Analytically Tractable Models for Decision Making under Present Bias

Yasunori Akagi[1], **Naoki Marumo**[2], **Takeshi Kurashima**[1]

[1]NTT Human Informatics Labratories, NTT Corporation,
[2]NTT Communication Science Laboratories, NTT Corporation
{yasunori.akagi, naoki.marumo, takeshi.kurashima}@ntt.com

Abstract

Time-inconsistency is a characteristic of human behavior in which people plan for long-term benefits but take actions that differ from the plan due to conflicts with short-term benefits. Such time-inconsistent behavior is believed to be caused by present bias, a tendency to overestimate immediate rewards and underestimate future rewards. It is essential in the research community of behavior economics to investigate the relationship between present bias and time-inconsistency. In this paper, we propose a model for analyzing agent behavior with present bias in tasks to make progress toward a goal over a period. Unlike previous models, the state sequence of the agent can be described analytically in our model. Based on this property, we analyze three crucial problems related to agents under present bias: task abandonment, optimal goal setting, and optimal reward scheduling. Extensive analysis reveals how present bias affects the condition under which task abandonment occurs and optimal intervention strategies. Our findings are meaningful for preventing task abandonment and intervening through incentives in the real world.

1 Introduction

People often do not achieve their goals because they change their plans in the middle of a task, even when nothing unexpected happens. For instance, some people may plan to stick to a diet for a month but end up indulging on the weekend. Similarly, a student may plan to do their assignments every day during vacation but procrastinate until the last day. Such behavior is known as *time-inconsistency*, which is a topic of active research in behavioral economics.

Time-inconsistent behavior is often caused by *present bias* (Frederick, Loewenstein, and O'donoghue 2002), a tendency to overestimate the value of immediate rewards and underestimate the value of future rewards. As an example, when planning a diet, the individual may not give much thought to treats on weekends, believing they can resist the temptation. However, when the weekend arrives, the desire for those treats becomes stronger, and the person may find it difficult to stick to their diet plan.

Researchers have been utilizing mathematical models to study the effects of present bias on human behavior, and recently Kleinberg and Oren (2014) proposed a new model which combines previous models. The Kleinberg–Oren (KO) model is based on graph theory and simulates the time-inconsistent behavior of a person affected by present bias with an agent moving on a graph. The model represents a task with a directed acyclic graph whose vertices correspond to the agent's states. Each edge has a cost for moving between vertices, and the goal vertex has a reward. On each vertex, the agent evaluates the value of each path to the goal and approaches the goal by following the path that seems most valuable to it. The value of each path is evaluated with a particular discounting scheme called quasi-hyperbolic discounting (Laibson 1997), which introduces present bias into the agent. The KO model effectively reproduces typical time-inconsistencies in behavioral economics, such as procrastination, task abandonment, and choice reduction. The model is also highly expressive and can represent various real-world tasks. We could use this model to predict undesirable future outcomes or determine interventions for agents to achieve better outcomes.

However, the high flexibility of the model also makes it challenging to analyze its properties. The agent's future behavior in the KO model cannot be derived in a closed form; simulation is necessary to determine the agent's behavior. Additionally, it is difficult to identify optimal interventions for guiding the agent to the goal. Tang et al. (2017) and Albers and Kraft (2019) considered intervention by adding intermediate rewards, while Albers and Kraft (2021) increased the cost of edges. It is shown in both settings that finding the optimal intervention is NP-hard, which demonstrates the computational difficulty of the KO model.

We introduce a model that is easy to analyze and compute while still being able to handle typical tasks in real life based on the KO model. We limit tasks to increasing a real number, which we refer to as *progress*, over a period. This type of task is commonly encountered in daily life; for example, completing a graduation thesis within six months (see the beginning of Section 3 for more examples). Our model shares the same assumptions on agent behavior as that of the KO model but differs in three ways.

- The agent's state is represented as a pair of time index and progress rather than as a vertex of a graph.
- The agent can take (uncountably) infinite states as we regard progress as a real number rather than as an integer.

- The cost of an agent's action is expressed as $c(\Delta)$ with a function $c : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, where Δ denotes the change in the progress between two adjacent time steps.

The main benefit of this model is that it simplifies its theoretical treatment. We demonstrate that for a specific class of c, the trajectory of states taken by an agent can be described analytically. This tractability is a significant advantage not found in existing models and allows us to perform further theoretical analyses, including finding optimal interventions. Based on such analytical descriptions, we analyze three problems related to agents under present bias.

The first issue we examine is *task abandonment*, which refers to when a person starts a long-term project but gives up on it before completion, despite no change in the costs or rewards associated with the project. We theoretically analyze the conditions that lead to task abandonment in our model. Our analysis shows that there is a threshold for the strength of present bias, which plays an essential role in the occurrence of task abandonment. Task abandonment never occurs when the present bias is weaker than the threshold. We further analyze the asymptotic behavior of the threshold as the period's length grows, obtaining analytical expressions for it. These results provide new insights into the conditions under which task abandonment occurs in real-world tasks.

The second problem is *goal optimization*; the goal has been fixed thus far, but from here, we control the goal to intervene in the agent's behavior. We consider the problem of setting the goal to maximize the progress achieved by the agent, given a period and reward amount. The crucial aspect in solving this problem is whether to allow exploitative rewards, i.e., rewards that influence the agent's behavior but are not claimed because the agent cannot satisfy the conditions for receiving the reward. Our analysis reveals the structure of the optimal solution in both cases of allowing and not allowing exploitative rewards. It shows that exploitative rewards can increase the progress of agents with strong present bias. This result suggests that people with strong present bias are easily controlled and deceived by exploitative rewards.

The third problem is *reward scheduling*, a more advanced intervention for the agent. Given a total period and reward budget, the goal is to maximize final progress by appropriately splitting the total period and setting goals and rewards for each period. We propose an efficient algorithm for finding the optimal solution and derive analytical optimal solutions for special cases. Our analysis shows that the optimal strategy depends on the strength of present bias. For agents with weak present bias that can adequately evaluate rewards and costs, offering rewards all at once is best. In contrast, frequent intermediate rewards increase progress for agents with strong present bias. This result indicates that we should vary the reward scheduling plan depending on the strength of the agents' present bias.

Note that proofs are deferred to Section 8.

2 Related Work

The relationship between present bias and time-inconsistent behavior has been a central topic in behavioral economics for many years, both experimentally and theoretically (Frederick, Loewenstein, and O'donoghue 2002; Camerer and Loewenstein 2004; Wilkinson and Klaes 2017). In these studies, a discounting scheme for future values plays a crucial role. Classical economics used exponential discounting (Samuelson 1937), which discounts value at a constant rate but can only lead to time-consistent behavior. To resolve this issue, hyperbolic discounting, in which the discount rate decreases with time, was proposed and succeeded in explaining the time-inconsistent behavior of people (Ainslie 1975). Quasi-hyperbolic discounting (Laibson 1997; Phelps and Pollak 1968) was proposed to relax the analytical intractability of hyperbolic discounting and is widely used. This study also utilizes quasi-hyperbolic discounting to introduce present bias into the model.

