Delta Matters: An Analytically Tractable Model for β–δ Discounting Agents

Yasunori Akagi[1], **Takeshi Kurashima**[1]

[1]NTT Human Informatics Laboratories
yasunori.akagi@ntt.com, takeshi.kurashima@ntt.com

Abstract

Humans exhibit time-inconsistent behavior, in which planned actions diverge from executed actions. Understanding time-inconsistency and designing appropriate interventions is a key research challenge in computer science and behavioral economics. Previous work focuses on progress-based tasks and derives a closed-form description of agent behavior, from which they obtain optimal intervention strategies. They model time-inconsistency using the β–δ discounting (quasi-hyperbolic discounting), but the analysis is limited to the case $\delta = 1$. In this paper, we relax that constraint and show that a closed-form description of agent behavior remains possible for the general case $0 < \delta \leq 1$. Based on this result, we derive the conditions under which agents abandon tasks and develop efficient methods for computing optimal interventions. Our analysis reveals that agent behavior and optimal interventions depend critically on the value of δ, suggesting that fixing $\delta = 1$ in many prior studies may unduly simplify real-world decision-making processes.

1 Introduction

Human decision making often requires intertemporal choices, and time preference, which measures how much future value is discounted, plays a crucial role (Frederick, Loewenstein, and O'donoghue 2002). Economics has traditionally modeled time preference using the exponential discounting framework, which discounts value at a constant rate (Samuelson 1937). In this framework, present and future values are evaluated consistently, yielding *time-consistent* behavior: decisions do not change as time elapses.

However, actual human behavior frequently exhibits *time-inconsistency*, in which planned actions diverge from executed actions. For example, consider an individual who is planning to diet. At the planning stage, they believe that by foregoing indulgent weekend meals they will succeed in their diet. However, as the weekend approaches, the immediate allure of the indulgent meal intensifies, making it impossible to resist; the individual indulges and consequently fails to adhere to the diet plan. This shift in valuation between planning and execution is known as time-inconsistency, and as the example illustrates, it can impede long-term goal achievement. Consequently, it has become an important research challenge in behavioral economics and computer science to model the mechanisms underlying time-inconsistent behavior mathematically, predict such behavior, and derive optimal interventions that support goal attainment.

One prominent example of this line of research is the Directed Acyclic Graph (DAG) model of time-inconsistent behavior proposed by Kleinberg and Oren (2014). In this model, tasks are represented as a DAG, and an agent repeatedly selects a path that minimizes cost computed under β–δ discounting (with $\delta = 1$ fixed), thereby reproducing prototypical behaviors such as procrastination and task abandonment while flexibly capturing diverse real-world task structures. Building on this, Akagi, Marumo, and Kurashima (2024) introduced an analytically tractable model on progress-based tasks, enabling rigorous analysis of task abandonment conditions and designing interventions that maximize progress.

These models represent time preferences using β–δ discounting (quasi-hyperbolic discounting) (Phelps and Pollak 1968; Laibson 1997). Under β–δ discounting, a payoff occurring at time $t > 0$ is multiplied by the factor $\beta\delta^t$, where $0 < \beta, \delta \leq 1$ are parameters that shape the discount. The parameter β governs the degree of *present bias*—the overweighting of immediate value relative to future value—while δ governs the rate at which future value is further discounted over time, capturing long-term patience. This discounting scheme contrasts with the classical exponential model, in which the discount factor is given by δ^t. Figure 1 illustrates the shapes of the discount functions.

Although β and δ are essential for modeling human time preference, prior work has predominantly focused on the effects of β, fixing $\delta = 1$ to isolate the impact of present bias. For example, Kleinberg and Oren (2014) analyze agent behavior under the assumption $\delta = 1$, and many subsequent studies have retained this assumption, discussing only the role of β (Kleinberg, Oren, and Raghavan 2016, 2017; Akagi, Marumo, and Kurashima 2024; Tang et al. 2017; Albers and Kraft 2019; Gravin et al. 2016). The assumption $\delta = 1$ implies that a reward received after 5 days is valued the same as one received after 365 days, which is clearly unrealistic (see Figure 1). Nevertheless, the influence of δ on agent behavior and the optimal design of interventions under β–δ discounting has received little attention.

