-preference model predicts exactly the opposite.

At this point, one might lodge the following objection: “Despite its inaccuracy, we are justified in continuing to use the time-preference model because discounting utility with time is the most straightforward and tractable procedure by which to model concern for future welfare.” In response, consider an analogy to charitable giving for contemporaries. There is presumably some correlation between spatial distance and people’s concern for others. Given this, government organizations could develop spatial-preference models that feature a spatial-discount rate, which captures the average effects of spatial distance on people’s concern for the welfare of others. This spatial-discount rate could then be part of policies concerning charitable giving. Organizations might include the spatial discount rate to capture the supposed psychological factors that exist for such tradeoffs. But in this case, it would be clear that the spatial-discount rate is a very crude approximation of what matters to people. Perceived social closeness has a large effect on charitable giving decisions, and in modern times such perceived closeness has become less and less correlated with spatial distance. There are also gross inaccuracies: sometimes people living in far away countries seem socially closer than the people living in countries nearby. And sometimes the closest
people are viewed as the most socially distant, as when humans “outgroup” members of rival political parties or nearby racial groups that they come to despise. If policymakers want charitable policies to partly reflect the psychological considerations of their citizens, then they would not be served well by adding a spatial-discount rate to their policies, despite its tractability.

I suggest that the same is true of the time-preference model. There may be some weak correlation between temporal distance and people’s concern for others, but a time-discount rate is a very crude approximation of what matters, and in some instances delivers verdicts that are in the opposite direction of what is plausible. It also pushes decisions in a harmful direction (from a temporally neutral perspective) unnecessarily, as the Cleopatra case helps to show. Specifically, the use of the time-preference model predicts a diminishment of concern for very distant future generations in comparison to merely distant generations, where there is no reason to think that this accurately reflects people’s preferences.

Considering this, it would be better to let time-discounting reflect only the things it does usefully approximate: factors like opportunity costs, probabilities, changes in consumption rates, or the marginal value of consumption. And then to grant that concern for future welfare may be diminished due to social preference. This is currently how government organizations view charitable giving for contemporaries. In essence, we should conceive of sacrifices for future generations as similar to charitable sacrifices for socially-distant contemporaries.\(^\text{16}\) Currently, such approximations of social preference are done on a case-by-case basis and with considerable guesswork, though we can expect our ability to quantify social distance to improve as the study of social discounting, which is currently in its infancy, improves. In any event, if government agencies adopted a spatial discount rate in an attempt to more precisely quantify concern for contemporaries, then the effect would be to make such judgments less precise instead of more so. The same is true, I contend, with the use of a time-discount rate to quantify concern for future welfare. Approximations of social preference are best done with the concept of social discounting or, alternatively, such approximations should be done on a case-by-case basis until our ability to determine social discount rates improves.

Finally, the social-preference model is an improvement over the time-preference model not only because it makes better reflective sense of the things that seem to matter to us, but also because it gives us a better picture of what is at stake in debates over intergenerational welfare tradeoffs. If people are not reacting to temporal properties in showing diminished concern for future welfare, but instead social properties, then when economists like Nicolas Stern [2007, 31] call for a rejection of discounting utility with time, what they really should be calling for is a rejection of social bias. If so, the debate over discounting future welfare is just like the debate over discounting the welfare of socially-distance contemporaries. Much of the work, for example, of consequentialists like Peter Singer\(^\text{17}\) in motivating the rejection of social bias in charitable giving decisions would be directly applicable to intergenerational welfare tradeoffs. Over just the last couple of decades, a breathtaking about of progress, or at least change, has occurred in common attitudes toward charitable giving and the justifiability of

---

\(^\text{16}\) Cf. Schelling, 2000, 834: “To invest resources now in reduced greenhouse gas emissions is to transfer consumption from present-day people—whoever those people are who are making these sacrifices—for the benefit of people in the distant future. It is very much like making sacrifices for people who are distant geographically or distant culturally.”

\(^\text{17}\) See, for example, his *The Expanding Social Circle* [1981].
social bias. If debates over the discounting of future welfare appear stagnate in comparison, perhaps it is because they focus on the wrong properties — their focus has been temporal relationships instead of social ones.

5. Conclusion

The time-preference model of intergenerational tradeoffs posits an intergenerational time preference, in which the utility of future welfare is discounted with time. When separated from confounds with the use of thought experiments, the time-preference model predicts widespread arbitrariness and implausible or even absurd reactions to tradeoffs. A more charitable interpretation posits that what matters is the social connections between people rather than their relative temporal positions. Correspondingly, the social-preference model, which focuses on social discounting instead of pure time discounting, offers more plausible explanations while staying rooted in recent psychological research.

Social preferences explain diminished concern for both future generations and for socially-distance contemporaries. In the latter case, economic decision makers eschew the use a “spatial discount rate” to quantify such concern, and instead rely directly on estimations of social preference, even though such judgments are more difficult to quantify. I suggest that the same should be true in the case of concern for future generations: time discount rates are useful for tracking economic factors — such as opportunity costs, probabilities, changes in consumption rates, or the marginal value of consumption — but they are not useful for tracking the psychological factors behind diminished concern for future welfare.

Finally, regardless of what is useful for economic decision making, I am most confident in concluding that the social-preference model is an improvement on the time-preference model for the purposes of debating the justifiability of discounting future welfare and for motivating people to care more about future people. For these purposes, it is best to assume the most charitable interpretation available that is consistent with the evidence.

References


Giles W. Story, Ivo Vlaev, Ben Seymour, Ara Darzi, and Raymond J. Dolan. Does temporal discounting explain unhealthy behavior? a systematic review and reinforcement learning