Numerous studies have analyzed human behavior on the basis of models with quasi-hyperbolic discounting. These studies covered topics such as consumption-saving behavior (Laibson 1997, 1998), addiction (O'Donoghue and Rabin 1999; Gruber and Köszegi 2001), and decisions in information acquisition (Carrillo and Mariotti 2000). Particularly relevant to our study is the analysis of procrastination and task abandonment in long-term projects by O'Donoghue and Rabin (2008). This study is similar to ours in that it analyzes the influence of present bias on long-term goal-achieving behavior. However, it differs from ours in that the model does not allow for an analytical description of the agent's state sequences, nor does it deal with the optimization problem of interventions.

As mentioned in Section 1, our model is inspired by the work of Kleinberg and Oren (2014). They investigated the graph-theoretic properties of the model, such as cost ratio, possible paths, and minimal motivating subgraphs. However, they do not give the analytical description of agents' action sequences and optimization algorithms for interventions to guide agents. Although subsequent studies have tackled the optimization problems of various interventions (Albers and Kraft 2019, 2021; Tang et al. 2017), they show the computational intractability of these problems and only propose approximation algorithms. Our study succeeds in analytically describing the agent's behavior and finding optimal intervention strategies by restricting the types of tasks.

The KO model has been extended in various directions to model real-world human behavior more accurately. Kleinberg, Oren, and Raghavan (2016) introduced a *sophisticated* agent in the model, which is an agent aware of its present bias and its influence (in contrast, the KO model assumes a naive agent unaware of its own present bias). Kleinberg, Oren, and Raghavan (2017) introduced *sunk-cost* bias (Arkes and Blumer 1985; Kahneman and Tversky 2013) into the model, which is the tendency for people to continue investing in something they have already invested in, even when it no longer makes rational sense. Gravin et al. (2016) proposed a model that draws the present-bias parameter from a fixed distribution in each round. Our model does not adopt such extensions and is based on the original KO model. Developing analytically tractable models that reflect these advanced factors will be future work.

Our model and reinforcement learning (Sutton and Barto

2018) are deeply related. The assumption that agents maximize their gains with time discounting is a commonality between reinforcement learning and our model. However, reinforcement learning and our model have significantly different objectives. Reinforcement learning explores what actions an agent should take to maximize its gains in unknown environments. In contrast, our model aims to gain insights into how the agent behaves under present bias and what interventions can improve the agent's behavior.

3 Proposed Model

Our study deals with tasks that involve increasing a real number, called progress, to reach a goal over a period. We assume that the progress never decreases. The following tasks fall into this category.

- Consider a person who sets a goal of exercising 30 hours in a month to improve his/her health condition. In this case, the period corresponds to one month, and the progress corresponds to the time spent exercising to date.
- Consider a salesperson trying to achieve a sales target of one million dollars in one year. In this case, the period corresponds to one year, and the progress corresponds to total sales so far.
- Consider a student who sets a goal to complete his/her graduation thesis within six months. In this case, the period corresponds to six months, and the progress corresponds to the degree of completion of the graduation thesis.

We explain how to model the behavior of an agent dealing with such a task under present bias.

3.1 Formulation

Let $T \in \mathbb{Z}_{>0}$, $\theta \in \mathbb{R}_{\geq 0}$, and $R \in \mathbb{R}_{\geq 0}$ denote the period, goal, and reward, respectively. Let a pair (t, x) denote the state of an agent, where $t \in \{0, 1, \dots, T\}$ is the time index, and $x \in \mathbb{R}_{\geq 0}$ represents the progress that the agent has achieved at that time. The agent is initially in the state $(0, 0)$ and follows the sequence $(1, x_1), (2, x_2), \dots, (T, x_T)$ as time passes. Transitioning from state (t, x_t) to state $(t + 1, x_{t+1})$ involves a cost. The cost is expressed as $c(x_{t+1} - x_t)$ with a cost function $c : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$.

Let us explain how an agent decides which state to take next. Suppose that the current state of the agent is $(t - 1, x_{t-1})$ for $t \geq 1$. The agent evaluates the cost to follow the state sequence $(t, y_t), \dots, (T, y_T)$ by

$$\mathcal{C}_t(y_t, \dots, y_T) := \frac{1}{\beta} c(y_t - x_{t-1}) + \sum_{i=t+1}^{T} c(y_i - y_{i-1}) - R \cdot \mathbf{1}[y_T \geq \theta], \tag{1}$$

where $\mathbf{1}[y_T \geq \theta] = 1$ if $y_T \geq \theta$, and $\mathbf{1}[y_T \geq \theta] = 0$ otherwise. The first term of (1) represents the transition cost from $(t - 1, x_{t-1})$ to (t, y_t), the second term represents the cost at subsequent times, and the third term represents the reward.

The first term is amplified by a coefficient of $\frac{1}{\beta}$ due to present bias, where $\beta \in (0, 1]$ is the *present-bias parameter*; the agent overestimates the cost currently faced because of present bias.[1] This formulation of present bias is called *quasi-hyperbolic discounting* (Laibson 1997) and is also used in the KO model. The third term of (1) means that the reward R is obtained if the final progress y_T is greater than or equal to θ. The reward has a negative sign because it has the opposite effect on the cost.

The agent computes the state sequence $(t, y_t^*), \dots, (T, y_T^*)$ that minimizes the cost (1) and transitions from state $(t - 1, x_{t-1})$ to (t, y_t^*). Formally, the agent's state sequence is defined by $x_0 := 0$ and

$$x_t := \operatorname*{argmin}_{y_t \in \mathbb{R}} \min_{y_{t+1}, \dots, y_T \in \mathbb{R}} \mathcal{C}_t(y_t, \dots, y_T) \tag{2}$$

for $t = 1, \dots, T$.

Let us verify that this model is a variant of the KO model. A vertex of the graph in the KO model corresponds to a state (t, x) in our model, and an edge of the graph corresponds to a transition between two states (t, x_t) and $(t + 1, x_{t+1})$. Under such correspondence, the rules of agent behavior are perfectly consistent between the two models. Note that our model differs from the KO model in that the possible state of the agent is a continuous quantity rather than a discrete quantity.

To make our model mathematically tractable, we assume that the cost function c can be written as

$$c(\Delta) = \begin{cases} \Delta^\alpha & \text{if } \Delta \geq 0, \\ +\infty & \text{otherwise}, \end{cases} \tag{3}$$

where $\alpha > 1$ is a parameter. Because we focus on tasks where the progress never decreases, we set $c(\Delta) = +\infty$ for $\Delta < 0$. Although this function may appear limited initially, it has the characteristics required for our desired modeling.

First, the cost function c should be convex. This is because tasks we deal with here are less labor-intensive if one does them steadily over a long time rather than all at once in a short time. For example, exercising one hour daily for ten days is less demanding than exercising ten hours at a time. Writing three pages of a graduation thesis per day over ten days is less burdensome than writing 30 pages in one day. The convexity of c expresses this property. Second, the cost function c should satisfy the property that $c(\Delta) = 0$ if and only if $\Delta = 0$. In the accumulation-type tasks considered in this paper, it is reasonable to suppose that there is no progress without effort and vice versa. Our cost function in (3) satisfies these two conditions and has α as a parameter, which can be adjusted to approximate the cost function for real-world tasks.

