PRESENT BIAS

ANUJIT CHAKRABORTY

Abstract. Present bias is the inclination to prefer a smaller present reward to a larger later reward, but reversing this preference when both rewards are equally delayed. This paper investigates and characterizes the most general class of present-biased temporal preferences. We show that any present-biased preference has a max-min representation, which can be cognitively interpreted as if the decision maker considers the most conservative present equivalents in the face of uncertainty about future tastes. We also discuss empirical anomalies that temporal models like beta-delta or hyperbolic discounting cannot account for, but the proposed general representation can accommodate.

For the most recent version of this paper, please visit this link.

Exponential discounting is extensively used in economics to study the trade-offs between alternatives that are obtained at different points in time. Under exponential discounting, the relative preference for early over later rewards depends only on the temporal distance between the rewards (stationarity). However, recent experimental findings have called the model into question. Specifically, experiments have shown that small rewards in the present are often preferred to larger rewards in the future, but this preference is reversed when the rewards are equally delayed. As an example, consider the following two choices:

Example 1.

A. \$100 today *vs* B. \$110 in a week

C. \$100 in 4 weeks *vs* D. \$110 in 5 weeks

Date: December 31, 2017.
* I wish to express my deepest gratitude towards my advisor Yoram Halevy for his guidance, encouragement, and support. I also wish to thank Mike Peters, Hao Li, Wei Li, Vitor Farinha Luz, Ali Lazrak, Severine Toussaert, David Freeman, seminar discussants at the University of Toronto and the Summer School of Econometric Society for useful comments, discussions, and suggestions. Financial support from SSHRC (FAS # F11-04991) is gratefully acknowledged.

Many decision makers choose A over B, and D over C. This behavior extends to the domain of primary rewards, as shown by the following choice pattern exhibited by thirsty subjects in an experiment by McClure et al. (2007):

A. Juice now	*vs*	B. Twice the amount of juice in 5 minutes
C. Juice in 20 minutes	*vs*	D. Twice the amount of juice in 25 minutes

60% of all decision makers chose A over B, but only 30% chose C over D, indicating that the prospect of getting twice the amount of juice at the cost of a waiting five more minutes became relatively less attractive when it was pitched against the alternative in the present. This specific pattern of choice reversal can be attributed to a bias we might have towards alternatives in the present, and hence is aptly called present bias or immediacy effect. This is one of the most well documented time preference anomalies (Thaler 1981; Loewenstein and Prelec 1992; Frederick et al. 2002), replicated in the domains of both primary and monetary rewards. If preferences are stable across decision-times and the decision-makers are unable to ward against against the behavior of their future selves, the same phenomenon creates dynamic inconsistency in behavior: People consistently fail to follow up on the plans they had made earlier, especially if the plans entail upfront costs but future benefits. Every year many people pledge to exercise more, eat healthier, become financially responsible or quit smoking starting next year but fail to follow through when the occasion arrives, to their own frustration.

There is a big literature on what kind of utility representations could rationalize choices made by a present-biased decision maker (DM), which we succinctly summarize in Table 3 in Appendix I. Though all of these models capture the behavioral phenomenon of present bias, none of them can be called *the* model of present-biased preferences. Instead they are all models of present bias and *some additional* temporal behavior that is idiosyncratic to the model.[1] Moreover, these

[1] For example, Quasi-Hyperbolic Discounting (called β-δ discounting interchangeably) additionally assumes Quasi-stationarity: violations of constant discounting happen *only* in the present period and the decision maker (DM)'s discounting between any two future periods separated by a fixed distance is always constant. On the other hand, Hyperbolic Discounting (which subsumes Proportional and Power discounting as special cases) captures the behavior of a decision maker whose discounting between any two periods separated by a fixed distance decreases as both periods are moved into the future.

additional behavioral features conflict across the models and are often not empirically well-founded. This raises the following natural question: What is the most general model of present-biased preferences? Or alternatively, what general class of utilities is consistent with present-biased behavior? Such a model would be able to represent present-biased preferences without imposing any extraneous behavioral assumption on the decision maker. This paper proposes a behavioral characterization for such general class of utilities. We start by introspecting about what exactly present-biased behavior implies in terms of choices over temporal objects. The following example provides the motivation for our "weak present bias" axiom.

Example 2. Suppose that a DM chooses (B) \$110 in a week over (A) \$100 today. What can we infer about his choice between (B′) \$110 in 5 weeks versus (A′) \$100 in 4 weeks, if we condition on the person being (weakly) present-biased?[2]

Note that B$\succsim$A, implies that a possible present-premium (\$100 is available at the present) and the early factor (\$100 is available 1 week earlier) are not enough to compensate for the size-of-the-prize factor (\$110>\$100). Equally delaying both alternatives preserves the early factor and the size-of-the-prize factor, but, the already inferior \$100 prize further loses its potential present-premium, which should only make the case for the previous preference stronger. Hence, B$\succsim$A must imply B′ $\succsim$A′ to be consistent with a weak notion of present bias.

We use this motivation to define a Weak Present Bias axiom, which relaxes stationarity by allowing for present bias but rules out any choice reversals inconsistent with present bias[3]. We then show that if a decision maker satisfies Weak Present Bias and some basic postulates of rationality, then, his preferences over receiving an alternative (x, t) (that is receiving prize x at time t) can be represented in the following way (henceforth called the *minimum representation*)

$$V(x,t) = \min{}_{u \in \mathcal{U}} u^{-1}(\delta^t u(x))$$

where $\delta \in (0, 1)$ is the discount factor, and $\mathcal{U}$ is a set of continuous and increasing utility functions. The minimum representation can be interpreted as if the

[2]The choice of monetary reward for this example is without loss of generality. The reader can replace monetary reward with a primary reward in the example, and the main message of this example would still go through undeterred.

[3]Note that we are assuming present-premium≥ 0, thus ruling out the case where it is negative, i.e, something that would be consistent with future bias. This "weak" inequality of present-premium is conceptually equivalent to a "weak" presence of present bias in preferences.

DM has not one, but a set $\mathcal{U}$ of potential future tastes or utilities. Each potential future taste (captured by a utility function $u \in \mathcal{U}$) suggests a different *present equivalent*[4] for the alternative (x, t). The DM resolves this multiplicity by considering the *most conservative or minimal present equivalent* as his final utility. Given that the present equivalent of any prize in the present is the prize itself, the minimum representation has no caution imposed on the present, thus treating present and future in fundamentally different ways. [5] This can also be interpreted as if, immediate alternatives are not evaluated through similar standards of conservativeness, as is expected of a DM with present bias. Moreover, the fact that all alternatives are procedurally reduced to present equivalents for evaluation and comparison, underlines the psychological salience of the present to the DM.

Our model of decision making nests all the popular models of present-biased discounting as special cases, as those models satisfy all the axioms imposed in our analysis. Moreover, there are several robust empirical phenomena discussed in Sections 3 and 9 which temporal models like β-δ or hyperbolic discounting cannot account for, but the current model can. For example, Keren and Roelofsma (1995) show that once all prizes under consideration are made risky, they are no longer subject to present-biased preference reversals anymore. In other words, once certainty is lost, present bias is lost too. None of the models of behavior that treat the time and risk components of an alternative separately (for example, any discounted expected or non-expected utility model) can accommodate such behavior. We extend our analysis to a richer domain of preferences over risky timed prospects and provide an extended minimum representation that can account for this puzzling behavioral phenomenon. In Section 10 we show how a benevolent social planner can use insights from time-risk behavior to improve the welfare of present-biased individuals. Another choice pattern that most temporal models fail to accommodate is the stake dependence of present bias. For example, a DM might have a bias for the present, but he might also expend considerably more cognitive effort to fight off this bias when the stakes are large. His large stake choices would satisfy stationarity, whereas he would appear to be present-biased

[4]*Present equivalent* of an alternative (x, t) is the immediate prize that the DM would consider equivalent to (x, t). For a felicity function u defined on the space of all possible prizes x, and a discount factor of δ, the discounted utility from (x, t) is $\delta^t u(x)$. Hence the corresponding present equivalent is $u^{-1}(\delta^t u(x))$.

[5]As, $\delta^0 = 1$, $u^{-1}(\delta^0 u(x)) = u^{-1}(u(x)) = x$ for all $u \in \mathcal{U}$, and hence, $\min_{u \in \mathcal{U}}(u^{-1}(\delta^0 u(x))) = x$.

in his choices over smaller stakes (see Halevy 2015 for supporting evidence). We show how our representation can accommodate such preferences in Section 9.

The subjective max-min feature of the functional form has been used previously by Cerreia-Vioglio et al. (2015) in the domain of risk preferences, though they had the minimum replaced by an infimum. In their paper, Cerreia-Vioglio et al. (2015) show that if we weaken the Independence axiom to account for the Certainty Effect (Allais 1953), we obtain a representation where a decision maker evaluates the certainty equivalent of each lottery with respect to a set of Bernoulli utility functions and then takes the infimum of those values as a measure of prudence. We discuss this connection in greater detail in Section 5 and describe how the techniques used in our paper can be used to provide an alternative derivation of their main result in a reduced domain.

The paper is arranged as follows: Section 1 defines the novel Weak Present Bias axiom and provides the main representation theorem of the paper. Section 2 builds on the main result to provide intuition about the separation of β-δ discounting from Hyperbolic discounting. Section 3 extends the main result to a richer domain with risk. Section 4 discusses extensions of the representation result to consumption streams. We provide an intuition of the inner workings of the proofs in Section 5. Section 6 comments on the uniqueness of the results. Section 7 surveys the literature closely related to this paper. Sections 8 and 9 discuss the testability, refutability and empirical content of our model. Sections 10-11 provide applications, policy implications and extensions of the main results of the paper. The proofs of the main theorems are included in Appendix II.

1. Model and The Main Result

A decision maker has preferences $\succsim$ defined on all timed alternatives $(x, t) \in \mathbb{X} \times \mathbb{T}$ where the first component could be a desireable prize (monetary or non-monetary) and the second component is the time at which the prize is received. Let $\mathbb{T} = \{0, 1, 2, ...\infty\}$ or $T = [0, \infty)$ and $\mathbb{X} = [0, M]$ for $M > 0$. We impose the following conditions on behavior.

A0: $\succsim$ is complete and transitive.

Completeness and transitivity are standard assumptions in the literature, though one can easily argue that they are more normative than descriptive in nature. The few instances of present-biased intransitive preferences studied in the economics literature, notably Read (2001), Rubinstein (2003) and Ok and Masatlioglu (2007) fall outside our domain of consideration due to (A0).

A1: CONTINUITY: $\succsim$ is continuous, that is the strict upper and lower contour sets of each timed alternative is open w.r.t the product topology.

Continuity is a technical assumption that is generally used to derive the continuity of the utility function over the relevant domain. When, $\mathbb{T} = \mathbb{R}_+$, the standard β-δ model does not satisfy continuity at $t = 0$.[6]

A2: DISCOUNTING: For $t, s \in \mathbb{T}$, if $t > s$ then $(x, s) \succ (x, t)$ for $x > 0$ and $(x, s) \sim (x, t)$ for $x = 0$. For $y > x > 0$, there exists $t \in \mathbb{T}$ such that, $(x, 0) \succsim (y, t)$.

The Discounting axiom has two components. The first part says that the decision maker always prefers any non-zero reward at an earlier date. The second part states that any reward converges to the zero reward (and hence, continually loses its value), as it is sufficiently delayed.

A3: MONOTONICITY: For all $t \in \mathbb{T}$ $(x, t) \succ (y, t)$ if $x > y$.

The Monotonicity axiom requires that at any point in time, larger rewards are strictly preferred to smaller ones. Finally, in light of Example 1, we formally define Weak Present Bias below.

A4: **WEAK PRESENT BIAS:** If $(y, t) \succsim (x, 0)$ then, $(y, t + t_1) \succsim (x, t_1)$ for all $x, y \in X$ and $t, t_1 \in \mathbb{T}$.

The stronger Stationarity axiom is stated below.

Stationarity: $(y, t_1) \succsim (x, t_2)$ if and only if, $(y, t+t_1) \succsim (x, t+t_2)$ for all $x, y \in X$ and $t, t_1, t_2 \in \mathbb{T}$.

Weak present bias, as defined in the fourth axiom is the most intuitive weakening of Stationarity in light of the experimental evidence about present bias or immediacy effect. It allows for choice reversals that are consistent with present-bias,

[6] Pan et al. (2015) axiomatize a model of Two Stage Exponential (TSE) discounting which incorporates the idea of β-δ discounting while maintaining continuity.

something that Stationarity does not allow. On the other hand, having an opposite bias for future consumption is ruled out . [7]Other than all the separable discounting models mentioned in Appendix I, this Weak Present Bias axiom is also satisfied by the non-separable models of present bias proposed by Benhabib et al. (2010) [8] and Noor (2011). This stands testimony to the fact that the Weak Present Bias axiom is able to capture the general behavioral property of present bias in a very succinct way. Now we present our main representation result.

Theorem 1. *The following two statements are equivalent:*

i) The relation $\succsim$ defined on $\mathbb{X} \times \mathbb{T}$ satisfies axioms A0-A4.

ii) For any $\delta \in (0,1)$, there exists a set $\mathcal{U}_\delta$ of monotonically increasing continuous functions such that

$$F(x,t) = \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u(x)) \tag{1}$$

represents the binary relation $\succsim$. The set $\mathcal{U}_\delta$ has the following properties: $u(0) = 0$ and $u(M) = 1$ for all $u \in \mathcal{U}_\delta$. $F(x,t)$ is continuous.

Note that for any timed alternative (x,t), $u^{-1}(\delta^t u(x))$ in (1) computes its "present equivalent", the amount in the present which the individual would deem equivalent to (x,t) if u were his utility function. For all present prizes, the present equivalents are trivially equal to the prize itself ($u^{-1}(\delta^0 u(x)) = x \ \forall u$) irrespective of the utility function under consideration, and thus there is no scope or need for prudence. Whereas for timed alternatives in the future, whenever $\mathcal{U}$ is not a singleton, the DM chooses the most conservative present equivalent due to the minimum functional, thus exhibiting prudence. This is the primary intuition of how this functional form treats the present differently from the future and thus incorporates present bias into it. A potential motivation for the minimum representation and differential treatment towards present and future, follows from Loewenstein (1996)'s visceral states argument: "*..immediately experienced visceral factors have a disproportionate effect on behavior and tend to crowd out virtually*

[7]Further, $(y,t) \succ (x,0)$ and $(y,t+t_1) \sim (x,t_1)$ is also not consistent with WPB, Continuity and Monotonicity. The reason being that, by Continuity, there would exist $y' < y$, $(y',t) \succ (x,0)$ and $(x,t_1) \succ (y',t+t_1)$. Whereas, $(y,t) \sim (x,0)$ and $(y,t+t_1) \succ (x,t_1)$ is allowed by the postulates A0-4.

[8]Benhabib et al. (2010) introduce the discount factor

$$\Delta(y,t) = \begin{cases} 1 & t = 0 \\ (1-(1-\theta)rt)^{(1-\theta)} - \frac{b}{y} & t > 0 \end{cases}$$

all goals other than that of mitigating the action, ...but.. people under weigh, or even ignore, visceral factors that they will experience in the future." The following example shows an easy application of the theorem to represent present-biased choices.

Example 3. Consider $\mathcal{U} = \{u_1, u_2\}$, where,

$$\begin{aligned} u_1(x) &= x^a \text{ for a>0} \\ u_2(x) &= 1 - \exp(-bx) \text{ for b>0} \end{aligned}$$

Also consider, $a = .99$, $b = .00021$, $\delta = .91$. One can easily check that a minimum representation with respect to this $\mathcal{U}$ would satisfy Weak Present Bias (also follows from Theorem 1). The minimum representation with respect to this $\mathcal{U}$ would assign the following utilities to the timed alternatives in Example 1.

$$\begin{aligned} V(100,0) &= \min(100,100) = 100 \\ V(110,1) &= \min(100.056, 99.995) = 99.995 \\ V(100,4) &= \min(68.317, 68.48) = 68.317 \\ V(110,5) &= \min(68.320, 68.344) = 68.320 \end{aligned}$$

Hence,

$$\begin{aligned} V(100,0) &> V(110,1) \\ V(100,4) &< V(110,5) \end{aligned}$$

Thus the minimum function with a simple $\mathcal{U}$ can be used to accommodate present biased choice reversals.

2. Special cases

This section applies Theorem 1 to a popular model of present bias, the β-δ model (Phelps and Pollak 1968; Laibson 1997). The $\beta - \delta$ model evaluates each alternative (x,t) as $U(x,t) = (\beta + (1-\beta).1_{t=0})\delta^t u(x)$, where u, δ, β have standard interpretation. $1_{t=0}$ is the indicator function that takes value of 1 if $t = 0$ and value 0 otherwise, thus assigning a special role to the present. Given that the $\beta - \delta$ model satisfies Weak Present Bias and all the other axioms included in Theorem 1 (for the discrete case) , any such $\beta - \delta$ representation must have an alternative minimum representation, as shown in Theorem 1.

Below, we consider the simplest possible $\beta - \delta$ representation with linear felicity function $u(x) = x$, $\mathbb{T} = \{0, 1, 2, ..\}$ and construct the corresponding Weak Present Bias representation.

Claim 1. β-δ representation with $u(x) = x$ has an alternative minimum representation.

Proof. Define the functions $u_y : \mathbb{R} \to \mathbb{R}_+$ for all $y \in \mathbb{R}_+$:

$$u_y(x) = \begin{cases} \dfrac{x}{\beta} & \text{for } x \leq \beta\delta y \\ \delta y + (x - \beta\delta y)\dfrac{1-\delta}{1-\beta\delta} & \text{for } \beta\delta y < x \leq y \\ x & \text{for } x > y \end{cases}$$

For any $y \in \mathbb{R}_+$, $x \leq u_y(x) \leq \dfrac{x}{\beta}$ for all $x \in \mathbb{R}_+$. As u_y is an increasing function, it must be that $x \geq u_y^{-1}(x) \geq \beta x$. Since, $x \leq u_y(x)$, we get $\delta^t u_y(x) \geq \delta^t x$, which implies,

$$u_y^{-1}(\delta^t u_y(x)) \geq u_y^{-1}(\delta^t x) \geq \beta\delta^t x$$

Finally, for $x = y$, $\delta^t u_y(x) = \delta^t x < \delta x$ and, hence, $u_y(\delta^t u_y(x)) = \beta\delta^t x$.

Therefore, $V(x,t) = \min_{y \in \mathbb{R}_+} u_y^{-1}(\delta^t u_y(x)) = (\beta + (1-\beta).1_{t=0})\delta^t x$, which finishes our proof.[9] ☐

This shows that if we start with a rich enough set of piece-wise linear utilities, the minimum representation with respect to that set, is enough to generate behavior consistent with β-δ discounting. In the example above, the set values taken by the set of functions is bounded above and below at each non-zero point x of the domain by $[\dfrac{x}{\beta}, x]$, and this brings us to our next result. Our next theorem characterizes the behavioral axiom necessary and sufficient for the functions in $\mathcal{U}_\delta$ to be similarly bounded.

We start by introducing two more axioms.

