An Attentional Model of Time Discounting

Zark Zijian Wang

June 05, 2024

1 Introduction

decision maker (DM)

Kullback–Leibler (KL) divergence (also called relative entropy)

hard attention

information avoidance

endogenous time preferences

optimal expectation

we present an axiomatic characterization of AAD with the optimal discounting framework

2 Model Setting

Assume time is discrete. Let $s_{0\to T} \equiv [s_0, s_1, ..., s_T]$ denote a reward sequence that starts delivering rewards at period 0 and ends at period T. At each period t of $s_{0\to T}$, a specific reward s_t is delivered, where $t \in \{0, 1, \ldots, T\}$. Throughout this paper, we only consider non-negative rewards and finite length of sequence, i.e. we set $s_t \in \mathbb{R}_{\geq 0}$ and $1 \leq T < \infty$. The DM's choice set is constituted by a range of alternative reward sequences which start from period 0 and end at some finite period. When making an intertemporal choice, the DM seeks to find a reward sequence $s_{0\to T}$ in her choice set, which has the highest value among all alternative reward sequences. To calculate the value of each reward sequence, we admit the additive discounted utility framework. The value of $s_{0\to T}$ is defined as $U(s_{0\to T}) \equiv \sum_{t=0}^{T} w_t u(s_t)$, where $u(s_t)$ is the instantaneous utility of receiving s_t, and w_t is the decision weight (sometimes called discount factors) assigned to s_t. We assume the function $u(s_t)$ is strictly increasing with s_t and for any $s_t > 0$, we have $u(s_t) > 0$. For convenience, we set $u(0) = 0$.

The determination of w_t is central to this paper. We believe that, due to the DM's limited attention and demand for information, the DM tends to overweight the large rewards and underweight the small rewards within the sequence. Specifically, we suggest w_t follow a (generalized) softmax function. We define any decision weight in this style as an *attention-adjusted discount* factors (AAD), as in Definition 1.

Definition 1: *Let $\mathcal{W} \equiv [w_0, ..., w_T]$ denote the decision weights for all specific rewards in $s_{0\to T}$. $\mathcal{W}$ is called attention-adjusted discount factors (AADs) if for any $t \in \{0, 1, \ldots, T\}$,*

$$w_t = \frac{d_t e^{u(s_t)/\lambda}}{\sum_{\tau=0}^{T} d_\tau e^{u(s_\tau)/\lambda}} \tag{1}$$

where $d_t > 0$, $\lambda > 0$, $u(.)$ is the utility function.

In intuition, how Definition 1 reflects the role of attention in valuating reward sequences can be explained with four points. First, each reward in a sequence could be viewed as an information source and we assume the DM allocates limited information-processing resources across those information sources. The AADs capture this notion by normalizing the discount factors, i.e. fixing the sum of w_t at 1. As a result, increasing the decision weight of one reward would reduce the decision weights of other rewards in the sequence, implying that focusing on one reward would make DM insensitive to the values of other rewards. Meanwhile, when there are more rewards in the sequence, DM needs to split attention across a wider range to process each of them, which may reduce the attention to, or decision weight of, each individual reward.

Second, w_t is strictly increasing with s_t, indicating that DM would pay more attention to larger rewards. This is consistent with many empirical studies that suggest people tend to pay more attention to information associated with larger rewards. For instance, people perform a "value-driven attentional capture" effect in visual search (Della Libera and Chelazzi, 2009; Hickey et al., 2010; Anderson et al., 2011; Chelazzi et al., 2013; Jahfari and Theeuwes, 2017). In one study (Hickey et al., 2010), researchers recruit participants to do a series of visual search trials, in each of which participants earn a reward after detecting a target object from distractors. If a target object is associated with a large reward in previous trials, it can naturally capture more attention. Therefore, in the next trial, presenting the object as a distractor slows down the target detection.[1] In addition, in financial decision making, investors usually perform an ostrich effect (Galai and Sade, 2006; Karlsson et al., 2009). One relevant evidence is that stock traders log in their brokerage accounts less frequently after market declines (Sicherman et al., 2016).

Third, w_t is "anchored" in a reference weight d_t. For a certain sequence of rewards, d_t could denote the initial weight that the DM would assign to a reward delivered at period t without knowing its realization. The determination of d_t is mediated by the difficulty to mentally represent a future event (Trope and Liberman, 2003) and the frequency of time delays in a global context (Stewart et al., 2006). The constraint on the deviation between w_t and d_t indicates that reallocating attention or acquiring new information is costly. The deviation of w_t from d_t depends on parameter λ, which as we discuss in the next section, can reflect the unit cost of attention adjustment. A large λ implies a low learning rate and a high cognitive cost in adapting the decision weights to the local context.

Fourth, we adopt the idea of Gottlieb (2012) and Gottlieb et al. (2013) that attention can be understood as an active information-sampling mechanism which selects information based on the perceived utility of information. For intertemporal choices, we assume the DM would selectively sample value information from each reward (information source) when processing a reward sequence, and the AAD can represent an approximately optimal sampling strategy. Note that the AADs follow a softmax function. Matějka and McKay (2015) and Maćkowiak

[1] Strotz (1955) shows that if, for any reward delivered at period t, the DM's discount factor is δ^t, then her preference will be stationary and consistent over time.

et al. (2023) claim that if a behavioral strategy conforms to this type of function, then it can be interpreted as a solution to some optimization problem under information constraints.

3 Interpretation

3.1 Information Maximizing Exploration

In this section, we provide two approaches to characterize AAD: the first is based on information maximizing exploration, and the second is based on optimal discounting. These approaches are closely related to the idea proposed by Gottlieb (2012), Gottlieb et al. (2013) and Sharot and Sunstein (2020), that people tend to pay attention to information with high *instrumental utility* (help identifying the optimal action), *cognitive utility* (satisfying curiosity), or *hedonic utility* (inducing positive feelings). It is worth mentioning that the well-known rational inattention theories are grounded in the instrumental utility of information.[2] Instead, in this paper, we draw on the cognitive and hedonic utility of information to build our theory of time discounting. Our first approach to characterizing AAD is relevant to the cognitive utility: the DM's information acquisition process is curiosity-driven. The model setting of this approach, similar with Gottlieb (2012) and Gottlieb et al. (2013), is based on a reinforcement learning framework. Specifically, we assume the DM seeks to maximize the information gain with a commonly-used exploration strategy. Our second approach is relevant to the hedonic utility: the DM consider the feelings of multiple selves and seeks to maximize their total utility under some cognitive cost. The theoretical background for the second approach is Noor and Takeoka (2022, 2024). We describe the first approach in this subsection and the second approach in Section 3.2.

For the information maximizing exploration approach, we assume that before having any information of a reward sequence, the DM perceives it has no value. Then, each reward in the sequence $s_{0\to T}$ is processed as an individual information source. The DM engages her attention to actively sample signals at each information source and update her belief about

[2] It is worth mentioning that if we make the "hidden zeros" explicit in LL and SS, adding a common delay under the AAD model would always yield a the common difference effect.

the sequence value accordingly. The signals are nosiy. For any $t \in \{0, 1, \ldots, T\}$, the signal sampled at information source s_t could be represented by $x_t = u(s_t) + \epsilon_t$, where each ϵ_t is i.i.d. and $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. The sampling weight for information source s_t is denoted by w_t.

The DM's belief about the sequence value $U(s_{0\to T})$ is updated as follows. At first, she holds a prior U_0, and given she perceives no value from the reward sequence, the prior could be represented by $U_0 \sim N(0, \sigma^2)$. Second, she draws a series of signals at each information source s_t. Note we define $U(s_{0\to T})$ as a weighted mean of instantaneous utilities, i.e. $U(s_{0\to T}) = \sum_{t=0}^{T} w_t u(s_t)$. Let $\bar{x}$ denote the mean sample signal and U denote a realization of $U(s_{0\to T})$. If there are k signals being sampled in total, we should have $\bar{x}|U, \sigma_\epsilon \sim N(U, \frac{\sigma_\epsilon^2}{k})$. Third, she uses the sampled signals to infer $U(s_{0\to T})$ in a Bayesian fashion. Let U_k denote the valuer's posterior about the sequence value after receiving k signals. According to the Bayes' rule, we have $U_k \sim N(\mu_k, \sigma_k^2)$ and

$$\mu_k = \frac{k^2\sigma_\epsilon^{-2}}{\sigma^{-2} + k^2\sigma_\epsilon^{-2}}\bar{x} \qquad , \qquad \sigma_k^2 = \frac{1}{\sigma^{-2} + k^2\sigma_\epsilon^{-2}}$$

We assume the DM takes μ_k as the valuation of reward sequence. It is clear that as $k \to \infty$, the sequence value will converge to the mean sample signal, i.e. $\mu_k \to \bar{x}$.

The DM's goal of sampling signals is to maximize her information gain. The information gain is defined as the KL divergence from the prior U_0 to the posterior U_k. In intuition, the KL divergence provides a measure for distance between distributions. As the DM acquires more information about $s_{0\to T}$, her posterior belief should move farther away from the prior. We let $p_0(U)$ and $p_k(U)$ denote the probability density functions of U_0 and U_k. Then, the information gain is

$$\begin{aligned} D_{KL}(U_k||U_0) &= \int_{-\infty}^{\infty} p_k(U) \log\left(p_k(U)/p_0(U)\right) dU \\ &= \frac{\sigma_k^2 + \mu_k^2}{2\sigma^2} - \log\left(\frac{\sigma_k}{\sigma}\right) - \frac{1}{2} \end{aligned} \tag{2}$$

Notably, in Equation (2), σ_k depends only on sample size k and μ_k is proportional to $\bar{x}$. Therefore, the problem of maximizing $D_{KL}(U_k||U_0)$ could be reduced to maximizing $\bar{x}$ (as each $u(s_t)$ is non-negative). The reason is that, drawing more samples can always increase the precision of the DM's estimate about $U(s_{0\to T})$, and a larger $\bar{x}$ implies more "surprises" in comparison to the DM's initial perception that $s_{0\to T}$ contains no value.

Maximizing the mean sample signal $\bar{x}$ under a limited sample size k is actually a multi-armed bandit problem (Sutton and Barto, 2018, Ch.2). On the one hand, the DM wants to draw more samples at information sources that are known to produce greater value signals (exploit). On the other hand, she wants to learn some value information from other information sources (explore). We assume the DM would take a softmax exploration strategy to solve this problem. That is,

$$w_t \propto d_t e^{\bar{x}_t/\lambda}$$

where $\bar{x}_t$ is the mean sample signal generated by information source s_t so far, $1/\lambda$ is the learning rate, and d_t is the initial sampling weight for s_t.[3] Note $\bar{x}_t$ cannot be calculated without doing simulations under a certain σ_ϵ. For researchers, modelling an intertemporal choice in this way requires conducting a series of simulations and then calibrating σ_ϵ for every choiceable option, which could be computationally expensive. Fortunately, according to the weak law of large numbers, as the sample size k gets larger, $\bar{x}_t$ is more likely to fall into a neighborhood of $u(s_t)$. Thus, the AAD which assumes $w_t \propto d_t e^{u(s_t)/\lambda}$ could be viewed as a proper approximation to the softmax exploration strategy.