Note that our model assumes that the agent is *naive*. *Naivete* and *sophistication* have been studied in relation to procrastination (O'Donoghue and Rabin 1999a, 2001). Naive agents plan for the future without considering that

[1] Note that a large present-bias parameter β means weak present bias. Although this may seem confusing, we adopt this notation in our paper for consistency with existing studies (Laibson 1997; O'Donoghue and Rabin 2008; Kleinberg and Oren 2014).

their future selves will be affected by present bias, while sophisticated agents make plans considering that present bias will also affect their future selves. Naive models themselves are crucial and have been extensively studied (Kleinberg and Oren 2014; Gravin et al. 2016; Albers and Kraft 2021) because they will be the basis of complex models, including sophisticated ones (Kleinberg, Oren, and Raghavan 2016, 2017). We assume naivete in this paper and leave extensions to sophisticated models as future work.

3.2 Analytical Solution

To simplify the notation, let

$$p_t \coloneqq \frac{T-t}{T-t+\beta^{\frac{1}{\alpha-1}}}. \tag{4}$$

The following lemma gives a recursive formula for the state sequence $((t, x_t))_{t=0}^{T}$ taken by the agent.

Lemma 1. The following holds for $t = 1, 2, \ldots, T$:

$$x_t = \begin{cases} \theta + p_t(x_{t-1} - \theta) & \text{if } x_{t-1} \geq \tilde{\theta}_t, \\ x_{t-1} & \text{otherwise,} \end{cases}$$

where $\tilde{\theta}_t \coloneqq \theta - R^{\frac{1}{\alpha}}\left(T - t + \beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}}$.

Because $\tilde{\theta}_t$ is increasing in t, Lemma 1 implies that there exists $t^* \in \{0, \ldots, T\}$ such that

$$x_t = \begin{cases} \theta + p_t(x_{t-1} - \theta) & \text{if } t \leq t^*, \\ x_{t-1} & \text{otherwise.} \end{cases} \tag{5}$$

If $t^* = T$, the agent achieves the goal θ without giving up; otherwise, the agent gives up at time t^*. In particular, the agent abandons the task in the middle if $1 \leq t^* < T$. This result indicates that our model can reproduce task abandonment, in which an agent starts a task but gives up in the middle without any changes in underlying costs and rewards.

The following theorem characterizes the abandonment time t^* and gives an analytical formula for the sequence $((t, x_t))_{t=1}^{T}$ taken by the agent.

Theorem 1. The abandonment time t^* is the smallest $t \in \{0, \ldots, T-1\}$ such that

$$\left(T - t - 1 + \beta^{\frac{1}{\alpha-1}}\right)^{1-\alpha} \prod_{i=1}^{t} p_i^{\alpha} > \frac{R}{\theta^{\alpha}} \tag{6}$$

if there exists such t; otherwise, $t^* = T$. Moreover, the following holds for $t = 1, \ldots, T$:

$$x_t = \theta\left(1 - \prod_{i=1}^{\min\{t, t^*\}} p_i\right). \tag{7}$$

The analytical formula (7) is a major advantage of our model over the KO model. Through the formula, we will investigate the properties of the agent's behavior under present bias in the following sections.

Figure 1 shows the agents' state sequences computed by (7) for $\alpha = 2$, $T = 10$, and $R = \theta = 1$. A very small β, say $\beta = 0.1$, leads to task abandonment, i.e., giving up the task without reaching the goal. When β is relatively small, say

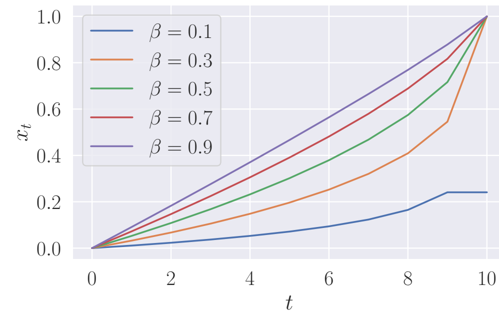

Figure 1: State sequences $((t, x_t))_{t=1}^{T}$ for $\alpha = 2$, $T = 10$, and $R = \theta = 1$.

$\beta = 0.3$, progress starts slowly and grows rapidly toward the end. This phenomenon can be interpreted as procrastination induced by present bias. In contrast, if β is as large as 0.9, the agent makes (nearly) constant progress over the whole period.

4 Task Abandonment and Present Bias

The previous section showed that our model reproduces task abandonment. This section analyzes the relationship between task abandonment and the present-bias parameter β.

4.1 Condition for Task Abandonment

Let us fix the period T. We introduce the concept of task-abandonment inducing for the present-bias parameter.

Definition 1. The present-bias parameter β is said to be *task-abandonment inducing* (TAI) if the abandonment time t^* in Theorem 1 satisfies $0 < t^* < T$ for some $\theta, R \in \mathbb{R}_{\geq 0}$.

An agent with a TAI β may abandon the task in the middle depending on the reward R or the goal θ. On the other hand, an agent with non-TAI β never abandons the task, i.e., either gives up the goal from the beginning or achieves the goal without giving up. Investigating the TAI condition on β helps us understand the relationship between the strength of present bias and the time-inconsistency of abandoning a task in the middle. The model's properties, including optimal intervention strategies, are greatly affected by whether or not β is TAI, as shown in Sections 5 and 6.

To simplify the notation, let q_t denote the left-hand side on condition (6):

$$q_t \coloneqq \left(T - t - 1 + \beta^{\frac{1}{\alpha-1}}\right)^{1-\alpha} \prod_{i=1}^{t} p_i^{\alpha}. \tag{8}$$

Theorem 1 implies that β is not TAI if and only if $\max_{0 \leq t < T} q_t = q_0$. To check if β is TAI, let us observe the properties of q_t.

Lemma 2. The following hold:

(a) if $\beta \leq (1 - \frac{1}{\alpha})^{\alpha-1}$, then $q_0 < q_1 < \cdots < q_{T-1}$,

(b) if $(1 - \frac{1}{\alpha})^{\alpha-1} < \beta < (1 - \frac{1}{\alpha})^{\frac{\alpha-1}{2}}$, then there exists $t \in \{0, \ldots, T-1\}$ such that $q_0 \geq q_1 \geq \cdots \geq q_t$ and $q_t < q_{t+1} < \cdots < q_{T-1}$,

(c) if $\beta \geq (1-\frac{1}{\alpha})^{\frac{\alpha-1}{2}}$, then $q_0 > q_1 > \cdots > q_{T-1}$.

A closer look at case (b) yields the following lemma.

Lemma 3. There exists $(1-\frac{1}{\alpha})^{\alpha-1} < \beta_0 < (1-\frac{1}{\alpha})^{\frac{\alpha-1}{2}}$ that depends only on T and α and satisfies the following:

(a) if $(1-\frac{1}{\alpha})^{\alpha-1} < \beta < \beta_0$, then $q_0 < q_{T-1}$,
(b) if $\beta = \beta_0$, then $q_0 = q_{T-1}$,
(c) if $\beta_0 < \beta < (1-\frac{1}{\alpha})^{\frac{\alpha-1}{2}}$, then $q_0 > q_{T-1}$.