A5: EVENTUAL STATIONARITY: For any $x > z > 0 \in \mathbb{X}$, there exists $t_1 \in \mathbb{T}$, such that for all $t \geq 0$, $(z, t) \succ (x, t + t_1)$ and $(z, 0) \succ (x_t, t_1 + t)$ for all x_t such that $(x, 0) \sim (x_t, t)$.

[9]This is not necessarily the only possible minimum-representation of the β-δ discounting.

A6: NON-TRIVIALITY: For any $x \in \mathbb{X}$, and $t \in \mathbb{T}$, there exists $z \in \mathbb{X}$, such that $(z, t) \succ (x, 0)$.

The last axiom basically means that the space of prizes is rich enough to have exceedingly better outcomes is imposed *only when* $\mathbb{X} = \mathbb{R}_+$ and not for $\mathbb{X} =[0, M]$. (Compare Theorem 2 and Corollary 1)

A5 is the more crucial axiom. That for any $x > z > 0 \in X$, there exists a sufficient delay $\tau_1 \in \mathbb{T}$, such that $(z, 0) \succ (x, \tau_1)$ is already implied by Discounting (A2). What has been added is the existence of delay t_1 for which we additionally have $(z, t) \succ (x, t + t_1)$ for all $t \geq 0$: This intuitively means once the later larger prize is "sufficiently" delayed, the relative rates at which the attractiveness of the earlier and later rewards fall with further delay (increasing values of t) are consistent with stationarity. This rules out certain preference reversals that were previously allowed under WPB. The last and third part of the axiom, $(z, 0) \succ (x_t, t_1 + t)$ for any x_t such that $(x, 0) \sim (x_t, t)$, also has the same interpretation. The A5 property provides a crucial separation between two popular classes of present-biased discounting functions: β-δ discounting and Hyperbolic discounting , as only the former satisfies it, but the latter does not. We show this more formally in Proposition 4 in Appendix II.

Theorem 2. *Let* $\mathbb{T} = \{0, 1, 2, ...\infty\}$ *and* $\mathbb{X} = \mathbb{R}_+$. *The following two statements are equivalent:*

i) The relation $\succsim$ *satisfies properties A0-A6.*

ii) There exists a set $\mathcal{U}_\delta$ *of monotonically increasing continuous functions such that*

$$F(x, t) = \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u(x)) \tag{2}$$

represents the binary relation $\succsim$. *The set* $\mathcal{U}_\delta$ *has the following properties:*$u(0) = 0$ *for all* $u \in \mathcal{U}$, $\sup_u u(x)$ *is bounded above,* $\inf_u u(x) > 0$ $\forall x > 0$, $\inf_u \dfrac{u(z)}{u(x)}$ *is unbounded in* z *for all* $x > 0$. $F(x, t)$ *is continuous.*

This theorem implies that any "minimum-representation" of hyperbolic discounting must require a set of functions which would take unbounded set values at some point of the domain. The immediate conclusion one can draw from here is that one cannot generate any variant of Hyperbolic discounting (coupled with any felicity function) with a minimum representation over a finite set $\mathcal{U}$ of utilities. This theorem also has a straightforward corollary, where we consider the prize domain $\mathbb{X}$ =$[0, M]$ and drop A6.

Corollary 1. *Let* $\mathbb{T} = \{0, 1, 2, ...\infty\}$ *and* $\mathbb{X} = [0, M]$*. The following two statements are equivalent:*

i) The relation $\succsim$ *satisfies properties A0-A5.*

ii) There exists a set $\mathcal{U}_\delta$ *of monotonically increasing continuous functions such that*

$$F(x, t) = \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u(x))$$

represents the binary relation $\succsim$*. The set* $\mathcal{U}_\delta$ *has the following properties:*$u(0) = 0$*,* $u(1) = 1$ *for all* $u \in \mathcal{U}_\delta$*,* $\inf_u u(x) > 0 \ \forall x$*.* $F(x, t)$ *is continuous.*

3. An extension to risky prospects

In this section, we extend the representation derived in Section 1 to risk. This extension serves the following three goals. First, it shows that the representation in Section 1 has a natural extension to simple binary lotteries, with zero being one of the lottery outcomes. Second, through the extended representation we are able to accommodate experimental evidence that is inconsistent with most previous temporal models of behavior. Finally, through this extension, we will be able to identify a unique discount factor δ for any DM satisfying certain postulates of behavior.

We start by presenting the experimental evidence from time-risk domain that our model would be able to accommodate, but, the temporal models from Appendix I would not. In the following text, we summarize each alternative by the triplet (x, p, t) where x is a monetary prize, p is the probability with which x is attained at time t. For the first three rows, x was offered in Euros, and in the next four, x was offered in Dutch Guilder, t was measured in months in Columns 1:3, and measured in weeks in Columns 4:7.

The data can be interpreted in the following way: People have an affinity for both certainty and immediacy. The loss in either certainty or immediacy has a similar disproportionate effect on preferences (compare rows 5 and 6 with row 4, or rows 2-3 with row 1). Further, this disproportional effect for a loss in certainty can only be observed when all choices are in the present, and that for the loss of

	Prospect A	Prospect B	% chosing A	% chosing B	N
1	(9,1,0)	(12,.8,0)	58%	42%	142
2	(9,.1,0)	(12,.08,0)	22%	78%	65
3	(9,1,3)	(12,.8,3)	43%	57%	221
4	(100,1,0)	(110,1,4)	82%	18%	60
5	(100,1,26)	(110,1,30)	37%	63%	60
6	(100,.5,0)	(110,.5,4)	39%	61%	100
7	(100,.5,26)	(110,.5,30)	33%	67%	100

TABLE 1. Rows 1-3 are from Baucells and Heukamp (2010), the rest are from Keren and Roelofsma (1995).

immediacy can only be observed when all choices are certain. For example, there is very little evidence of present-biased reversals over risky prospects (compare rows 6-7, with rows 4-5). It is the latter finding that is at odds with most temporal models of behavior. In fact it rules out *all* discounted expected or non-expected utility functional forms which are separable in the temporal and risk components.[10]

We will consider preferences over triplets $(x, p, t) \in \mathbb{X} \times \mathbb{P} \times \mathbb{T}$, which describe the prospect of receiving a reward $x \in \mathbb{X}$ at time $t \in \mathbb{T}$ with a probability$p \in [0, 1]$. $\mathbb{X} = [0, M]$ is a positive reward interval, $\mathbb{P} = [0, 1]$ is the unit interval of probability, and $\mathbb{T} = [0, \infty)$ is the time interval. We impose the following conditions on behavior.

B0: $\succsim$ is complete and transitive.

B1: CONTINUITY: $\succsim$ is continuous, that is the strict upper and lower contour sets of each risky timed alternative are open w.r.t the product topology.

B2: DISCOUNTING: For $t, s \in \mathbb{T}$, if $t > s$ then $(x, p, s) \succ (x, p, t)$ for $x, p > 0$ and $(x, p, s) \sim (x, p, t)$ for $x = 0$ or $p = 0$. For $y > x > 0$, there exists $T \in \mathbb{T}$ such that, $(x, q, 0) \succsim (y, 1, T)$.

B3: PRIZE AND RISK MONOTONICITY: For all $t \in \mathbb{T}$, $(x, p, t) \succsim (y, q, t)$ if $x \geq y$ and $p \geq q$. The preference is strict if at least one of the two following inequalities is strict.

Note that the first four axioms are just extensions of A0-A3.

[10]Rows 1 and 3 also imply the same.

B4: WEAK PRESENT BIAS: If $(y, 1, t) \succsim (x, 1, 0)$ then, $(y, 1, t + t_1) \succsim (x, 1, t_1)$ for all $x, y \in \mathbb{X}$, $\alpha \in [0, 1]$ and $t, t_1 \in \mathbb{T}$.

B5: PROBABILITY-TIME TRADEOFF: For all $x, y \in \mathbb{X}$, $p, q, \theta \in (0, 1]$, and $t, s, D \in \mathbb{T}$, $(x, p\theta, t) \succsim (x, p, t + D) \implies (y, q\theta, s) \succsim (y, q, s + D)$.

The fifth axiom (used previously in Baucells and Heukamp 2012) says that passage of time and introduction of risk have similar effects on behavior, and there is a consistent way in which time and risk can be traded off across the domain of behavior. This axiom additionally implies calibration properties that we will utilize to pin down a unique discount factor δ for any DM. True to the spirit of the experimental findings discussed above, the Weak Present Bias axiom is only imposed for prizes with certainty.

Additionally, (B4) when combined with (B5) captures a decision maker's joint bias towards certainty as well as the present, i.e, it embeds Weak Present Bias *as well as Weak Certainty Bias*[11] in itself. This underlines the insight that once risk and time can be traded-off, Weak Present Bias and Weak Certainty Bias are behaviorally equivalent. Similar relations between time and risk preferences have been elaborated on previously by Halevy (2008), Baucells and Heukamp (2012), Saito (2009), Fudenberg and Levine (2011), Epper and Fehr-Duda (2012) and Chakraborty and Halevy (2015). In Section 5, we will discuss how the Weak Certainty Bias postulate connects the current work to previous literature on risk preferences.

We are now ready for our next result.

Theorem 3. *The following two statements are equivalent:*

i) The relation $\succsim$ on $\mathbb{X} \times \mathbb{P} \times \mathbb{T}$ satisfies properties B0-B5.

ii) There exists a unique $\delta \in (0, 1)$ and a set $\mathcal{U}$ of monotonically increasing continuous functions such that $F(x, p, t) = \min_{u \in \mathcal{U}}(u^{-1}(p\delta^t u(x)))$ represents the relation $\succsim$. For all the functions $u \in \mathcal{U}$, $u(M) = 1$ and $u(0) = 0$. Moreover, $F(x, p, t)$ is continuous.

The next example shows a potential application of this representation in light of Keren and Roelofsma (1995)'s experimental results.

[11] Weak Certainty Bias can be defined on $\mathbb{X} \times \mathbb{P}$ in the following fashion: If $(y, p) \succsim (x, 1)$ then, $(y, p\alpha) \succsim (x, \alpha)$ for all $x, y \in \mathbb{X}$ and $\alpha \in [0, 1]$.

Example 4. Consider the set of functions $\mathcal{U}$ and parameters considered in Example 3. When applied to the representation derived in Theorem 3, they predict the following choice pattern.

$$\begin{aligned} V(100,1,0) &> V(110,1,1) \\ V(100,1,4) &< V(110,1,5) \end{aligned}$$

$$\begin{aligned} V(100,.5,0) &< V(110,.5,1) \\ V(100,.5,4) &< V(110,.5,5) \end{aligned}$$

Note that this is exactly the choice pattern obtained in the original Keren and Roelofsma (1995) experiment: time and risk affect choices in similar ways, and once certainty is removed present bias disappears.

4. Extension to consumption streams

In this section, we extend the representation derived in Section 1 to deterministic consumption streams. The DM's preferences $\succsim$ are defined over $[0,\infty)^T$, the set of all consumption streams of finite length $T > 1$. We impose the following conditions on behavior.

D0: $\succsim$ is complete and transitive.

D1: CONTINUITY: $\succsim$ is continuous, that is the strict upper and lower contour sets of each consumption stream are open w.r.t the product topology.

D2: DISCOUNTING: If $0 \leq s < t \leq T-1$, then $(0,..\underbrace{y}_{\text{in period } s},..,0) \succsim (0,..\underbrace{y}_{\text{in period } t},..,0)$ for $y \geq 0$ with the relation being strict if and only if $y > 0$. Further, for $y_0 > x > 0$, and for any sequences $(y^1, y^2, y^3, ..y^m)$ and $(n^1, n^2, .., n^m)$, where, $(0,..0,\underbrace{y^{i-1}}_{\text{in period } n^i},0..,0) \succsim (y^i,0,..,0)$ $\forall i \in \{1,2,...,m\}$, $0 < n^i \leq T-1$ and $\sum_1^m n^i = t$, there exists $t \in \mathbb{N}$ such that, $y_m \leq x$.

D3: MONOTONICITY: For any $(x_0, x_1, ..x_{T-1})$, $(y_0, y_1, ..y_{T-1}) \in [0,\infty)^T$, $(x_0, x_1, ..x_{T-1}) \succsim (y_0, y_1, ..y_{T-1})$ if $x_t \geq y_t$ for all 0≤t $\leq T-1$. The preference is strict if at least one of the inequalities is strict.

D4: WEAK PRESENT BIAS: If $(0,..\underbrace{y}_{\text{in period } t},..,0) \succsim (x,0,..,0)$ then, $(0,..\underbrace{y}_{\text{in period } t+t_1},..,0) \succsim (0,.\underbrace{x}_{\text{in period } t_1}.,0)$ for all $x,y \in \mathbb{X}$ and $t,t_1 \in \mathbb{T}$.

Note that the first five axioms are alternative restatements of A0-A4 in the current domain, but the Discounting axiom warrants some independent discussion. As before, the second part of the Discounting axiom states that any period-consumption keeps falling arbitrarily in present-equivalent value, as one increases the total discounting it is subjected to. Due to the added restriction that the DM can only consider time delays of upto $T-1$ periods for $T \geq 2$, we have approximated arbitrary delays by a sequence of delays, none greater than $T-1$. But, the restatement is also a stronger version of the former (under WPB) as it also imposes path independence (by stating the axiom for arbitrary sequences $(n^i)_{i=1}^m$ of delays instead of requiring it to hold for a particular sequence of delays, that sum to t) while achieving this total discounting. This is necessary while working with the non-compact prize space of $[0,\infty)$.

D5: STRONG ADDITIVITY: For any pair of orthogonal[12] consumption bundles $(x_0,x_1,..x_{T-1}),\ (y_0,y_1,..y_{T-1}) \in [0,\infty)^T$, if, $(x_0,x_1,..x_{T-1}) \sim (z_0,0,..,0)$ and $(y_0,y_1,..y_{T-1}) \sim (z_0',0,..,0)$, then,
$(x_0+y_0,x_1+y_1,..x_{T-1}+y_{T-1}) \sim (z_0+z_0',0,..,0)$.

Orthogonality of consumption vectors imply that $x_t > 0$ only if $y_t = 0$, and $y_t > 0$ only if $x_t = 0$ for all t. The fifth axiom implies the standard notion of Additivity used in axiomatizations of additive representation of streams, and is hence named Strong Additivity.

We are now ready for our next result.

Theorem 4. *The following two statements are equivalent:*

i) The relation $\succsim$ on $[0,\infty)^T$ satisfies properties D0-D5.

ii) For any $\delta \in (0,1)$, there exists a set $\mathcal{U}_\delta$ of monotonically increasing continuous functions such that

$$F(x_0,x_1,..,x_{T-1}) = x + \sum_{1}^{T-1} \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u(x_t))$$

12

Two vectors are orthogonal if their dot product is zero.

represents the binary relation $\succsim$. The set $\mathcal{U}_\delta$ has the following properties: $u(0) = 0$ and $u(M) = 1$ for all $u \in \mathcal{U}_\delta$. $F(.)$ is continuous.

5. An outline of the proofs

This section outlines the proofs of Theorems 1-3 chronologically and places the methodology used in the proofs in the context of recent literature.

We will provide the outline for the case of $\mathbb{T} \in [0, \infty)$, as it is less technical but conveys the main idea behind the proofs nonetheless. For any timed alternative (z, τ), there exists $x \in \mathbb{X}$ such that $(z, \tau) \sim (x, 0)$. This follows from monotonicity, continuity, connectedness of the prize-domain and this guarantees that any (timed) alternative has a well defined present equivalent with respect to $\succsim$. It is easy to see that when $\tau = 0$, one must have $z = x$. Given the present equivalents with respect to $\succsim$ are well defined, one possible utility representation $V : \mathbb{X} \times \mathbb{T} \to \mathbb{R}_+$ is the function that assigns to every alternative (z, τ), the present equivalent according to the relation $(z, \tau) \sim (x, 0)$. The crux of the remaining proof lies in showing that there exists a set of utilities $\mathcal{U}_\delta$ such that the previously defined V function can be rewritten as

$$V(z, \tau) = x \quad = \quad \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^\tau u(z))$$

The proof is constructive. For any point $x^* \in (0, M)$, we construct a function $u_{x^*}(.)$ in the following steps.

i) We assign $u_{x^*}(0) = 0$, $u_{x^*}(x^*) = 1$.

ii) For any $x \in (x^*, M]$, we find $t > 0$ such that $(x, t) \sim (x^*, 0)$. Define, $u_{x^*}(x) = \delta^{-t}$ (for any $\delta \in (0, 1)$ under consideration) and re-label x as x_t.

iii) For $y \in (0, x^*)$, define $u_{x^*}(y) = \min\{\delta^\tau : \ (x_t, t + \tau) \sim (y, 0)\}$ for some t from step (ii).

We show that the minimum is well defined in step (iii), and the constructed $u_{x^*}()$ is strictly increasing, continuous, and has the following crucial property: If $(z, t) \sim (x, 0)$ then, $\delta^t u_{x^*}(z) \geq u_{x^*}(x)$ and subsequently, $u_{x^*}^{-1}(\delta^t u_{x^*}(z)) \geq x$, with the weak inequality replaced by equality if $x = x^*$. The asymmetric construction of $u_{x^*}()$ on the left and right of x^* is crucial for this to hold.

Next we define $\mathcal{U}_\delta = \{u_{x^*}(.) : \ x^* \in (0, M)\}$. It readily follows from the aforementioned property of constructed utility functions that $\min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u_x(z)) = x$ whenever $(z, t) \sim (x, 0)$.

Theorem 2 builds on these methods and insights of Theorem 1. Eventual Stationarity gurantees that the functions in $\mathcal{U}$ can be constrcuted in a way such that for any two points $x < y$ there exists t_1 for which $u(x) > \delta^{t_1} u(y)$ for all $u \in \mathcal{U}$. Now when one normalizes, $u(1) = 1$ for all $u \in \mathcal{U}$, using the condition mentioned in the previous sentence, one additionally obtains that $\sup_u u(x)$ is bounded above and $\inf_u u(x) > 0 \; \forall x > 0$.

Theorem 3 connects time and risk in the following way: Given the Probability-Time Tradeoff axiom, the $\mathbb{X} \times \mathbb{P} \times \mathbb{T}$ domain is isomorphic to either of the reduced domains of $\mathbb{X} \times \mathbb{P}$ or $\mathbb{X} \times \mathbb{T}$. For example, there exists unique $\delta \in (0, 1)$ such that $(x, p, t) \sim (x, p\delta^t, 0)$ and $(x, p, t) \sim (x, 1, t + \log_\delta p)$ for all $x \in \mathbb{X}$ and $p \in \mathbb{P}$. This theorem restricts its domain to $\mathbb{T} = \mathbb{R}_+$, unlike Theorem 1, which holds equally for $\mathbb{T} = \mathbb{N}_0$ as well. The axioms on $\mathbb{X} \times \mathbb{P} \times \mathbb{T}$ domain imply completeness, transitivity, continuity, risk monotonicity (Discounting respectively), Weak Certainty Bias (Weak Present Bias respectively) for a preference defined on the reduced domain of $\mathbb{X} \times \mathbb{P}$ ($\mathbb{X} \times \mathbb{T}$ respectively for $\mathbb{T} = \mathbb{R}_+$). Proving Theorem 3, now reduces to proving that there is a minimum representation on $\mathbb{X} \times \mathbb{P}$ or $\mathbb{X} \times \mathbb{T}$ of the forms $\min_{u \in \mathcal{U}}(u^{-1}(pu(x)))$ or $\min_{u \in \mathcal{U}}(u^{-1}(\delta^t u(x)))$ respectively. Additionally, proving any one of the representations from the implied axioms on the relevant domain is equivalent to proving all of the representations on the respective domains. This flexibility is allowed by the Probability Time Tradeoff axiom. In the Appendix, we prove how the reduction from the richer domain to $\mathbb{X} \times \mathbb{P}$ or $\mathbb{X} \times \mathbb{T}$ works, and then prove that a relation on $\mathbb{X} \times \mathbb{P}$ satisfies completeness, transitivity, continuity, risk monotonicity and Weak Certainty Bias *if and only if* the relation on $\mathbb{X} \times \mathbb{P}$ can be represented by the functional form of $\min_{u \in \mathcal{U}}(u^{-1}(pu(x)))$.