Those who familiar with reinforcement learning algorithms may notice that here $u(s_t)$ is a special case of action-value function (assuming that the learner only cares about the value of current reward in each draw of sample). The AAD thus can be viewed as a specific version of the soft Q-learning or policy gradient method for solving the given multi-armed bandit problem (Haarnoja et al., 2017; Schulman et al., 2017). Such methods are widely used (and sample-efficient) in reinforcement learning. Moreover, one may argue that the applicability of softmax exploration strategy is subject to our model assumptions. Under alternative assumptions, the strategy may not be ideal. We acknowledge this limitation and suggest that researchers interested in modifying our model consider different objective functions or

[3] For theoretical analysis about intertemporal correlation aversion, please see Epstein (1983), Epstein and Zin (1989), Weil (1990), Bommier (2005), and Bommier et al. (2017). The AAD model takes a similar form to the class of models defined in Epstein (1983). A key feature of such models is that the discount factor for future utilities is dependent on the utility achieved in the current period.

different families of noises. For example, if the DM aims to minimize the regret rather than maximizing $\bar{x}$, the softmax exploration strategy can produce suboptimal actions and one remedy is to use the Gumbel–softmax strategy (Cesa-Bianchi et al., 2017). If noises $\epsilon_0, ..., \epsilon_T$ do not follow an i.i.d normal distribution, the information gain $D_{KL}(U_k||U_0)$ may be complex to compute, thus one can use its variational bound as the objective (Houthooft et al., 2016). Compared to these complex settings, the model setting in this subsection aims to provides a simple benchmark for understanding the role of attention in mental valuation of a reward sequence.

Two strands of literature can justify the information maximizing approach to characterizing AAD. First, our assumption that the DM updates decision weights toward a greater $D_{KL}(U_k||U_0)$ is compatible with the well-known finding that Bayesian surprises attract visual attention (Itti and Baldi, 2009). Second, the softmax exploration strategy is widely used by neuroscientists in reinforcement learning studies (Daw et al., 2006; Niv et al., 2012; FitzGerald et al., 2012; Collins and Frank, 2014; Niv et al., 2015; Leong et al., 2017). For instance, Daw et al. (2006) find the softmax strategy can characterize humans' exploration behavior better than other classic strategies (e.g. ϵ-greedy). Collins and Frank (2014) show that models based on the softmax strategy exhibit a good performance in explaining the striatal dopaminergic system's activities (which is central in brain's sensation of pleasure and learning of rewarding actions) in reinforcement learning tasks.

3.2 Optimal Discounting

The second approach to characterize AAD is based on the optimal discounting model (Noor and Takeoka, 2022, 2024). In one version of this model, the authors assume that the DM has a limited capacity of attention (or in their term, "empathy"), and before evaluating a reward sequence $s_{0\to T}$, she naturally focuses on the current period. The instantaneous utility $u(s_t)$ represents the well-being that the DM's self of period t can obtain from the reward sequence. For valuating $s_{0\to T}$, the DM needs to split attention over T time periods to consider the feeling of each self. This re-allocation of attention is cognitive costly. The DM seeks to find a balance between improving the overall well-being of multiple selves and reducing the incurred cognitive cost. Noor and Takeoka (2022, 2024) specify an optimization problem to capture this decision. In this paper, we adopt a variant of their original model. The formal definition of the optimal discounting problem is given by Definition 2. [4]

Definition 2: *Given reward sequence $s_{0\to T} = [s_0, ..., s_T]$, the following optimization problem is called an optimal discounting problem for $s_{0\to T}$:*

$$
\begin{aligned}
\max_{\mathcal{W}} \quad & \sum_{t=0}^{T} w_t u(s_t) - C(\mathcal{W}) \\
s.t. \quad & \sum_{t=0}^{T} w_t \leq M \\
& w_t \geq 0 \text{ for all } t \in \{0, 1, ..., T\}
\end{aligned}
$$

where $M > 0$, $C(\mathcal{W}) \geq 0$. For any $t \in \{0, 1, \ldots, T\}$, $u(s_t) < \infty$. $C(\mathcal{W})$ *is the cognitive cost function and is constituted by time-separable costs, i.e. $C(\mathcal{W}) = \sum_{t=0}^{T} f_t(w_t)$, where for all $w_t \in (0, 1)$, $f_t(w_t)$ is differentiable. $f'_t(w_t)$ is continuous and* stric*tly increasing, and* $\lim_{w_t \to 0} f'_t(w_t) = -\infty$.

Here w_t reflects the attention paid to consider the feeling of t-period self. The DM's objective function is the attention-weighted sum of utilities obtained by the multiple selves minus the cognitive cost of attention re-allocation. As is illustrated by Noor and Takeoka (2022, 2024), a key feature of this model is that decision weight w_t is increasing with s_t, indicating the DM tends to pay more attention to larger rewards. Moreover, it is easy to validate that if the following three conditions are satisfied, the solution to the optimal discounting problem will take an AAD form:

(i) The constraint on sum of decision weights is always tight. That is, $\sum_{t=0}^{T} w_t = M$. Without loss of generality, we can set $M = 1$.

(ii) There exists a realization of decision weights $\mathcal{D} = [d_0, ..., d_T]$ such that $d_t > 0$ for all $t \in \{0, \ldots, T\}$ and the cognitive cost is proportional to the KL divergence from $\mathcal{D}$ to

[4] To validate, suppose $u(x_s) = 5$, $u(x_l) = 10$, $u(y_s) = 1$, $u(y_l) = 3$. The results of each lottery contain only two periods, t_1 and t_2. The reference weights are uniformly distributed, i.e. $d_{t_1} = d_{t_2}$. In this case, setting $\lambda = 1$ would generate intertemporal correlation seeking, while setting $\lambda = 100$ would generate intertemporal correlation aversion.

the DM's strategy $\mathcal{W}$ where applicable. That is, $C(\mathcal{W}) = \lambda \cdot D_{KL}(\mathcal{W}||\mathcal{D})$, where $\lambda > 0$.

Here d_t sets a reference for determining the decision weight w_t, the parameter λ indicates how costly the attention re-allocation process is, and $D_{KL}(\mathcal{W}||\mathcal{D}) = \sum_{t=0}^{T} w_t \log(\frac{w_t}{d_t})$. The solution to the optimal discounting problem under condition (i)-(ii) can be derived in the same way as Theorem 1 in Matějka and McKay (2015). Note this solution is equivalent to that of a bounded rationality model: assuming the DM wants to find a $\mathcal{W}$ that maximizes $\sum_{t=0}^{T} w_t u(s_t)$ but can only search for solutions within a KL neighborhood of $\mathcal{D}$. Related models can be found in Todorov (2009).

We interpret the implications of condition (i)-(ii) with behavioral axioms. Note if each s_t is an independent option and $\mathcal{W}$ simply represents the DM's choice strategy across options, then these condition can be characterized by rational inattention theories, e.g. Caplin et al. (2022). However, here $\mathcal{W}$ is a component of sequence value $U(s_{0\to T})$, and the DM is assumed to choose the option with highest sequence value. Thus, the behavioral implications of condition (i)-(ii) should be derived in different ways. To illustrate, let $\succsim$ denote the preference relation between two reward sequences.[5] For any reward sequence $s_{0\to T} = [s_0, ..., s_T]$, we define $s_{0\to t} = [s_0, ..., s_t]$ as a sub-sequence of it, where $1 \le t \le T$.[6] We first introduce two axioms for $\succsim$:

Axiom 1: $\succsim$ *has the following properties:*

(a) *(complete order)* $\succsim$ *is complete and transitive.*

(b) *(continuity) For any reward sequences* s, s' *and reward* $c \in \mathbb{R}_{\ge 0}$*, the sets* $\{\alpha \in (0,1) | \alpha \cdot s + (1-\alpha) \cdot c \succsim s'\}$ *and* $\{\alpha \in (0,1) | s' \succsim \alpha \cdot s + (1-\alpha) \cdot c\}$ *are closed.*

(c) *(state-independent) For any reward sequences* s, s' *and reward* $c \in \mathbb{R}_{\ge 0}$*,* $s \succsim s'$ *implies for any* $\alpha \in (0,1)$*,* $\alpha \cdot s + (1-\alpha) \cdot c \sim \alpha \cdot s' + (1-\alpha) \cdot c$*.*

[5] If $a \succsim b$ and $b \succsim a$, we say $a \sim b$ ("a is the same good as b"). If $a \succsim b$ does not hold, we say $b \succ a$ ("b is better than a"). $\succsim$ can also characterize the preference relation between single rewards as the single rewards can be viewed as one-period sequences.

[6] Notably, every sub-sequence starts with period 0.

(d) *(reduction of compound alternatives) For any reward sequences* s, s', y *and rewards* $c_1, c_2 \in \mathbb{R}_{\geq 0}$*, if there exist* $\alpha, \beta \in (0, 1)$ *such that* $s \sim \alpha \cdot y + (1-\alpha) \cdot c_1$*, then* $s' \sim \beta \cdot s + (1-\beta) \cdot c_2$ *implies* $s' \sim \beta\alpha \cdot y + \beta(1-\alpha) \cdot c_1 + (1-\beta) \cdot c_2$.

Axiom 2: *For any* $s_{0\to T}$ *and any* $\alpha_1, \alpha_2 \in (0, 1)$*, there exists* $c \in \mathbb{R}_{\geq 0}$ *such that* $\alpha_1 \cdot s_{0\to T-1} + \alpha_2 \cdot s_T \sim c$.

The two axioms are almost standard in decision theories. The assumption of complete order implies preferences between reward sequences can be characterized by an utility function. Continuity and state-independence ensure that in a stochastic setting where the DM can receive one reward sequence under some states and receive a single reward under other states, her preference can be characterized by expected utility (Herstein and Milnor, 1953). Reduction of compound alternatives ensures that the DM's valuation on a reward sequence is constant across states. Axiom 2 is an extension of the Constant-Equivalence assumption in Bleichrodt et al. (2008). It implies there always exists a constant that can represent the value of a linear combination of sub-sequence $s_{0\to T}$ and the end-period reward s_T, as long as the weights lie in $(0, 1)$.

For a given $s_{0\to T}$, the optimal discounting model can generate a sequence of decision weights $[w_0, ..., w_T]$. Furthermore, the model assumes the DM's preference for $s_{0\to T}$ can be characterized by the preference for $w_0 \cdot s_0 + w_1 \cdot s_1 + ... + w_T \cdot s_T$. We use Definition 3 to capture this assumption.[7]

Definition 3: *Given reward sequence* $s_{0\to T} = [s_0, ..., s_T]$ *and* $s'_{0\to T'} = [s'_0, ..., s'_{T'}]$*, the preference relation* $\succsim$ *has an optimal discounting representation if*

$$s_{0\to T} \succsim s'_{0\to T'} \quad \Longleftrightarrow \quad \sum_{t=0}^{T} w_t \cdot s_t \succsim \sum_{t=0}^{T'} w'_t \cdot s'_t$$

where $\{w_t\}_{t=0}^{T}$ *and* $\{w'_t\}_{t=0}^{T'}$ *are solutions to the optimal discounting problems for* $s_{0\to T}$ *and* $s'_{0\to T'}$ *respectively.*

Furthermore, if Definition 3 is satisfied and $\{w_t\}_{t=0}^{T}$ as well as $\{w'_t\}_{t=0}^{T'}$ takes the AAD

[7] Noor and Takeoka (2022) refer the "optimal discounting representation" in Definition 3 as Costly Empathy representation.

form, we say $\succsim$ has an *AAD representation*. Now we specify two behavioral axioms that are key to characterize the AAD functions.

