Progress Report for Annual Review 22/23 - v2

Zark Zijian Wang

August 05, 2023

1 Introduction

In this document, I introduce three pieces of work.

The first is a model I term "attentional discounted utility". I leverage attentional mechanism to explain the anomalies in intertemporal choice. The model is developed based on expected discounted utility framework.

The second is an experimental test on whether (and how) attentional mechanism can influence intertemporal choice.

The third is an experiment aiming to test how people valuate risky outcomes and efforts. One prominent application of the findings from this experiment is that to help us validate a key hypothesis in rational inattention theories. I have designed the experimental program.

2 Attentional Discounted Utility

2.1 The Model

Expected discounted utility framework has been widely used in behavioral and economic research. Given a sequence of rewards $X_T = [x_0, x_1, ..., x_T]$, which yields reward x_t in time period t,[1] $t \in \{0, 1, ..., T\}$, the expected discounted utility of X_T is then calculated by

$$EDU(X_T) = \mathbb{E}\left[\sum_{t=0}^{T} d_t u(x_t)\right]$$

where d_t is the discounting factor of time period t, reward level x_t is a random variable defined on $\mathbb{R}_{\geq 0}$, $u(.)$ denote the decision maker's instantaneous utility function, where $u' > 0$ and $u'' < 0$. The time length of this sequence, denoted by T, is finite.

I define a class of models that incorporates attentional mechanism into such a valuation process of sequential rewards. In such models, a decision maker allocates attention across sequential outcomes when processing information about rewards, and attends more to the outcomes with larger rewards. I refer to this class of models as *Attentional Discounted Utility* (ADU). Previous studies exploring the ADU setting include the works by Gershman and Bhui (2020) and Noor and Takeoka (2022). Specifically, I focus on a particular variant of ADU, which I term as *ADU with Shannon cost function* (ADUS). ADUS retains the architecture of expected discounted utility, but modifies the conventional discounting factor d_t to w_t, which I refer to as attention weight hereafter, where

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

In Equation (1), w_t adopts a form resembling logistic function. It indicates three properties of attention. First, the sum of all w_t equals 1, which indicates the attention available to be allocated across sequential outcomes is limited. Second, w_t is increasing in x_t, indicating the decision maker is naturally inclined to allocate more attention to periods with larger potential rewards. This is in line with existing evidence, suggesting people often pay much attention to pleasant information and avoid to know unpleasant information (e.g. ostrich effect). Third, it appears that w_t is anchored in d_t. This indicates that d_t represents the original attention weight assigned to each period, and the attention adjustment process (converting d_t to w_t) is costly. The parameter λ quantifies the cost of attention adjustment.

[1] I use uppercase letters to represent a sequence and lowercase letters to represent each element within the sequence.

I provide two rationales for Equation (1). The first is based on rational inattention theories. The second is an axiomatic theory. I present the first here and discuss the second in Section 2.3.

The necessary notations and definitions are described as follows. Let $S_T = [s_0, s_1, ..., s_T]$ be a potential realization of X_T, and $\mathcal{S}(X_T)$ be the support of X_T, i.e. the smallest set containing any potentially realized sequence S_T, where $\mathcal{S}(X_T) \subseteq \mathbb{R}^{T+1}_{\geq 0}$. In each S_T, the attention weight assigned to period t is decided by a weight function $w(s_t)$. I follow Noor and Takeoka (2022; 2023) to define the process that generates $w(s_t)$ as *constrained optimal discounting*. The key assumption of constrained optimal discounting process is that, when evaluating a reward sequence, the decision maker wants to find an allocation policy for attention weights, in order to maximize the anticipatory utility that she can obtain from the sequence. However, this attention re-allocation process triggers a cognitive cost. She needs to balance the benefit of focusing on periods containing larger rewards and the cognitive cost of shifting attention. The formal definition is given by Definition 1.

Definition 1: Let W be the support of all possible weight functions. Given a stochastic reward sequence X_T, the following optimization problem is called the *constrained optimal discounting* problem for X_T:

$$\begin{aligned}
\max_{w \in W} \quad & \sum_{S_T \in \mathcal{S}(X_T)} \sum_{t=0}^{T} w(s_t)u(s_t) - C(w) \\
s.t. \quad & \sum_{S_T \in \mathcal{S}} \sum_{t=0}^{T} w(s_t) = m \\
& w(s_t) \geq 0, \forall t \in \{0, 1, \dots, T\}
\end{aligned}$$

where m is a constant, $C : [0, m]^{T+1} \rightarrow \mathbb{R}_{>0}$ is called *information cost* function, $\partial C / \partial w(s_t) > 0$ and $\partial^2 C / \partial w(s_t)^2 > 0$. That is, the information cost is increasing and convex in $w(s_t)$.

In Definition 1, the cognitive cost of shifting attention is decided by an information cost function $C(.)$. This implies that the cognitive cost is generated by focusing on certain reward information. A well-known specification of $C(.)$ is the Shannon cost function, proposed by Matějka and McKay (2015). The Shannon cost function was originally used to justify the multinominal logit model in discrete choice analysis, and so far has been topical in rational inattention literature. To construct this style of information cost function, Matějka and McKay (2015) introduce three assumptions. The first is that the sum of all weights $m \equiv 1$. The second assumption is, before acquiring any information, the decision maker establishes an initial weight allocation for different attributes (outcomes), which remains invariant across states. The weights are then updated in a manner consistent with Bayes rule. In ADU setting, it means $d_t = \sum_{S_T \in \mathcal{S}(X_T)} w(s_t)$. The third assumption is, the information cost is linear to the information gains, measured by Shannon mutual information. That is,

$$C(w) = \lambda \sum_{S_T \in \mathcal{S}(X_T)} \sum_{t=0}^{T} w(s_t) \log \left(\frac{w_t(S_T)}{d_t} \right)$$

where λ is a parameter denoting unit cost of information ($\lambda > 0$). With this Shannon cost function, the constrained optimal discounting problem can be easily solved by Lagrangian method, and the solution is the same as Equation (1).

2.2 Implications in Intertemporal Choice

Here I document six implications of the ADUS model in intertemporal choice. Each of them can be precisely proved.

1. The model is consistent with "hidden zero effect".
2. The requirement of magnitude effect on curvature of utility function can be relaxed under ADUS.
3. The model specifies a novel condition for common difference effect.
4. The model suggests that the discount function may be concave for the near future and convex for the far future.
5. The model offers an alternative account for S-shaped value function.
6. In dynamic budget allocation, decision makers who exihibit greater patience at the beginning will perform less inconsistent behaviors in subsequent periods.

2.3 Axiomatic Characterization

2.4 The Role of Attention Updating in Inconsistent Planning

2.5 Empirical Analysis

3 Limited and Motivated Attention over Sequential Outcomes

4 Valuation of Risk and Effort

5 Time Schedule

Reply to Comments

Gershman, S. J. and Bhui, R. (2020). Rationally inattentive intertemporal choice. *Nature Communications*, 11(1):3365. Number: 1 Publisher: Nature Publishing Group.

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.

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

Noor, J. and Takeoka, N. (2023). Constrained Optimal Discounting.