This result on the reduced $\mathbb{X} \times \mathbb{P}$ domain brings us to a very interesting connection that the present work has with Cerreia-Vioglio et al. (2015). In that paper, the authors consider preferences over lotteries ($\mathcal{L}$) defined over a compact real interval $[w, b]$ of outcomes. To account for violations of the Independence Axiom[13] based on a DM's bias towards certainty or sure prizes[14], they relax it in favor of Negative Certainty Independence (NCI) axiom defined below.

NCI: (Dillenberger 2010) For $p, q \in \mathcal{L}$, $x \in [w, b]$, and $\lambda \in (0, 1)$,

$$q \trianglerighteq L_x \implies \lambda p + (1-\lambda) q \trianglerighteq \lambda p + (1-\lambda) L_x$$

[13]For $p, q, r \in L$, and $\lambda \in (0, 1)$, $p \trianglerighteq q$ if and only if $\lambda p + (1-\lambda) r \trianglerighteq \lambda q + (1-\lambda) r$.
[14]We denote the lottery that gives the outcome $x \in [w, b]$ for sure as $L_x \in \mathcal{L}$.

Cerreia-Vioglio et al. (2015) show that if $\trianglerighteq$ satisfies NCI and some basic rationality postulates, then there exists a set of continuous and strictly increasing functions $\mathcal{W}$, such that the relation $\trianglerighteq$ can be represented by a continuous function $V(p) = \inf_{u \in \mathcal{W}} c(p,u)$, where $c(p,u)$ is the certainty equivalent of the lottery p with respect to $u \in \mathcal{U}$. The proof of their theorem has the following steps: From $\trianglerighteq$, they construct a partial relation $\trianglerighteq'$ which is the largest sub-relation of the original preference $\trianglerighteq$ that satisfies the Independence axiom. By Cerreia-Vioglio (2009), $\trianglerighteq'$ is reflexive, transitive (but possibly incomplete), continuous and satisfies Independence. Next, following Dubra et al. (2004) [15], there exists a set $\mathcal{W}$ of continuous functions on $[w, b]$ that constitutes an Expected Multi-Utility representation of $\trianglerighteq'$, that is, $p \trianglerighteq' q$ if and only if $E_v(p) \geq E_v(q)$ for all $v \in \mathcal{W}$. Now taking an infimum of the present equivalents with respect to all the functions in $\mathcal{W}$ yields a representation that assigns to each lottery its certainty equivalent implied by the relation $\trianglerighteq$.

This NCI axiom when reduced to the domain of binary lotteries on $\mathbb{X} \times \mathbb{P}$, conveys the same behavior as the Weak Certainty Bias axiom we have discussed above and have used in the proof of our theorem. Our representation over $\mathbb{X} \times \mathbb{P}$ is a minimum representation that is an exact parallel of the infimum representation obtained by Cerreia-Vioglio et al. (2015). This is no coincidence: we provide an alternative derivation of Cerreia-Vioglio et al. (2015)'s result in a reduced domain of lotteries for similar behavior and show that their infimum representation can be replaced with a minimum representation under the implied axioms in our domain. Our proof is essentially constructive, as illustrated in Claim 1, and it does not use any intermediate results (for example, results from Dubra et al. (2004)).

The similarity in functional forms naturally prompts the question: Could the proof in Cerreia-Vioglio et al. (2015) be applied directly to our representation theorems? The answer to the question is negative for the following two reasons. Firstly, when an NCI-like axiom (Weak Certainty Bias) is imposed on my restricted domain of binary lotteries, the results from Cerreia-Vioglio et al. (2015) no longer follow as corollaries of their main theorem due to the reduced strength of the implied axioms. This follows the usual relation between size of domain

[15] Dubra et al. (2004) define a convex cone in the linear space generated by the lotteries related by $\trianglerighteq'$ and then apply an infinite-dimensional version of the separating hyperplane theorem to establish the existence of $\mathcal{W}$.

and strength of axiom. Secondly, there is no way of starting with an appropriately defined axiom of present bias on consumption streams (instead of timed payments) and reaching a present-biased utility representation on streams by using the route (Present Bias)⇔(NCI)⇔(Multi EU) ⇔ (Present-biased representation), under any equivalence of time and implicit risk necessary for the first and last steps.

6. Uniqueness

The uniqueness results discussed here are formulated keeping the main representation theorem of the paper in mind, but they apply equally to the other representation theorems with minor adjustments. We start with a crucial question about the representation: Could we have come across an alternative representation for the same preferences without the exponential discounting part inside the present equivalents? For example, could we have ended up with a representation of the form:

$$V'(x,t) = \min_{u\in\mathcal{U}} u^{-1}(\Delta(t)u(x)) \tag{3}$$

where $\Delta(t)$ is some time-decreasing *discount function* other than exponential discounting, for example the hyperbolic one? Note that this is an interesting question, as a positive answer would open the door to representations where the present equivalents are taken with respect to hyperbolic or quasi-hyperbolic discounting. However, the answer is negative. If we start with any $\Delta(t)$ such that $\dfrac{\Delta(t+t_1)}{\Delta(t)} \neq \Delta(t_1)$ for some t, t_1 , there would either 1) exist some binary relation which satisfies all the axioms in this paper, but cannot be represented by the representation in (3), or 2) the representation in (3) with a permissible set of utilities $\mathcal{U}$ would represent preferences which do not satisfy at least one of the axioms in this paper, thus breaking the two-way relation between the axioms and representation.

Proposition 1. *Given the axioms A0-4, the representation in (3) is unique in the discounting function $\Delta(t) = \delta^t$ inside the present equivalent function.*

Proof. See Appendix II. □

One of the limitations of representations over $\mathbb{X} \times \mathbb{T}$ space (the domain used in Sections 1 and 2) is the lack of uniqueness in terms of the discount factor δ. We

inherit the non-uniqueness of δ in Theorems 1-2 from Fishburn and Rubinstein (1982). Fishburn and Rubinstein (1982) impose A0-A3 along with Stationarity on preferences to derive a exponential discounting representation. In their representation, given those conditions on preferences, and given $\delta \in (0,1)$, there exists a continuous increasing function f such that (x,t) is weakly preferred to (y,s) if and only if $\delta^t u(x) \geq \delta^s u(y)$. They have the following result: if (u,δ) is a representation for a preference $\succsim$ then so is (v,β) where $\beta \in (0,1)$ and $v = u^{\frac{\log \beta}{\log \delta}}$. Same holds for our representations in Theorems 1-2: if $(\delta, \mathcal{U})$ is a representation of $\succsim$, then so is $(\alpha, \mathcal{F})$,where $\mathcal{F}$ is constructed by the functions $v = u^{\frac{\log \beta}{\log \delta}}$ for $u \in \mathcal{U}$. Obviously this is a restriction imposed by working on the prize-time domain and we can no longer consider δ as a measure of impatience. To put things in perspective, in a seminal paper Koopmans (1972) instead considers the richer domain of consumption streams, and under the additional assumptions of separability and stationarity, he derives a time-separable additive exponential discounting representation of behavior. In Theorem 3 we provide a representation over a richer domain where the discount factor $\delta \in (0,1)$ is unique.
Next, we show that the set of functions in the representation in (1) is unique up to its convex closure. Define

$$\mathcal{F} = \{u : [0, M] \to \mathbb{R}_+ : u(0) = 0,\ u \text{ is strictly increasing and continuous}\}$$

Define the topology of compact convergence on the set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$. Also, let $co(A)$ and $\bar{A}$ define the convex hull and closure of the set A (with respect to the defined topology), and $\bar{co}(A)$ define the convex closure of the set A.

Proposition 2. *If $\mathcal{U}, \mathcal{U}' \subset \mathcal{F}$ are such that $\bar{co}(\mathcal{U}) = \bar{co}(\mathcal{U}')$, and the functional form in (1) allows for a continuous minimum representation for both of those sets, then,* $\min_{u \in \mathcal{U}} u^{-1}(\delta^t u(x)) = \min_{u \in \mathcal{U}'} u^{-1}(\delta^t u(x))$.

Proof. See Appendix II. □

Proposition 3. *i) If there exists a concave function $f \in \mathcal{U}$, and if $\mathcal{U}_1$ is the subset of convex functions in $\mathcal{U}$, then* $\min_{u \in \mathcal{U}}(u^{-1}(\delta^t u(x))) = \min_{u \in \mathcal{U} \setminus \mathcal{U}_1}(u^{-1}(\delta^t u(x)))$.

ii) If $u_1, u_2 \in \mathcal{U}$ and u_1 is concave relative to u_2, then, $\min_{u \in \mathcal{U}} u^{-1}(\delta^t u(x)) = \min_{u \in \mathcal{U} \setminus \{u_2\}} u^{-1}(\delta^t u(x))$.

Proof. See Appendix II. □

The Uniqueness results in this section would be well-served by a result of the following form: For any preference relation $\succsim$ that satisfies axioms A0-A4, there exists a set $\mathcal{U}^*_\delta$, such that if $\mathcal{U}_\delta$ constitutes a minimum representation of $\succsim$then $\mathcal{U}^*_\delta \subseteq \bar{co}(\mathcal{U}_\delta)$, and the set $\mathcal{U}^*_\delta$is unique up to its closed convex hull. Such a set $\mathcal{U}^*_\delta$ could potentially be derived as a multi-discounted utility representation for the largest sub-relation of $\succsim$ on which stationarity is satisfied, and if the resulting set of utilities in the multi-utility representation is also unique up to convex closures.

7. Related Literature

This paper is closely linked to the literature that explores the conditions under which a "rational" person may have present-biased preferences. Sozou (1998), Dasgupta and Maskin (2005) and Halevy (2008) explain particular uncertainty conditions that can give rise to present-biased behavior. While telling an uncertainty story *sufficient* to explain present bias, all these models explicitly assume the particular structure of risk or uncertainty with a relevant risk attitude, and these assumptions are central to establishing behavior consistent with present bias in the respective models. In this paper we deviate from this norm: we do not explicitly assume any uncertainty framework or uncertainty attitude. But we still obtain a subjective state space representation that *is necessary and sufficient* for present bias. The set of future tastes $\mathcal{U}$ can be considered to be the subjective state-space, and the decision maker considers the most conservative state dependent utility $\min_{u\in\mathcal{U}} u^{-1}(\delta^t u(x))$ to evaluate each timed alternative.

Our representation looks similar to the max-min expected utility representation of Gilboa and Schmeidler (1989) used in the uncertainty or ambiguity aversion literature, though there is no objective state space or uncertainty defined in our set-up. We have already discussed the connection of our paper with Cerreia-Vioglio et al. (2015) in terms of the similarity in representation. There are other variants of the minimum or infimum functional in previous literature, for example, Cerreia-Vioglio (2009) and Maccheroni (2002), used in different contexts.

There is also a sizable literature on the behavioral characterizations of temporal preferences, that the current project adds to. Olea and Strzalecki (2014), Hayashi (2003) and Pan et al. (2015) characterize the behavioral conditions necessary and sufficient for β-δ discounting, Loewenstein and Prelec (1992) characterize Hyperbolic discounting, and, Koopmans (1972), Fishburn and Rubinstein (1982) do the same for exponential discounting. Gul and Pesendorfer (2001) study a two-period model where individuals have preferences over sets of alternatives that represent second-period choices. Their axioms provide a representation that identifies the decision maker's commitment ranking, temptation ranking and cost of self-control.

8. Properties of the representation

We propose an alternative notion of "present premium" comparison below. The present premium can be considered as the maximal amount of future consumption one is willing to forego to have the residual moved to the present. For example, if $(y,t) \sim (x,0)$, then the present premium of (y,t) is $(y-x) \geq 0$.

Consider the following partial relation defined on the set of binary relations $\succsim$ over $\mathbb{X} \times \mathbb{T}$.

Definition 1. $\succsim_1$ allows a higher premium to the present than $\succsim_2$ if for all $x, y \in \mathbb{X}$ and $t \in \mathbb{T}$

$$(x,t) \succsim_1 (y,0) \quad \Longrightarrow \quad (x,t) \succsim_2 (y,0)$$

The next result connects this notion of comparative present premia to our representation.

Theorem 5. *Let $\succsim_1$ and $\succsim_2$ be two binary relations which allow for minimum representation w.r.t sets $\mathcal{U}_{\delta,1}$ and $\mathcal{U}_{\delta,2}$ respectively. The following two statements are equivalent:*

i) $\succsim_1$ allows a higher premium to the present than $\succsim_2$.

ii) Both $\mathcal{U}_{\delta,1}$ and $\mathcal{U}_{\delta,1} \cup \mathcal{U}_{\delta,2}$ provide minimum representations of $\succsim_1$.

Proof. See Appendix II. □

One might wonder if there could also be a representation theorem similar to Theorem 1 for an appropriately defined Weak Future Bias axiom. Below we define Weak Future Bias, and provide a corresponding representation.

A4*: **WEAK FUTURE BIAS:** If $(x,0) \succsim (y,t)$ then, $(x,t_1) \succsim (y, t+t_1)$ for all $x, y \in X$ and $t, t_1 \in \mathbb{T}$.

This is an alternative relaxation of Stationarity that is complementary to WPB. Weak Present Bias, when combined with Weak Future Bias yields the Stationarity Axiom. We now present the following result.

Theorem 6. *Let $\mathbb{T} = [0, \infty)$ and $\mathbb{X} = [0, M]$. The following two statements are equivalent:*

i) The relation $\succsim$ satisfies properties A0-A3 and A4.*

ii) There exists a set $\mathcal{U}_\delta$ of monotonically increasing continuous functions such that

$$F(x, t) = \max_{u \in \mathcal{U}} u^{-1}(\delta^t u(x))$$

represents the binary relation $\succsim$. The set $\mathcal{U}_\delta$ has the following properties: $u(0) = 0$ and $u(M) = 1$ for all $u \in \mathcal{U}_\delta$. $F(x, t)$ is continuous.

As expected Weak Future Bias is characterized by a weakly optimistic attitude towards the future. The proof is similar to that of Theorem 1, and is hence omitted.

Testable Implications. The major testable condition in the paper comes from the Weak Present Bias axiom: If $(y, t) \succsim (x, 0)$ then, $(y, t + t_1) \succsim (x, t_1)$ for all $x, y \in X$ and $t, t_1 \in \mathbb{T}$. Stated in terms of the contra-positive, If $(x, t_1) \succ (y, t+t_1)$ for some $x, y \in X$ and $t, t_1 \in \mathbb{T}$, $t, t_1 > 0$, then, $(x, 0) \succ (y, t)$. Intuitively speaking, this model only allows preference reversals that arise from present bias (as restricted by the Weak Present Bias axiom). So any temporal preference that stems from any other behavioral phenomenon would refute the model.

9. Stake Dependent Present Bias

Consider a decision maker who makes the following 2 pairs of choices.

Example 5.

\$100 today	$\succ$	\$110 in a week
\$110 in 5 weeks	$\succ$	\$100 in 4 weeks
\$11 in a week	$\sim$	\$10 today
\$11 in 5 weeks	$\sim$	\$10 in 4 weeks

Both pairs of choices are consistent with Weak Present Bias, but there is a classical choice reversal (or a local violation of Stationarity) only in the first pair[16]. This kind of choice is at odds with all the models of present bias that

[16]This kind of behavior closely parallels the "magnitude effect": in studies that vary the outcome sizes, subjects appear to exhibit greater patience toward larger rewards. For example, Thaler (1981) finds that respondents were on average indifferent between \$15 now and \$60 in a year,

we have mentioned other than the one in this paper, but not necessarily at odds with economic intuition. For example, if a DM's present bias is driven by the psychological fear of future uncertainty, the higher the stake, the higher would be the manifestation of this fear, and the more present-biased he would appear. The opposite phenomenon, when a subject appears strictly present-biased for smaller stakes but appears stationary at larger stakes (for the same set of temporal values) can happen, if the subjects get better at temporal decisions at higher stakes due to cognitive optimization. None of the models in Appendix I can account for the behavior in Example 5,[17] whereas, the simple minimum function mentioned in Example 3 can account for such choices. There is scope to run future experiments to test for such stake dependent behavior. The closest precedent for such an experimental design appears in Halevy (2015) where the author finds evidence of stake dependent present bias.

10. Application to a Timing Game

In this section we are going to study dynamic decision-making games for a present-biased DM whose preferences are consistent with the time-risk relations outlined in Keren and Roelofsma (1995). Present-biased preferences, when extended to a dynamic context[18], create time inconsistent preferences, which in turn results in inefficient decision making and loss in long-term welfare. The goal of this section is to convince the reader about the importance of axiomatization of risk-time relations, by showing that risk-time relations have important welfare implications for such a present-biased individual.

Consider the following game adopted from O'Donoghue and Rabin (1999). Suppose a DM gets a coupon to watch a free movie, over the next four Saturdays. He has to redeem the coupon an hour before the movie starts. His free ticket is issued subject to availability of tickets, and if there are no available tickets, the coupon is wasted. Hence there is some risk while redeeming the coupon. The movies on offer are of increasing quality- the theater is showing a mediocre movie this week, a good movie next week, a great movie in two weeks and Forrest Gump

\$250 now and \$350 in a year, and \$3000 now and \$4000 in a year, suggesting a (yearly) discount factor of 0.25, 0.71 and 0.75 respectively.

[17]For example, if one tries to fit a β-δ model to this data, the second pair of choices immediately suggest $\beta = 1$, which in turn is inconsistent with the first pair of choices.