In the remaining part of the paper, let β_0 denote the β_0 in Lemma 3. Since β is not TAI if and only if $\max_{0\leq t<T} q_t = q_0$, Lemmas 2 and 3 yield the following theorem.

Theorem 2. The parameter β is TAI if and only if $\beta < \beta_0$.

Theorem 2 shows that if β is large, i.e., the present bias is weak, the agent will accomplish the goal if it has decided to achieve the goal at $t = 0$. If β is small, the agent may abandon the task in the middle. These results are consistent with the past findings on present bias and time-inconsistency (Kleinberg and Oren 2014). The threshold β_0 will be crucial in the following sections.

As in Lemma 3, β_0 is lower-bounded as $\beta_0 > (1-\frac{1}{\alpha})^{\alpha-1}$. At the same time, elementary calculus shows that $(1-\frac{1}{\alpha})^{\alpha-1} > \frac{1}{e}$ for all $\alpha > 1$ and $\lim_{\alpha\to\infty}(1-\frac{1}{\alpha})^{\alpha-1} = \frac{1}{e}$, where e is Euler's number. Therefore, a small $\beta \leq \frac{1}{e} \approx 0.368$ is consistently TAI in our model; interestingly, the threshold involves Euler's number.

4.2 Asymptotic Formula for β_0

The following theorem gives an asymptotic formula for β_0 when $T \to \infty$.

Theorem 3. The following holds:

$$\beta_0 = \left(1-\frac{1}{\alpha}\right)^{\alpha-1}\left(1+\frac{\alpha\log\Gamma(1-\frac{1}{\alpha})+\log(1-\frac{1}{\alpha})}{\log T}\right) + o\left(\frac{1}{\log T}\right),$$

where Γ is the gamma function. In particular, if $\alpha = 2$, then

$$\beta_0 = \frac{1}{2} + \frac{\log(\pi/2)}{2\log T} + o\left(\frac{1}{\log T}\right).$$

Theorem 3 gives several insights about the threshold β_0.

- The threshold β_0 converges to $\left(1-\frac{1}{\alpha}\right)^{\alpha-1}$ as $T \to \infty$. The value $\left(1-\frac{1}{\alpha}\right)^{\alpha-1}$ is equal to the lower bound on β_0 given in Lemma 3.
- The convergence speed is slow as $\beta_0 - \left(1-\frac{1}{\alpha}\right)^{\alpha-1} = \Theta(1/\log T)$.
- When $\alpha = 2$, the asymptotic formula for β_0 involves Pi, interestingly.

5 Optimal Goal Setting

In this section, we consider the problem of setting a goal θ to maximize the final progress x_T:

$$\max_{\theta\geq 0}\ x_T, \tag{9}$$

given the period T and reward R. This problem arises naturally in real-world scenarios. For example, when a company's CEO establishes a sales goal for an employee and offers a bonus upon goal achievement. In this context, if the CEO sets an excessively ambitious goal, the employee may perceive them as unattainable and consequently lose motivation, resulting in reduced overall sales performance. Conversely, if the goal is set too low, the employee may settle for minimal sales to attain the rewards, leading to suboptimal sales outcomes. The question then arises: How should the CEO set the goal to maximize attained sales?

The optimal solution to this problem varies depending on whether *exploitative rewards* are allowed. Exploitative rewards are rewards placed to motivate the agent but are not claimed because the agent never reaches the target. Note that a biased agent may give up a task in the middle, but even in that case, the agent has achieved a certain amount of progress by then. Specifically, a biased agent may make greater progress when we set an unreachable but high goal (i.e., a goal with an exploitative reward) than when we set a reachable but low goal. However, using exploitative rewards can raise ethical concerns and decrease human motivation in real-world scenarios, so they should be used cautiously. The importance of considering whether to allow exploitative reward in the KO model is highlighted in previous studies (Kleinberg and Oren 2014; Tang et al. 2017; Albers and Kraft 2019), but has yet to be examined in goal-setting problems. This section will examine both scenarios where exploitative rewards are allowed and not allowed. We use the notation below:

$$\Lambda(t) := \left(T-t+\beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}}\frac{\Gamma(T-t+1)\Gamma(T+\beta^{\frac{1}{\alpha-1}})}{\Gamma(T)\Gamma(T-t+1+\beta^{\frac{1}{\alpha-1}})}.$$

5.1 Optimal Solutions

When β is not TAI, no reward setting can be exploitative. In other words, the optimal solution remains the same whether or not exploitative rewards are allowed. The following theorem provides an explicit formula for the optimal solution.

Theorem 4. Suppose that β is not TAI. Regardless of whether or not exploitative rewards are allowed, the optimal solution to problem (9) is

$$\theta = R^{\frac{1}{\alpha}}\Lambda(1).$$

The optimal value is the same as θ.

When β is TAI, the optimal solution can vary depending on whether or not exploitative rewards are allowed. The following theorem gives an explicit formula for when exploitative rewards are not allowed. For cases where they are allowed, we can reduce the continuous optimization problem (9) to a discrete one, though the optimal solution is difficult to express in a closed form.

Theorem 5. Suppose that β is TAI.

(a) Suppose that exploitative rewards are not allowed. Then the optimal solution to problem (9) is

$$\theta = R^{\frac{1}{\alpha}}\Lambda(T)$$

The optimal value is the same as θ.

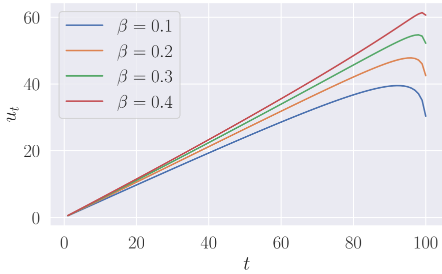

Figure 2: The plots of u_t defined by (10) when $\alpha = 10$, $T = 100$, and $R = 1$.

(b) Suppose that exploitative rewards are allowed. If the agent abandons the reward at time $t \in \{1, \dots, T\}$, the maximum final progress is

$$u_t := R^{\frac{1}{\alpha}} \left(\Lambda(t) - \frac{T - t}{\left(T - t + \beta^{\frac{1}{\alpha-1}}\right)^{1/\alpha}} \right). \tag{10}$$

Problem (9) is reduced to

$$\max_{t \in \{1,\dots,T\}} u_t. \tag{11}$$

For the optimal solution t^* to problem (11), the optimal solution to (9) is written by

$$\theta = R^{\frac{1}{\alpha}} \Lambda(t^*). \tag{12}$$

As we can compute $u_1, \dots, u_T$ in (11) in $O(T)$ time, the optimal solution to problem (9) is obtained in $O(T)$ time from Theorem 5(b) when exploitative rewards are allowed.

5.2 Discussion

Figure 2 shows u_t defined by (10) when $\alpha = 10$, $T = 100$, and $R = 1$. We observe the following from Figure 2.

