Pure and Impure Time Preferences

Kristie Millera, James Nortonb, and Andrew Lathamc

aUniversity of Sydney, bUniversity of Iceland), cAarhus University

<kristie\_miller@yahoo.com>

ABSTRACT

This paper investigates two assumptions of the exponential discounted utility theory (EDU) to which Callender draws our attention: namely that we can cleanly distinguish pure from impure temporal preferences, and that past discounting can be ignored. Drawing on recent empirical work in this area, we argue that insofar as one might have thought that past-directed preferences are more pure than future ones, then there is evidence that people’s pure preferences (insofar as we can make sense of that notion) show more interpersonal variation than has previously been thought.

**ARTICLE HISTORY** Received 1 Jan 2020; Accepted 1 Feb 2020

KEYWORDS preferences; pure time preferences; temporal discounting; past discounting

In ‘The Normative Standard for Future Discounting’, Callender [2022] carefully formulates the ‘master argument’ that has tacitly underpinned the widespread acceptance of Samuelson’s [1937] *exponential discounted utility theory* (EDU). The idea behind EDU is that a rational agent, insofar as she discounts future goods, will do so exponentially—for example, she will equally value $20 now, $10 in one month, $5 in two months, $2.50 in three months, and so on. Such an agent does not open herself up to the kind of exploitation to which a non-exponential discounter is potentially exposed. By carefully constructing the master argument for EDU, Callender intends to draw out a variety of dubious assumptions on which the argument rests.

The purpose of this discussion is to further investigate, in the light of recent empirical work, two of these assumptions, namely that ‘we can cleanly distinguish pure and impure temporal preferences (without the purification making your preferences alien)’ and that ‘past discounting can be ignored’ [Callender 2022: XXX]. We will argue that the kind of agent who could have pure time preferences is not the kind of agent who makes decisions, and thus that whatever norms might govern such preferences are completely useless to decision-making agents such as us. Moreover, we will show that if past-directed preferences are thought to be more pure than (and perhaps shed light on) future-directed preferences, there is evidence that people’s pure preferences show much more interpersonal variation than has been previously supposed.

EDU is a normative standard only for *pure* time preferences. Thus, its relevance depends on our capacity to distinguish pure from impure preferences. What are pure time preferences? According to Lowry and Peterson [2011: 490], these are ‘preference[s] for something to come at one point in time rather than another, not because this will make the benefit greater or more certain, but merely because of when it occurs in time’. The idea is that while there are a variety of factors motivating people’s preferences for things to occur at one time or another, one such factor is simply the respective temporal locations of the goods on offer. These are preferences that persist even if we ‘hold all else equal’ regarding the two options.

Imagine you like bushwalks. I offer you a bushwalk tomorrow, or a bushwalk next Sunday. You prefer the bushwalk tomorrow. Whether or not you are displaying a pure time preference depends on the reasons for your preference. If the only reason you prefer the bushwalk tomorrow is because you’re already bushwalking next Saturday and thus you’ll be sore on Sunday, this is not a pure time preference: tomorrow’s offer is of a superior bushwalking experience. Likewise, if your preference is only based upon uncertainty that the bushwalk will really happen next week. Maybe I’ve been flaky in the past. Then you prefer a more certain bushwalk over a less certain bushwalk, and the relative levels of uncertainty are merely contingently correlated with where in time those bushwalks occur. Somewhat more obscurely to those not steeped in the relevant philosophical tradition, if you only prefer to bushwalk tomorrow because you believe that the earlier in your life you bushwalk, the better, or because you believe that you will, next week, be yourself to a lesser degree[[1]](#footnote-1) (i.e., be less similar to your current psychology) your preferences are impure.

It has typically been thought that if, after holding fixed all *these* factors, you still have a preference, then you have a pure time preference. Moreover, it has been assumed that agents like us exhibit pure time preferences, and that while in most decision-making scenarios these preferences will be modulated by a variety of other factors, we can control for these factors in order to experimentally reveal pure time preferences. For example, to reduce uncertainty regarding a potential future payoff, instead of presenting participants with the option of being given money later—an offer on which participants might doubt they will follow though—psychologists offer participants cash cards *now* that are set to activate at the agreed time (see, e.g., the methodology of Kable and Glimcher [2007]). However, Callender rejects the orthodox position that by controlling for these factors we can reveal pure time preferences, and we agree.

Even if we have pure time preferences, Callender is sceptical that they can be revealed; for he thinks that there are sources of impurity that psychologists, economists and philosophers have been failing to recognise, or to take seriously. Given the nature of the world in which we live—in particular, considering the thermodynamic asymmetries we observe around here—in any decision between a sooner or later payoff, one must choose between the sooner payoff at age N or the later payoff at age N+. In worlds like ours where the direction (or apparent direction) of time is associated with (or reducible to) the entropic gradient, the recipients of these respective payoffs are different, and so a preference between these options can never straightforwardly reveal a pure time preference.