[18]We are imposing Time Invariance of preferences following Halevy (2015). We will make precise assumptions about sophitication/ naivete as we go.

in three weeks. Our DM perceives the quality of these movies as 30, 40, 60 and 90 on a scale of $0-100$. In our problem, the DM can make a decision maximum 4 times, at $\tau = 1, 2, 3, 4$ (measured in weeks). The DM's utility at calendar time τ from watching a movie of quality x with probability p at calendar time $t+\tau$(in weeks) is given by:

$$U^{\tau}(x, p, \tau + t) = \begin{cases} p^{100}\alpha^{t}x & \text{for } p^{100}\alpha^{t} \geq \alpha^{\frac{1}{2}} \\ \left(\dfrac{\alpha}{\beta}\right)^{\frac{1}{2}} p\beta^{t}x & \text{for } p^{100}\alpha^{t} < \alpha^{\frac{1}{2}} \end{cases}$$

Where, $\beta = .99$, $\alpha = (.99)^{100} \approx .36$. This utility function (which is inspired by Pan et al. (2015)'s Two Stage Exponential discounting model) has the following interpretation: The DM has a long run weekly discount factor of .99 that sets in after a delay of half a week for $p = 1$. Before reaching the cut-off, the DM is extremely impatient, with a smaller discount factor of $\alpha = \beta^{100} \approx .36$, and hence is biased towards the present and very short-run outcomes. Similarly, the DM also proportionally undervalues probabilities close to 1. *The utility function(s) U^{τ} define a preference that satisfies all the axioms in Section 3, and hence have a minimum representation.* The DM is time-inconsistent, as his preferences between future options differ between any two decision periods τ_1 and τ_2 for $\tau_1, \tau_2 \in \{1, 2, 3, 4\}$. Let us assume that the DM is aware of his future preferences, that is she is sophisticated, a notion pioneered by Pollak (1968). We are going to use the following notion of equilibrium for this game.

Definition 2. (O'Donoghue and Rabin (1999)) A Perception Perfect Strategy for *sophisticates* is a strategy $s^s = (s^s_1, s^s_2, s^s_3, s^s_4)$, such that such that for all $t < 4$, $s^s_t = Y$ if and only if $U^t(t) \geq U^t(\tau')$ where $\tau' = \min_{\tau > t}\{s^s_{\tau} = Y\}$.

In any period, sophisticates correctly calculate when their future selves would redeem the coupon if they wait now. They then decide on redeeming the coupon if and only if doing it now is preferred to letting their future selves do it. We consider the following two cases:

Case 1: Suppose, there is not much demand for movie tickets in that city, and the DM knows that he can always book a ticket through his coupon and $p = 1$ for all alternatives under consideration.

In this case, the unique Perception Perfect Strategy is $s^s = (Y, Y, Y, Y)$. The knowledge that the future selves are going to be present biased creates an unwinding effect: The period 2 sophisticate would choose to use the coupon towards the good movie as he knows that the period 3 sophisticate would end up using the coupon for the great movie instead of going for Forrest Gump due to present bias. The period 1 sophisticate in turns correctly understands that waiting now would only result in watching the good movie and hence decides to see the mediocre movie right now instead.

Case 2: Suppose, due to persistent demand for movie tickets in that city, and the DM knows that redeeming a coupon results in a movie ticket in only 99% of cases.

The unique Perception Perfect Strategy is $s^s = (N, N, N, Y)$. The unwinding from the previous case does not happen here due to the risk involved in redeeming the coupon. Once the present is risky (equivalent to having a front end delay due to Probability Time Tradeoff), the bias previously assigned to the present vanishes, stopping the unraveling. The DM waits until the final period to cash in his coupon when the expected returns are the highest to the long run self.

		t				s^s_τ
		1	2	3	4	
τ	4				90	Y
	3			60	54.2	Y
	2		40	36.1	53.6	Y
	1	30	24	35.8	53	Y

		t				s^s_τ
		1	2	3	4	
τ	4				54.2	Y
	3			36.1	53.6	N
	2		24	35.8	53	N
	1	18	24	35.8	52.57	N

TABLE 2. The Left table is for Case 1 ($p = 1$), the right table is for Case 2 ($p = .9$). The entries in the table provide $U^\tau(x, p, t)$. The sophisticated DM compares the quantities in red row-wise for each τ when making a decision.

It would be instructive to compare the two cases in terms of welfare implications. Since present-biased preferences are often used to model self-control problems rooted in the pursuit of immediate gratification, we would compare welfare from the long run perspective. This outcome in Case 1 is consistent with the following general result in O'Donoghue and Rabin (1999): When benefits are immediate, the sophisticates "preprorate", i.e, they do it earlier than it might be optimal. For example, considering the long term self's interests, given a long term weekly discount factor of .99 for movie quality, the equilibrium outcome of watching the mediocre movie (quality of 30) in the first week, instead of Forrest Gump (quality of 90) definitely results in sub-optimal welfare in Case 1. For example, considering the choices from a $\tau = 0$ self gives $U^0(30, 1, 1) = 18$, and $U^0(90, 1, 4) = 53$. On the other hand, the introduction of a small amount of risk in Case 2, stops the unraveling in terms of "preprorating" (preponing consumption), thus helping the DM attain the most efficient outcome in equilibrium, thus reversing the O'Donoghue and Rabin (1999) result. In fact, not only is the highest level of available welfare achieved in Case 2 after the introduction of risk, *the equilibrium welfare improves from Case 1 to Case 2 in the absolute sense, even though apriori Case 2 seems to be worse than Case 1 for the DM*!

$$U^0(30, 1, 1) = 18 \quad < \quad U^0(90, .99, 4) = 52$$

This is an interesting application of how introducing a dominated menu of choices can result in absolute welfare improvement.

What would happen if the DM had the same preferences $U^\tau()$, but, instead was unaware that his preferences were dynamically inconsistent? Let us consider the extreme case (popularly called "naïveté" in the literature) where the DM thinks that his future selves' preferences would be identical to his current selves'. We will call such a DM naive, and use the following equilibrium notion to characterize their behavior.

Definition 3. A Perception Perfect Strategy for naifs is a strategy $s^n = (s^n_1, s^n_2, s^n_3, s^n_4)$, such that such that for all $t < 4$, $s^n_t = Y$ if and only if $U^t(t) \geq U^t(\tau)$ for all $\tau > t$.

The naive DM, acting under his false belief of time consistency, redeems the coupon in the current period if and only if it yields him the highest payoff among the remaining periods. Table 2 tells us that in Case 1, $s^n = (N, N, Y, Y)$, and in Case 2, $s^n = (N, N, N, Y)$. Thus the introduction of risk in this example also helps a naive DM make the most efficient choice in equilibrium.

11. Choice over Timed Bads

Most of the discussion on Present Bias till now has been centered around timed prizes or consumption, in general objects which are desirable. The central result of this paper is that Present Bias (as defined in A4 in Section 1) over such outcomes, can be represented by a minimum representation. This section would provide us the answers to the following two natural follow-up questions: 1) What would Present Bias look like when timed undesireable-goods or bads (for example, effort) are concerned? 2) What would be a utility representation of such preferences?

We would consider the richer domain that includes risk, without loss of generality. The DM has preferences over triplets (x, p, t), which describe the prospect of receiving an undesirable good $x \in \mathbb{X}$ at time $t \in \mathbb{T}$ with a probability$p \in [0, 1]$. We impose the following conditions on behavior.

C0: $\succsim$ is complete and transitive.

C1: CONTINUITY: $\succsim$ is continuous, that is the strict upper and lower contour sets of each timed alternative are open w.r.t the product topology.

The first two axioms are identical to axioms B0 and B1 used in Section 3.

C2: DISCOUNTING: For $t, s \in \mathbb{T}$, if $s > t$ then $(x, p, s) \succ (x, p, t)$ for $x, p > 0$ and $(x, p, s) \sim (x, p, t)$ for $x = 0$ or $p = 0$. For $x > y > 0$, there exists $T \in \mathbb{T}$ such that, $(x, q, 0) \succsim (y, 1, T)$.

C3: PRIZE AND RISK MONOTONICITY: For all $t \in \mathbb{T}$, $(x, p, t) \succsim (y, q, t)$ if $y \geq x$ and $q \geq p$. The first binary relation is strict if at least one of the 2 following relations are strict and if $y, q > 0$.

Discounting and Monotonicity have been adapted in the most intuitive way. People want to delay bad outcomes and they prefer when bad outcomes are less likely. Also when bad outcomes are concerned, more is worse.

C4: WEAK PRESENT BIAS: If $(x, 1, 0) \succsim (y, 1, t)$ then, $(x, 1, t_1) \succsim (y, 1, t + t_1)$ for all $x, y \in X$ and $t, t_1 \in \mathbb{T}$.

The Weak Present Certainty Bias requires that given the present and certainty are special, a DM would try to avoid bad outcomes which are in the present and are certain. Moreover, loss of certainty or immediacy can only make bad outcomes better.

C5: PROBABILITY-TIME TRADEOFF: For all $x, y \in \mathbb{X}$, $p \in (0, 1]$, and $t, s \in \mathbb{T}$, $(x, p\theta, t) \succsim (x, p, t + D) \implies (y, q\theta, s) \succsim (y, q, s + D)$.

The Probability-Time tradeoff axiom is unchanged and has the same interpretation as before.

Theorem 7. *The following two statements are equivalent:*

i) The relation $\succsim$ *on* $\mathbb{X} \times \mathbb{P} \times \mathbb{T}$ *satisfies properties C0-C5.*

ii) There exists a unique $\delta \in (0,1)$ *and a set* $\mathcal{U}$ *of monotinically decreasing continuous functions such that*

$$F(x,p,t) = \max_{u \in \mathcal{U}} -u^{-1}(p\delta^t u(x)) = -\min_{u \in \mathcal{U}} u^{-1}(p\delta^t u(x))$$

represents the relation $\succsim$*. For all the functions* $u \in \mathcal{U}$*,* $u(M) = -1$ *and* $u(0) = 0$*. Moreover, F(x,p,t) is continous.*

Conclusion

This paper provides an intuitive behavioral definition of (Weak) Present Bias and characterizes a general class of utility functions consistent with such behavior. Our utility representation can be interpreted as if a DM is unsure about future tastes and present bias arises as an outcome of his cautious behavior in the face of uncertainty about future tastes. Given most of the previous models of present bias have extraneous behavioral assumptions over and above present bias which are often empirically unsupported, we believe that our representation theorem is an important theoretical development in this literature. Having a more general representation for present bias, also helps us accommodate empirical phenomenon (for example, stake dependent present biased behavior) that previous models could not account for. We have extended the model to incorporate time-risk relations in behavior and provided an example where this relation can be utilized for welfare improving policy design. Given the axiomatic nature of our work, we provide simple testable conditions necessary and sufficient for our utility representations. These conditions can be easily taken to the laboratory or field to be empirically tested. We hope that this paper generates further interest in theoretical and applied work directed towards forming a better understanding of intertemporal preferences.

APPENDIX I

MODELS OF PRESENT BIAS

Consider the general separable discounted utility mode defined over timed prospects (x, t)

$$V(x,t) = \Delta(t)u(x)$$

Here, $\Delta(t)$ is the discount factor, and $u(x)$ is the felicity function. Below[19], we give a brief summary of the literature on different discounting models which accommodate present bias, in terms of the discount functions they propose. We also include the exponential discounting model as a point of reference.

	Model	Author(s)	$\Delta(t)$
0	Exponential discounting	Samuelson (1937)	$(1+g)^{-t}, g > 0$
1	Quasi-hyperbolic discounting	Phelps and Pollak (1968)	$(\beta + (1-\beta)_{t=0})(1+g)^{-t}, \beta < 1, g > 0$
2	Proportional discounting	Herrnstein (1981)	$(1+gt)^{-1}, g > 0$
3	Power discounting	Harvey (1986)	$(1+t)^{-\alpha}, \alpha > 0$
4	Hyperbolic discounting	Loewenstein and Prelec (1992)	$(1+gt)^{-\alpha/\gamma}, \alpha > 0, g > 0$
5	Constant sensitivity	Ebert and Prelec (2007)	$\exp[-(at)^b], a > 0, 1 > b > 0$

TABLE 3. Models of temporal behavior

APPENDIX II

Theorem 1: *Let* $\mathbb{T} = \{0, 1, 2, ...\infty\}$ *or* $\mathbb{T} =[0, \infty)$ *and* $\mathbb{X} = [0, M]$ *for* $M > 0$. *The following two statements are equivalent:*

i) The relation $\succsim$ *defined on* $\mathbb{X} \times \mathbb{T}$ *satisfies properties A0-A4.*

ii) For any $\delta \in (0, 1)$ *there exists a set* $\mathcal{U}_\delta$ *of monotinically increasing continuous functions such that*

$$F(x,t) = \min_{u \in \mathcal{U}_\delta}(u^{-1}(\delta^t u(x)))$$

represents the binary relation $\succsim$. *Moreover,* $u(0) = 0$ *and* $u(M) = 1$ *for all* $u \in \mathcal{U}_\delta$. $F(x, t)$ *is continuous.*

Proof: We start by showing (ii) implies (i). To show Weak Present Bias, we follow the following steps

[19]We take the idea of tabular presentation from Abdellaoui et al (2010).

$$
\begin{array}{rcl}
& & (y,t) \succsim (x,0) \\
\implies & & \min_{u\in\mathcal{U}_\delta}(u^{-1}(\delta^t u(y))) \geq \min_{u\in\mathcal{U}}(u^{-1}(u(x))) \\
\implies & & \min_{u\in\mathcal{U}_\delta}(u^{-1}(\delta^t u(y))) \geq x \\
\implies & & u^{-1}(\delta^t u(y)) \geq x \qquad \forall u \in \mathcal{U}_\delta
\end{array}
$$

$$
\begin{array}{lcl}
\implies & \delta^t u(y) \geq u(x) & \forall u \in \mathcal{U}_\delta \\
\implies & \delta^{t+t_1} u(y) \geq \delta^{t_1} u(x) & \forall u \in \mathcal{U}_\delta \\
\implies & u^{-1}(\delta^{t+t_1} u(y)) \geq u^{-1}(\delta^{t_1} u(x)) & \forall u \in \mathcal{U}_\delta \\
\implies & \min_{u\in\mathcal{U}_\delta}(u^{-1}(\delta^{t+t_1} u(y))) \geq \min_{u\in\mathcal{U}_\delta}(u^{-1}(\delta^{t_1} u(x))) & \\
\implies & (y, t+t_1) \succsim (x, t_1) &
\end{array}
$$

To show Monotonicity and Discounting, let us show $(x,t) \succ (y,s)$, when, either $x > y$ and $t = s$, or, $x = y$ and $t < s$. As all the functions $u \in \mathcal{U}_\delta$ are strictly increasing, and $\delta \in (0,1)$,

$$
\begin{array}{rcl}
\delta^t u(x) & > & \delta^s u(y) \quad \forall u \in \mathcal{U}_\delta \\
\iff u^{-1}(\delta^t u(x)) & > & u^{-1}(\delta^s u(y)) \quad \forall u \in \mathcal{U}_\delta \\
\iff \min\limits_{u\in\mathcal{U}_\delta} u^{-1}(\delta^t u(x)) & > & \min\limits_{u\in\mathcal{U}_\delta} u^{-1}(\delta^s u(y)) \\
\iff (x,t) & \succ & (y,s)
\end{array}
$$

For proving the second statement under Discounting, start with any $u_1 \in \mathcal{U}_\delta$. For $z > x > 0$, and $\delta \in (0,1)$ there must exist t big enough such that

$$
\begin{array}{rcl}
u_1(x) & > & \delta^t u_1(z) \\
\iff u_1^{-1}(u_1(x)) & > & u_1^{-1}(\delta^t u_1(z)) \\
\iff x & > & \min\limits_{u\in\mathcal{U}_\delta} u^{-1}((\delta^t u(z))
\end{array}
$$

Hence, there exists t big enough such that $(x,0) \succ (z,t)$.
That $\succsim$ satisfies continuity follows directly from the definition of continuity on the utility function.

Now, we will prove the other direction of the representation theorem. We will first deal with the case of $\mathbb{T}$ =$[0,\infty)$. A similar proof technique would be used in the proof of Theorem 3.

Proof for the case when $\mathbb{T}$ =$[0,\infty)$.

Proof. For every $x^* \in (0,M)$, we are going to provide an increasing utility function u_{x^*} on $[0,M]$ which would have $\delta^\tau u_{x^*}(x) \geq u_{x^*}(y)$ if $(x,\tau) \succsim (y,0)$. Additionally it would also have $\delta^t u_{x^*}(x_t) = u_{x^*}(x^*)$ for all $(x^*,0) \sim (x_t,t)$.
Fix $u_{x^*}(x^*) = 1$, $u_{x^*}(0) = 0$.
For any $x \in (x^*,M]$, by Discounting there exists a delay T large enough, such that $(x^*,0) \succ (x,T)$. Hence, it must be true that $(x,0) \succ (x^*,0) \succ (x,T)$. By Continuity there must exist $t(x) \in \mathbb{T}$ such that, $(x,t(x)) \sim (x^*,0)$. Define the utility at x as

$$u_{x^*}(x) = \delta^{-t(x)} \tag{4}$$

It would be helpful to introduce the following additional notation to move seamlessly between prizes and time in terms of indifference of time-prize pairs w.r.t $(x^*,0)$. For $t > 0$, define x_t as the value in $(x^*,M]$ such that $(x_t,t) \sim (x^*,0)$. Using continuity, we can say that all points in the interval $(x^*,M]$ can be written as x_t for some $t > 0$. This notation essentially implements the inverse of the $t(x)$ function defined in the previous paragraph.
Now, for $x \in (0,x^*)$, define

$$u_{x^*}(x) = \inf\{\delta^\tau : \text{ There exists } t \text{ such that } (x_t, t+\tau) \sim (x,0)\} \tag{5}$$

Firstly, we will show that the infimum in (5) can be replaced by minimum. Let the infimum be obtained at a value $I = \delta^{\tau^*}$. Consider a sequence of delays $\{\tau_n\}$ that converge above to τ^*, and $(x_{t_n}, t_n + \tau_n) \sim (x,0)$. Note that $t_n \in [0,t_{max}]$ where $(x^*,0) \sim (M,t_{max})$. Hence, $\{t_n\}$ must lie in this compact interval, and must have a convergent subsequence $\{t_{n_k}\}$. If t^* is the corresponding limit of $\{t_{n_k}\}$, we know that $t^* \in [0,t_{max}]$. Similarly, x_t can be considered a continuous function in t (this also follows from the continuity of $\succsim$). Therefore, $x_{t_{n_k}} \to x_{t^*}$ when $t_{n_k} \to t^*$. Thus, we have $(x_{t_{n_k}}, t_{n_k} + \tau_{n_k}) \sim (x,1)$ for all elements of $\{n_k\}$. As, $n_k \to \infty$, $x_{t_{n_k}} \to x_{t^*}$, $t_{n_k} + \tau_{n_k} \to t^* + \tau^*$. Then, using the continuity of $\succsim$, $(x_{t^*}, t^* + \tau^*) \sim (x,1)$. Hence, the infimum can be replaced by a minimum.

Now we will show that the utility defined in (4) and (5) has the following properties : 1) It is increasing. 2) $\delta^t u_{x^*}(x_t) = u_{x^*}(x^*)$ for all $(x^*, 0) \sim (x_t, t)$. 3) $(x, \tau) \succsim (y, 0)$ implies $\delta^\tau u_{x^*}(x) \geq u_{x^*}(y)$, 4) u is continuous. The first two properties are true by definition of u. We will show the third and fourth in some detail.

Consider $(x, \tau) \succsim (y, 0)$. In the case of interest, $\tau > 0$ and hence, $x > y$.