Axiom 3 (sequential outcome-betweenness): *For any $s_{0\to T}$, there exists $\alpha \in (0,1)$ such that $s_{0\to T} \sim \alpha \cdot s_{0\to T-1} + (1-\alpha) \cdot s_T$.*

Axiom 4 (sequential bracket-independence): *Suppose $T \geq 2$. For any $s_{0\to T}$, if there exist $\alpha_1, \alpha_2, \beta_0, \beta_1, \beta_2 \in (0,1)$ such that $s_{0\to T} \sim \alpha_1 \cdot s_{0\to T-1} + \alpha_2 \cdot s_T$ and $s_{0\to T} \sim \beta_0 \cdot s_{0\to T-2} + \beta_1 \cdot s_{T-1} + \beta_2 \cdot s_T$, then we must have $\alpha_2 = \beta_2$.*

Axiom 3 implies that for a reward sequence $s_{0\to T-1}$, if we add a new reward s_T at the end of the sequence, then the value of the new sequence should lie between the original sequence $s_{0\to T-1}$ and the newly added reward s_T. Notably, Axiom 3 is consistent with the empirical evidence about *violation of dominance* (Scholten and Read, 2014; Jiang et al., 2017) in intertemporal choice. Suppose the DM is indifferent between a small-sooner reward (SS) “receive £75 today” and a large-later reward (LL) “receive £100 in 52 weeks”. Scholten and Read (2014) find when we add a tiny reward after the payment in SS, e.g. changing SS to “receive £75 today and £3 in 52 weeks”, the DM would be more likely to prefer LL over SS. Jiang et al. (2017) find the same effect can apply to LL. That is, if we add a tiny reward after the payment in LL, e.g. changing LL to “receive £100 in 52 weeks and £3 in 53 weeks”, the DM may be more likely to prefer SS over LL.

Axiom 4 implies that no matter how the DM brackets the rewards into sub-sequences (or how the sub-sequences get further decomposed), the decision weights for rewards outside them should not be affected. Specifically, suppose we decompose reward sequence $s_{0\to T}$ and find its value is equivalent to a linear combination of $s_{0\to T-1}$ and s_T. We also can further decompose $s_{0\to T-1}$ to a linear combination of $s_{0\to T-2}$ and s_{T-1}. But no matter how we operate, as long as the decomposition is carried out inside $s_{0\to T-1}$, the weight of s_T in the valuation of $s_{0\to T}$ will always remain the same. This axiom is an analog to independence of irrelevant alternatives in discrete choice problems (which is a key feature of softmax choice function).

We show in Proposition 1 that the optimal discounting model plus Axiom 1-4 can exactly produce AAD.

Proposition 1: *Suppose $\succsim$ has an optimal discounting representation, then it satisfies Axiom 1-4 if and only if has an AAD representation.*

The necessity ("only if") is easy to see. We present the proof of sufficiency ("if") in Appendix A. The sketch of the proof is as follows. First, by recursively applying Axiom 3 and Axiom 1 to each sub-sequence of $s_{0\to T}$, we can obtain $s_{0\to T} \sim w_0 \cdot s_0 + ... + w_T \cdot s_T$, where $\sum_{t=0}^{T} w_t = 1$ and $w_t > 0$. Second, by the FOC of the optimal discounting problem, we have $f_t'(w_t) = u(s_t) + \theta$, where θ is the Lagrangian multiplier. Given $f_t'(.)$ is continuous and strictly increasing, we can set $w_t = \phi_t(u(s_t) + \theta)$, where $\phi_t(.)$ is also a continuous and strictly increasing function. Third, Axiom 4 indicates that the relative decision weights between rewards within a reward (sub-)sequence is irrelevant to the rewards outside it. Imagine that we add a new reward s_{T+1} to the end of $s_{0\to T}$ and the decision weights for s_{T+1} are denoted by $\{w_t'\}_{t=0}^{T+1}$. Then, by Axiom 4 (jointly with Axiom 1-3), we should have $\frac{w_0}{w_0'} = \frac{w_1}{w_1'} = ... = \frac{w_T}{w_T'}$. Thus, $w_t \propto w_t'$. Suppose $w_t' = \phi_t(u(s_t) - \eta)$, we can obtain $w_t \propto e^{\ln \phi_t(u(s_t)-\eta)}$. Fourth, we can adjust s_{T+1} arbitrarily to get different realizations of η. Suppose under some s_{T+1}, $w_t' = \phi_t(u(s_t))$. Then, we also have $w_t \propto e^{\ln \phi_t(u(s_t))}$. By comparing the two proportional relations, we can find that there exists $\kappa > 0$ such that $\ln \phi_t(u(s_t)) = \ln \phi_t(u(s_t) - \eta) + \kappa\eta$ for all $t \in \{0, 1, \dots, T\}$, which indicates $\ln \phi_t(.)$ is linear under some given η. Finally, we show that the linear condition holds when $\eta \in [0, u_{\max} - u_{\min}]$, where $u_{\max}, u_{\min}$ are the maximum and minimum instantaneous utilities in $s_{0\to T}$. So, we can rewrite $\ln \phi_t(u(s_t))$ as $\ln \phi_t(u_{\min}) + \kappa(u(s_t) - u_{\min})$. Setting $d_t = \phi_t(u_{\min})$, $\lambda = 1/\kappa$, and reframing the utility function, we obtain $w_t \propto d_t e^{u(s_t)/\lambda}$, which is AAD.

4 Implications for Decision Making

4.1 Hidden Zero Effect

Empirical evidence suggests the subjective discount factor of a reward is dependent not just on delay but also on the framing of reward sequences. In this subsection, we discuss the evidence of (asymmetric) hidden zero effect (Magen et al., 2008; Radu et al., 2011; Read et al., 2017). Similar with violation of dominance, this effect also provides a justification for our assumption that the sum of AADs is a fixed amount.

To illustrate, suppose the DM is indifferent between "receive £100 today" (SS) and "receive £120 in 25 weeks" (LL). The hidden zero effect suggests that people are more likely to choose LL when SS is framed as a sequence rather than as a single-period reward. In other words, if we frame SS as "receive £100 today and £0 in 25 weeks" (SS1), the DM would prefer LL to SS1. Moreover, Read et al. (2017) find that framing LL as "receive £0 today and £120 in 25 weeks" (LL1) has no effect on preference.

The hidden zero effect can be explain by the AAD model. When a DM valuates a reward sequence $s_{0\to T}$, the AAD model assumes that she splits a fixed amount of attention over T periods. For the given example, the DM may perceive the time length of SS as "today" and perceives the time length of SS1 as "25 weeks". In the former case, she can focus her attention on the current period when she can get £100. While in the latter case, she have to spend some attention to future periods in which no reward is delivered, which also decreases the decision weight assigned to the current period. As a result, she values SS1 lower than SS. By contrast, the DM may perceive the time length of both LL and LL1 as "25 weeks". When she valuate LL, she has already paid some attention to periods earlier than "25 weeks". Therefore, changing LL to LL1 does not change the choice.

4.2 Relation to Hyperbolic Discounting

Most of the intertemporal choice studies only involve comparisons between single-period rewards (SS and LL). Here we derive the discount factor for SS/LL under the AAD model and use that to illustrate how attention allocation can account for the anomalies in such decision settings. For simplicity, we assume the reference weight $d_t = \delta^t$, $\delta \in (0, 1]$, i.e. the DM initial discount factor (before learning information about $s_{0\to T}$) is exponential.[8]

Consider a reward sequence $s_{0\to T}$ where for all $t \leq T$, $u(s_t) = 0$ and only $u(s_T) > 0$. This implies the DM receives nothing until period T. In this case, the DM's valuation of $s_{0\to T}$

[8] Strotz (1955) shows that if, for any reward delivered at period t, the DM's discount factor is δ^t, then her preference will be stationary and consistent over time.

is $U(s_{0\to T}) = w_T u(s_T)$. Let $v(x) = u(x)/\lambda$. By Definition 1, we can derive that w_T is a function of s_T:

$$w_T = \frac{1}{1 + G(T)e^{-v(s_T)}} \tag{3}$$

where

$$G(T) = \begin{cases} \frac{1}{1-\delta}(\delta^{-T} - 1)\ , & 0 < \delta < 1 \\ T\ , & \delta = 1 \end{cases}$$

This w_T can represent the discount function for a single reward s_T, delivered at period T. Interestingly, when $\delta = 1$, $w_T(s_T)$ takes a form similar with hyperbolic discounting. In recent years, some studies have attempted to provide a rational account for hyperbolic discounting. For instance, Gabaix and Laibson (2017) propose a model with similar assumptions to our information maximizing exploration approach to AAD: the DM's perception of utility is noisy and the DM updates her beliefs about utility with the Bayes' rule. Nevetherless, Gabaix and Laibson (2017) account for hyperbolic discounting with an assumption that the variance of utility signal is proportional to delay, whereas we propose the DM seeks to maximize her information gain when learning information about utilities. Meanwhile, when $v(s_t) = \ln(\beta s_t + 1)$, where $\beta > 0$, the discount function w_T takes a similar form to Gershman and Bhui (2020). Therefore, some remarks about such models can also be generated by the AAD in a special case.

In the following three subsections, we use Equation (3) to explain three decision anomalies: the common difference effect (and its reverse), timing risk aversion, and S-shaped value function.

4.3 Common Difference Effect

The common difference effect (Loewenstein and Prelec, 1992) implies that, when the DM faces a choice between LL and SS, adding a common delay to both options can increase her preference for LL. For example, suppose the DM is indifferent between "receive £120 in 25 weeks" (LL) and "receive £100 today" (LL). Then, she would prefer "receive £120 in 40 weeks" to "receive £100 in 15 weeks".

Let (v_l, t_l) denote a reward of utility v_l, delivered at period t_l and (v_s, t_s) denote a reward of utility v_s, delivered at period t_s. We set $v_l > v_s > 0$, $t_l > t_s > 0$. So, (v_l, t_l) can represent a LL and (v_s, t_s) can represents a SS. We represent the discount factors for LL and SS by $w_{t_l}(v_l)$ and $w_{t_s}(v_s)$. Suppose $w_{t_l}(v_l) \cdot v_l = w_{t_s}(v_s) \cdot v_s$, then the common difference effect implies that $w_{t_l+\Delta t}(v_l) \cdot v_l > w_{t_s+\Delta t}(v_s) \cdot v_s$, where $\Delta t > 0$. Given Equation (3), the conditions for the common difference effect are derived in Proposition 2.

Proposition 2: *The following statements are true:*

(a) *If $\delta = 1$, the common difference effect always holds.*

(b) *If $0 < \delta < 1$, i.e. the DM is initially impatient, the common difference effect holds when and only when$v_l - v_s + \ln(v_l/v_s) > (t_l - t_s)\ln(1/\delta)$.*

The proof of Proposition 2 is in Appendix B. The part (b) of Proposition 2 yields a novel prediction about the common difference effect. That is, for an impaient DM, to make this effect hold, the relative and absolute differences in reward utility between LL and SS must be significantly larger than their aboslute difference in time delay. In the opposite, if the difference in delay is significantly larger than the difference in reward utility, we may observe a reversed common difference effect.[9] Figure 1 demonstrates an example for this reversed effect.

When the DM is impatient, adding a common delay would naturally make v_l and v_s more discounted, i.e. less attention is paid to the corresponding rewards. Since the sum of decision weights is fixed, this implies the the DM frees up some attention and she needs to reallocate it across the periods in each reward sequence (LL and SS). There are three mechanisms jointly determining whether we could observe the common difference effect or not.