- Let $t^* = \mathrm{argmax}_{t \in \{1,\dots,T\}} u_t$. Since u_t represents the maximum final progress when giving up at time t, it is optimal not to use an exploitative reward when $t^* = T$, and optimal to use it when $t^* \neq T$. Thus, using exploitative rewards is optimal in all the settings of Figure 2.
- The smaller β is, the larger u_{t^*}/u_T is. Because u_{t^*} and u_T are the maximum values of final progress with and without exploitative rewards, respectively, large u_{t^*}/u_T means a significant effect of the exploitative rewards. Thus, the smaller β is, the more significant the effect of the exploitative rewards.

These observations indicate that exploitative rewards boost the final progress of agents with strong present bias; they get lured easily by exploitative rewards.

Comparison with the literature. Several studies have analyzed optimal goal setting under present bias via mathematical models (Koch and Nafziger 2011; Koch et al. 2014; Hsiaw 2013). However, these studies have limited the agent's possible actions at each time step to a few discrete options (e.g., binary choices like "perform a task" or "do not perform a task"). In contrast, our model allows agents to select continuous progress at each time step. This study is the first to consider the optimization problem of goal-setting in continuous-state settings.

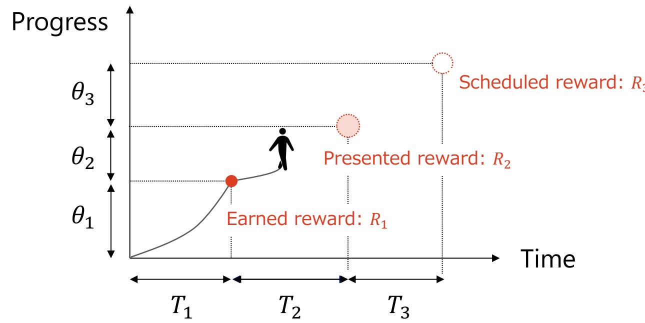

Figure 3: An example of reward scheduling when $k = 3$.

6 Optimal Reward Scheduling

Next, we consider a reward scheduling problem that aims to maximize the agent's progress when the reward can be presented multiple times. Our analysis will show that presenting rewards multiple times can increase the sum of progress compared to presenting all rewards at once. We will also see that optimal reward scheduling varies greatly depending on the present-bias parameter β.

For the problem setup, we focus on the setting where exploitative rewards are not allowed to maintain the agent's motivation. Given the total period $T \in \mathbb{Z}_{>0}$ and the total reward $R \in \mathbb{R}_{\geq 0}$, we divide them into k periods $T_1, \dots, T_k \in \mathbb{Z}_{>0}$ and k rewards $R_1, \dots, R_k \in \mathbb{R}_{\geq 0}$, where $\sum_{i=1}^k T_i = T$, $\sum_{i=1}^k R_i = R$, and $k \in \{1, \dots, T\}$ is arbitrary. At the beginning of the i-th period, the agent is offered the reward R_i and works toward the reward. If the agent increases the progress by $\theta_i \in \mathbb{R}_{\geq 0}$ during the period, it receives the reward R_i. Note that agents always earn their rewards since exploitative rewards are not allowed. Note also that since a reward R_i is offered only after the previous reward R_{i-1} has been earned, multiple rewards are never offered simultaneously. Our goal is to maximize the sum of progress over the period under the constraint that the agent earns all rewards; this constraint derives from the setting where exploitative rewards are not allowed. Note that the sum of progress equals to $\sum_{i=1}^k \theta_i$, because exploitative rewards are not allowed. For this purpose, we seek optimal reward scheduling: k, $(T_i)_{i=1}^k$, $(R_i)_{i=1}^k$, and $(\theta_i)_{i=1}^k$. Figure 3 illustrates an example of reward scheduling in this setting.

6.1 Optimal Solutions

The following theorems characterize the optimal solution.

Theorem 6. If β is not TAI, the optimal reward scheduling is

$$k = 1, \ \ T_1 = T, \ \ R_1 = R, \ \ \theta_1 = R^{\frac{1}{\alpha}} \left(T - 1 + \beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}}.$$

Theorem 7. If β is TAI, the optimal reward scheduling is

$$k = k^*, \quad T_i = T_i^*, \quad R_i = R\frac{F(T_i^*)}{\sum_{j=1}^{k} F(T_j^*)},$$
$$\theta_i = (\beta R_i)^{\frac{1}{\alpha}} \frac{\Gamma(T_i^* + \beta^{\frac{1}{\alpha-1}})}{\Gamma(T_i^*)\Gamma(1+\beta^{\frac{1}{\alpha-1}})},$$

where k^* and $(T_i^*)_{i=1}^{k^*}$ is the optimal solution of

$$\max_{k,\,(T_i)_{i=1}^k} \sum_{i=1}^{k} F(T_i), \quad \text{s.t.} \quad \sum_{i=1}^{k} T_i = T \tag{13}$$

and

$$F(x) \coloneqq \left(\frac{\Gamma(x+\beta^{\frac{1}{\alpha-1}})}{\Gamma(x)}\right)^{\frac{\alpha}{\alpha-1}}.$$

Theorem 6 shows that with non-TAI β, it is optimal to give the reward in a lump sum without splitting it. Theorem 7 indicates that we can obtain the optimal reward scheduling by solving the problem (13). Although the problem (13) is difficult to solve explicitly, we can compute the optimal solution by the following algorithmic procedure.

Let v_T be the optimal value of problem (13) and let $v_0 \coloneqq 0$. Considering the case of $T_1 = 1, 2, \ldots, T$ separately yields a recursive formula:

$$v_T = \max_{1\le t\le T}\{F(t) + v_{T-t}\}. \tag{14}$$

After computing $v_1, \ldots, v_T$ according to this formula, find $t \in \{1, \ldots, T\}$ such that $v_T = F(t) + v_{T-t}$. Such t implies that $T_1 = t$ in an optimal solution to problem (13), and thus the problem size is reduced from T to $T - t$. Repeating this reduction gives us the optimal solution to problem (13). The above procedure can be performed in $O(T^2)$ time.

We can also observe the effect of splitting the reward by examining the two extreme cases: offering all rewards at once ($k = 1$) and offering rewards every time ($k = T$).

Theorem 8. Let $\mathcal{P}(k)$ be the maximum total progress achieved by the agent when k is fixed. If β is TAI,

$$\frac{\mathcal{P}(T)}{\mathcal{P}(1)} = \Gamma(1+\beta^{\frac{1}{\alpha-1}})T^{1-\frac{1}{\alpha}-\beta^{\frac{1}{\alpha-1}}} + O\big(T^{-\frac{1}{\alpha}-\beta^{\frac{1}{\alpha-1}}}\big).$$

If the present-bias parameter β is as small as $\beta \le (1 - \frac{1}{\alpha})^{\alpha-1}$, this ratio diverges to infinity as $T \to \infty$. The smaller β, the faster the divergence. The results suggest that frequent rewards are effective for agents with strong present bias.

6.2 Special Case: TAI β and $\alpha = 2$

Assuming $\alpha = 2$ enables us to analyze problem (13) in more detail. For convenience, let $F(0) \coloneqq \lim_{x\searrow 0} F(x) = 0$.

Theorem 9. Suppose that $\alpha = 2$.