Indeed, Callender [2022: XXX] argues that while the existence of an entropic gradient is ‘logically detachable from time itself’ (and hence that preferences due to changes in entropy ought to be considered impure), nevertheless, entropic increase (and hence ageing) ‘may well be nomological implications of the best package of laws of nature for our universe’. Hence, he concludes, worlds in which macroscopic objects don’t age might be counter-nomic. In turn, worlds in which we don’t find this particular ‘impurity’ in agents’ preferences are nomically impossible. If that’s so, why should agents located in our world care about pure time preferences, or the norms that might govern them?

We think a somewhat different argument can be marshalled here. Suppose that some nomically possible worlds lack an entropy gradient. Even so, it’s pretty clear that the presence of such a gradient—at least around here—is what makes deliberation possible in the first place. We deliberate, in part, on the basis of information about the past: memories, records, and so on. Indeed, without any past records at all, it would be impossible to deliberate, insofar as we would have no more reason to think that, for instance, eating chocolate would satiate our desire for sugar, than we would to think that it would cause our legs to fall off, or our heads to explode. All deliberation occurs against a backdrop of beliefs about the past, and, in particular, beliefs about reliable connections between actions and effects.

But it has been convincingly argued that a local entropic gradient is necessary for there to be such records (see, for example, Kutach [2013], Albert [2000], and Loewer [2007]). The reason we have records of the past, (and not the future) is down to this local entropic gradient. But if that’s right, then it follows that, of nomic necessity, any region in which there are deliberators, is a region in which there is an entropic gradient, and in which there is ageing towards what those deliberators call the future. So even if it’s not nomically necessary that there be such a gradient, it is nomically necessary that there is such a gradient where there are deliberators, and hence it is nomically necessary that if we are the sorts of things that can deliberate, then we are, of nomic necessity, the sorts of things that age. So, it is nomically necessary that this sort of impurity cannot be removed.

Moreover, our own research has suggested new and unforeseen sources of impurity that reinforce Callender’s concerns. In particular, it appears that whether or not our experimental methodology encourages participants to think of themselves as agents who can act in the world to bring about various events (or not) can drastically change the reported pattern of preferences. As there are reasons to think that in this case the *more* agential preferences are *less* pure, this reinforces the case against the usefulness and relevance of pure time preferences, and the norms that govern them.

One recent study we performed [Greene et al. 2020] confirmed the common assumption that most people are *hedonically* *future-biased:* we prefer that positive experiences be in our future and prefer that negative experiences be in our past. Greene et al. elicited future-biased preferences by asking people how much they agreed that, for instance, they would prefer to learn that they will eat their favourite/most disliked meal tomorrow rather than yesterday. By contrast, another recent study we performed [Greene et al., 2022] found that when people were asked to indicate their preferences about the temporal location of the very same events *using a slider* that ranged from a preference for yesterday to a preference for tomorrow, via ‘no preference between these options’, the vast majority of people preferred that *both their favourite and most disliked meal* be located in the future.

This difference in preferences regarding negative experiences, arising from such a small methodological change, is striking. Greene et al. [ibid.] explain these divergent results by suggesting that the former methodology inclines people to think as though the future is fixed, while the act of pushing the slider inclines people to think of the future as open, and thus to engage with the possibility of mitigating the ‘badness’ of (or even avoiding) eating their most disliked meal. Insofar as one thinks that this option is on the table (as it were), it makes sense to prefer negative events in one’s future, facilitating the opportunity to mitigate or avoid, and as a result maximize one’s total utility. This sort of reasoning, of course, renders these preferences impure: time-neutral reasons such as total utility and uncertainty are playing a substantial role in determining participants overall temporal preferences.

There are two upshots to be considered here. The first is simply that experimentally probing pure time preferences is no easy task. Given that such a small methodological change can make such a difference to reported preferences—and render them impure—we ought to be sceptical that any of our methodologies can isolate pure time preferences.

The second is that in order to isolate something more worthy of being called pure time preferences, Greene et al [2020] needed to supress participants’ tendency to think of themselves as agents who are able to effect change on what future events will occur. But this simply reinforces Callender’s concerns about the usefulness of thinking about pure time preferences and the norms that govern them. Callender [2022: XXX] points out that the kinds of beings that would have pure preferences are ‘those that don’t associate temporal duration with any character at all (e.g., ageing), don’t know that life is finite, and more’. He then wonders why we ought to care about the preferences of such strange creatures. Our research suggests a further way in which these creatures are strange: they don’t see themselves as capable of acting in such a way as to make the future better or worse. They simply see two things that might, passively, happen to them.