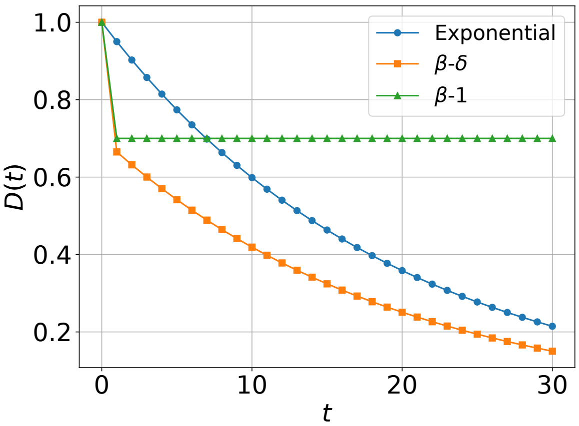

Figure 1: Example of discount functions. For $\beta = 0.7$ and $\delta = 0.95$, we plot the discount factor $D(t)$ at time t for exponential discounting with factor δ, β-δ discounting, and β-1 discounting (β–δ discounting with $\delta = 1$).

This paper analyzes how the parameter δ affects agent behavior and optimal interventions. To this end, we build on the analytically tractable model for progress-based tasks introduced by Akagi, Marumo, and Kurashima (2024). First, we extend their model to the general case $\delta \neq 1$ and show that the agent's behavior still admits a closed-form description. Based on this closed-form, we derive theoretically the conditions under which an agent abandons a task and examine which combinations of (β, δ) make abandonment more likely. We then address two optimal intervention problems, goal optimization and reward scheduling, design efficient algorithms to solve them, and use these algorithms to analyze how optimal interventions vary with the parameters (β, δ).

Our analysis yields the following insights:

- *Task abandonment.* Whether an agent may abandon a task midway depends not only on β but on the combination (β, δ). In particular, a smaller δ makes abandonment more likely. This contrasts with exponential discounting, where abandonment never occurs for any δ.
- *Goal-setting optimization.* In the problem of deciding the optimal goal to maximize final progress, δ plays a critical role. Notably, exploitative rewards (announced rewards that the agent cannot actually obtain) are most effective when β is small and δ is large. Though powerful, exploitative rewards raise ethical concerns, making our analysis important for guiding their appropriate use.
- *Reward-scheduling optimization.* In the problem of optimally timing rewards to maximize overall progress, δ has a major influence. In realistic parameter ranges, empirically estimated as $\beta \approx$0.5–0.9, $\delta \approx$0.90–0.99 (Laibson et al. 2024; Cheung, Tymula, and Wang 2021), the optimal schedule structure is highly sensitive to δ.

These results demonstrate that δ influences agent behavior and optimal interventions in a fundamentally distinct manner from β, and suggest that fixing $\delta = 1$ as in much prior work may excessively simplify real-world decision processes.

Proofs are deferred to Section 9.

2 Related Work

Various time discount functions have been proposed to model human time preferences. Representative examples include exponential discounting (Samuelson 1937), long employed in economics; hyperbolic discounting (Ainslie 1975), introduced to capture time-inconsistent preferences; and the β–δ discounting (quasi-hyperbolic discounting), which plays a central role in this study. Subsequent research has further generalized these discount functions, such as generalized hyperbolic discounting (Loewenstein and Prelec 1992) and generalized Weibull discounting (Takeuchi 2011).

Numerous studies have explored time discounting, time-inconsistency, task abandonment, and optimal intervention. One of the latest pivotal works is DAG model of time-inconsistent behavior proposed by Kleinberg and Oren (2014). This model has gathered widespread attention for its combination of modeling flexibility (capturing a broad class of real-world tasks) and tractability (leveraging graph-theoretic and algorithmic techniques). Subsequent extensions include more complicated biases (Kleinberg, Oren, and Raghavan 2016, 2017; Gravin et al. 2016) and deriving optimal interventions (Tang et al. 2017; Albers and Kraft 2019, 2021; Halpern and Saraf 2023; Belova et al. 2024). Notably, all of these works fix $\delta = 1$ in β–δ discounting.

The work most closely related to ours is the analytically tractable model for present-biased agents in progress-based tasks introduced by Akagi, Marumo, and Kurashima (2024). Their model can be interpreted as a combination of the "cumulative procrastination" framework of O'Donoghue and Rabin (2006) and the DAG-based agent model of Kleinberg and Oren (2014). It has also been extended to a continuous time model (Akagi, Kim, and Kurashima 2025), providing a versatile application framework. Our principal contribution lies in generalizing this model to the case $\delta \neq 1$ and analytically characterizing how varying δ influences agent behavior and optimal intervention. For a detailed comparison between the findings of Akagi, Marumo, and Kurashima (2024); Akagi, Kim, and Kurashima (2025) and ours, see Section 7.