(a) If $\frac{1}{2} \le \beta < \beta_0$, the optimal reward scheduling is

$$k = 1,\; T_1 = T,\; R_1 = R,\; \theta_1 = \sqrt{\beta R}\frac{\Gamma(T+\beta)}{\Gamma(T)\Gamma(1+\beta)}.$$

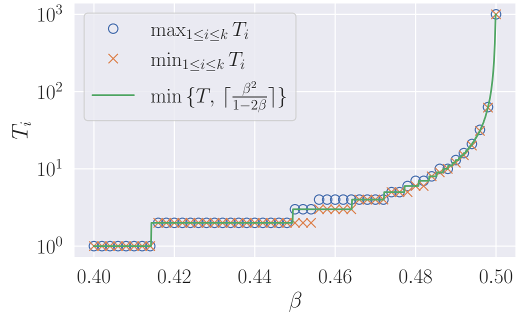

Figure 4: Optimal reward interval with $\alpha = 2$ and $T = 1000$. The markers indicate the maximum and minimum lengths of the periods in the optimal solution $(T_i)_{i=1}^k$ computed with (14). The green line shows the nearly optimal reward interval (15) derived from the theoretical analysis.

(b) If $\beta \le \sqrt{2} - 1$, the optimal reward scheduling is

$$k = T, \quad T_1 = \cdots = T_k = 1,$$
$$R_1 = \cdots = R_k = \frac{R}{T}, \quad \theta_1 = \cdots = \theta_k = \sqrt{\frac{\beta R}{T}}.$$

The case of $\sqrt{2} - 1 < \beta < \frac{1}{2}$ is the most nontrivial. The value $G(x) \coloneqq F(x)/x$ represents the increase in the objective function value per unit of time when the reward is placed at time x. Intuitively, placing a reward every x^* units of time is expected to be nearly optimal, where x^* is $x \in \{1, \ldots, T\}$ that maximizes $G(x)$, and we have

$$x^* = \min\left\{T,\; \left\lceil\frac{\beta^2}{1-2\beta}\right\rceil\right\}. \tag{15}$$

Figure 4 compares the nearly optimal reward interval (15) with the optimal reward interval computed with (14). It suggests that the analytical approximate solution (15) agrees with the exact solution.

6.3 Discussion

The results indicate that when the present-bias parameter β is large, the agent's progress can be maximized by giving the reward in a lump sum, but when β is small, giving the reward multiple times enhances the progress. This is because agents with strong present bias are more likely to abandon a task due to procrastination. Appropriate intermediate rewards can prevent them from abandoning the task and increase their progress. This finding is also significant when using incentives in the real world. It suggests that it is desirable to tailor the reward plan to each person's present bias rather than providing rewards uniformly to everyone.

Comparison with the literature. The relationship between present bias and rewards (or incentives) is one of the critical topics in behavioral economics. For example, O'Donoghue and Rabin (1999b) consider how rewards should be set when agents are paid according to task completion time. However, only some studies have handled the problem of appropriately allocating a predetermined amount of rewards to help agents achieve their long-term goals. Tang et al. (2017); Albers and Kraft (2019) provide studies on this problem based on the KO model to prove its NP-hardness. We proposed a closed-form optimal solution to this problem and a simple and exact optimization algorithm by applying appropriate restrictions on the graph structure and cost function. Furthermore, we provide the first mathematical clarification of the relationship between the effectiveness of intermediate rewards and the present bias.

7 Conclusion

This paper proposed a model for understanding the goal-achieving behavior of agents with present bias. Although our model is restricted to tasks that accumulate progress, the agent's behavior can be solved analytically. Based on the analytical solution, we analyzed three problems: task abandonment, goal optimization problem, and reward scheduling problem. We obtained meaningful insights into the relationship between the strength of present bias and these problems and their application in the real world.

8 Proofs

Due to space limitations, only proofs related to the main results are provided. Please refer to the full version (Akagi, Marumo, and Kurashima 2023) for the proof of Theorems 3, 8 and 9.

8.1 Proof of Lemma 1

Proof. For $t = T$, the desired result can be verified easily. We assume that $t < T$ below.

Let us consider the min-operation on (2). Assuming $y_T < \theta$, the minimum is achieved at $y_t = y_{t+1} = \cdots = y_T$. Otherwise, the minimum is achieved when $y_T = \theta$. Therefore, the minimum can be evaluated as

$$\begin{aligned}
&\min_{y_{t+1},\dots,y_T\in\mathbb{R}} \mathcal{C}_t(y_t,\dots,y_T)\\
&= \min\left\{0,\ \min_{y_{t+1},\dots,y_{T-1}\in\mathbb{R}} \mathcal{C}_t(y_t,\dots,y_{T-1},\theta)\right\}\\
&= \min\left\{0,\ \frac{1}{\beta}c(y_t-x_{t-1}) + (T-t)c\Big(\frac{\theta-y_t}{T-t}\Big) - R\right\},
\end{aligned}$$

where we use Jensen's inequality for the last equality. Furthermore, we bound the second term as

$$\begin{aligned}
&\frac{1}{\beta}c(y_t-x_{t-1}) + (T-t)c\Big(\frac{\theta-y_t}{T-t}\Big)\\
&= \frac{1}{\beta}(y_t-x_{t-1})^\alpha + (T-t)^{1-\alpha}(\theta-y_t)^\alpha\\
&\geq (\theta-x_{t-1})^\alpha(\beta^{\frac{1}{\alpha-1}} + T-t)^{1-\alpha},
\end{aligned}$$

where we have used (3) and Hölder's inequality:

$$\begin{aligned}
\theta - x_{t-1} &= (y_t - x_{t-1}) + (\theta - y_t)\\
&\leq \Big(\beta^{\frac{1}{\alpha-1}} + (T-t)\Big)^{\frac{\alpha-1}{\alpha}}\\
&\quad\cdot\Big(\frac{1}{\beta}(y_t-x_{t-1})^\alpha + (T-t)^{1-\alpha}(\theta-y_t)^\alpha\Big)^{\frac{1}{\alpha}}.
\end{aligned}$$

The equality holds when $\beta^{\frac{1}{\alpha-1}}(T-t)^{1-\alpha}(\theta-y_t)^\alpha = (T-t)\frac{1}{\beta}(y_t-x_{t-1})^\alpha$, or equivalently

$$y_t = \theta + p_t(x_{t-1} - \theta). \tag{16}$$

Hence, if $(\theta - x_{t-1})^\alpha(\beta^{\frac{1}{\alpha-1}} + T - t)^{1-\alpha} - R \leq 0$, the minimum on (2) is achieved at (16). Otherwise, the minimum is achieved at $y_t = x_{t-1}$, which completes the proof. □

8.2 Proof of Theorem 1

Proof. The formula (7) follows from (5) by induction on t. The t^* is the smallest $t \in \{0,\dots,T-1\}$ such that $x_t < \theta - R^{\frac{1}{\alpha}}(T-t-1+\beta^{\frac{1}{\alpha-1}})^{\frac{\alpha-1}{\alpha}}$ if there exists such t. A simple calculation shows with (7) show the equivalence between the condition on t and (6). □