First, the existing periods with no reward delivered would grab some attention. That is, the DM would attend more to the rewards of zero utility, delivered in duration $[0, t_l)$ for LL and in duration $[0, t_s]$ for SS. Given $t_l > t_s$, the relevant duration in LL may naturally

[9] It is worth mentioning that if we make the "hidden zeros" explicit in LL and SS, adding a common delay under the AAD model would always yield a the common difference effect.

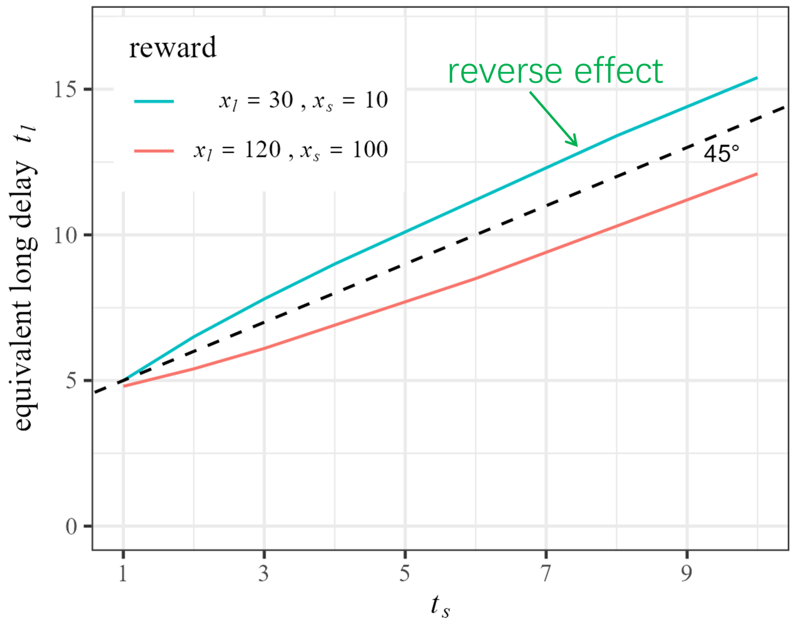


Figure 1: The common difference effect and its reverse

Note: x_l and x_s are the positive reward levels for LL and SS. The values of LL and SS are calculated based on Equation (3). $d_t = 0.75^t$, $u(x) = x^{0.6}$, $\lambda = 2$. For each certain t_s, we identify the delay t_l that makes the value of LL equivalent to SS. If the common different effect is valid, for one unit increase in t_s, the resultant t_l should increase by a level smaller than one unit.

capture more attention than that in SS. In other words, the common delay makes the DM focus more on the waiting time in LL than in SS, which decreases her preference for LL.

Second, the newly added time intervals also grab some attention. That is, the DM needs to pay some attention to rewards (of zero utility as well) delivered in duration $(t_l, t_l + \Delta t]$ in LL and in duration $(t_s, t_s + \Delta t]$ in SS. For LL, there are already plenty of periods over which DM has to split her attention. So, the duration $(t_l, t_l + \Delta t]$ in LL can capture less attention than its counterpart in SS. This increases the DM's preference for LL.

Third, the only positive reward, delivered in t_l for LL and in t_s for SS, may draw some attention back. Given that the DM in general tends to pay more attention to larger rewards, the positive reward in LL can capture more "free" attention than that that in SS. This also increases the preference for LL. If the latter two mechanisms override the first mechanism, we would observe a common difference effect in DM's choices.

4.4 Concavity of Discount Function

Many time discounting models, such as exponential and hyperbolic discounting, assume the discount function is convex in time delay. This style of discount function predicts DM is *risk seeking over time lotteries*. To illustrate, suppose a reward of level x is delivered at period t_l with probability π and is delivered at period t_s with probability $1-\pi$, where $0<\pi<1$. Meanwhile, another reward of the same level is delivered at period t_m, where $t_m=\pi t_l+(1-\pi)t_s$. Under such discount functions, the DM should prefer the former reward to the latter reward. For instance, she may prefer receiving an amount of money today or in 20 weeks with equal chance, rather than receiving it in 10 weeks with certainty. However, experimental studies suggest that people are often *risk averse over time lotteries*, i.e. they prefer the reward to be delivered at a certain time (Onay and Öncüler, 2007; DeJarnette et al., 2020).

One way to accommodate the evidence about risk aversion over time lotteries is to make the discount function concave in terms of delay. Notably, Onay and Öncüler (2007) find that people are more likely to be risk averse over time lotteries when π is small, and to be risk seeking when π is large. Given that when π gets larger, t_m is also larger, we can conclude that the discount function may be concave in delay for the near future but convex for the far future. That is, the discount function is of inverse-S shape. Takeuchi (2011) also find evidence that supports this shape of discount function.

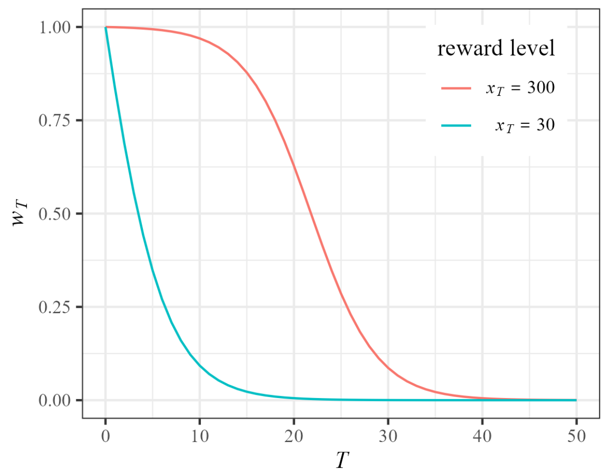
In Proposition 3, we apply Equation (3) and show that the AAD is compatible with this shape of discount function as long as the DM is impatient and the reward level x is large enough.

Proposition 3: *If $\delta=1$, then the discount function w_T is convex in T. If $0<\delta<1$, there exist a reward threshold $\underline{x}>0$ and a time threshold $\underline{T}>0$ such that:*

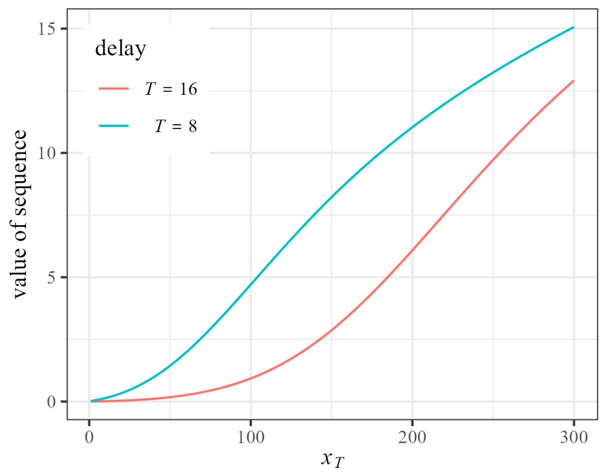
(a) *when $x\leq\underline{x}$, the discount function is convex in T;*

(b) *when $x>\underline{x}$, the discount function is convex in T given $T\geq\underline{T}$, and it is concave in T given $0<T<\underline{T}$.*

The proof of Proposition 3 is in Appendix C. Figure 2(a) demonstrates the convex discount function (blue line) and the inverse-S shaped discount function (red line) that could be yielded by Equation (3).



(a) discount function



(b) value function

Figure 2: Discount function and value function for a delayed reward

Note: A reward of level x_T is delivered at period T. The discount function and value function are calculated based on Equation (3). $d_t = 0.75^t$, $u(x) = x^{0.6}$, $\lambda = 2$.

4.5 S-Shaped Value Function

In decision theories, it is commonly assumed that the utility function $u(.)$ satisfies $u'' < 0$. This usually suggests the value function of a reward is concave. However, empirical evidence suggests that the value functions are often S-shaped. Such S-shaped value functions can be generated by various sources, such as reference dependence (Kahneman and Tversky, 1979) and efficient coding of numbers (Louie and Glimcher, 2012). Through the AAD model, we provide a novel account for S-shaped value function based on the insight that larger rewards capture more attention.

Consider a reward of level x delivered at period T. Its value function can be represented by $U(x, T) = w_T(x)u(x)$. We assume $u' > 0$, $u'' < 0$, and w_T is determined by Equation (3). w_T is increasing with x as the DM tends to pay more attention to larger rewards. Both functions $u(x)$ and $w_T(x)$ are concave in x; so when x is small, they both grow fast. At some conditions, it is possible that the product of the two functions is convex in x when x is small enough. We derive the conditions for the S-shaped value function in Proposition 4.

Proposition 4: *Suppose $T \geq 1$, $\frac{d}{dx}\left(\frac{1}{v'(x)}\right)$ is continuous in $(0, +\infty)$, then:*

(a) *There exists a threshold $\bar{x} \in \mathbb{R}_{\geq 0}$ such that $U(x, T)$ is strictly concave in x when $x \in [\bar{x}, +\infty)$;*

(b) *If $\frac{d}{dx}\left(\frac{1}{v'(x)}\right)$ is right-continuous at $x = 0$ and $\frac{d}{dx}\left(\frac{1}{v'(0)}\right) < 1$, there exists a threshold x^* in $(0, \bar{x})$ such that, for any $x \in (0, x^*)$, $U(x, T)$ is strictly convex in x.*

(c) *There exists an unit cost of attention adjustment λ^* and an interval (x_1, x_2) such that, if $\lambda < \lambda^*$, for any $x \in (x_1, x_2)$, $U(x, T)$ is strictly convex in x, where $\lambda^* > 0$ and $(x_1, x_2) \subset (0, \bar{x})$.*

The proof of Proposition 4 is in Appendix D. Proposition 4 implies, if the derivative of $\frac{1}{v'(x)}$ converges to a small number when $x \to 0^+$, or the unit cost of attention adjustment λ is small enough, the value function $U(x, T)$ will be an S-shaped in some interval of x. Figure 2(b) demonstrates examples of this S-shaped value function.

4.6 Intertemporal Correlation Aversion

Consider a DM facing two lotteries. For one lottery, she can receive £100 today and £100 in 30 weeks with probability 1/2, and receive £3 today and £3 in 20 weeks with probability 1/2. For the other lottery, she can receive £3 today and £100 in 30 weeks with probability 1/2, and receive £100 today and £3 in 20 weeks with probability 1/2. In the former lottery, rewards delivered at different periods are positively correlated, whereas in the latter lottery, those rewards are negatively correlated. The expected discounted utility theory predicts the DM is indifferent between the two lotteries. However, reccent studies find the evidence of *intertemporal correlation aversion*(Andersen et al., 2018; Rohde and Yu, 2023). That is, people often prefer the latter lottery than the former one.[10]

[10]For theoretical analysis about intertemporal correlation aversion, please see Epstein (1983), Epstein and Zin (1989), Weil (1990), Bommier (2005), and Bommier et al. (2017). The AAD model takes a similar

For the above example, intertemporal correlation aversion can be generated by the AAD model as follows. The AAD model assumes the DM allocates decision weights within each specific reward sequence, which implies she would aggregate values over time in each state. For simplicity, suppose there are only two periods. In the state that the DM receives £3 in two periods, suppose the DM allocates decision weight w to the first period and $1-w$ to the second period. Note in Definition 1, when $u(s_0) = u(s_1) = ... = u(s_T)$, the decision weight w_t for every period t within the reward sequence $s_{0\to T}$ will remain the same as the reference weight d_t. So, in the state that the DM receives £100 in two periods, the allocation of decision weights is the same as w and $1-w$. In the state that the DM can receive £100 in the first period and £3 in the second period, the reward of £100 can capture more attention so that its decision weight, say w', is greater than w. Similarly, in the state that the DM receives £3 earlier and then £100, the decision weight for the reward of £100, say $1-w''$, is also greater than $1-w$. Therefore, the value of the lottery in which rewards are positively correlated, can be represented by $0.5 \cdot u(3) + 0.5 \cdot u(100)$. By contrast, for the lottery in which rewards are negatively correlated, the value can be represented by $0.5(1-w'+w'') \cdot u(3) + 0.5(1-w''+w') \cdot u(100)$. Given $(1-w'')+w' > 1-w+w > 1$, the decision weight assigned to $u(100)$, which is $0.5(1-w''+w')$, should be greater than 0.5. As a result, the DM prefer the latter lottery than the former lottery.