Although many studies have underscored the importance of the exponential discount factor δ in the context of exponential discounting, relatively few have focused on the role of δ within the β–δ framework. For example, Meier and Sprenger (2008) conducted a field experiment using credit reports and found that δ, more so than β, strongly correlates with individuals' default behavior and FICO credit scores. Similarly, Burks et al. (2012) reported that, in a field study of truck driver trainees, including δ as a predictor improved forecasts of outcomes such as smoking behavior, credit scores, BMI, and job performance, compared to using β alone. Although our theoretical results do not directly validate these empirical findings, they share the common direction of highlighting the significance of δ. We believe that future work must validate relationships between these studies through both theoretical analysis and experimental investigation.

3 Model

3.1 Progress-based Task

This study considers *progress-based tasks* proposed by Akagi, Marumo, and Kurashima (2024). In this type of task, an agent accumulates a quantity called *progress* over a fixed period and receives a reward if a pre-specified progress target is achieved within the time limit. Such tasks frequently occur in everyday life. For example, the task "exercise for 30 hours within one month to improve one's health" falls into this category; here, the period is one month, and the progress corresponds to the cumulative hours of exercise performed. We assume that progress is non-decreasing over time[1].

We treat time as a discrete quantity. We denote the total length of the period by $T \in \mathbb{Z}_{>0}$, the goal progress by $\theta \in \mathbb{R}_{>0}$, and the reward by $R \in \mathbb{R}_{>0}$. The agent's state is represented by the tuple (t, x), where t is the time step and x is the progress.

3.2 Model Definition

We describe the decision-making model of an agent in a progress-based task. The proposed model with $\delta = 1$ corresponds to the model proposed by Akagi, Marumo, and Kurashima (2024); thus, the proposed model is a generalization of their model.

An agent in state $(t-1, x_{t-1})$ computes the cost of the states sequence $(t, y_t), \ldots, (T, y_T)$ taken from time t onward as follows:

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

where $\mathbf{1}[\cdot]$ is the indicator function, taking the value 1 if $y_T \geq \theta$ and 0 otherwise, and $c(\Delta)$ denotes the cost required to generate progress Δ. In this study, we assume

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

where $\alpha > 1$ is a parameter determining the shape of the cost function, and $c(\Delta) = +\infty$ for $\Delta < 0$ reflects the model assumption that progress cannot decrease.

The cost function $\mathcal{C}_t(y_t, \ldots, y_T)$ is designed according to quasi-hyperbolic discounting (Phelps and Pollak 1968; Laibson 1997). In quasi-hyperbolic discounting, the parameters β and δ satisfy $0 < \beta, \delta \leq 1$ and determine the form of temporal discounting. When an agent is in state $(t-1, x_{t-1})$, the immediate cost $c(y_t - x_{t-1})$ is not discounted, whereas a cost incurred at time $i\ (> t)$, $c(y_i - y_{i-1})$, is discounted by the factor $\beta\delta^{i-t}$. The parameter β captures the weight given to all future costs relative to the present, and δ captures how the discounting of future costs increases over time. The third term represents the reward: if the agent's final progress y_T meets or exceeds the goal θ, the agent obtains the discounted reward $\beta\delta^{T-t+1}R$. The preceding minus sign indicates that the reward plays the inverse role of costs. Figure 2 illustrates an example of computation of the cost of the states sequence $(t, y_t), \ldots, (T, y_T)$.

[1] In a dieting task where weight loss is progress, that progress can decrease (i.e., weight regain). Tasks in which progress may decline fall outside the scope of this work and are left for future research.

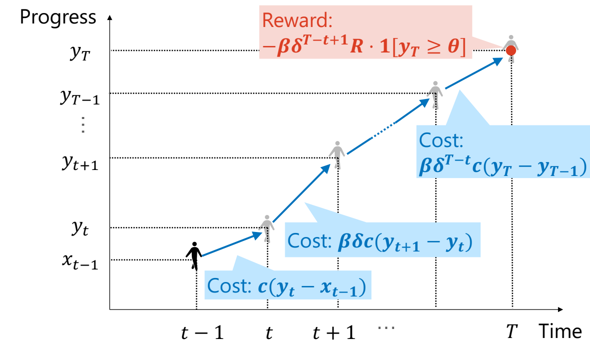

Figure 2: An example of the cost of the states sequence $(t, y_t), \ldots, (T, y_T)$.