8.3 Proof of Lemma 2

Proof. Let $\gamma := \beta^{\frac{1}{\alpha-1}}$. From definitions (4) and (8) of p_t and q_t, we have

$$\begin{aligned}
&\log\Big(\frac{q_{t-1}}{q_t}\Big)\\
&= (1-\alpha)\log\Big(\frac{T-t+\gamma}{T-t-1+\gamma}\Big) - \alpha\log\Big(\frac{T-t}{T-t+\gamma}\Big)\\
&= \log\Big(1+\frac{\gamma}{T-t}\Big) + (\alpha-1)\log\Big(1-\frac{1-\gamma}{T-t}\Big)\\
&= f(T-t),
\end{aligned}$$

where

$$f(x) := \log\Big(1+\frac{\gamma}{x}\Big) + (\alpha-1)\log\Big(1-\frac{1-\gamma}{x}\Big).$$

Let us investigate the function f. We have

$$\lim_{x\searrow 1-\gamma} f(x) = -\infty,\qquad \lim_{x\to+\infty} f(x) = 0,$$
$$f'(x) = \frac{\alpha\gamma(1-\gamma) + (\alpha(1-\gamma)-1)x}{x(x+\gamma)(x+\gamma-1)}.$$

Case (a): $\alpha(1-\gamma) \geq 1$**, or equivalently** $\beta \leq (1-\frac{1}{\alpha})^{\alpha-1}$**.** The function $f(x)$ is increasing for $x > 1-\gamma$. Hence, $f(x) < 0$ for $x \geq 1$, which yields Lemma 2(a).

Case (b): $\alpha(1-\gamma) < 1 < \alpha(1-\gamma^2)$**, or equivalently** $(1-\frac{1}{\alpha})^{\alpha-1} < \beta < (1-\frac{1}{\alpha})^{\frac{\alpha-1}{2}}$**.** The function $f(x)$ is increasing for $1-\gamma < x < \frac{\alpha\gamma(1-\gamma)}{1-\alpha(1-\gamma)}$ and is decreasing for $x > \frac{\alpha\gamma(1-\gamma)}{1-\alpha(1-\gamma)}$. Hence, there exists $a > 1-\gamma$ such that $f(x) < 0$ for $x < a$ and $f(x) \geq 0$ for $x \geq a$, which yields Lemma 2(b).

Case (c): $\alpha(1-\gamma^2) \leq 1$**, or equivalently** $\beta \geq (1-\frac{1}{\alpha})^{\frac{\alpha-1}{2}}$**.** A simple calculation shows that $\alpha(1-\gamma^2) \leq 1$ is equivalent to $\frac{\alpha\gamma(1-\gamma)}{1-\alpha(1-\gamma)} \leq 1$, and thus $f(x)$ is decreasing for $x \geq 1$. Hence, $f(x) > 0$ for $x \geq 1$, which yields Lemma 2(c). □

8.4 Proof of Lemma 3

Proof. Let $\gamma := \beta^{\frac{1}{\alpha-1}}$. We have

$$\frac{q_{T-1}}{q_0} = \left(1 + \frac{T-1}{\gamma}\right)^{\alpha-1} \prod_{t=1}^{T-1} \left(\frac{T-t}{T-t+\gamma}\right)^{\alpha},$$

and the value is decreasing in γ, or equivalently in β. We also know from Lemma 2 that $\frac{q_{T-1}}{q_0} > 1$ for $\beta = (1-\frac{1}{\alpha})^{\alpha-1}$ and $\frac{q_{T-1}}{q_0} < 1$ for $\beta = (1-\frac{1}{\alpha})^{\frac{\alpha-1}{2}}$, which completes the proof. □

8.5 Proof of Theorem 4

Proof. Remember that $\max_{0 \le t < T} q_t = q_0$, as mentioned in Section 4. From (6), if $q_0 > \frac{R}{\theta^\alpha}$, then $t^* = 0$, and therefore $x_T = 0$; otherwise, we have $t^* = T$. Thus, the optimal solution to problem (9) is the largest θ such that $q_0 \le \frac{R}{\theta^\alpha}$, i.e., $\theta = (R/q_0)^{\frac{1}{\alpha}} = R^{\frac{1}{\alpha}}(T-1+\beta^{\frac{1}{\alpha-1}})^{\frac{\alpha-1}{\alpha}}$. □

8.6 Proof of Theorem 5

Proof of Theorem 5(a). The agent does not abandon the reward if and only if $q_{T-1} \le \frac{R}{\theta^\alpha}$, since $\max_{0 \le t < T} q_t = q_{T-1}$ as discussed in Section 4. Therefore, the optimal solution to problem (9) is the largest θ such that $q_{T-1} \le \frac{R}{\theta^\alpha}$, i.e., $\theta = (R/q_{T-1})^{\frac{1}{\alpha}}$. A simple calculation with definitions (4) and (8) of p_t and q_t leads to

$$\theta = (R/q_{T-1})^{\frac{1}{\alpha}} = (\beta R)^{\frac{1}{\alpha}} \frac{\Gamma(T + \beta^{\frac{1}{\alpha-1}})}{\Gamma(T)\Gamma(1+\beta^{\frac{1}{\alpha-1}})}.$$

□

Proof of Theorem 5(b). Fix the time t at which the agent abandons the reward. Then Theorem 1 gives

$$q_{t-1} \le \frac{R}{\theta^\alpha} < q_t \tag{17}$$

and $x_T = \theta(1 - \prod_{i=1}^t p_i)$. Thus, the optimal θ is the largest θ that satisfies (17), i.e., $\theta = (R/q_{t-1})^{\frac{1}{\alpha}}$, and then

$$\begin{aligned} x_T &= \left(\frac{R}{q_{t-1}}\right)^{\frac{1}{\alpha}} \left(1 - \prod_{i=1}^t p_i\right) \\ &= R^{\frac{1}{\alpha}} \left(T - t + \beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}} \\ &\cdot \left(\frac{\Gamma(T-t+1)\Gamma(T+\beta^{\frac{1}{\alpha-1}})}{\Gamma(T)\Gamma(T-t+1+\beta^{\frac{1}{\alpha-1}})} - \frac{T-t}{T-t+\beta^{\frac{1}{\alpha-1}}}\right). \end{aligned}$$

Problem (9) is now reduced to (11). Eq. (12) follows from $\theta = (R/q_{t-1})^{\frac{1}{\alpha}}$. □

8.7 Proof of Theorem 6

Proof. From Theorem 4, the optimal θ_i is

$$\theta_i = R_i^{\frac{1}{\alpha}} \left(T_i - 1 + \beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}}, \tag{18}$$

given R_i and T_i. Thus, the reward scheduling problem is reduced to

$$\begin{aligned} &\max_{k,\, (T_i)_{i=1}^k,\, (R_i)_{i=1}^k} \sum_{i=1}^k R_i^{\frac{1}{\alpha}} \left(T_i - 1 + \beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}}, \\ &\quad \text{s.t.} \quad \sum_{i=1}^k T_i = T, \quad \sum_{i=1}^k R_i = R. \end{aligned} \tag{19}$$