Now let $x > y > x^*$. Let, $u(y) = \delta^{-t_1}$, which means, $(y, t_1) \sim (x^*, 0)$. Given $(x, \tau) \succsim (y, 0)$, we must have

$$(x, \tau + t_1) \succsim (y, t_1) \sim (x^*, 0)$$

Hence, if $(x, t_2) \sim (x^*, 0)$, then,

$$\begin{aligned} t_2 &\geq \tau + t_1 \\ \iff u_{x^*}(x) = \delta^{-t_2} &\geq \delta^{-(\tau+t_1)} \\ \iff \delta^\tau u_{x^*}(x) &\geq \delta^{-t_1} = u_{x^*}(y) \end{aligned}$$

If, $x > x^* > y$, the proof follows from the way the utility has been defined.

Let $y < x < x^*$. Let, $u_{x^*}(x) = \delta^{t_1}$, which means, $(x_t, t + t_1) \sim (x, 0)$ for some $x_t \in [x^*, M]$. Given $(x, \tau) \succsim (y, 0)$, we must have

$$(x_t, t + t_1 + \tau) \succsim (x, \tau) \succsim (y, 0)$$

Hence, $u_{x^*}(y) \leq \delta^{\tau+t_1} = \delta^\tau u_{x^*}(x)$.

Now we turn to proving the continuity of u_{x^*}. The continuity at x^* from the right, or on $(x^*, M]$ is easy to see.

Next, for any $r = \delta^s \in (0, 1)$, define

$$f(r) = \sup\{y : \ (x_t, t + s) \sim (y, 0)\} = \hat{y} \tag{6}$$

The supremum can be replaced by a maximum, and the proof is similar to the one before. Suppose there is a sequence of $\{y_n\}$ that converges up to a value $\hat{y}$, and, $(x_{t_n}, t_n + s) \sim (y_n, 0)$. Note that t_n lies in a compact interval $[0, t_{max}]$, and hence has a convergent subsequence t_{n_k} that converges to a point in that interval $\hat{t} \in [0, t_{max}]$. Now, x_t is continuous in t (in the usual sense), and hence, x_{t_n} also converges to $x_{\hat{t}}$. Further, $y_{n_k} \to \hat{y}$ as $n_k \to \infty$. Therefore, using, $(x_{t_{n_k}}, t_{n_k} + s) \sim (y_{n_k}, 0)$, as, $n_k \to \infty$, it must be that $(x_{\hat{t}}, \hat{t} + s) \sim (\hat{y}, 0)$. Hence, the supremum in (6) must have been attained from $x_{\hat{t}}$, and hence the supremum can be replaced by a maximum. Further given this is a maximum, we can say

that $\hat{y} \in (0, x^*)$. The f function is well defined, strictly increasing and is the inverse function of u_{x^*} from $(0,1)$ to $(0, x^*)$, in the sense that, $u(f(r)) = r$. This function can be used to show the continuity of u at the point x^*.
Finally, the function u can be easily normalized to have $u_{x^*}(M) = 1$. (By dividing the function from before by $u_{x^*}(M)$.)
Now, consider $\mathcal{U}_\delta = \{u_{x^*}() : \ x^* \in (0, M]\}$. By construction of the functions, it must be that

$$\begin{aligned}
(x,t) \succsim (y,0) \quad &\iff \quad \delta^t u(x) \geq u(y) \ \forall u \in \mathcal{U}_\delta \\
(x,t) \sim (y,0) \quad &\iff \quad \delta^t u(x) \geq u(y) \ \forall u \in \mathcal{U}_\delta \\
& \qquad \text{and } \delta^t u_y(x) = u_y(y) \text{ for some } u_y \in \mathcal{U}_\delta
\end{aligned}$$

For any (z, τ), consider the sets $\{(y,0) \in \mathbb{X} \times \mathbb{T} : \ (y,0) \succsim (z,\tau)\}$ and $\{(y,0) \in \mathbb{X} \times \mathbb{T} : \ (z,\tau) \succsim (y,0)\}$. Both are non-empty, as $(z,0)$ belongs to the first one and $(0,0)$ in the second one. Both sets are closed in the product topology. Their union is connected, and hence there exists an element in their intersection, i.e, there exists a $y_1 \in \mathbb{X}$ such that $(y_1, 0) \sim (x,t)$. By monotonicity this y_1 must be unique. Therefore there must exist a continuous present equivalent utility representation for $\succsim$. We show this formally in the next two paragraphs.
Given $\succsim$ is complete, transitive and satisfies continuity, there exists a continuous function $\bar{F} : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ such that $\bar{F}(a) \geq \bar{F}(b)$ if and only if $a \succsim b$ for $a, b \in \mathbb{X} \times \mathbb{T}$. (Following Theorem 1, Fishburn and Rubinstein (1982)).
We define $G : \mathbb{X} \to \mathbb{R}$ as $G(x) = \bar{F}(x, 0)$. The function G would be strictly monotonic and continuous. Also define $F : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ as $F(x,t) = G^{-1}(\bar{F}(x,t))$. As any alternative has a unique present equivalent, F is well defined, is a monotonic continuous transformation of $\bar{F}$ (hence represents $\succsim$) and $F(x, 0) = x$ for all $x \in \mathbb{X}$. By definition the function F assigns to every alternative its present equivalent as the corresponding utility. Therefore, the present equivalent utility representation is continuous.
We will show that the function W defined below also assigns to every alternative (z, τ) an utility exactly equal to its present equivalent.

$$W(x,t) \quad = \quad \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u(x)) = F(x,t)$$

Consider any $(z,\tau) \sim (y_1, 0)$. By definition of $\mathcal{U}_\delta$ and by construction of its constituent functions, it must be that for all $u \in \mathcal{U}_\delta$, $\delta^\tau u(z) \geq u(y_1)$ and there exists a function u_{y_1} such that $\delta^\tau u_{y_1}(z) = u(y_1)$. This is equivalent to the following statement: For all $u \in \mathcal{U}_\delta$, $u^{-1}(\delta^\tau u(z)) \geq y_1$ and there exists a function u_{y_1} such that $u_{y_1}^{-1}(\delta^\tau u_{y_1}(z)) = y_1$.
Therefore, $W(z,\tau) = \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^\tau u(z))$ is continuous utility representation of the relation $\succsim$.

$\square$

Proof for the case of $\mathbb{T}$ =$\{0, 1, 2, ...\}$.
This proof would be more technical and we will break down the proof of this case into the following Lemmas.

Lemma 1. *Under Axioms A0-A4, for a fixed x_0, and any x_t and t such that $(x_t, t) \sim (x_0, 0)$, there exists a continuous strictly increasing function u such that $\delta^t u(x_t) = u(x_0)$ and $\delta^t u(z_1) \geq u(z_0)$ for all $(z_1, t) \succsim (z_0, 0)$. Further, $u(0) = 0$, $u(M) = 1$.*

Proof. By the Discounting axiom, we know that there exists a smallest integer $n \geq 1$ such that $(x_0, 0) \succsim (M, n)$. Choose $x_0^* = x_0$. For $0 < t < n$, find x_t^* such that $(x_0^*, 0) \sim (x_t^*, t)$. If $(x_0, 0) \succ (M, n)$, choose $x_n = M$.
We define x_{-1}^* in the following way

$$x_{-1}^* = \min\{x \in \mathbb{X} : (x, 0) \succsim (x_j^*, j+1),\ \ j = 0, 1, 2, ...n\}$$

The idea is to look at the present equivalents of $(x_j^*, j+1)$ and take the maximum of those present equivalents. The alternative way to express the same is to look at the intersection of the weak upper counter sets of $(x_j^*, j+1)$ on $\mathbb{X} \times \{0\}$, and then take the minimal value from that set.
Next we will use this to define x_{-2}^*, then use x_{-1}^* and x_{-2}^* to define x_{-3}^*. In general, for $i \in \{-1, -2, -3...\}$ define x_i^* recursively as the minimum of the set

$$\{x \in \mathbb{X} : (x, 0) \succsim (x_j^*, j-i), j = i+1, i+2, ...n\}$$

The definition uses the same idea as before. We consider the intersection of the weak upper counter sets of $(x_j^*, j-i)$ on $\mathbb{X} \times \{0\}$ and take its minimum. The set is non-empty (x_0^* belongs to it, for example), closed and the minimum exists due to the continuity, monotonicity and discounting properties.

Next we show that for every x_i^* with $i \leq -1$, there exists $j \in \{0, 1, ..n\}$ such that $(x_j^*, j - i) \succsim (x_i^*, 0)$. The proof is by induction. For $i = -1$, it is immediate from the definition. Suppose, it holds for all $i \geq -m$. Consider x_{-i-1}^*. By construction, there must exist $k \in \{-m, -m+1, ..n\}$ such that $(x_{-i-1}^*, 0) \sim (x_k^*, k+i+1)$. If $k \in \{0, 1, ..n\}$ we are done already. If not, by the induction hypothesis, there exists $j \in \{0, 1, ..n\}$ such that $(x_j^*, j-k) \succsim (x_k^*, 0)$, which gives, $(x_j^*, j+i+1) \succsim (x_k^*, k+i+1)$, and hence, $(x_j^*, j+i+1) \succsim (x_{-i-1}^*, 0)$, completing the proof.

Next we will show that the sequence $\{...x_{-2}^*, x_{-1}^*, x_0^*, x_1^*, x_2^*, ..\}$ is converges below to 0. Suppose not (we are going for a proof by contradiction), that is there exists $w > 0$ such that $x_i \geq w$ for all $i \in \mathbb{Z}$. As, $M > z > 0$, there must exist t_1 big enough such that $(z, 0) \succ (M, t_1)$. Consider the element $x_{-t_1}^*$ from the sequence in consideration. Using the result from the previous paragraph, it must be true that there exists $j \in \{0, 1, .., n\}$, such that $(x_j^*, j+t_1) \succsim (x_{-t_1}^*, 0)$. Now, as $M \geq x_j^*$, we must have, $(M, t_1) \succsim (x_{-t_1}^*, 0) \succ (z, 0)$, which provides a contradiction.

Consider any $y_0 \in (x_0^*, x_1^*)$.

We are going to find a $y_1, y_2, ..y_{n-1}$ recursively.

Finding y_1: For each point $y \in (x_1, x_2]$, take reflections of length 1, i.e, find x_y such that $(y, 1) \sim (x_y, 0)$. Note that, $(x_1^*, 0) \succ (y, 1) \succ (x_0^*, 0)$. Hence, $x_y \in (x_0^*, x_1^*)$. Let, x_{x_2} be the reflection for the point x_2. For any $y \in (x_1^*, x_2^*]$, $f(y) = x_0^* + (x_y - x_0^*)\dfrac{(x_1^* - x_0^*)}{(x_{x_2} - x_0^*)}$. Now, for $y_0 \in (x_0^*, x_1^*)$, define y_1 as $f^{-1}(y_0)$.

We can check that this method satisfies the 2 following conditions:

1) Consider two such sequences, one starting from y_0^1, and another from y_0^2, with $y_0^1 > y_0^2$. We will have $y_1^1 > y_1^2$.

2) All points in intervals (x_1^*, x_2^*) are included by some y_1 from the sequence. This follows from monotonicity and discounting too.

Now, the recursive step:

For each point $y \in (x_i^*, x_{i+1}^*]$, take reflections of length $j \in \{i, i-1.., 1\}$ conditional on those reflections being in the corresponding $(x_{i-j}^*, x_{i+1-j}^*]$ intervals. For any y, at least one of these reflections must exist, and in particular the one with length i always exists, as $(x_1^*, 0) \succ (x_{i+1}^*, i) \succsim (y, i)$ and $(y, i) \succ (x_i^*, i) \sim (x_0^*, 0)$.

Now, for each such reflection, find the corresponding sequence of $\{y_0, y_1, ..y_{i-1}\}$

it belongs to, and denote the smallest y_0 from that collection of sequences as $x_y \in [x_0^*, x_1^*]$. Note that $x_{x_{i+1}} \leq x_1^*$. Define the $1:1$ strictly increasing function f from $(x_{n-1}, x_n]$ to $(x_i^*, x_{i+1}^*]$ in the following way: For any $y \in (x_i^*, x_{i+1}^*]$, $f(y) = x_0^* + (x_y - x_0^*)\dfrac{(x_1^* - x_0^*)}{(x_{x_{i+1}} - x_0^*)}$. Now, define y_i as $f^{-1}(y_0)$. The conditions mentioned above are still satisfied for the extended sequence.

For $i \leq -1$, define y_i recursively in the following way. Start by finding y_i' as the minimum of the set

$$\{y \in \mathbb{X} : (y, 0) \succsim (y_j, j - i), j = i + 1, i + 2, ...n\}$$

Define x'_{-i} as the minimum of the set

$$\{y \in \mathbb{X} : (y, 0) \succsim (y_j, j - i), j = i + 1, i + 2, ...n - 1\}$$

Finally, define

$$y_i = x_{i+1}^* - (x_{i+1}^* - y_i')\frac{(x_{i+1}^* - x_i^*)}{(x_{i+1}^* - x_i')} \tag{7}$$

Given $y_0^1 > y_0^2$ determines the order of $y_t^1 > y_t^2$, for $t \in \{1, 2, ..n-2\}$, our inductive procedure make sure this holds true for all $t \leq -1$ too.

One can check for covering properties of the sequences by induction. Suppose all points in the intervals (x_i^*, x_{i+1}^*) are covered by y_i for some sequence, for $i \geq j$ for some integer j. We are going to show that all points in (x_{j-1}^*, x_j^*) are also covered by y_{j-1} for some sequence. Take any point $y \in (x_{j-1}^*, x_j^*)$, and consider its corresponding y' as defined in Equation 7. Consider the reflections from point y' of sizes $1, ..n - j + 1$, i.e, the points at those temporal distances which are indifferent to it, conditional on being in the corresponding intervals. By the induction hypothesis, each of those reflection end points must be coming from some $y_0 \in (x_0^*, x_1^*)$. Take the sequence with smallest y_0, and that sequence would result in having $y \in (x_{j-1}^*, x_j^*)$ as its next element.

Now, define u on $\mathbb{X}$ as follows: Set $u(x_n^*) = u(x_n^*) = 1$. For the sequence $..., x_{-2}^*, x_{-1}^*, x_0^*, x_1^*, ..$, let $u(x_i^*) = \delta^{i-n}$ for all positive and negative integers i. Next, let us define u on (x_{n-1}^*, x_n^*) as any continuous and increasing function with $\inf_{(x_{n-1}^*, x_n^*)} u(x) = \delta = u(x_{n-1}^*)$ and $\sup_{(x_{n-1}^*, x_n^*)} u(x) = 1 = u(M)$. We can extend each dual sequence with some as $u(y_i) = \delta^{i-n} u(y_0)$. This finishes the construction of a u that satisfies the conditions mentioned in the Lemma. □

Lemma 2. *Under Axioms A0-A4, there exists a continuous present equivalent utility function $F : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ that represents $\succsim$. Moreover, for $\delta \in (0, 1)$, $F(z, \tau) = \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^\tau u(z))$ for some set $\mathcal{U}_\delta$ of strictly monotonic, continuous functions, $u(0) = 0$ and $u(M) = 1$ for all $u \in \mathcal{U}_\delta$.*

Proof. Consider the set $\mathcal{U}_\delta$ of all strictly monotonic, continuous functions u such that $\delta^t u(z_1) \geq u(z_0)$ for all $(z_1, t) \succsim (z_0, 0)$, $u(0) = 0$ and $u(M) = 1$. By the previous Lemma, this set is non-empty, and for any $(z_1, t) \sim (z_0, 0)$ includes a function u, such that $\delta^t u(z_1) = u(z_0)$. By construction of the functions, it must be that

$$
\begin{aligned}
(x,t) \succsim (y,0) \quad &\iff \quad \delta^t u(x) \geq u(y) \ \forall u \in \mathcal{U}_\delta \\
(x,t) \sim (y,0) \quad &\iff \quad \delta^t u(x) \geq u(y) \ \forall u \in \mathcal{U}_\delta \\
&\qquad\quad \text{and } \delta^t u_y(x) = u_y(y) \text{ for some } u_y \in \mathcal{U}_\delta
\end{aligned}
$$

For any (z, τ), consider the sets $\{(y, 0) \in \mathbb{X} \times \mathbb{T} : (y, 0) \succsim (z, \tau)\}$ and $\{(y, 0) \in \mathbb{X} \times \mathbb{T} : (z, \tau) \succsim (y, 0)\}$. Both are non-empty, as $(z, 0)$ belongs to the first one and $(0, 0)$ in the second one. Both sets are closed in the product topology. Their union is connected, and hence there exists an element in their intersection, i.e, there exists a $y_1 \in \mathbb{X}$ such that $(y_1, 0) \sim (x, t)$. By monotonicity this y_1 must be unique. Therefore there must exist a continuous present equivalent utility representation for $\succsim$. We show this formally in the next two paragraphs.

Given $\succsim$ is complete, transitive and satisfies continuity, there exists a continuous function $\bar{F} : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ such that $\bar{F}(a) \geq \bar{F}(b)$ if and only if $a \succsim b$ for $a, b \in \mathbb{X} \times \mathbb{T}$. (Following Theorem 1, Fishburn and Rubinstein (1982)).

We define $G : \mathbb{X} \to \mathbb{R}$ as $G(x) = \bar{F}(x, 0)$. The function G would be strictly monotonic and continuous. Also define $F : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ as $F(x, t) = G^{-1}(\bar{F}(x, t))$. As any alternative has a unique present equivalent, F is well defined, is a monotonic continuous transformation of $\bar{F}$ (hence represents $\succsim$) and $F(x, 0) = x$ for all $x \in \mathbb{X}$. By definition the function F assigns to every alternative its present equivalent as the corresponding utility. Therefore, the present equivalent utility representation is continuous.

We will show that the function W defined below also assigns to every alternative

(z, τ) an utility exactly equal to its present equivalent.

$$W(x,t) = \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u(x)) = F(x,t)$$

Consider any $(z, \tau) \sim (y_1, 0)$. By definition of $\mathcal{U}_\delta$ and by construction of its constituent functions, it must be that for all $u \in \mathcal{U}_\delta$, $\delta^\tau u(z) \geq u(y_1)$ and there exists a function u_{y_1} such that $\delta^\tau u_{y_1}(z) = u(y_1)$. This is equivalent to the following statement: For all $u \in \mathcal{U}_\delta$, $u^{-1}(\delta^\tau u(z)) \geq y_1$ and there exists a function u_{y_1} such that $u_{y_1}^{-1}(\delta^\tau u_{y_1}(z)) = y_1$.
Therefore, $W(z, \tau) = \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^\tau u(z)) = F(z, \tau)$ is a continuous utility representation of the relation $\succsim$. □

Proposition 1: *Given the axioms A0-4, the representation form in (3) is unique in the discounting function $\Delta(t) = \delta^t$ inside the present equivalent function.*

Proof. We start with the case where $\Delta(t)$ is such that $\dfrac{\Delta(t+t_1)}{\Delta(t)} < \Delta(t_1)$ for some t, t_1. Consider any singleton $\mathcal{U} = \{u\}$.

$$\begin{aligned}
(y,t) &\sim (x,0) \\
\implies u^{-1}(\Delta(t)u(y)) &= x \\
\implies \Delta(t)u(y) &= u(x) \\
\implies \Delta(t+t_1)u(y) &= \frac{\Delta(t+t_1)}{\Delta(t)}u(x) < \Delta(t_1)u(x) \\
\implies u^{-1}(\Delta(t+t_1)u(y)) &< u^{-1}(\Delta(t_1)u(x)) \\
\implies (x,t_1) &\succ (y,t+t_1)
\end{aligned}$$

Hence, the relation implied by the representation contradicts Weak Present Bias.