At state $(t-1, x_{t-1})$, the agent proceeds as follows:

- For each candidate sequence of future states $(t, y_t), \ldots, (T, y_T)$, compute the cost $\mathcal{C}_t(y_t, \ldots, y_T)$ using (1) and let $(t, y_t^*), \ldots, (T, y_T^*)$ be the sequence that minimizes this cost. If there are multiple sequences that minimize the cost, we choose the one which maximizes y_t.
- Move from $(t-1, x_{t-1})$ to (t, y_t^*).

That is, the sequence of states is determined by $x_0 = 0$ and

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

This behavioral model reflects that, at each time step, the agent evaluates the cost of each candidate future sequence under β–δ discounting and follows the sequence that yields the minimal cost.

3.3 Closed Form Progress

Theorem 1. *The agent's progress x_t can be written as*

$$x_t = \theta \left(1 - \prod_{i=1}^{\min\{t, t^*\}} p_i \right), \tag{3}$$

where

$$p_t := \frac{\sum_{i=1}^{T-t} \bar{\delta}^i}{\sum_{i=1}^{T-t} \bar{\delta}^i + \bar{\beta}\bar{\delta}^{T-t+1}},$$
$$\bar{\beta} := \beta^{\frac{1}{\alpha-1}}, \quad \bar{\delta} := \delta^{\frac{1}{\alpha-1}},$$

and t^* is defined as the smallest $t \in \{0, \ldots, T-1\}$ satisfying

$$\left(\sum_{i=1}^{T-t-1} \bar{\delta}^i + \bar{\beta}\bar{\delta}^{T-t} \right)^{1-\alpha} \prod_{i=1}^{t} p_i^\alpha > \frac{R}{\theta^\alpha}, \tag{4}$$

if such a t exists; otherwise, $t^* = T$.

Theorem 1 indicates that an agent's progress admits a closed-form expression. The quantity t^* corresponds to the time step at which the agent abandons the task, and $t^* = T$ signifies that the agent never abandons the task and completes it. This result extends the $\delta = 1$ case obtained in (Akagi, Marumo, and Kurashima 2024) to the general case $\delta \leq 1$. Following, we analyze the behavior of β–δ agents based on this expression.

4 Task Abandonment

This section investigates the relationship between the conditions of task abandonment of agents and the discount parameters. Understanding this relationship enables the prediction of task abandonment and helps design intervention strategies to maximize the agent's final progress.

First, we extend the concept of the *Task-Abandonment Inducing (TAI)* parameter proposed in (Akagi, Marumo, and Kurashima 2024) to the case of β–δ discounting.

Definition 1. *For a β–δ agent, if there exist $\theta, R \in \mathbb{R}_{\geq 0}$ such that $t^* \neq 0$ and $t^* \neq T$, then the discount parameters (β, δ) are said to be* Task-Abandonment Inducing *(TAI).*

Here, t^* is the time step at which the agent abandons the task, as defined in Theorem 1. If $t^* = 0$, the agent abandons the task from the outset, i.e., never begins. If $t^* = T$, the agent never abandons and completes the task. Hence, an agent with TAI parameters may abandon the task midway for some θ and R. Conversely, an agent without TAI parameters will never abandon midway for any θ or R, and will either never start or see the task through to completion. The TAI property thus serves not only as an indicator of propensity for task abandonment, but also has significant implications for optimal intervention strategies (see Sections 5 and 6).

Define

$$q_t := \left(\sum_{i=1}^{T-t-1} \bar{\delta}^i + \bar{\beta}\bar{\delta}^{T-t} \right)^{1-\alpha} \prod_{i=1}^{t} p_i^{\alpha}, \tag{5}$$

which is the left-hand side of (4). By Theorem 1, t^* is the smallest t satisfying $q_t > \frac{R}{\theta^\alpha}$. Therefore, the necessary and sufficient condition for (β, δ) to be TAI is

$$\max_{t \in \{0,1,\ldots,T-1\}} q_t \neq q_0. \tag{6}$$

We derive the necessary and sufficient condition for (β, δ) to be TAI. To present this result, we prepare some preliminaries. First, define