In a more general setting, whether the AAD model can continuously produce intertemporal correlation aversion is modulated by the unit cost of attention adjustment λ. To illustrate, we adopt the same theoretical setting as Bommier (2005). Let (s_1, s_2) denote the result of a lottery in which the DM can receive reward s_1 in period t_1 and then reward s_2 in period t_2, where $t_2 > t_1 \geq 0$. There are two lotteries, $L1$ and $L2$. The results of each lottery is of the same length of sequence. $L1$ generates (x_s, y_s) and (x_l, y_l) with equal chance, L_2 generates (x_s, y_l) and (x_l, y_s) with equal chance, $x_l > x_s > 0$, $y_l > y_s > 0$. By Proposition 5, we show that in this setting, we can always find a λ that makes the DM intertemporal correlation averse.

Proposition 5: *Suppose $U(L1), U(L2)$ are the values of $L1$ and $L2$. For any $x_l > x_s > 0$,*

form to the class of models defined in Epstein (1983). A key feature of such models is that the discount factor for future utilities is dependent on the utility achieved in the current period.

$y_l > y_s > 0$, *any reference weights, and any time length of lottery results, there exists a threshold* λ^{**} such that for any unit co*st of attention adjustment* $\lambda > \lambda^{**}$, *the DM would perform intertemporal correlation aversion, i.e.* $U(L1) < U(L2)$.

The proof of Proposition 5 is in Appendix E. The threshold λ^{**} is jointly determined by x_l, y_l, y_s, as well as the reference weights for rewards delivered at t_1 and t_2. Notably, when $\lambda < \lambda^{**}$, the DM may be intertemporal correlation seeking under some conditions.[11] This suggests a potentially new mechanism for intertemporal correlation aversion, that is, DM performs intertemporal correlation aversion perhaps because she attends more to larger rewards while attention adjustment is very costly.

4.7 Learning and Inconsistent Planning

Suppose a DM has budget m ($m > 0$) and is considering how to spend it over different time periods. We can use a reward sequence x to represent this decision problem, where the DM's spending in period t is x_t. In period 0, she wants to find a x such that

$$\max_x \sum_{t=0}^{T} w_t u(x_t) \quad s.t. \sum_{t=0}^{T} x_t = m \tag{3}$$

where w_t is the attention-adjusted discounting factor in period t. I assume $w_t = \delta^t e^{u(x_t)/\lambda} / \sum_{t=\tau}^{T} \delta^\tau e^{u(x_\tau)/\lambda}$ and there is no risk under this setting.

5 Discussion

5.1 Relation to Other Models of Intertemporal Choice

The theory most similar to AAD is the salience theory (Bordalo et al., 2012, 2013, 2020).

rational inattention

[11] To validate, suppose $u(x_s) = 5$, $u(x_l) = 10$, $u(y_s) = 1$, $u(y_l) = 3$. The results of each lottery contain only two periods, t_1 and t_2. The reference weights are uniformly distributed, i.e. $d_{t_1} = d_{t_2}$. In this case, setting $\lambda = 1$ would generate intertemporal correlation seeking, while setting $\lambda = 100$ would generate intertemporal correlation aversion.

focus-weighted utility

bayesian updating and discounting

optimal precision

Relation with money/delay trade-off

5.2 Limitation

attention biases learning: learning rate is high for attended reward

sum of decision weights

6 Conclusion

Reference

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2018). Multiattribute utility theory, intertemporal utility, and correlation aversion. *International Economic Review*, 59(2):537–555.

Anderson, B. A., Laurent, P. A., and Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, 108(25):10367–10371.

Bleichrodt, H., Rohde, K. I., and Wakker, P. P. (2008). Koopmans' constant discounting for intertemporal choice: A simplification and a generalization. *Journal of Mathematical Psychology*, 52(6):341–347.

Bommier, A. (2005). Risk aversion, intertemporal elasticity of substitution and correlation aversion. Technical report, ETH Zurich.

Bommier, A., Kochov, A., and Le Grand, F. (2017). On monotone recursive preferences. *Econometrica*, 85(5):1433–1466.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, 127(3):1243–1285.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention, and choice. *The Quarterly journal of economics*, 135(3):1399–1442.

Caplin, A., Dean, M., and Leahy, J. (2022). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130(6):1676–1715.

Cesa-Bianchi, N., Gentile, C., Lugosi, G., and Neu, G. (2017). Boltzmann exploration done right. *Advances in neural information processing systems*, 30.

Chelazzi, L., Perlato, A., Santandrea, E., and Della Libera, C. (2013). Rewards teach visual selective attention. *Vision research*, 85:58–72.

Collins, A. G. and Frank, M. J. (2014). Opponent actor learning (opal): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological review*, 121(3):337.

Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–879.

DeJarnette, P., Dillenberger, D., Gottlieb, D., and Ortoleva, P. (2020). Time lotteries and stochastic impatience. *Econometrica*, 88(2):619–656.

Della Libera, C. and Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological science*, 20(6):778–784.

Epstein, L. G. (1983). Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, 31(1):133–152.

Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica*, 57:937–969.

FitzGerald, T. H., Friston, K. J., and Dolan, R. J. (2012). Action-specific value signals in reward-related regions of the human brain. *Journal of Neuroscience*, 32(46):16417–16423.

Gabaix, X. and Laibson, D. (2017). Myopia and discounting. Technical report, National bureau of economic research.

Galai, D. and Sade, O. (2006). The "ostrich effect" and the relationship between the liquidity and the yields of financial assets. *The Journal of Business*, 79(5):2741–2759.

Gershman, S. J. and Bhui, R. (2020). Rationally inattentive intertemporal choice. *Nature communications*, 11(1):3365.

Gottlieb, J. (2012). Attention, learning, and the value of information. *Neuron*, 76(2):281–295.

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., and Baranes, A. (2013). Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in cognitive sciences*, 17(11):585–593.

Haarnoja, T., Tang, H., Abbeel, P., and Levine, S. (2017). Reinforcement learning with deep energy-based policies. In *International conference on machine learning*, pages 1352–1361. PMLR.

Herstein, I. N. and Milnor, J. (1953). An axiomatic approach to measurable utility. *Econometrica, Journal of the Econometric Society*, pages 291–297.

Hickey, C., Chelazzi, L., and Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience*, 30(33):11096–11103.

Houthooft, R., Chen, X., Duan, Y., Schulman, J., De Turck, F., and Abbeel, P. (2016). Vime: Variational information maximizing exploration. *Advances in neural information processing systems*, 29.

Itti, L. and Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision research*, 49(10):1295–1306.

Jahfari, S. and Theeuwes, J. (2017). Sensitivity to value-driven attention is predicted by how we learn from value. *Psychonomic bulletin & review*, 24(2):408–415.

Jiang, C.-M., Sun, H.-M., Zhu, L.-F., Zhao, L., Liu, H.-Z., and Sun, H.-Y. (2017). Better is worse, worse is better: Reexamination of violations of dominance in intertemporal choice. *Judgment and Decision Making*, 12(3):253–259.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Karlsson, N., Loewenstein, G., and Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty*, 38:95–115.

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., and Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, 93(2):451–463.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597.

Louie, K. and Glimcher, P. W. (2012). Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences*, 1251(1):13–32.

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2023). Rational inattention: A review. *Journal of Economic Literature*, 61(1):226–273.

Magen, E., Dweck, C. S., and Gross, J. J. (2008). The hidden zero effect: Representing a single choice as an extended sequence reduces impulsive choice. *Psychological science*, 19(7):648.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298.

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., and Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience*, 35(21):8145–8157.

Niv, Y., Edlund, J. A., Dayan, P., and O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience*, 32(2):551–562.

Noor, J. and Takeoka, N. (2022). Optimal discounting. *Econometrica*, 90(2):585–623.

Noor, J. and Takeoka, N. (2024). Constrained optimal discounting. *Available at SSRN 4703748*.

Onay, S. and Öncüler, A. (2007). Intertemporal choice under timing risk: An experimental approach. *Journal of Risk and Uncertainty*, 34:99–121.

Radu, P. T., Yi, R., Bickel, W. K., Gross, J. J., and McClure, S. M. (2011). A mechanism for reducing delay discounting by altering temporal attention. *Journal of the experimental analysis of behavior*, 96(3):363–385.

Read, D., Olivola, C. Y., and Hardisty, D. J. (2017). The value of nothing: Asymmetric attention to opportunity costs drives intertemporal decision making. *Management Science*, 63(12):4277–4297.

Rohde, K. I. and Yu, X. (2023). Intertemporal correlation aversion—a model-free measurement. *Management Science*.

Scholten, M. and Read, D. (2014). Better is worse, worse is better: Violations of dominance in intertemporal choice. *Decision*, 1(3):215.

Schulman, J., Chen, X., and Abbeel, P. (2017). Equivalence between policy gradients and soft q-learning. *arXiv preprint arXiv:1704.06440*.

Sharot, T. and Sunstein, C. R. (2020). How people decide what they want to know. *Nature Human Behaviour*, 4(1):14–19.

Sicherman, N., Loewenstein, G., Seppi, D. J., and Utkus, S. P. (2016). Financial attention. *The Review of Financial Studies*, 29(4):863–897.

Stewart, N., Chater, N., and Brown, G. D. (2006). Decision by sampling. *Cognitive psychology*, 53(1):1–26.

Strotz, R. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Takeuchi, K. (2011). Non-parametric test of time consistency: Present bias and future bias. *Games and Economic Behavior*, 71(2):456–478.

Todorov, E. (2009). Efficient computation of optimal actions. *Proceedings of the national academy of sciences*, 106(28):11478–11483.

Trope, Y. and Liberman, N. (2003). Temporal construal. *Psychological review*, 110(3):403.

Weil, P. (1990). Nonexpected utility in macroeconomics. *The Quarterly Journal of Economics*, 105(1):29–42.

Appendix

A. Proof of Proposition 1

We present the proof of sufficiency here. That is, if $\succsim$ has an optimal discounting representation and satisfies Axiom 1-4, then it has an AAD representation.