While it’s possible to have preferences over states of affairs over which one has no control—we prefer that Australia not go up in flames, but there’s little that we can do about that—insofar as preferences serve as input into deliberations, the very fact of their doing so undermines their purity. If we are deliberate about whether to P, or not P, then we must take it as open to us that we can either P, or not P: we cannot deliberate about that which we take to be fixed and immutable (see Price [1996]). But insofar as we see ourselves as agents with respect to P, and hence as being able to deliberate about whether to P, this will make our P-wise preferences impure. To put things another way, pure preferences, even if we could isolate them, would be completely useless to decision-making agents such as us. Not only would agents who only expressed pure preferences be alien in the ways Callender points out, such agents would not deliberate. That makes it quite mysterious how the norms that might govern those agents’ preferences could be relevant to deliberators like us.

So while psychologists, economists and philosophers often claim to observe pure time preferences, albeit modulated by non-temporal factors, it is worth emphasising again that all we ever *really* observe are impure time preferences. Given what we are required to abstract away from in order to get at pure time preferences, we should never expect to be able to directly observe them. Nor is there much to learn from thinking about such preferences, even if we could discern them, about what preferences *we* ought to have.

The second assumption behind Callender’s master argument that we wish to further investigate is that past-directed preferences can be ignored. While these have been largely neglected in the literature, one might wonder if probing such preferences might be a more effective way to reveal pure time preferences in agents like us. There are a couple of reasons for this. For one, the (local) past is much more certain than the future. Second, these preferences are *not intended* to be useful for decision-making; on the plausible assumption that we cannot causally intercede in the past (backwards causation aside), we cannot deliberate about what to do in the past. Of course, not all of the factors that make preferences impure are automatically eliminated by looking to past-directed preferences. If the ageing process makes preferences impure, then such considerations apply to past-directed preferences too. Similarly, if concerns about whether person-stages at other times are oneself to a high degree play a role here, then the same will be true for past person-stages as for future ones. Nevertheless, there is at least some scope to think that investigation of past-directed preferences might get us closer to pure time preferences.

As Callender notes, following Hedden [2015], EDU has some surprising upshots for past-directed preferences, which are in tension with the common presupposition noted above that we are *hedonically* *future-biased.* That’s because an exponential discounter will value the near past more than the near future, and value the far past more than the near past. Consider again the exponential discounter we introduced in our introductory paragraph, for whom the value of a dollar is halved for every month of delay. She equally values $20 now, and $5 in two months. But when we project her discount curve into the past, we find that she places the same value on $40 a month ago, and $80 two months ago. EDU, then, predicts and prescribes a pattern of past-directed preferences that is quite divorced from the descriptive realities [Callender 2022; Hedden 2015]. Who would equally value $80 two months past and $5 two months future?

It could be that the prescriptions of EDU regarding past-directed preferences are so different from future-biased preferences because future-biased preferences are impure.

Recent work sheds light on this mismatch. We recently performed a study [Greene, Holcombe, Latham, Miller and Norton (2021) which found that as philosophers have predicted, on average, people have *near-biased* preferences in both temporal directions. That is, overall people prefer positive events in their near rather than far future, and near rather than far past, and negative events to be in their far rather than near future, and far rather than near past. In this way, the average participant strikingly fails to comply with the prescriptions of EDU. However, it is notable that a large proportion of people expressed no preferences, and indeed a large proportion of people expressed the opposite preference—a *far-biased* preference. This was so in both temporal directions. So the data is not best represented by the average near-biased response, but as a trimodal distribution, with substantial groups expressing each kind of preference, and a substantial group expressing no preference.

There are four interesting upshots of these results. First, in each temporal direction, we had a substantial group respond with time-neutral preferences. They didn’t mind whether their favourite or most disliked meal was near or far future, or near or far past. Notably, a discount rate of zero is exponential and thus rational by the lights of EDU. Discovering that there are people with these preferences is interesting because prior work has assumed that people discount—at least towards the future—and then the questions of interest have been the rate at which people discount, which personality traits correlate with different discount rates, what kind of discounting is rational, and so on.

Secondly, it is notable that some people are far-biased towards the future: there are people who prefer positive things in their distant rather than near future, and people who prefer negative things to happen sooner rather than later. Perhaps this is because of the value of positive waiting (enjoying waiting for a positive event) and the disvalue of negative waiting (suffering while waiting for a negative event) (see Loewenstein [1987; 1996]). Such factors, of course render these preferences impure. Thirdly, some people in fact report the strange past-directed preferences predicted by EDU—what we might call a far-bias towards the past. This cannot be explained in terms of positive or negative waiting (see Greene et al. 2021 for discussion).