$$h_1(x) := \left(\frac{(\alpha-1)\left(1 - x^{\frac{1}{\alpha-1}}\right)}{x^{\frac{1}{\alpha-1}}} \right)^{\alpha-1},$$

$$h_2(x) := \left(\frac{(\alpha-1)\left(1 - x^{\frac{1}{\alpha-1}}\right)}{x^{\frac{1}{\alpha-1}}\left(1 - \alpha(1 - x^{\frac{1}{\alpha-1}})\right)} \right)^{\alpha-1},$$

$$\gamma_1 := \left(1 - \tfrac{1}{\alpha}\right)^{\alpha-1}, \quad \gamma_2 := \left(1 - \tfrac{1}{\alpha}\right)^{\frac{\alpha-1}{2}}.$$

Then we have:

(a) $h_1(\gamma_1) = 1$, $h_2(\gamma_2) = 1$, and $h_1(1) = h_2(1) = 0$,

(b) h_1 is strictly decreasing on $[\gamma_1, 1]$, and h_2 is strictly decreasing on $[\gamma_2, 1]$.

Hence h_1 and h_2 admit inverse functions on $[\gamma_1, 1]$ and $[\gamma_2, 1]$, respectively; denote these inverses by h_1^{-1} and h_2^{-1}. With these preparations, we state the following theorem.

Theorem 2. *Fix $\delta \in (0, 1]$. Then there exists a threshold $\beta_0(\delta)$ satisfying $h_1^{-1}(\delta) < \beta_0(\delta) < h_2^{-1}(\delta)$, and (β, δ) is TAI if and only if $\beta < \beta_0(\delta)$.*

Theorem 2 implies that for a fixed δ, the TAI property is determined by whether β is smaller than the threshold $\beta_0(\delta)$. Although deriving an analytic form for $\beta_0(\delta)$ is difficult, it is guaranteed to lie between $h_1^{-1}(\delta)$ and $h_2^{-1}(\delta)$. These inverse functions can be expressed explicitly as functions of δ as follows:

$$h_1^{-1}(\delta) = \left(\frac{\alpha-1}{\alpha - 1 + \delta^{\frac{1}{\alpha-1}}} \right)^{\alpha-1},$$

$$h_2^{-1}(\delta) = \left(\frac{g^-(\delta)}{2\alpha\delta^{\frac{1}{\alpha-1}}} \right)^{\alpha-1} = \left(\frac{2(\alpha-1)}{g^+(\delta)} \right)^{\alpha-1},$$

where

$$g^-(\delta) := \sqrt{(\alpha-1)^2\left(1 - \delta^{\frac{1}{\alpha-1}}\right)^2 + 4\alpha(\alpha-1)\delta^{\frac{1}{\alpha-1}}} - (\alpha-1)\left(1 - \delta^{\frac{1}{\alpha-1}}\right),$$

$$g^+(\delta) := \sqrt{(\alpha-1)^2\left(1 - \delta^{\frac{1}{\alpha-1}}\right)^2 + 4\alpha(\alpha-1)\delta^{\frac{1}{\alpha-1}}} + (\alpha-1)\left(1 - \delta^{\frac{1}{\alpha-1}}\right).$$

Figure 3 illustrates the relationship among h_1, h_2, and β_0 in the β–δ plane for the case $\alpha = 5$ and $T = 100$. The green dashed line plots the numerically computed $\beta_0(\delta)$.[2] Moreover, Figure 4 plots the same relationship for $\alpha = 1.1, T = 100$ and $\alpha = 10, T = 100$.

From Figure 3, we observe:

- The β–δ space is divided into two connected regions: one where the parameters are TAI and one where they are not.
- Smaller values of β tend to induce the TAI property.
- Smaller values of δ also tend to induce the TAI property.

Furthermore, Figure 4 indicates that the TAI region expands as α decreases. This observation is supported by the proposition below.

Proposition 1.

$$\lim_{\alpha \to 1^+} h_1^{-1}(\delta) = \lim_{\alpha \to 1^+} h_2^{-1}(\delta) = 1,$$

$$\lim_{\alpha \to \infty} h_1^{-1}(\delta) = \frac{1}{e}, \lim_{\alpha \to \infty} h_2^{-1}(\delta) = \frac{1}{\sqrt{e}}$$

for $0 < \delta \leq 1$, where e is Euler's number.