Lemma 1: *If Axiom 1 and 3 hold, for any $s_{0\to T}$, there exist $w_0, w_1, \ldots, w_T > 0$ such that $s_{0\to T} \sim w_0 \cdot s_0 + ... + w_T \cdot s_T$, where $\sum_{t=0}^{T} w_t = 1$.*

Proof: If $T = 1$, Lemma 1 is a direct application of Axiom 3. If $T \geq 2$, for any $2 \leq t \leq T$, there should exist $\alpha_t \in (0,1)$ such that $s_{0\to t} \sim \alpha_t \cdot s_{0\to t-1} + (1-\alpha_t) \cdot s_t$. By state-independence and reduction of compound alternatives, we can recursively apply this equivalence relation as follows:

$$\begin{aligned}
s_{0\to T} &\sim \alpha_{T-1} \cdot s_{0\to T-1} + (1 - \alpha_{T-1}) \cdot s_T \\
&\sim \alpha_{T-1}\alpha_{T-2} \cdot s_{0\to T-2} + \alpha_{T-1}(1 - \alpha_{T-2}) \cdot s_{T-1} + (1 - \alpha_{T-1}) \cdot s_T \\
&\sim ... \\
&\sim w_0 \cdot s_0 + w_1 \cdot s_1 + ... + w_T \cdot s_T
\end{aligned}$$

where $w_0 = \prod_{t=0}^{T-1} \alpha_t$, $w_T = 1 - \alpha_{T-1}$, and for $0 < t < T$, $w_t = (1 - \alpha_{t-1}) \prod_{\tau=t}^{T-1} \alpha_\tau$. It is easy to show the sum of $w_0, \ldots, w_T$ is equal to 1. *QED*.

Therefore, if Axiom 1 and 3 hold, for any reward sequence $s_{0\to T}$, we can always find a convex combination of all its elements, such that the DM is indifferent between the reward sequence and this convex combination. If $s_{0\to T}$ is a constant sequence, i.e. all its elements are constant, then we can directly assume $\mathcal{W}$ is AAD-style. Henceforth, we discuss whether AAD can also apply to non-constant sequences.

By Lemma 2, we show adding a new reward to the end of $s_{0\to T}$ has no impact on the relative decision weights of rewards in the original reward sequence.

Lemma 2: *For any $s_{0\to T+1}$, if $s_{0\to T} \sim \sum_{t=0}^{T} w_t \cdot s_t$ and $s_{0\to T+1} \sim \sum_{t=0}^{T+1} w'_t \cdot s_t$, where* $w_t, w'_t > 0$ and $\sum_{t=0}^{T} w_t = 1$, $\sum_{t=0}^{T+1} w'_t = 1$, *then when Axiom 1-4 hold, we can obtain*

$\frac{w_0'}{w_0} = \frac{w_1'}{w_1} = \ldots = \frac{w_T'}{w_T}$.

Proof: According to Axiom 3, for any $s_{0\to T+1}$, there exist $\alpha, \zeta \in (0,1)$ such that

$$\begin{aligned} s_{0\to T} &\sim \alpha \cdot s_{0\to T-1} + (1-\alpha)\cdot s_T \\ s_{0\to T+1} &\sim \zeta \cdot s_{0\to T} + (1-\zeta)\cdot s_{T+1} \end{aligned} \tag{A1}$$

On the other hand, we drawn on Lemma 1 and set

$$s_{0\to T+1} \sim \beta_0 \cdot s_{0\to T-1} + \beta_1 \cdot s_T + (1-\beta_0-\beta_1)\cdot s_{T+1} \tag{A2}$$

where $\beta_0, \beta_1 > 0$. According to Axiom 4, $1-\zeta = 1-\beta_0-\beta_1$. So, $\beta_1 = \zeta - \beta_0$. This also implies $\zeta > \beta_0$.

According to Axiom 2, we suppose there exists a reward sequence s such that $s \sim \frac{\beta_0}{\zeta} \cdot s_{0\to T-1} + (1-\frac{\beta_0}{\zeta})\cdot s_1$. By Equation (A2) and reduction of compound alternatives, we have $s_{0\to T+1} \sim \zeta \cdot s + (1-\zeta)\cdot s_{T+1}$. Combining Equation (A2) with the second line of Equation (A1) and applying transitivity and state-independence, we obtain $s_{0\to T} \sim \frac{\beta_0}{\zeta}\cdot s_{0\to T-1} + (1-\frac{\beta_0}{\zeta})\cdot s_1$.

We aim to prove that for any $s_{0\to T+1}$, we can obtain $\alpha = \frac{\beta_0}{\zeta}$. To do this, we first assume (without loss of generality) that $\alpha > \frac{\beta_0}{\zeta}$.

Consider the case that $s_{0\to T-1} \succ s_T$. By state-independence, for any $c \in \mathbb{R}_{\geq 0}$, we have $(\alpha - \frac{\beta_0}{\zeta})\cdot s_{0\to T-1} + (1-\alpha+\frac{\beta_0}{\zeta})\cdot c \succ (\alpha - \frac{\beta_0}{\zeta})\cdot s_T + (1-\alpha+\frac{\beta_0}{\zeta})\cdot c$. By Axiom 2, there exists $z \in \mathbb{R}_{\geq 0}$ such that $(1-\alpha)\cdot s_T + \frac{\beta_0}{\zeta}\cdot s_{0\to T-1} \sim z$. Given c is arbitrary, we set $(1-\alpha+\frac{\beta_0}{\zeta})\cdot c \sim z$. By reduction of compound alternatives, we can derive that

$$(\alpha - \frac{\beta_0}{\zeta})\cdot s_{0\to T-1} + (1-\alpha)\cdot s_T + \frac{\beta_0}{\zeta}\cdot s_{0\to T-1} \succ (\alpha - \frac{\beta_0}{\zeta})\cdot s_T + (1-\alpha)\cdot s_T + \frac{\beta_0}{\zeta}\cdot s_{0\to T-1}$$

where the LHS can be rearranged to $\alpha \cdot s_{0\to T-1} + (1-\alpha)\cdot s_T$, and the RHS can be rearranged to $\frac{\beta_0}{\zeta}\cdot s_{0\to T-1} + (1-\frac{\beta_0}{\zeta})\cdot s_1$. They both should be indifferent from $s_{0\to T}$. This results in a contradiction. Similarly, in the case that $s_T \succ s_{0\to T-1}$, we can also derive a contradiction. Meanwhile, when $s_{0\to T} \sim s_T$, α and $\frac{\beta_0}{\zeta}$ can be any number within $(0,1)$. So, we can directly set $\alpha = \frac{\beta_0}{\zeta}$.

Thus, we have $\alpha = \frac{\beta_0}{\zeta}$ for any $s_{0\to T+1}$, which indicates $\frac{\beta_0}{\alpha} = \frac{\beta_1}{1-\alpha} = \zeta$. We can recursively apply this equality to any sub-sequence $s_{0\to t}$ ($t \leq T$) of $s_{0\to T+1}$, so that the lemma will be proved. *QED.*

Now we move on to prove Proposition 1. The proof contains six steps.

First, we add the constraints $\sum_{t=0}^{T} w_t = 1$ and $w_t > 0$ to the optimal discounting problem for $s_{0\to T}$ so that the problem is compatible with Lemma 1. According to the FOC of its solution, for all $t = 0, 1, \ldots, T$, we have

$$f_t'(w_t) = u(s_t) + \theta \tag{A3}$$

where θ is the Lagrange multiplier. Given that $f_t'(w_t)$ is strictly increasing, w_t is increasing with $u(s_t) + \theta$. We define the solution as $w_t = \phi_t(u(s_t) + \theta)$.

Second, we add a new reward s_{T+1} to the end of $s_{0\to T}$ and apply Lemma 2 as a constraint to optimal discounting problem. Look at the optimal discounting problem for $s_{0\to T+1}$. For all $t \leq T$, the FOC of its solution will take the same form as Equation (A3). So, if importing s_{T+1} changes some w_t to w_t' ($w_t' \neq w_t$, where w_t is the solution to optimal discounting problem for $s_{0\to T}$), the only way is through changing the multiplier θ. Suppose importing s_{T+1} changes θ to $\theta - \Delta\theta$, we have $w_t' = \phi_t(u(s_t) + \theta - \Delta\theta)$.

By Lemma 2, we know $\frac{w_0}{w_0'} = \frac{w_1}{w_1'} = \ldots = \frac{w_T}{w_T'}$. In other words, for $t = 0, 1, \ldots, T$, we have $w_t \propto \phi_t(u(s_t) + \theta - \Delta\theta)$. We can rewrite w_t as

$$w_t = \frac{\phi_t(u(s_t) + \theta - \Delta\theta)}{\sum_{\tau=0}^{T} \phi_\tau(u(s_\tau) + \theta - \Delta\theta)} \tag{A4}$$

Third, we show that in $s_{0\to T}$, if we change each s_t to z_t such that $u(z_t) = u(s_t) + \Delta u$, the decision weights $w_0, \ldots, w_T$ will remain the same. Note $\sum_{t=0}^{T} \phi_t(u(s_t) + \theta) = 1$. It is clear that $\sum_{t=0}^{T} \phi_t(u(z_t) + \theta - \Delta u) = 1$. Suppose changing every s_t to z_t moves θ to θ' and $\theta' < \theta - \Delta u$. Then, we must have $\phi_t(u(z_t) + \theta') < \phi_t(u(z_t) + \theta - \Delta u)$ since $\phi_t(.)$ is strictly increasing. This results in $\sum_{t=0}^{T} \phi_t(u(z_t) + \theta') < 1$, which contradicts with the constraint that the sum of all decision weights is 1. The same contradiction can apply to the case that

$\theta' > \theta - \Delta u$. Therefore, changing every s_t to z_t must move θ to $\theta - \Delta u$, and each w_t can only be moved to $\phi_t(u(z_t) + \theta - \Delta u)$, which is exactly the same as the original decision weight.

A natural corollary of this step is that, subtracting or adding a common number to all intantaneous utilities in a reward sequence has no effect on decision weights. What actually matters for determining the decision weights is the difference between instantaneous utilities. This indicates, for convenience, we can subtract or add an arbitrary number to the utility function.

In other words, for a given $s_{0\to T}$ and s_{T+1}, we can define a new utility function $v(.)$ such that $v(s_t) = u(s_t) + \theta - \Delta\theta$. So, Equation (A4) can be re-written as

$$w_t = \frac{\phi_t(v(s_t))}{\sum_{\tau=0}^{T} \phi_\tau(v(s_\tau))} \tag{A5}$$

If w_t takes the AAD form under the utility function $v(.)$, i.e. $w_t \propto d_t e^{v(s_t)/\lambda}$, then it should also take the AAD form under the original utility function $u(.)$.

Fourth, we show that in Equation (A4), $\Delta\theta$ has two properties: (i) $\Delta\theta$ is strictly increasing with $u(s_{T+1})$; (ii) suppose $\Delta\theta = \underline{\theta}$ when $u(s_{T+1}) = \underline{u}$ and $\Delta\theta = \bar{\theta}$ when $u(s_{T+1}) = \bar{u}$, where $\underline{u} < \bar{u}$, then for any $l \in (\underline{\theta}, \bar{\theta})$, there exists $u(s_{T+1}) \in (\underline{u}, \bar{u})$ such that $\Delta\theta = l$.

The property (i) can be shown by contradiction. Given $w_0, \ldots, w_{T+1}$ a sequence of decision weights for $s_{0\to T+1}$. Suppose $u(s_{T+1})$ is increased but $\Delta\theta$ is constant. In this case, each of $w'_0, \ldots, w'_T$ should also be constant. However, w'_{T+1} must increase as it is strictly increasing with $u(s_{T+1}) + \theta - \Delta\theta$ (θ is determined by the optimal discounting problem for $s_{0\to T}$; thus, any operations on s_{T+1} should have no effect on θ). This contradicts with the constraint that $\sum_{t=0}^{T+1} w'_t = 1$. The only way to avoid such contradictions is to set $\Delta\theta$ strictly increasing with s_{T+1}, so that $w'_0, \ldots, w'_T$ are decreasing with $u(s_{T+1})$.