Now assume the opposite, let there exists some $t, t_1 > 0$ such that $\dfrac{\Delta(t+t_1)}{\Delta(t)} > \Delta(t_1)$. Now suppose we started with a relation $\succsim$ which has $(y,t) \sim (x,0)$ as well as $(y, t+t_1) \sim (x, t_1)$ for all t, t_1 and some x, y. (This does not necessarily mean that the person's preferences satisfy stationarity in the broader sense as we do not ask this from all x, y.) We will show below that such preferences cannot be represented by the functional form we started with for any set of functions $\mathcal{U}$.

$$
\begin{aligned}
(y,t) &\sim (x,0) \\
\implies \min_{u\in\mathcal{U}}(u^{-1}(\Delta(t)u(y))) &\geq \min_{u\in\mathcal{U}}(u^{-1}(u(x))) = x \\
\implies \Delta(t)u(y) &\geq u(x) \;\; \forall u \in \mathcal{U} \\
\implies \Delta(t+t_1)u(y) &\geq \frac{\Delta(t+t_1)}{\Delta(t)}u(x) > \Delta(t_1)u(x) \;\; \forall u \in \mathcal{U} \\
\implies u^{-1}(\Delta(t+t_1)u(y)) &> u^{-1}(\Delta(t_1)u(x)) \;\; \forall u \in \mathcal{U} \\
\implies \min_{u\in\mathcal{U}}(u^{-1}(\Delta(t+t_1)u(y))) &> \min_{u\in\mathcal{U}}(u^{-1}(\Delta(t_1)u(x))) \\
\implies (y, t+t_1) &\succ (x, t_1)
\end{aligned}
$$

This completes our proof. □

Proposition 2: *If $\mathcal{U}$,$\mathcal{U}' \subset \mathcal{F}$ are such that $\bar{co}(\mathcal{U}) = \bar{co}(\mathcal{U}')$, and the functional form in (1) allows for a continuous minimum representation for both of those sets, then,* $\min_{u\in\mathcal{U}} u^{-1}(\delta^t u(x)) = \min_{u\in\mathcal{U}'} u^{-1}(\delta^t u(x))$.

Proof. We will prove this in 2 steps.
First we will show that for any set A, $\min_{u\in A} u^{-1}(\delta^t u(x)) = \min_{u\in \bar{A}} u^{-1}(\delta^t u(x))$, where $\bar{A}$ is the closure of the set A.
It is easy to see the direction that $\min_{u\in A} u^{-1}(\delta^t u(x)) \geq \min_{u\in \bar{A}} u^{-1}(\delta^t u(x))$.
We will prove the other direction by contradiction. Suppose, $\min_{u\in A} u^{-1}(\delta^t u(x)) > \min_{u\in \bar{A}} u^{-1}(\delta^t u(x))$. This would imply that there exists $v \in \bar{A}\backslash A$ and some $\epsilon > 0$, such that $v^{-1}(\delta^t v(x)) + \epsilon < u^{-1}(\delta^t u(x))$ for all $u \in A$. By definition of the topology of compact convergence and given that v belongs to the set of limit points of A, there must exist a sequence of functions $\{v_n\} \subset A$ which converges to v in the topology of compact convergence, i.e, for any compact set $K \subset \mathbb{R}_+$, $\lim_{n\to\infty} \sup_{x\in K} |v_n(x) - v(x)| = 0$. It can be shown that under this condition, $v_n^{-1}(\delta^t v_n(x))$ would also converge to $v^{-1}(\delta^t v(x))$ where $v_n \in \mathcal{U}$.[20] This constitutes a violation of $v^{-1}(\delta^t v(x)) + \epsilon < u^{-1}(\delta^t u(x))$ for all $u \in A$. Hence, it must be $\min_{u\in A} u^{-1}(\delta^t u(x)) = \min_{u\in \bar{A}} u^{-1}(\delta^t u(x))$.
As a second part of this proof, we will show that for any set A, $\min_{u\in A}(u^{-1}(\delta^t u(x))) = \min_{u\in co(A)}(u^{-1}(\delta^t u(x)))$.
It is easy to see that $\min_{u\in A}(u^{-1}(\delta^t u(x))) \geq \min_{u\in co(A)} u^{-1}(\delta^t u(x))$, as $A \subset$

[20] As, $v_n \to v$ in the topology of compact convergence, $v_n \to v$ point wise, hence, $\delta^t v_n(x) \to \delta^t v(x)$. Now, as $v_n^{-1} \to v^{-1}$ compact convergence (proof later in the appendix), $v_n^{-1}(\delta^t v_n(x)) \to v^{-1}(\delta^t v(x))$.

$co(A)$. We will again use proof by contradiction to show the opposite direction. We assume that there exists a $\bar{u} \in \bar{co}(A)$ and $(x,t) \in \mathbb{X} \times \mathbb{T}$, such that $\bar{u} = \sum_{i=1}^{n} \lambda_i u_i$, $\sum_{i=1}^{n} \lambda_i = 1$ and $\bar{u}^{-1}(\delta^s \bar{u}(y)) < \min_i u_i^{-1}(\delta^s u_i(y))$. This would imply that $u_i(\bar{u}^{-1}(\delta^s \bar{u}(y))) < \delta^s u_i(y)$ for all i.
Now,

$$\begin{aligned} \delta^s \bar{u}(y) &= \delta^s \sum_i \lambda_i u_i(y) \\ &= \sum_i \lambda_i \delta^s u_i(y) \\ &> \sum_i \lambda_i u_i(\bar{u}^{-1}(\delta^s \bar{u}(y))) \\ &= \bar{u}(\bar{u}^{-1}(\delta^s \bar{u}(y))) \\ &= \delta^s \bar{u}(y) \end{aligned}$$

This gives us a contradiction. Note that the equality right after the inequality comes from the definition of $\bar{u}$.
Hence, we have, $\min_{u \in A} u^{-1}(\delta^t u(x)) = \min_{u \in co(A)} u^{-1}(\delta^t u(x))$. □

Proposition 3: *i) If there exists a concave function $f \in \mathcal{U}$, and if $\mathcal{U}_1$ is the subset of convex functions in $\mathcal{U}$, then* $\min_{u \in \mathcal{U}} u^{-1}(\delta^t u(x)) = \min_{u \in \mathcal{U} \setminus \mathcal{U}_1} u^{-1}(\delta^t u(x))$.

ii) If $u_1, u_2 \in \mathcal{U}$ and u_1 is concave relative to u_2, then, $\min_{u \in \mathcal{U}} u^{-1}(\delta^t u(x)) = \min_{u \in \mathcal{U} \setminus \{u_2\}} u^{-1}(\delta^t u(x))$.

Proof. If a function u is convex,

$$\begin{aligned} u^{-1}(\delta^t u(x)) &= u^{-1}(\delta^t u(x) + (1-\delta^t) u(0)) \\ &\geq u^{-1}(u(\delta^t x + (1-\delta^t) 0)) \\ &= \delta^t x \end{aligned}$$

Similarly for concave f , we would have, $f^{-1}(\delta^t f(x)) \leq \delta^t x$ which completes the proof of part (i). Note that this result is expected given concave functions give rise to more conservative present equivalents.
For part (ii), note that

$$\begin{aligned} u_1^{-1}(\delta^t u_1(x)) &= u_1^{-1}(\delta^t u_1(u_2^{-1}(u_2(x)))) \\ &\leq u_1^{-1}(u_1(u_2^{-1}(\delta^t u_2(x)))) \\ &= u_2^{-1}(\delta^t u_2(x)) \end{aligned}$$

Where the inequality arises from the fact that u_1 is concave relative to u_2. $\square$

Proposition 4. *Eventual stationarity is satisfied by β-δ discounting, but not hyperbolic discounting.*

Now for any $x > z > 0 \in X$, choose $t_1 > \log_{\frac{1}{\delta}}\left(\frac{u(x)}{u(z)}\right)$.

$$\begin{aligned} t_1 &> \log_{\frac{1}{\delta}}\left(\frac{u(x)}{u(z)}\right) \\ \iff \left(\frac{1}{\delta}\right)^{t_1} &> \frac{u(x)}{u(z)} \\ \implies u(z) &> \delta^{t_1} u(x) > \beta\delta^{t_1} u(x) \\ \implies \beta\delta^t u(z) &> \beta\delta^{t+t_1} u(x) \\ (z,t) &\succ (x, t+t_1) \end{aligned}$$

Also, $(x,0) \sim (x_t, t)$ implies, $u(x) = \beta\delta^t u(x_t)$, which implies,

$$\begin{aligned} u(z) &> \delta^{t_1} u(x) = \beta\delta^{t+t_1} u(x_t) \\ (z,0) &\succ (x_t, t+t_1) \end{aligned}$$

This shows that $\beta - \delta$ does indeed satisfy A5.

Now consider the simple variant of Hyperbolic discounting model when $\alpha = \gamma = 1$. Fix any felicity function u and $x > z > 0 \in X$. We will show that there does not exist t_1,such that $(z,t) \succ (x, t+t_1)$ for all $t \geq 0$.

$$\begin{aligned} (z,t) &\succ (x, t+t_1) \text{ for all } t \geq 0 \\ \iff \frac{u(z)}{1+t} &> \frac{u(x)}{1+t+t_1} \text{ for all } t \geq 0 \\ \iff \frac{1+t+t_1}{1+t} &> \frac{u(x)}{u(z)} \text{ for all } t \geq 0 \\ \iff 1 + \frac{t_1}{1+t} &> \frac{u(x)}{u(z)} \text{ for all } t \geq 0 \end{aligned}$$

Note that the last statement is not possible, as for fixed t_1 the LHS$\downarrow 1$ as $t \uparrow \infty$, whereas, the RHS is always a fixed number, that is strictly greater than one. Hence, hyperbolic discounting does not satisfy A5.

Theorem 2: *The following two statements are equivalent:*

i) The relation $\succsim$ *satisfies properties A0-A6.*

ii) There exists a set $\mathcal{U}_\delta$ *of monotinically increasing continuous functions such that*

$$F(x,t) = \min_{u \in \mathcal{U}} u^{-1}(\delta^t u(x))$$

represents the binary relation $\succsim$. *The set* $\mathcal{U}$ *has the following properties:* $u(0) = 0$ *for all* $u \in \mathcal{U}$, $\sup_u u(x)$ *is bounded above,* $\inf_u u(x) > 0 \ \forall x$, $\inf_u \dfrac{u(z)}{u(x)}$ *is unbounded in* z *for all* $x > 0$.

Proof : Going from (ii) to (i) :

That (ii) implies Monotonicity, Discounting, Weak Present Bias and Continuity has already been shown in the proof of Theorem 1.

Showing Eventual Stationarity: Given $\sup_u u(x)$ is bounded above and $\inf_u u(x) > 0$, for any choice of $x, z > 0$ and $\delta \in (0,1)$ there exists $t_1 > 0$ big enough such that $\inf_u u(z) > \delta^{t_1} \inf_u u(x)$. This would imply that, *for all* $u \in \mathcal{U}$,

$$u(z) \quad > \quad \delta^{t_1} u(x)$$

and, hence, $(z, 0) \succ (x, t_1)$.

Now, for $t > 0$ consider x_t such that $(x_t, t) \sim (x, 0)$. By the representation, this implies that there exists $u_1 \in \mathcal{U}$ such that

$$\begin{aligned} \delta^t u_1(x_t) &= u_1(x) \\ \implies \delta^{t+t_1} u_1(x_t) &= \delta^{t_1} u_1(x) < u_1(z) \\ \implies \min_u u^{-1}(\delta^{t+t_1} u_1(x_t)) &\leq u_1^{-1}(\delta^{t+t_1} u_1(x_t)) < u_1^{-1}(u_1(z)) = \min_u u^{-1}(u(z)) \end{aligned}$$

Hence, $(z, 0) \succ (x_t, t + t_1)$.

Similarly, *for all* $u \in \mathcal{U}$,

$$\begin{aligned} \delta^t u(z) &> \delta^{t+t_1} u(x) \\ \implies \min_u u^{-1}(\delta^t u(z)) &> \min_u u^{-1}(\delta^{t+t_1} u(x)) \end{aligned}$$

Hence, $(z, t) \succ (x, t + t_1)$.
Showing Non-triviality: We have that $\inf_u \dfrac{u(z)}{u(x)}$ is unbounded in z for all $x > 0$. Therefore, for any x, and $t \in \mathbb{T}$, there exists z, such that

$$\begin{aligned}
\inf_u \frac{u(z)}{u(x)} &> \delta^{-t} \\
\implies \frac{u(z)}{u(x)} &> \delta^{-t} \quad \forall u \in \mathcal{U} \\
\implies \delta^t u(z) &> u(x) \quad \forall u \in \mathcal{U} \\
\implies u^{-1}(\delta^t u(z)) &> u^{-1}(u(x)) \quad \forall u \in \mathcal{U} \\
\implies \min_u u^{-1}(\delta^t u(z)) &> \min_u u^{-1}(u(x)) \\
(z, t) &\succ (x, 0)
\end{aligned}$$

To go from the direction (i) to (ii) of Theorem 2, one needs to follow Lemma 3-5.

Lemma 3. *Under Axioms A1-A6, for any $(x_0, t), (x_t, 0)$ such that $(x_0, t) \sim (x_t, 0)$ in the original relation, there exists $u \in \mathcal{U}$ such that $\delta^t u(x_t) = u(x_0)$ and $\delta^t u(z_1) \geq u(z_0)$ for all $(z_1, t) \geq (z_0, 0)$. Moreover, u is strictly monotonic, continuous, and $u(0) = 0$, $u(1) = 1$.*

Proof. We will prove it for $t = 1$, $x_0, x_t > 0$ and then show the general guideline for a general t.
We define the following procedure: Choose $x_0^* = 1$. Find x_1^* such that $(x_0^*, 0) \sim (x_1^*, 1)$. We can do it because of the Non-Triviality assumption. Clearly, $x_1^* = x_1$. Next find $x_{-1}^* = \max\{x_{-1}, x'_{-1}\}$ where $(x_0^*, 1) \sim_v (x_{-1}, 0)$ and $(x_1^*, 2) \sim_v (x'_{-1}, 0)$. The value $x_{-1} > 0$ exists because, $(x_0^*, 0) \succ (x_0^*, 1) \succ (0, 1)$, coupled with the fact that $\succsim$ is continuous. Same with x'_{-1}.
Note that $x_0^* > x_{-1}^*$ by discounting. Next going in the opposite direction, we find $x_2^* = \min\{x_2, x'_2, x''_2\}$, where, $(x_1^*, 0) \sim_v (x_2, 1)$, $(x_0^*, 0) \sim_v (x'_2, 2)$ and $(x_{-1}^*, 0) \sim_v (x''_2, 3)$. Next we find $x_{-2}^*, x_3^*, x_{-3}^*, x_4^*, \ldots$ sequentially. Thus one can find a sequence $\ldots x_{-3}^* < x_{-2}^* < x_{-1}^* < x_0^* < x_1^* < x_2^* \ldots$
We will show that this sequence is unbounded above and converges below to 0 . Consider any $z < x_0^*$. By A5, there must exist t_1 such that $(z, 0) \succ (x_0^*, t_1)$. and given for any $t > 0$, $(x_0, 0) \succsim (x_t^*, t)$, by monotonicity, it must hold that $(z, 0) \succ (x_t^*, t + t_1)$. By definition of x_{-1}^*, either $(x_{-1}^*, 0) \sim (x_0^*, 1)$ or $(x_{-1}^*, 0) \sim (x_1^*, 2)$, if not both. So, by WPB, either $(x_0^*, t_1) \succsim (x_{-1}^*, t_1 - 1)$ or $(x_1^*, t_1 + 1) \succsim (x_{-1}^*, t_1 - 1)$,

and hence, either $(z, 0) \succ (x^*_{-1}, t_1 - 1)$. One can use the construction of the sequence, and induction, here on, to show that, for any general $0 < i < t_1$, $(z, 0) \succ (x^*_{-i}, t_1 - i)$. Hence, it must be that $x^*_{-t_1} \leq z$, which proves that the sequence converges below to zero. To show that the sequence is unbounded above, one uses a similar trick. Consider $z > x^*_0$. There must exist t_2 such that $(x^*_0, t) \succ (z, t + t_2)$ for all $t \geq 0$, and given for any $t > 0$, $(x^*_{-t}, 0) \succsim (x^*_0, t)$, by monotonicity, it must hold that $(x^*_{-t}, 0) \succsim (x^*_0, t) \succ (z, t + t_2)$. By definition of x^*_1, $(x^*_1, 1) \sim (x^*_0, 0) \succ (z, t_2)$. So, by WPB, it must be that $(x^*_1, 0) \succ (z, t_2 - 1)$. ($z < x^*_1$ is trivial and hence neglected). One can use the construction of the sequence, and induction, here on, to show that, for any general $0 < i < t_2$, $(x^*_i, 0) \succ (z, t_2 - i)$. Hence, it must be that $x^*_{t_2} \geq z$, which proves that the sequence diverges to infinity.

Consider any $y_0 \in (x^*_0, x^*_1)$. We find y'_{-1} such that $(y'_{-1}, 0) \sim (y^*_0, 1)$. Finally,

$$y^*_{-1} = x^*_0 - (x^*_0 - y'_{-1}) \frac{(x^*_0 - x^*_{-1})}{(x^*_0 - x_{-1})} \in (x^*_{-1}, x^*_0).$$

The upper bound on y^*_{-1} comes from the fact that $(x^*_0 > y'_{-1})$ and the lower bound comes from the fact that y'_{-1} is bounded below by x_{-1}. Note that for $y^*_0, \hat{y}_0 \in (x^*_0, x^*_1)$, $y^*_0 > \hat{y}_0$ if and only if $y^*_{-1} > \hat{y}_{-1}$. And finally, for any $y^*_{-1} \in (x^*_{-1}, x^*_0)$ there exists a $y^*_0 \in (x^*_0, x^*_1)$ corresponding to it.

Next we will define an inductive procedure to find the other points in such sequences. Let $\mathcal{S}$ be the set of all such sequences. The induction hypothesis is that for every $y^*_0 \in (x^*_0, x^*_1)$ we have already defined a corresponding chain[21] $\mathcal{S}_i = y^*_{-i} < ...y^*_{-3} < y^*_{-2} < y^*_{-1} < y^*_0 < y^*_1 < y^*_2.. < y^*_{i-1}$, $i \geq 2$ such that i) $y^*_n \in (x^*_n, x^*_{n+1})$ for all the elements of all the chains. ii) If we compare the n^{th} elements of 2 chains they are always similarly ranked, regardless of the value of n. iii) If the last element constructed is y^*_i for $i \in \mathbb{N}$ then, any point in (x_n, x_{n+1}) for $n \in \{-i, ..i - 1\}$ is part of exactly one chain in $\mathcal{S}_i$.