Finally, and surprisingly, there was no association at all between people’s future- and past-directed preferences. That is, the fact that someone is (for example) near-biased towards the future, gives you no information about their past-directed preferences. Without any further information, they are just as likely to be near-biased or far-biased regarding past events. Indeed, there are some subjects in this experiment who are near-biased towards the future, and far-biased towards the past, just as EDU prescribes.

What does this tell us about EDU and pure time preferences? Firstly, it is unclear whether investigating someone’s past-directed preferences is a good way to discover their *pure* future-directed preferences. On one hand, the lack of correlation between *reported* past- and future-directed preferences suggests that investigating someone’s past-directed preferences is not a good methodology for extracting their pure future-directed preferences. On the other hand, this very lack of correlation might be due to the relative lack of purity of reported future-directed preferences, which are being modulated in different ways for different participants. In other words, it is consistent with a lack of correlation in reported past- and future-directed preferences that reported past-directed preferences nonetheless correlate with pure future-directed preferences, but we cannot know whether this is the case.

Secondly, if we set such worries aside and simply assume that experimentally probing past-directed preferences sheds light on people’s *pure* future-directed preferences, these data tell us that there is an amazing amount of interpersonal variation in people’s pure time preferences, in both temporal directions. This is significant because, as Callender [2022: XXX] notes, the Standard Model has it that we are merely ‘known to depart from the rational standard’. These wildly varying results suggests a very different picture according to which, no matter what the rational standard, many of us are flouting it.

If we may be permitted to stretch the metaphor to breaking point, we suggest in conclusion that our recent empirical research serves as a vigorous thump to the table on which Callender has built this ‘high stakes game of Jenga’ [ibid.: XXX]. If agents like us can have pure time preferences, our research suggests that these vary wildly, and thus that EDU is woefully detached from the descriptive realities. Worse, there are reasons to think that the kind of agent who could have pure time preferences is not the kind of agent who makes decisions, and thus that whatever norms might govern such preferences are completely useless to decision-making agents like us.

# References

Albert, D. 2000. *Time and Chance*, Cambridge, MA: Harvard University Press.

Braddon-Mitchell, D, and Miller, K. (2020). “Surviving, to a degree”. *Philosophical Studies* <https://doi.org/10.1007/s11098-019-01410-0> 177:3905-3831.

Callender, C. 2022. The Normative Standard for Future Discounting, *Australasian Philosophical Review* 5/3: XXX-XXX.

Greene, P., Latham, A. J., Holcombe, A., Miller, K., and Norton, J. 2021. “The Rationality of Near Bias towards both Future and Past Events”. Review of Philosophy and Psychology. DOI: 10.1007/s13164-020-00518-1

Greene, P., A.J. Latham, K. Miller, and J. Norton 2020. Hedonic and Non-hedonic Bias Towards the Future, *Australasian Journal of Philosophy* 99/1: 148–63.

Greene, P., A.J. Latham, K. Miller, and J. Norton, J.2022. Why Are People So Darn Past-biased? in *Temporal Asymmetries in Philosophy and Psychology*, eds Christoph Hoerl, Teresa McCormack and Alison Fernandes, Oxford University Press. DOI:10.1093/oso/9780198862901.003.0006.

Hedden, B. 2015. Reasons without Persons: Rationality, Identity, and Time, Oxford University Press.

Kable, J.W., and P.W. Glimcher 2007. The Neural Correlates of Subjective Value during Intertemporal Choice, *Nature Neuroscience* 10/12: 1625.

Kutach, D. 2013. Causation and its Basis in Fundamental Physics. Oxford University Press.

Loewenstein, G. 1996. Out of Control: Visceral Influences on Behavior, *Organizational Behavior and Human Decision Processes* 65/3: 272–92.

Loewenstein, G. 1987. Anticipation and the Valuation of Delayed Consumption, *The Economic Journal* 97/387: 666–84.

Loewer, B. 2007. Counterfactuals and the Second Law, in *Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited*, eds Huw Price and Richard Corry,Oxford University Press: 293–326.

Lowry, R., and M. Peterson 2011. Pure Time Preference, *Pacific Philosophical Quarterly* 92: 490–508

Price, H. 1996. Time’s Arrow and Archimedes’ Point: New Directions for the Physics of Time. Oxford University Press.

Samuelson, P. 1937. A Note on Measurement of Utility, *Review of Economic Studies* 4: 155–61.

1. For a view of this kind see Braddon-Mitchell and Miller [2020]. [↑](#footnote-ref-1)