For property (ii), note that for any reward sequence $s_{0\to T+1}$ and a given θ, $\Delta\theta$ is defined as the solution to $\sum_{t=0}^{T+1} \phi_t(u(s_t) + \theta - \Delta\theta) = 1$. Given an arbitrary number $l \in (\underline{\theta}, \bar{\theta})$, the proof of property (ii) consists of two stages. First, for $t = 0, 1, \ldots, T$, we need to show that $u(s_t) + \theta - l$ is still in the domain of $\phi_t(.)$. Second, for period $T + 1$, we need to show for any $\omega \in (0, 1)$, there exists $u(s_{T+1}) \in \mathbb{R}$ such that $\phi_{T+1}(u(s_{T+1}) + \theta - l) = \omega$.

For the first stage, note $\phi_t(.)$ is the inverse function of $f_t'(.)$. Suppose when $\Delta\theta = \bar{\theta}$, we have $f_t'(w_t^a) = u(s_t) + \theta - \bar{\theta}$, and when $\Delta\theta = \underline{\theta}$, we have $f_t'(w_t^b) = u(s_t) + \theta - \underline{\theta}$. For any $l \in (\underline{\theta}, \bar{\theta})$, we have $u(s_t) + \theta - l \in (f_t'(w_t^a), f_t'(w_t^b))$. Given that $f_t'(.)$ is continuous and strictly increasing, there must be $w_t \in (w_t^a, w_t^b)$ such that $f_t'(w_t) = u(s_t) + \theta - l$. So, $u(s_t) + \theta - l$ is in the domain of $\theta_t(.)$. For the second stage, given an arbitrary $\omega \in (0, 1)$, we can set $u(s_{T+1}) = f'(\omega) - \theta + l$, so that the desired condition is satisfied.

A corollary of this step is that we can manipulate $\Delta\theta$ in Equation (A4) at any level in $[\underline{\theta}, \bar{\theta}]$ by changing a hypothetical s_{T+1}.

Fifth, we show $\ln \phi_t(.)$ is linear under some condition. To do this, let us add a hypothetical s_{T+1} to the end of s_T and let $w_t' = \phi_t(v(s_t))$ denote the decision weights for $t = 0, 1, \ldots, T+1$ in $s_{0\to T+1}$. We change the hypothetical s_{T+1} within the set $\{s_{T+1} | v(s_{T+1}) \in [\underline{v}, \bar{v}]\}$ and see what will happen to the decision weights from period 0 to period T. Suppose this changes each w_t' to $\phi_t(v(s_t) - \eta)$. Set $\eta = \underline{\eta}$ when $u(s_{T+1}) = \underline{v}$ and $\eta = \bar{\eta}$ when $u(s_{T+1}) = \bar{v}$. By Equation (A5), we have

$$\frac{\phi_t(v(s_t))}{\sum_{\tau=0}^{T} \phi_\tau(v(s_\tau))} = \frac{\phi_t(v(s_t) - \eta)}{\sum_{\tau=0}^{T} \phi_\tau(v(s_\tau) - \eta)} \tag{A6}$$

For each $t = 0, 1, ..., T$, we can rewrite $\phi_t(v(s_t))$ as $e^{\ln \phi_t(v(s_t))}$. For the LHS of Equation (A6), multiplying both the numerator and the denominator by a same number will not affect the value. Therefore, Equation (A6) can be rewritten as

$$\frac{e^{\ln \phi_t(v(s_t)) - \kappa\eta}}{\sum_{\tau=0}^{T} e^{\ln \phi_\tau(v(s_\tau)) - \kappa\eta}} = \frac{e^{\ln \phi_t(v(s_t) - \eta)}}{\sum_{\tau=0}^{T} e^{\ln \phi_\tau(v(s_\tau) - \eta)}}$$

where κ can be any real-valued constant. By properly selecting κ, we for all $t = 0, 1, ..., T$, can obtain

$$\ln \phi_t(v(s_t)) - \kappa\eta = \ln \phi_t(v(s_t) - \eta) \tag{A7}$$

as long as $\eta \in [\underline{\eta}, \bar{\eta}]$. And given $\ln \phi_t(.)$ is strictly increasing, for any $\eta \neq 0$, we have $\kappa > 0$.

Finally, we denote the maximum and minimum of $\{v(s_t)\}_{t=0}^{T}$ by $v_{\max}$ and $v_{\min}$, and show that Equation (A7) can hold if $\eta = v_{\max} - v_{\min}$. That is, $v_{\max} - v_{\min} \in [\underline{\eta}, \bar{\eta}]$, where $\underline{\eta}, \bar{\eta}$ are the realizations of η when $v(s_{T+1}) = \underline{v}$ and $v(s_{T+1}) = \bar{v}$. Obviously, $\underline{\eta}$ can take the value $\underline{\eta} = 0$. Thus, we focus on whether $\bar{\eta}$ can take a value $\bar{\eta} \geq v_{\max} - v_{\min}$.

The proof is similar with the fourth step and consists of two stages. First, we show that for $t = 0, 1, \ldots, T$, $v(s_t) - v_{\max} + v_{\min}$ is in the domain of $\phi_t(.)$. That is, under some w_t, we have $f_t'(w_t) = v(s_t) - v_{\max} + v_{\min}$. Note $v_{\max} - v_{\min} \in [0, +\infty)$. On the one hand, there exists $w_t \in (0,1)$ for $f_t'(w_t) = v(s_t)$, which is the solution to Equation (A5). On the other hand, by Definition 2, we know $\lim_{w_t \to 0} f_t'(w_t) = -\infty$. Given $f_t'(w_t)$ is continuous and strictly increasing, there must be a solution w_t for $f_t'(w_t) = v(s_t) - v_{\max} + v_{\min}$. Second, we show that for any $\omega \in (0,1)$, there exists some $v(s_{T+1})$ such that $\phi_{T+1}(v(s_{T+1}) - v_{\max} + v_{\min}) = \omega$. This can be achieved by setting $v(s_{T+1}) = f_{T+1}'(\omega) + v_{\max} - v_{\min}$.

As a result, for any period t in $s_{0\to T}$, by Equation (A7), we have $\ln \phi_t(v(s_t)) = \ln \phi_t(v(s_t) - \eta) + \kappa\eta$ as long as $\eta \in [0, v_{\max} - v_{\min}]$, where $\kappa > 0$. We can rewrite each $\ln \phi_t(v(s_t))$ as $\ln \phi_t(v_{\min}) + \kappa(v(s_t) - v_{\min})$. Therefore, we have

$$w_t \propto \phi_t(v_{\min}) \cdot e^{\kappa(v(s_t) - v_{\min})} \tag{A8}$$

and $\sum_{t=0}^{T} w_t = 1$. In Equation (A8), setting $\phi_t(v_{\min}) = d_t$, $\lambda = 1/\kappa$, and apply the corollary of the third step, we can conclude that $w_t \propto d_t e^{u(s_t)/\lambda}$, which is of the AAD form.

B. Proof of Proposition 2

Note the instantaneous utilities of LL and SS are v_l and v_s, and the delays for LL and SS are t_l and t_s. According to Equation (3), the common difference effect implies that, if

$$\frac{v_s}{1 + G(t_s)e^{-v_s}} = \frac{v_l}{1 + G(t_l)e^{-v_l}} \tag{B1}$$

then for any $\Delta t \geq 0$, we have

$$\frac{v_s}{1 + G(t_s + \Delta t)e^{-v_s}} < \frac{v_l}{1 + G(t_l + \Delta t)e^{-v_l}} \tag{B2}$$

If $G(T) = T$, we have $G(t + \Delta t) = G(t) + \Delta t$. In this case, combining Equation (B1) and (B2), we can obtain

$$\frac{\Delta t e^{-v_s}}{v_s} > \frac{\Delta t e^{-v_l}}{v_l} \tag{B3}$$

Given that function $\psi(v) = e^{-v}/v$ is decreasing with v so long as $v > 0$, Equation (B3) is valid.

If $G(T) = \frac{1}{1-\delta}(\delta^{-T} - 1)$, we have

$$1 + G(t + \Delta t)e^{-v} = \delta^{-\Delta t}[1 + G(t)e^{-v}] + (\delta^{-\Delta t} - 1)(\frac{e^{-v}}{1-\delta} - 1)$$

Thus, combining Equation (B1) and (B2), we can obtain

$$(\delta^{-\Delta t} - 1)\frac{\frac{e^{-v_s}}{1-\delta} - 1}{v_s} > (\delta^{-\Delta t} - 1)\frac{\frac{e^{-v_l}}{1-\delta} - 1}{v_l} \tag{B4}$$

Given that $0 < \delta < 1$, we have $\delta^{-\Delta t} > 1$. So, Equation (B4) is valid if and only if

$$\frac{1}{v_s} - \frac{1}{v_l} < \frac{1}{1-\delta}(\frac{e^{-v_s}}{v_s} - \frac{e^{-v_l}}{v_l}) \tag{B5}$$

By Equation (B1), we know that

$$\frac{1}{v_s} - \frac{1}{v_l} = \frac{1}{1-\delta}\left[\frac{(\delta^{-t_l} - 1)e^{-v_l}}{v_l} - \frac{(\delta^{-t_s} - 1)e^{-v_s}}{v_s}\right] \tag{B6}$$

Combining Equation (B5) and (B6), we have

$$\delta^{-t_l}\frac{e^{-v_l}}{v_l} < \delta^{-t_s}\frac{e^{-v_s}}{v_s} \Longleftrightarrow v_l - v_s + \ln\left(\frac{v_l}{v_s}\right) > -(t_l - t_s)\ln\delta$$

C. Proof of Proposition 3

Suppose a positve reward x is delivered at period T. By Equation (3), if w_T is convex in T, we should have $\frac{\partial^2 w_T}{\partial T^2} \geq 0$. This implies

$$2G'(T)^2 \geq (G(T) + e^{v(x)})G''(T) \tag{C1}$$

If $\delta = 1$, then $G(T) = T$. We have $G'(T) = 1$, $G''(T) = 0$. Thus, Equation (C1) is always valid.

If $0 < \delta < 1$, then $G(T) = (1-\delta)^{-1}(\delta^{-T} - 1)$. We have $G'(T) = (1-\delta)^{-1}(-\ln\delta)\delta^{-T}$, $G''(t) = (-\ln\delta)G'(T)$. Thus, Equation (C1) is valid when

$$\delta^{-T} \geq (1-\delta)e^{v(x)} - 1 \tag{C2}$$

Given $T > 0$, Equation (C2) holds true in two cases. The first case is $1 \geq (1-\delta)e^{v(x)} - 1$, which implies that $v(x)$ is no greater than a certain threshold $v(\underline{x})$, where $v(\underline{x}) = \ln(\frac{2}{1-\delta})$. The second case is that $v(x)$ is above $v(\underline{x})$ and T is above a threshold $\underline{t}$. In the second case, we can take the logarithm on both sides of Equation (C2). It yields $\underline{t} = \frac{\ln[(1-\delta)\exp\{v(x)\}-1]}{\ln(1/\delta)}$.

D. Proof of Proposition 4

For convenience, we use v to represent $v(x) = u(x)/\lambda$, and use U represent $U(x,T)$. Set $g = G(T)$. The first-order derivative of U with respect to x can be written as

$$\frac{\partial U}{\partial x} = v' \frac{e^v + U}{e^v + g} \tag{D1}$$

If U is strictly concave in x, we should have $\frac{\partial^2 U}{\partial x^2} < 0$. By Equation (D1), we calculate the second-order derivative of U with respect to x, and rearrange this second-order condition to

$$2\zeta(v) + \frac{1}{1 + v\zeta(v)} - 1 < \frac{-v''}{(v')^2} \equiv \frac{d}{dx}\left(\frac{1}{v'}\right) \tag{D2}$$

where $\zeta(v) = g/(g + e^v)$. Since $v'' < 0$, the RHS of Equation (D2) is clearly positive.