Finding y^*_i where $i \geq 1$: Note that we can write $x^*_i = \min\{x^1_i, x^2_i, x^3_i ... x^{2i}_i\}$[22], where $(x^1_i, 1) \sim (x^*_{i-1}, 0)$,$(x^2_i, 2) \sim (x^*_{i-2}, 0)$..,$(x^{2i-1}_i, 2i-1) \sim (x^*_{-i+1}, 0)$. Similarly, $x^*_{i+1} = \min\{x^1_{i+1}, x^2_{i+1}, x^3_{i+1} ... x^{2i+1}_{i+1}\}$. Define, $x'_{i+1} = \min\{x^1_{i+1}, x^2_{i+1}, x^3_{i+1} ... x^{2i}_{i+1}\} \geq$

[21] A set paired with a total order.

[22] We are using one extra comparison than that existed in the original construction of the sequence, and this is to make sure that x^*_i has $2i$ comparisons in its construction, just like y^*_i. Given the structure of the sequence we can always add more comparisons than the original, but never have fewer comparisons.

x^*_{i+1}. Define $y'_i = \max\{y^1_i, y^2_i, y^3_i, ...y^{2i}_i\}$ where $(y^1_i, 1) \sim (y^*_{i-1}, 0)$,.., $(y^{2i}_i, 2i) \sim (y^*_{-i}, 0)$. Finally, $y^*_i = x^*_i + (y'_i - x^*_i)\dfrac{(x^*_{i+1} - x^*_i)}{(x'_{i+1} - x^*_i)} \in (x^*_i, x^*_{i+1})$. By monotonicity, $y^n_i \in (x^n_i, x^n_{i+1})$ for all $n \in \{1, 2, .., 2i\}$. Therefore, $y'_i \in (x^*_i, x'_{i+1})$. Therefore, $y^*_i \in (x^*_i, x^*_{i+1})$, the upper bound comes from the fact that $x'_{i+1} > y'_i$ and the lower bound comes from the fact that y'_i is bounded below by x^*_i. Note that for $y^*_0, \hat{y}^*_0 \in (x^*_0, x^*_1)$, $y^*_0 > \hat{y}^*_0$ if and only if $y^*_i > \hat{y}^*_i$. And finally, for any $\hat{y}^*_i \in (x^*_i, x^*_{i+1})$ there exists a $\hat{y}^*_0 \in (x^*_0, x^*_1)$ corresponding to it. The last part can be shown constructively.

Finding y^*_{-i-1} where $i \geq 1$: Note that $x^*_{-i} = \max\{x^1_{-i}, x^2_{-i}, x^3_{-i}...x^{2i+1}_{-i}\}$[23], where $(x^1_{-i}, 0) \sim_v (x^*_{-i+1}, 1)$,$(x^2_{-i}, 0) \sim_v (x^*_{-i+2}, 2)$.., $(x^{2i}_{-i}, 0) \sim_v (x^*_i, 2i)$. Similarly, $x^*_{-i-1} = \max\{x^1_{-i-1}, x^2_{-i-1}, ..\ x^{2i+1}_{-i-1}, x^{2i+2}_{-i-1}\}$.

Define, $x'_{-i-1} = \max\{x^1_{-i+1}, x^2_{-i+1}, x^3_{-i+1}, ...x^{2i+1}_{-i+1}\} \leq x^*_{-i-1}$.

Define $y'_{-i-1} = \max\{y^1_{-i+1}, y^2_{-i+1}, y^3_{-i+1}, ...y^{2i+1}_{-i+1}\}$ where $(y^1_{-i+1}, 0) \sim_v (y^*_{-i+2}, 1)$, ..., $(y_{-i+1}, 0) \sim_v (y^*_i, 2i+1)$. Finally, $y^*_{-i-1} = x^*_{-i} - (x^*_{-i} - y'_{-i-1})\dfrac{(x^*_{-i} - x^*_{-i-1})}{(x^*_{-i} - x'_{-i-1})} \in (x^*_{-i-1}, x^*_{-i})$. By monotonicity, $y^n_{-i-1} \in (x^n_{-i-1}, x^n_{-i})$ for all $n \in \{1, 2, .., 2i+1\}$. Therefore, $y'_{-i-1} \in (x'_{-i-1}, x^*_{-i})$. Therefore, $y^*_{-i-1} \in (x^*_{-i-1}, x^*_{-i})$, the upper bound comes from the fact that $x^*_{-i} > y'_{-i-1}$ and the lower bound comes from the fact that y'_{-i-1} is bounded below by x'_{-i-1}. Note that for $y^*_0, \hat{y}_0 \in (x^*_0, x^*_1)$, $y^*_0 > \hat{y}_0$ if and only if $y^*_{-i-1} > \hat{y}_{-i-1}$. And finally, for any $\hat{y}_{-i-1} \in (x^*_{-i-1}, x^*_{-i})$ there exists a $\hat{y}_0 \in (x^*_0, x^*_1)$ corresponding to it. The last part can be shown inductively. Fix $\hat{y}'_{-i-1}$. Find the points (whenever possible) $z_n \in (x^*_n, x^*_{n+1})$ for $n \in \{-i, -i+1, -i+2, ..i\}$ such that $(\hat{y}'_{-i-1}, 0) \sim_v (z_n, n+i+1)$. Note that we can always find atleast one such z_n.[24] Next, using the induction hypothesis we can map all the z_n's to a $y^*_0 \in (x^*_0, x^*_1)$. We take the maximum of all such y^*_0s and define it as $\hat{y}^*_0$. One can check that starting from this $(\hat{y}_{-i+1}, ..\hat{y}_0, \hat{y}_1..\hat{y}_i)$ would indeed result in ending with the $\hat{y}'_{-i-1}$ we started with. [25]

---

[23]As before, we are using one extra comparison than that existed in the original construction of the sequence.

[24]There exists k such that $(x^*_{-i}, 0) \sim_v (x^*_{-i+k}, k)$. In general, $(x^*_{-i}, 0) \succsim_v (x^*_{-i+k}, k)$.This implies $(\hat{y}'_{-i}, 0) \succ_v (x^*_{-i+k}, k)$ and $(x^*_{-i+1+k}, k) \succ_v (\hat{y}'_{-i}, 0)$. Hence, there exists $z_{-i+k} \in (x^*_{-i+k}, x^*_{-i+k+1})$ such that $(\hat{y}'_{-i}, 0) \sim_v (z_n, n+i)$.

[25]Suppose not. Given our definition of $\hat{y}_0$, starting from this $(\hat{y}_{-i+1}, ..\hat{y}_0, \hat{y}_1..\hat{y}_i)$ would give us $\hat{y}''_{-i} \geq \hat{y}'_{-i}$. Let, $\hat{y}''_{-i} > \hat{y}'$ and $(\hat{y}''_{-i}, 0) \sim (\hat{y}_{-i+k}, k)$ for $\hat{y}_{-i+k} \in (x_{-i+k}, x_{-i+k+1})$, this being the relation that binds while defining $\hat{y}''_{-i}$. Given, $(\hat{y}_{i+k}, k) \succ_v (\hat{y}'_{-i}, 0)$ and $(\hat{y}'_{-i}, 0) \succ (x'_{-i}, 0) \succsim_v (x^*_{-i+k}, k)$, there would exist $(\hat{y}''_{-i}, 0) \sim (\hat{y}^*_{-i+k}, k)$ for $\hat{y}_{-i+k} \in (x^*_{-i+k}, x^*_{-i+k+1})$ and $\hat{y}_{-i+k} < \hat{y}'_{-i+k}$ which would be a contradiction.

Now, define u on $\mathbb{X}$ as follows: Set $u_1(x_0^*) = 1$. For the sequence $.., x_{-1}^*, x_0^*, x_1^*, ..$, let $u(x_i^*) = \delta^i$ for all positive and negative integers i. Next, let us define $u_!$ on $(x_{-1}^*, 1)$ as any continuous and increasing function with $\inf_{(x_{-1}^*,1)} u_1(x) = \delta = u(x_{-1}^*)$ and $\sup_{(x_{-1}^*,1)} u_1(x) = 1 = u(1)$. We can extend each dual sequence with some $y_0 \in (x_{-1}^*, 1)$ as $u(y_i) = \delta^i u(y_0)$. Finally, define $U(x,t) = \delta^t \dfrac{u_1(x)}{u_1(1)}$ to ensure $u_1(1) = 1$ (note that $u_1(1) > 0$).

It is important to note here that the utility defined retains all the monotonicity, discounting and present bias properties. Consider any $(y,t) \succsim (x,0)$ in the original relation. The element x must belong to one of the sequences defined above. If x_t is the corresponding element to the right in that sequence separated by a distance of t, then, by construction we must have $u(x) = \delta^t u(x_t)$ and $(x,0) \succsim (x_t, t)$. By monotonicity, it would be true that $y > x_t$ and hence, $u(x) < \delta^t u(y)$.

Now we will extend the logic above to a more general case of $(x_0, t), (x_t, 0)$ such that $(x_0, t) \sim (x_t, 0)$ for $t > 1$.

For $i \in \{1, ..t\}$, let x_i be such that $(x_0, 0) \sim (x_i, i)$. We define the following procedure: Choose $x_0^* = x_0$, the same x_0 that was provided in the statement of this Lemma. Find x_1^* such that $(x_0^*, 0) \sim (x_1^*, 1)$. Of course, $x_1^* = x_1$. Next use the iterative format used in Lemma 2 to find $x_2^*, x_3^*, ..x_t^*$.

At each of these steps, by WPB, one would get, $x_i^* = x_i$, ending with $x_t^* = x_t$. We provide a brief outline for this, the proof requires induction.

Let, $x_2^* = \min\{x_2, x_2'\}$, where, $(x_2, 2) \sim_v (x_0^*, 0)$ and $(x_2', 1) \sim_v (x_1^*, 0)$. By WPB, the latter implies, $(x_2', 2) \succsim_v (x_1^*, 1)$. By definition of x_1^*, $(x_0^*, 0) \sim_v (x_1^*, 1)$. Putting it all together,

$$(x_2', 2) \succsim_v (x_1^*, 1) \sim {}_v (x_0^*, 0) \sim_v (x_2, 2)$$

Hence, $x_2' \geq x_2$, and $x_2^* = x_2$.

Similarly, let $x_3^* = \min\{x_3, x_3', x_3''\}$, where, $(x_3, 3) \sim_v (x_0^*, 0)$, $(x_3', 2) \sim_v (x_1^*, 0)$ and $(x_3'', 1) \sim_v (x_2^*, 0)$.

$$(x_3', 3) \succsim_v (x_1^*, 1) \sim {}_v (x_0^*, 0) \sim_v (x_3, 3)$$

Also,

$$(x_3'', 3) \succsim_v (x_2^*, 2) \sim {}_v (x_0^*, 0) \sim_v (x_3, 3)$$

And so on. Note that the sequence in which the elements are being found till now has been different that that in Lemma 2. Here on, find the sequence elements in the following order $x^*_{-1}, x^*_{t+1}, x^*_{-2}, x^*_{t+2}, ...$ using the iterative procedure as Lemma 2.

For any $y_0 \in (x^*_0, x^*_1)$, find similar sequences in the same order as we derived the sequence x^*.

Now, define u on $\mathbb{X}$ as before to finish the proof. Note that any u such constructed is strictly monotonic, continuous, and $u(0) = 0$, $u(1) = 1$. $\square$

Lemma 4. *Under Axioms A1-A6, there exists a set of functions $\mathcal{U}$ such that, for all $u \in \mathcal{U}$, u is strictly monotonic, continuous, and $u(0) = 0$, $u(1) = 1$, and $\delta^t u(z_1) \geq u(z_0)$ for all $(z_1, t) \geq (z_0, 0)$. Moreover, i) for any $(x, t) \sim (y, 0)$, there existsu $\in \mathcal{U}$ such that $\delta^t u(x_t) = u(x_0)$. ii) For $x > 0$, $\inf_{u \in \mathcal{U}} u(x) > 0$,* $\sup_{u \in \mathcal{U}} u(x) < \infty$

Proof. Consider the set $\mathcal{U}$ consisting of all functions u constructed from all the indifference relations $\sim$ in (3). It would suffice to show that $\inf_{u \in \mathcal{U}} u(x) > 0$, $\sup_{u \in \mathcal{U}} u(x) < \infty$.

First we will show that $\inf_{u \in \mathcal{U}} u(x) > 0$. This is trivial for points above $x = 1$. Consider $0 < x < 1$. Suppose we are constructing a function that would respect the relation $(x_0, 0) \sim (x_t, t)$.

By A5, there exists t_1 such that $(x, t) \succ (1, t + t_1)$ for all $t \geq 0$ and for any y_i such that such that $(1, 0) \sim (y_i, i)$ for $i \geq 0$, $(x, 0) \succ (y_i, t_1 + i)$. Consider the following cases:

CASE 1: Consider $x_0 < x < 1$. By A5, there exists $t \geq 1$, such that in the sequence constructed, $x_{t-1} < x \leq x_t$. Note that given the construction of the sequence for $(x, 0) \sim (x_i, i)$, it must be that for any (x_p, x_q), $p < q$, $(x_p, 0) \succsim (x_q, q - p))$. By monotonicity, using $x_{t-1} < x \leq x_t$, for any point x_i in the sequence , $|i| \leq t$, one has $(x_i, 0) \succsim (x_t, t - i) \succsim (x, t - i)$. Hence, for any element x_i of the sequence with $i \leq 0$, $(x_i, 0) \succsim (x, t - i) \succ (1, t_1 + t - i)$, with the last inequality coming from A5.[26] Hence, the $x_{(t+t_1)}$th element of the sequence must be weakly to the right of 1. Thus, $u(x) \geq \dfrac{1}{\delta^{t_1+1}}$.

[26]The property we are using implicitly without proving is the following: In our constructed sequences, x_i always is a direct reflection from $\{x_0, x_{-1}, x_{-2}, ..\}$ when i is positive, and a direct reflection of $\{x_0, x_1, x_2..\}$when i is negative. This follows from WPB.

CASE 2: Consider $x < x_0 < 1$. By construction of the dual sequence $\{..x_{-1}, x_0, x_1, ..\}$, it must be that $x_{-t_1} \leq x$ and $x_{t_1} \geq 1$. Thus, $u(x) \geq \frac{1}{\delta^{2t_1}}$. [27]
Hence, $u(x) \geq \frac{1}{\delta^{2t_1}}$ for all $u \in \mathcal{U}$.
Now, showing that $\sup_{u \in \mathcal{U}} u(x) < \infty$. This is trivial for points $x \leq 1$. Consider $x > 1$. By A4, there exists t_1 such that $(1, 0) \succ (x, t_1)$ and for any y such that $(x, 0) \sim (y, i)$, $i \geq 0$,$(1, 0) \succ (y, t_1 + i)$. Suppose we are constructing a function that would respect the relation $(x_0, 0) \sim (x_t, t)$, and in the process construction a dual sequence $\{..x_{-1}, x_0, x_1, ..\}$. There are two cases as before.

CASE 1: Consider $x_0 > x > 1$. By A4, there exists $t \geq 1$, such that in the sequence constructed, one has $x_{-t} \leq x < x_{-t+1}$. As before, given the construction of the sequence for $(x, 0) \sim (x_i, i)$, it must be that for any (x_p, x_q), $p < q$, $(x_p, 0) \succsim (x_q, q - p)$.) By monotonicity, using $x_{-t} \leq x$, for any point x_i in the sequence , $|i| \leq t$, $(x, 0) \succsim (x_{-t}, 0) \succsim (x_i, i + t)$. Hence, for any element x_i of the sequence with $i \geq 0$, $(1, 0) \succ (x_i, t_1 + i + t)$. Thus, $u(x) \leq \frac{1}{\delta^{t_1+1}}$.

CASE 2: Consider $x > x_0 > 1$. By construction of the dual sequence $\{..x_{-1}, x_0, x_1, ..\}$, it must be that $x_{-t_1} \leq 1$ and $x_{t_1} \geq x$. Thus, $u(x) \leq \frac{1}{\delta^{2t_1}}$.
Hence, $u(x) \leq \frac{1}{\delta^{2t_1}}$for all $u \in \mathcal{U}$. □

Lemma 5. *Under Axioms A1-A6, there exists a continuous present equivalent utility function $F : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ that represents $\succsim$. F is monotonically increasing in x and monotonically decreasing in t.*

Proof. The first part of this proof is very similar to Lemma 2, and we will omit it here. By construction of the set $\mathcal{U}$, $V(x, t) = \min_{v \in \mathcal{U}} v^{-1}(\delta^t v(x))$. Moreover, for all $u \in \mathcal{U}$, $u(0) = 0$, $u(1) = 1$, $\inf_{u \in \mathcal{U}} u(x) > 0$, $\sup_{u \in \mathcal{U}} u(x) < \infty$ for $x > 0$.
Finally, from A6, for any $x > 0$, and $t \in \mathbb{T}$, there exists z such that $(z, t) \succ (x, 0)$.

$$\begin{aligned} \delta^t u(z) &> u(x) \quad \forall u \in \mathcal{U} \\ \implies \frac{u(z)}{u(x)} &> \delta^{-t} \quad \forall u \in \mathcal{U} \\ \implies \inf_u \frac{u(z)}{u(x)} &\geq \delta^{-t} \quad \forall u \in \mathcal{U} \end{aligned}$$

[27]One can make the bound tighter.

But we had started with arbitrary t. Hence, $\inf_u \dfrac{u(z)}{u(x)}$ is unbounded above for any $x > 0$.

$\square$

Theorem 3: The following two statements are equivalent:

i) The relation $\succsim$ satisfies properties B0-B5.

ii) There exists a continuous function $F : \mathbb{X} \times \mathbb{P} \times \mathbb{T} \to \mathbb{R}$ such that $(x, p, t) \succsim (y, q, s)$ if and only if $F(x, p, t) \geq F(y, q, s)$. The function F is continuous, increasing in x, p and decreasing in $t \in \mathbb{T}$. There exists a unique $\delta \in (0, 1)$ and a set $\mathcal{U}$ of monotinically increasing continuous functions such that $F(x, p, t) = \min_{u \in \mathcal{U}} u^{-1}(p\delta^t u(x))$ and $u(0) = 0$ for all $u \in \mathcal{U}$.

Proof. Showing that (ii) implies (i) :
Continuity and monotonicity of $\succsim$ follow from the continuity and monotonicity of F. Weak Present Bias follows as before.
B5 can be shown in the following way:

$$\begin{aligned}
(x, p\theta, t) &\succsim (x, p, t + D) \\
\implies \min_u u^{-1}(p\theta\delta^t u(x)) &\geq \min_u u^{-1}(p\delta^{t+D} u(x)) \\
\implies \theta &\geq \delta^D \\
\implies \min_u u^{-1}(q\theta\delta^s u(y)) &\geq \min_u u^{-1}(q\delta^{s+D} u(y)) \\
\implies (y, q\theta, s) &\succsim (y, q, s + D)
\end{aligned}$$

We will prove the direction (i) to (ii) in the following three steps.