Note that $T_i \in \mathbb{Z}_{>0}$ but $R_i \in \mathbb{R}_{\ge 0}$. Hölder's inequality gives an upper bound on the objective function of problem (19):

$$\begin{aligned} &\sum_{i=1}^k R_i^{\frac{1}{\alpha}} \left(T_i - 1 + \beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}} \\ &\le \left(\sum_{i=1}^k R_i\right)^{\frac{1}{\alpha}} \left(\sum_{i=1}^k \left(T_i - 1 + \beta^{\frac{1}{\alpha-1}}\right)\right)^{\frac{\alpha-1}{\alpha}} \\ &= R^{\frac{1}{\alpha}} \left(T - k\left(1 - \beta^{\frac{1}{\alpha-1}}\right)\right)^{\frac{\alpha-1}{\alpha}} \\ &\le R^{\frac{1}{\alpha}} \left(T - 1 + \beta^{\frac{1}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}}, \end{aligned}$$

and the upper bound is achieved when $k = 1$, $T_1 = T$, and $R_1 = R$. Hence, the optimal solution is $k = 1$, $T_1 = T$, and $R_1 = R$. The optimal θ is obtained from (18). □

8.8 Proof of Theorem 7

Proof. From Theorem 5(a), the optimal θ_i is

$$\theta_i = (\beta R_i)^{\frac{1}{\alpha}} \frac{\Gamma(T_i + \beta^{\frac{1}{\alpha-1}})}{\Gamma(T_i)\Gamma(1+\beta^{\frac{1}{\alpha-1}})},$$

given R_i and T_i. Thus, the reward scheduling problem is reduced to

$$\begin{aligned} &\max_{k,\, (T_i)_{i=1}^k,\, (R_i)_{i=1}^k} \sum_{i=1}^k (\beta R_i)^{\frac{1}{\alpha}} \frac{\Gamma(T_i + \beta^{\frac{1}{\alpha-1}})}{\Gamma(T_i)\Gamma(1+\beta^{\frac{1}{\alpha-1}})}, \\ &\quad \text{s.t.} \quad \sum_{i=1}^k T_i = T, \quad \sum_{i=1}^k R_i = R. \end{aligned} \tag{20}$$

Let us first fix k and $(T_i)_{i=1}^k$. Then the optimal $(R_i)_{i=1}^k$ is obtained as follows: Hölder's inequality gives

$$\begin{aligned} &\sum_{i=1}^k R_i^{\frac{1}{\alpha}} \frac{\Gamma(T_i + \beta^{\frac{1}{\alpha-1}})}{\Gamma(T_i)} \\ &\le \left(\sum_{i=1}^k R_i\right)^{\frac{1}{\alpha}} \left(\sum_{i=1}^k \left(\frac{\Gamma(T_i + \beta^{\frac{1}{\alpha-1}})}{\Gamma(T_i)}\right)^{\frac{\alpha}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}} \\ &= R^{\frac{1}{\alpha}} \left(\sum_{i=1}^k F(T_i)\right)^{\frac{\alpha-1}{\alpha}}, \end{aligned}$$

where $F(x) := \left(\Gamma(x+\beta^{\frac{1}{\alpha-1}})/\Gamma(x)\right)^{\frac{\alpha}{\alpha-1}}$, and the equality holds when $R_i = R \cdot \left(F(T_i)/\sum_{j=1}^k F(T_j)\right)$ for all $1 \le i \le k$.

Next, we optimize k and $(T_i)_{i=1}^k$. Problem (20) is now reduced to (13). □

References

Ainslie, G. 1975. Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82(4): 463.

Akagi, Y.; Marumo, N.; and Kurashima, T. 2023. Analytically Tractable Models for Decision Making under Present Bias. arXiv:2302.08132.

Albers, S.; and Kraft, D. 2019. Motivating time-inconsistent agents: A computational approach. *Theory of Computing Systems*, 63(3): 466–487.

Albers, S.; and Kraft, D. 2021. On the value of penalties in time-inconsistent planning. *ACM Transactions on Economics and Computation*, 9(3): 1–18.

Arkes, H. R.; and Blumer, C. 1985. The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1): 124–140.

Camerer, C. F.; and Loewenstein, G. 2004. Behavioral economics: Past, present, future. *Advances in Behavioral Economics*, 1: 3–51.

Carrillo, J. D.; and Mariotti, T. 2000. Strategic ignorance as a self-disciplining device. *The Review of Economic Studies*, 67(3): 529–544.

Frederick, S.; Loewenstein, G.; and O'donoghue, T. 2002. Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2): 351–401.

Gravin, N.; Immorlica, N.; Lucier, B.; and Pountourakis, E. 2016. Procrastination with Variable Present Bias. In *Proceedings of the 17th ACM Conference on Economics and Computation*, 361–361.

Gruber, J.; and Köszegi, B. 2001. Is addiction "rational"? Theory and evidence. *The Quarterly Journal of Economics*, 116(4): 1261–1303.

Hsiaw, A. 2013. Goal-setting and self-control. *Journal of Economic Theory*, 148(2): 601–626.

Kahneman, D.; and Tversky, A. 2013. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, 99–127. World Scientific.

Kleinberg, J.; and Oren, S. 2014. Time-inconsistent planning: a computational problem in behavioral economics. In *Proceedings of the 15th ACM Conference on Economics and Computation*, 547–564.

Kleinberg, J.; Oren, S.; and Raghavan, M. 2016. Planning problems for sophisticated agents with present bias. In *Proceedings of the 17th ACM Conference on Economics and Computation*, 343–360.

Kleinberg, J.; Oren, S.; and Raghavan, M. 2017. Planning with multiple biases. In *Proceedings of the 18th ACM Conference on Economics and Computation*, 567–584.

Koch, A. K.; and Nafziger, J. 2011. Self-regulation through goal setting. *Scandinavian Journal of Economics*, 113(1): 212–227.

Koch, A. K.; Nafziger, J.; Suvorov, A.; and van de Ven, J. 2014. Self-rewards and personal motivation. *European Economic Review*, 68: 151–167.

Laibson, D. 1997. Goden eggs and hyperbolic discounting. *Quarterly Journal of Enonomics*, 2(112): 443–477.

Laibson, D. 1998. Life-cycle consumption and hyperbolic discount functions. *European Economic Review*, 42(3-5): 861–871.

O'Donoghue, T.; and Rabin, M. 1999a. Doing it now or later. *American Economic Review*, 89(1): 103–124.

O'Donoghue, T.; and Rabin, M. 1999b. Incentives for procrastinators. *The Quarterly Journal of Economics*, 114(3): 769–816.

O'Donoghue, T.; and Rabin, M. 2001. Choice and procrastination. *The Quarterly Journal of Economics*, 116(1): 121–160.

O'Donoghue, T.; and Rabin, M. 1999. Addiction and self-control. *Addiction: Entries and Exits*, 169206.

O'Donoghue, T.; and Rabin, M. 2008. Procrastination on long-term projects. *Journal of Economic Behavior & Organization*, 66(2): 161–175.

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

Samuelson, P. A. 1937. A note on measurement of utility. *The Review of Economic Studies*, 4(2): 155–161.

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

Tang, P.; Teng, Y.; Wang, Z.; Xiao, S.; and Xu, Y. 2017. Computational issues in time-inconsistent planning. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, 3665–3671.

Wilkinson, N.; and Klaes, M. 2017. *An Introduction to Behavioral Economics*. Bloomsbury Publishing.