To prove the first part of Proposition 4, we can show that when x is large enough, the LHS of Equation (D2) will be non-positive. To make the LHS non-positive, we require

$$\zeta(v) + \frac{1}{v} \leq \frac{1}{2} \tag{D3}$$

Note that $\zeta(v)$ is decreasing in v, and v is increasing in x. Hence, $\zeta(v) + \frac{1}{v}$ is decreasing in x. Besides, when $x \to 0$, it approaches to $+\infty$, and when $x \to +\infty$, it approaches to 0. When $\frac{d}{dx}\left(\frac{1}{v'(x)}\right)$ is continuous, there must be a unique realization of x in $(0, +\infty)$, say $\bar{x}$, making the equality in Equation (D3) hold. When $x \geq \bar{x}$, Equation (D3) is always valid, and therefore, $U(x, T)$ is concave in x.

To prove the second part, first note that when $x = 0$, the LHS of Equation (D2) will become $\frac{2g}{g+1}$. If $\frac{d}{dx}\left(\frac{1}{v'(0)}\right)$ is smaller than this value, then the LHS of Equation (D2) should be greater than the RHS at the point of $x = 0$. From the first part of proposition 4, we know the LHS is smaller than the RHS at the point of $x = \bar{x}$. Thus, given $\frac{d}{dx}\left(\frac{1}{v'(x)}\right)$ is continuous in $[0, \bar{x}]$, there must also be a point within $[0, \bar{x}]$, such that the LHS equals the RHS. Let x^* denote the minimum of x that makes the equality valid. Then, for any $x \in (0, x^*)$, we must have that the LHS of Equation (D2) is greater than the RHS, which implies $U(x, T)$ is convex in x. Given that $T \geq 1$, we have $g \geq 1$ and thus $\frac{2g}{g+1} \geq 1$. Thus, when $\frac{d}{dx}\left(\frac{1}{v'(0)}\right) < 1$, $U(x, t)$ can be convex in x for any $x \in (0, x^*)$, regardless of g.

The prove the third part, note that $v(x) = u(x)/\lambda$. So,

$$\frac{d}{dx}\left(\frac{1}{v'}\right) = \lambda \frac{d}{dx}\left(\frac{1}{u'}\right)$$

We arbitrarily draw a point from $(0, \bar{x})$ and derive the range λ relative to this point. For simplicity, we choose $x = \ln g$. In this case, the LHS of Equation (D2) becomes $\frac{2}{2+\ln g}$. Define a function $\xi(x)$, where ξ is the value of the LHS of Equation (D2) minus its RHS. Note $\xi(x)$ is continuous at $x = \ln g$. Therefore, for any positive real number b, there must exist a positive real number c such that, when $x \in (\ln g - c, \ln g + c)$, we have

$$\xi(\ln g) - b < \xi(x) < \xi(\ln g) + b \tag{D4}$$

If $\xi(\ln g) - b \geq 0$, then $\xi(x)$ will keep positive for all $x \in (\ln g - c, \ln g + c)$, which implies the LHS of Equation (D2) is always greater than its RHS.

Finally, we derive the condition for $\xi(\ln g) - b \geq 0$. Suppose when $x = \ln g$, $\frac{d}{dx}\left(\frac{1}{u'}\right) = a$ (note at this point we have $\frac{d}{dx}\left(\frac{1}{u'}\right) < +\infty$). Combining with Equation (D3), we know that

$\xi(\ln g) - b = \frac{2}{2+\ln g} - \lambda a - b$. Letting this value be non-negative, we obtain

$$\lambda \leq \frac{2}{a(2+\ln g)} - \frac{b}{a} \tag{D5}$$

Given $T \geq 1$, we have $g \geq 1$ and thus $\frac{2}{2+\ln g}$ should be positive. Meanwhile, given that $u' > 0$ and $u'' < 0$, a should also be positive. Since b can be any positive number, Equation (D5) holds if $\lambda < \frac{2}{a(2+\ln g)}$. That is, when λ is positive but smaller than a certain threshold, there must be an interval $(\ln g - c, \ln g + c)$ such that the LHS of Equation (D2) is greater than the RHS. Set $x_1 = \max\{0, \ln g - c\}$, $x_2 = \min\{\bar{x}, \ln g + c\}$. When $x \in (x_1, x_2)$, function $U(x, T)$ must be convex in x.

E. Proof of Proposition 5

The proof consists of four steps. First, we write the expressions for $U(L1)$ and $U(L2)$. Given a lottery result (s_1, s_2), we denote the decision weight of each positive reward by w_{t_1} and w_{t_2}. Suppose the time length of each lottery result is T. For a period τ at which no reward is delivered, the utility is zero. Let Ω denote the set of all such period τ, then $\Omega = \{\tau \,|\, 0 \leq \tau \leq T,\ \tau \neq t_1, t_2\}$. For all $j, k \in \{s, l\}$, we define $\phi_j = d_{t_1} e^{v(x_j)}$ and $\eta_k = d_{t_2} e^{v(y_k)}$, where $v(s) = u(s)/\lambda$, and d_t represents the reference weight for any reward delivered at period t. By the definition of AAD, we have

$$w_{t_1} = \frac{\phi_j}{\phi_j + \eta_k + D} \quad , \quad w_{t_2} = \frac{\eta_k}{\phi_j + \eta_k + D}$$

where $j, k \in \{s, l\}$, $D = \sum_{\tau \in \Omega} d_\tau \geq 0$. The value of a lottery L can be written as $U(L) = w_{t_1} u(s_1) + w_{t_2} u(s_2)$. Hence,

$$\begin{aligned} U(L1) &= 0.5\frac{\phi_s u(x_s) + \eta_s u(y_s)}{\phi_s + \eta_s + D} + 0.5\frac{\phi_l u(x_l) + \eta_l u(y_l)}{\phi_l + \eta_l + D} \\ U(L2) &= 0.5\frac{\phi_s u(x_s) + \eta_l u(y_l)}{\phi_s + \eta_l + D} + 0.5\frac{\phi_l u(x_l) + \eta_s u(y_s)}{\phi_l + \eta_s + D} \end{aligned} \tag{E1}$$

We observe that, when $x_l = x_s$, we have $U(L1) = U(L2)$.

Second, suppose we increase x_l from x_s by an increment. This increases both $U(L1)$ and $U(L2)$ (either by a positive or a negative number). To make $U(L1) < U(L2)$, this increment should increase $U(L2)$ by a greater number than $U(L1)$. Specifically, we assume $U(L2)$ is increasing faster than $U(L1)$ at any level of x_l. That is, the partial derivative of $U(L2)$ in terms of x_l is always greater than that of $U(L1)$. Given ϕ_l is increasing in x_l, to see this, we can take partial derivatives in terms of ϕ_l.

In each line of Equation (E1), note only the second term contains x_l. Thus, we focus on the difference between the the second terms. The second term of the $U(L1)$ is influenced by y_l, while that of the $U(L2)$ is influenced by y_s, where $y_l > y_s$. Thus, we can construct a function ξ such that

$$\xi(\phi_l, \eta) = \frac{\phi_l \cdot v(x_l) + \eta \cdot v(y)}{\phi_l + \eta + D}$$

where $\eta = d_{t_2} e^{v(y)}$. In reverse, we can define $v(x_l) = \ln(\phi_l/d_{t_1})$ and $v(y) = \ln(\eta/d_{t_2})$. The function ξ is similar to the second term of each line, but note we replace $u(.)$ by $v(.)$. When $y = y_l$, ξ is proportional to the second term of $U(L1)$. When $y = y_s$, ξ is proportional to the second term of $U(L2)$ (by the same proportion). Thus, to show that the partial derivative of $U(L2)$ in terms of x_l is greater than that of $U(L1)$, we just need to show $\partial\xi/\partial\phi_l$ is decreasing with y (or η).

Third, we take the first- and second-order partial derivatives of $\xi(\phi_l, \eta)$ in terms of ϕ_l. The partial derivative of ξ in terms of ϕ_l is

$$\frac{\partial\xi}{\partial\phi_l} = \frac{(v(x_l) + 1)\eta - v(y)\eta + \phi_l + D(v(x_l) + 1)}{(\phi_l + \eta + D)^2}$$

We need to show that for $y \in [y_s, y_l]$, we can obtain $\partial^2\xi/\partial\phi_l\partial\eta < 0$. This implies that

$$(v(x_l) + v(y) + 2)D - (\phi_l - \eta)(v(x_l) - v(y)) + 2(\phi_l + \eta) > 0 \tag{E2}$$

We want Equation (E2) to hold for any $D > 0$. Given the LHS is increasing with D, we have

$$2(\phi_l + \eta) > (\phi_l - \eta)(v(x_l) - v(y)) \tag{E3}$$

Define $\kappa = d_{t_2}/d_{t_1}$, $\alpha = v(x_l) - v(y)$. Note $\kappa \in \mathbb{R}_{>0}$, $\alpha \in \mathbb{R}$. Equation (E3) can be rewritten as

$$(\alpha - 2)\kappa^{-1}e^{\alpha} - \alpha - 2 < 0 \tag{E4}$$

Fourth, based on Equation (E4), we construct a function $h(\alpha) = (\alpha - 2)\kappa^{-1}e^{\alpha} - \alpha - 2$. We aim to examine whether there exists some α that can make $h(a) < 0$. Obviously, $\alpha = -2$ and $\alpha = 2$ satisfy this condition. Moreover, note $h(\alpha)$ is decreasing in α when $(\alpha - 1)e^{\alpha} \leq \kappa$ and is increasing in α otherwise. When either $\alpha \to -\infty$ or $\alpha \to +\infty$, $h(\alpha) \to +\infty$. Thus, there must be a limited intertval (α_1, α_2) such that $h(a) < 0$ so long as $\alpha \in (\alpha_1, \alpha_2)$, and $[-2, 2] \subset (\alpha_1, \alpha_2)$.

For a given positive number κ, the points α_1, α_2 are determined by the solution to $\frac{\alpha-2}{\alpha+2}e^{\alpha} = \kappa$. Since $v(s) = u(s)/\lambda$, for any x_l and $y \in [y_s, y_l]$, we have $\frac{u(x_l)-u(y)}{\lambda} \in (\alpha_1, \alpha_2)$. In other words, for any $x_l > x_s > 0$, $y_l > y_s > 0$, any time length of lottery results and reference weights (which determining D and κ), there exists some λ that makes $U(L1) < U(L2)$. Specifically, all $\lambda > \lambda^{**} = \max\{\frac{u(x_l)-u(y_l)}{\alpha_1}, \frac{u(x_l)-u(y_s)}{\alpha_2}\}$ satisfy the target condition.

Notably, if $\lambda \leq \lambda^{**}$, we have $h(a) > 0$, which by Equation (E2)(E3), indicates that under some conditions such as $D = 0$, there will be $\partial^2\xi/\partial\phi_l\partial\eta \geq 0$ for all $y \in [y_s, y_l]$. In that case, at each level of x_l, the partial derivative of $U(L1)$ in terms of x_l is greater than $U(L2)$. Increasing x_l by an increment may induce a greater increase in $U(L1)$ than in $U(L2)$. So, the DM may perform intertemporal correlation seeking.