Step 1: Recall the Probability Time Tradeoff axiom: for all $x, y \in \mathbb{X}$, $p, q \in (0, 1]$,and $t, s \in \mathbb{T}$, $(x, p\theta, t) \succsim (x, p, t + \Delta) \implies (y, q\theta, s) \succsim (y, q, s + \Delta)$.
This axiom has calibration properties that we will use. Given Monotinicity, $(x, 1, 0) \succ (x, 1, 1) \succ (x, 0, 0)$ for any $x \in \mathbb{X}$. By continuity, there must exist $\delta \in (0, 1)$ such that $(x, \delta, 0) \sim (x, 1, 1)$. Note that Probability-Time Tradeoff Axiom helps us write $(x, \delta, \tau + 1) \sim (x, 1, \tau)$ for all $x \in \mathbb{X}$ and $\tau \in \mathbb{T}$, and further extend it to $(x, q, t) \sim (x, q\delta^t, 0)$. For integer t's this follows by induction.
For any integer $b > 0$, there exists $\Delta(\frac{1}{b}) = \delta_1 \in \mathbb{P}$ such that $(x, \delta_1, 0) \sim (x, 1, \dfrac{1}{b})$. Now applying Probability Time Tradeoff (PTT) repeatedly b times, $(x, 1, 1) \sim (x, \delta_1^b, 0)$, which implies, $\delta_1 = \delta^{\frac{1}{b}}$. For any ratio of positive integers (rational number) $t = \dfrac{a}{b}$, $\Delta(\frac{a}{b}) = \delta^{\frac{a}{b}}$. This argument can be extended to all real $t > 0$. This

crucially helps us pin down δ as the discount factor.
Henceforth, we are going to concentrate on finding a representation of the reduced domain of $\mathbb{X} \times [0,1]$. Note that this reduced domain can also be conceptually seen as the set of all binary lotteries that have zero as one of the outcomes.

Step 2: The rest of the proof will have a similar flavor to the ones the reader has already encountered. For every $x^* \in \mathbb{X}$, we are going to provide an increasing utility function u on $[0, M]$ which would respect all the relations of the form $(x,p) \succsim (y,1)$, i.e, have $pu(x) \geq u(y)$ and also have $pu(y) = u(x^*)$ for all $(x^*,1) \sim (y,p)$.
Fix x^*, $u(0) = 0$ and $u(x^*) = 1$. For $x \in (x^*, M]$, define

$$u(x) = \{\frac{1}{p} : (x,p) \sim (x^*,1)\} \tag{8}$$

and,

$$x_q \quad = \quad \{x : (x,q) \sim (x^*,1)\} \tag{9}$$

The expressions in (8) and (9) exist due to the continuity of $\succsim$.
Now, for $x \in (0, x^*)$, define

$$u(x) = \inf\{p(q) : (x_q, qp(q)) \sim (x,1), q \leq 1\} \tag{10}$$

First, we will show that the infimum in (10) can be replaced by minimum. Consider a sequence of probabilities $\{p_n\}$ that converge below to p^*, and $(x_{q_n}, p_n q_n) \sim (x,1)$. Note that $q_n \in [q_{max}, 1]$ where $(x^*,1) \sim (M, q_{max})$. Hence, $\{q_n\}$ must be bounded by this closed interval, and must have a convergent subsequence $\{q_{n_k}\}$. Let q^* be the corresponding limit, and we know that $q^* \geq q_{max}$. Similarly, x_q can be considered continuous in q (this also follows from the continuity of $\succsim$). Therefore, $x_{q_{n_k}} \to x_{q^*}$ as $q_{n_k} \to q^*$. Also, it must be that $p_{n_k} \to p^*$ as $q_{n_k} \to q^*$. Thus, we have $(x_{q_{n_k}}, p_{n_k} q_{n_k}) \sim (x,1)$ for all elements of $\{n_k\}$. Then, using the continuity of $\succsim$, $(x_{q^*}, p^* q^*) \sim (x,1)$.

$$u(x) = \inf\{p : (x_q, pq) \sim (x,1)\} = \min\{p : (x_q, pq) \sim (x,1)\} = p^*$$

Now we will show that the utility defined in (8) and (10) has the following properties : 1) It is increasing. 2) $p_1 u(x_1) = u(y_1)$ 3) $(x,p) \succsim (y,1)$, implies $pu(x) \geq u(y)$ 4) u is continuous. The first two properties are true by definition of u. We will show the third in some detail. Consider $(x,p) \succsim (y,1)$. In the case of interest,

$p < 1$ and hence, $x > y$. Now let $x > y > x^*$. Let, $u(y) = 1/p_1$, which means, $(y, p_1) \sim (x^*, 1)$. Given $(x, p) \succsim (y, 1)$, we must have $\square$

$$(x, pp_1) \succsim (y, p_1) \sim (x^*, 1)$$

Hence,

$$\begin{aligned} u(x) &\geq \frac{1}{pp_1} \\ \iff pu(x) &\geq \frac{1}{p_1} = u(y) \end{aligned}$$

If, $x > x^* > y$, the proof follows from the way the utility has been defined.
Let $y < x < x^*$. Let, $u(x) = p_1$, which means, $(x_q, p_1 q) \sim (x, 1)$ for some x_q. Given $(x, p) \succsim (y, 1)$, we must have

$$(x_q, p_1 qp) \succsim (x, p) \succsim (y, 1)$$

Hence, $u(y) \leq pp_1$.
Now we turn to proving the continuity of u. The continuity at x^* from the right is easy to see.
Next, for any $r \in (0, 1)$, define

$$f(r) = \sup\{x \in [0, x^*) : (x_q, qr) \sim (x, 1)\} \tag{11}$$

The supremum can be replaced by a maximum, and the proof is similar to the one before. Suppose there is a sequence of $\{x_n\}$ that converges up to a value $\hat{x}$, and, $(x_{q(n)}, q(n)r) \sim (x_n, 1)$. Note that $q(n)$ lies in a closed interval, and hence has a convergent subsequence that converges to a point in that interval. Let us call this point $\hat{q}$. Now, x is continuous in q (in the usual sense), and hence, x_n and $x_{q(n)}$ also has a convergent subsequence. The convergent subsequence $\{x_{q(n)}\}$ and $\{x_n\}$ must have the same limit point, let us call it $x_{\hat{q}}$, a point in $[x^*, M]$. Hence, the supremum in (11) must have been attained from $x_{\hat{q}}$, and hence the supremum can be replaced by a maximum. The f function is well defined, strictly increasing and is the inverse function of u over $r \in (0, 1)$. This function can be used to show the continuity of u at the point x^*.
Finally, the function u can be easily normalized to have $u(M) = 1$.
Step 3: In this step, we construct the $\mathcal{U}$ set as in Theorem theorem **??**, to complete the proof.

Theorem 4: *The following two statements are equivalent:*

i) The relation $\succsim$ on $[0,\infty)^T$ satisfies properties D0-D5.

ii) For any $\delta \in (0,1)$, there exists a set $\mathcal{U}_\delta$ of monotinically increasing continuous functions such that

$$F(x_0, x_1, .., x_{T-1}) = x + \sum_{1}^{T-1} \min_{u \in \mathcal{U}_\delta} u^{-1}(\delta^t u(x_t))$$

represents the binary relation $\succsim$. The set $\mathcal{U}_\delta$ has the following properties: $u(0) = 0$ and $u(M) = 1$ for all $u \in \mathcal{U}_\delta$. $F(.)$ is continuous.

Proof: Going from (ii) to (i), we will show how the representation satisfies D5 and the second property in D2.
Suppose, $(x_0, x_1, .., x_{T-1})$ and $(y_0, y_1, .., y_{T-1})$ are orthogonal. Therefore,

$$\begin{aligned} F(x_0 + y_0, x_1 + y_1, .., x_{T-1} + y_{T-1}) &= F(x_0, x_1, .., x_{T-1}) + F(y_0, y_1, .., y_{T-1}) \\ &= F(z_0, 0, .., 0) + F(z_0', 0, .., 0) \\ &= z_0 + z_0' \\ &= F(z_0 + z_0', 0, .., 0) \end{aligned}$$

To see how Discounting can be derived, start by assuming $y^0 > x > 0$, and choose a function $u_1 \in \mathcal{U}$. As $\delta \in (0,1)$, there must exist t such that $\delta^t u_1(y^0) < u_1(x)$ and hence, $u_1^{-1}(\delta^t u_1(y^0)) < x$. For any sequences $(y^1, y^2, y^3, ..y^m)$ and $(n^1, n^2, .., n^m)$, where, $(0, ..0, \underbrace{y^{i-1}}_{\text{in period } n_i}, 0.., 0) \succsim (y^i, 0, .., 0)\ \forall i \in \{1, 2, ..., m\}$, one must have $\delta^{n_i} u_1(y^i) \geq u_1(y^{i-1})\ \forall i \in \{1, 2, ..., m\}$.
Multiplying all these inequalities gives us,

$$\begin{aligned} \delta^{\sum n_i} u_1(y^0) &\geq u_1(y^m) \\ \iff y_m &\leq u_1^{-1}(\delta^{\sum n_i} u_1(y^0)) \\ &= u_1^{-1}(\delta^t u_1(y^0)) \\ &< u_1^{-1}(u_1(x)) \\ &= x \\ \implies y_m &\leq x \end{aligned}$$

Now to show the proof for the direction (i) to (ii), we start by following the same steps we used in the proof of Theorem 1 to derive the set $\mathcal{U}_\delta$. There are two points to be noted during the construction of functions in $\mathcal{U}_\delta$. First, only comparisons upto lengths of $T-1$ periods need to be considered. Secondly, in the construction of each function $u \in \mathcal{U}_\delta$, the fact that the interative construction spans over $\mathbb{R}_{\geq 0}$ is guaranteed by the second part of the Discounting axiom. The additive representation across periods follows from using induction and the D5 axiom.

Theorem 5: *Let $\succsim_1$ and $\succsim_2$ be two binary relations which allow for minimum representation w.r.t sets $\mathcal{U}_{\delta,1}$ and $\mathcal{U}_{\delta,2}$ respectively. The following two statements are equivalent:*
i) $\succsim_1$ allows a higher premium to the present than $\succsim_2$.
ii) Both $\mathcal{U}_{\delta,1}$ and $\mathcal{U}_{\delta,1} \cup \mathcal{U}_{\delta,2}$ provide minimum representations for $\succsim_1$.

Proof. The direction from (i) to (ii): Consider any $(x,t) \in \mathbb{X} \times \mathbb{T}$ such that $(x,t) \sim_1 (y,0)$. Using (i), we must have, $(x,t) \succsim_2 (y,0)$.
Hence,

$$\begin{aligned} \min_{u \in \mathcal{U}_{\delta,2}} u^{-1}(\delta^t u(x)) &\geq y \\ \implies \min_{u \in \mathcal{U}_{\delta,1} \cup \mathcal{U}_{\delta,2}} u^{-1}(\delta^t u(x)) &= y \end{aligned}$$

Hence,

$$\min_{u \in \mathcal{U}_{\delta,1} \cup \mathcal{U}_{\delta,2}} u^{-1}(\delta^t u(x)) = \min_{u \in \mathcal{U}_{\delta,1}} u^{-1}(\delta^t u(x)) \tag{12}$$

To go in the opposite direction, let us assume, $(x,t) \succsim_1 (y,0)$.
Given, (12), it must be that

$$\begin{aligned} \min_{u \in \mathcal{U}_{\delta,1} \cup \mathcal{U}_{\delta,2}} u^{-1}(\delta^t u(x)) &= \min_{u \in \mathcal{U}_{\delta,1}} u^{-1}(\delta^t u(x)) \geq y \\ \implies u^{-1}(\delta^t u(x)) &\geq y \quad \forall u \in \mathcal{U}_{\delta,1} \cup \mathcal{U}_{\delta,2} \\ \implies u^{-1}(\delta^t u(x)) &\geq y \quad \forall u \in \mathcal{U}_{\delta,2} \\ \implies \min_{u \in \mathcal{U}_{\delta,2}} u^{-1}(\delta^t u(x)) &\geq y \end{aligned}$$

Hence, $(x,t) \succsim_2 (y,0)$, which completes the proof. □

Proposition 5. *Let f_n be a set of bijective, increasing, continuous functions. Let $f_n \to f$ "locally uniformly"/ compactly (equivalent notions in $\mathbb{R}^n$.), where f is bijective, increasing, continuous. Then, $f_n^{-1} \to f^{-1}$ compactly.*

Proof. Consider the composite function $g_n = f_n o f^{-1}$. Note that g_n is also bijective, increasing, continuous. As f_n converges locally uniformly to f, g_n converges locally uniformly to the identity function $g(x)$.
To see this, note that for any $\epsilon_1 > 0$

$$\begin{aligned}\sup_{x\in[c,d]} |g_n(x) - g(x)| &= \sup_{x\in[c,d]} |f_n(f^{-1}(x)) - f(f^{-1}(x))| \\ &= \sup_{y\in[f^{-1}(c),f^{-1}(d)]} |f_n(y) - f(y)| \\ &\leq \epsilon_1\end{aligned}$$

for $n \geq N_0$ for some N_0.
Choose an interval $[a, b]$. Now, there would exist n_1, n_2 such that $g_n(a-1) \leq a$ and $g_n(b+1) > b$ for $n \geq n_1$ and $n \geq n_2$ respectively. Similarly, there exists n_3 such that $\sup_{x\in[a-1,b+1]} |g_n(x) - g(x)| < \epsilon$ for $n \geq n_3$.
Finally, for $N \geq max\{n_1, n_2, n_3\}$

$$\begin{aligned}\sup_{x\in[a,b]} |g_n^{-1}(x) - g(x)| &\leq \sup_{x\in[g_n(a-1),g_n(b+1)]} |g_n^{-1}(x) - x| \\ &= \sup_{t\in[a-1,b+1]} |g_n^{-1}(g_n(t)) - g(t)| \\ &= \sup_{t\in[a-1,b+1]} |t - g(t)| \\ &< \epsilon\end{aligned}$$

Therefore, $g_n^{-1} = f o f_n^{-1}$ converges locally uniformly to the identity function. Therefore, f_n^{-1} converges locally uniformly to f^{-1}. □

References

Allais, M., "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 1953, *21* (4), 503–546.

Baucells, Manel and Franz H Heukamp, "Common ratio using delay," *Theory and Decision*, 2010, *68* (1-2), 149–158.

______ **and** ______, "Probability and time trade-off," *Management Science*, 2012, *58* (4), 831–842.

Benhabib, Jess, Alberto Bisin, and Andrew Schotter, "Present-bias, quasi-hyperbolic discounting, and fixed costs," *Games and Economic Behavior*, 2010, *69* (2), 205 – 223.

Cerreia-Vioglio, Simone, "Maxmin expected utility on a subjective state space: convex preferences under risk," *preprint*, 2009.

———, **David Dillenberger, and Pietro Ortoleva**, "Cautious expected utility and the certainty effect," *Econometrica*, 2015, *83* (2), 693–728.

Chakraborty, Anujit and Yoram Halevy, "Allais Meets Strotz: Remarks on the Relation Between Present Bias and the Certainty Effect," Technical Report, Vancouver School of Economics 2015.

Dasgupta, Partha and Eric Maskin, "Uncertainty and Hyperbolic Discounting," *American Economic Review*, 2005, *95* (4), 1290–1299.

Dillenberger, David, "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior," *Econometrica*, 2010, *78* (6), 1973–2004.

Dubra, Juan, Fabio Maccheroni, and Efe A Ok, "Expected utility theory without the completeness axiom," *Journal of Economic Theory*, 2004, *115* (1), 118–133.

Epper, Thomas and Helga Fehr-Duda, "The missing link: Unifying risk taking and time discounting," *University of Zurich Department of Economics Working Paper*, 2012, (96).

Fishburn, Peter C. and Ariel Rubinstein, "Time Preference," *International Economic Review*, 1982, *23* (3), 677–694.

Frederick, Shane, George Loewenstein, and Ted O'donoghue, "Time discounting and time preference: A critical review," *Journal of economic literature*, 2002, *40* (2), 351–401.

Fudenberg, Drew and David K. Levine, "Risk, Delay, and Convex Self-Control Costs," *American Economic Journal: Microeconomics*, August 2011, *3* (3), 34–68.

Gilboa, Itzhak and David Schmeidler, "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 1989, *18* (2), 141–153.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and Self-Control," *Econometrica*, 2001, *69* (6), 1403–1435.

Halevy, Yoram, "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review*, 2008, *98* (3), 1145–62.

———, "Time consistency: Stationarity and time invariance," *Econometrica*, 2015, *83* (1), 335–352.

Hayashi, Takashi, "Quasi-stationary cardinal utility and present bias," *Journal* *of Economic Theory*, 2003, *112* (2), 343–352.

Keren, Gideon and Peter Roelofsma, "Immediacy and certainty in intertemporal choice," *Organizational Behavior and Human Decision Processes*, 1995, *63* (3), 287–297.

Koopmans, Tjalling C., *Decision and Organization* 1972.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

Loewenstein, George, "Out of control: Visceral influences on behavior," *Organizational behavior and human decision processes*, 1996, *65* (3), 272–292.

________ **and Drazen Prelec**, "Anomalies in intertemporal choice: Evidence and an interpretation," *The Quarterly Journal of Economics*, 1992, pp. 573–597.

Maccheroni, Fabio, "Maxmin under Risk," *Economic Theory*, 2002, *19* (4), 823–831.

McClure, Samuel M., Keith M. Ericson, David I. Laibson, George Loewenstein, and Jonathan D. Cohen, "Time Discounting for Primary Rewards," *Journal of Neuroscience*, 2007, *27* (21), 5796–5804.

Noor, Jawwad, "Intertemporal choice and the magnitude effect," *Games and Economic Behavior*, 2011, *72*, 255 – 270.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, March 1999, *89* (1), 103–124.

Ok, Efe A and Yusufcan Masatlioglu, "A theory of (relative) discounting," *Journal of Economic Theory*, 2007, *137* (1), 214–245.

Olea, JL Montiel and Tomasz Strzalecki, "Axiomatization and measurement of quasi-hyperbolic discounting," *The Quarterly Journal of Economics*, 2014, *129* (3), 1449–1499.

Pan, Jinrui, Craig S Webb, and Horst Zank, "An extension of quasi-hyperbolic discounting to continuous time," *Games and Economic Behavior*, 2015, *89*, 43–55.

Phelps, Edmund S and Robert A Pollak, "On second-best national saving and game-equilibrium growth," *The Review of Economic Studies*, 1968, *35* (2), 185–199.

Pollak, Robert A, "Consistent planning," *The Review of Economic Studies*, 1968, *35* (2), 201–208.

Read, Daniel, "Is time-discounting hyperbolic or subadditive?," *Journal of risk and uncertainty*, 2001, *23* (1), 5–32.

Rubinstein, Ariel, "Economics and Psychology? The Case of Hyperbolic Discounting," *International Economic Review*, 2003, *44* (4), 1207–1216.

Saito, Kota, "A relationship between risk and time preferences," *The Center for Mathematical Studies in Economics and Management Sciences, Northwestern University, Discussion Paper*, 2009, (1477).

Sozou, Peter D, "On hyperbolic discounting and uncertain hazard rates," *Proceedings of the Royal Society of London B: Biological Sciences*, 1998, *265* (1409), 2015–2020.

Thaler, Richard, "Some empirical evidence on dynamic inconsistency," *Economics Letters*, 1981, *8* (3), 201–207.