[2] By Lemma 3 in Section 9.2, $\beta_0(\delta)$ is the value of β satisfying $q_0 = q_{T-1}$. Because Lemma 3 guarantees that $q_0 < q_{T-1}$ for $h_1^{-1}(\delta) \leq \beta < \beta_0(\delta)$ and $q_0 > q_{T-1}$ for $\beta_0(\delta) < \beta \leq h_2^{-1}(\delta)$, for fixed δ, we can find $\beta_0(\delta)$ numerically via binary search.

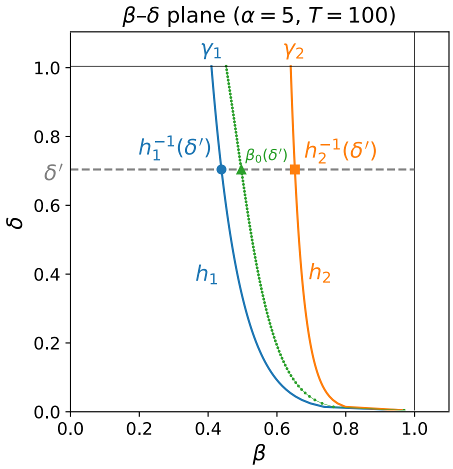

Figure 3: The relationship among h_1, h_2, and β_0 in the β–δ plane for $\alpha = 5$ and $T = 100$. The green dashed line shows the numerically computed values of $\beta_0(\delta)$.

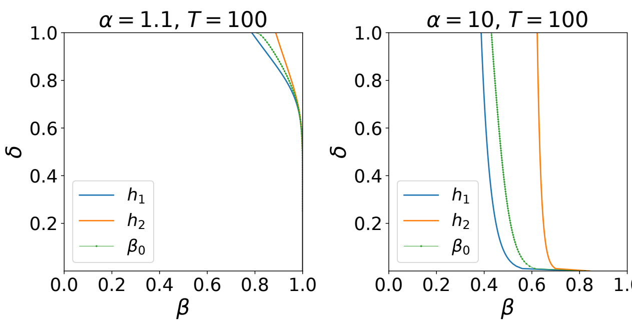

Figure 4: Left: the relationship among h_1, h_2, and β_0 for $\alpha = 1.1$, $T = 100$; Right: the same for $\alpha = 10$, $T = 100$.

Intuitively, when α is small, the cost of generating a large amount of progress in a short time is low, making the strategy of "catching up at the end" more attractive and thus increasing the tendency to procrastinate, resulting in task abandonment.

These results demonstrate that whether the parameter pair is TAI depends critically not only on β but also on δ. Therefore, analyses of agent behavior should pay careful attention to both discount parameters.

5 Goal Optimization

This section examines how to determine the goal θ to maximize the agent's final progress x_T, given the period T and reward R. Concretely, we consider the optimization problem

$$\max_{\theta \geq 0} x_T.$$

This problem differs fundamentally depending on whether *exploitative* rewards are permitted. An exploitative reward refers to a decoy incentive the agent can never obtain, since the goal is inherently unreachable. A time-inconsistent agent may initially believe the goal is attainable and accumulate progress, only to abandon the task later. In such cases, although no reward is paid, the agent accumulates partial progress toward the goal. One can induce the agent to work without ever paying the reward, and in some settings, exploitative rewards can yield greater final progress than non-exploitative schemes.

Exploitative rewards raise several ethical concerns: (i) Because they presuppose the agent's failure, repeated use increases the number of failures the agent experiences, potentially undermining the agent's confidence and reducing their self-efficacy. (ii) They may erode the trust between the agent and the intervention designer, making the agent more reluctant to make progress. (iii) Since exploiting agents with strong present bias is beneficial for the intervention designer, certain agents may incur disproportionately large losses, thereby creating inequality. Thus, intervention designers must be careful when using exploitative rewards in practice.

This study analyzes the optimal goal-setting problem when exploitative rewards are allowed and when not. Mathematically, disallowing exploitative rewards corresponds to adding the constraint $x_T \geq \theta$ to the original optimization problem.

Theorem 3 (Non-Exploitative Case)**.** *Suppose exploitative rewards are* not *permitted.*

(a) *If* (β, δ) *is* not *TAI, then the optimal goal is*

$$\theta = \left(\frac{R}{q_0}\right)^{\frac{1}{\alpha}} = R^{\frac{1}{\alpha}} \left(\sum_{i=1}^{T-1} \bar{\delta}^i + \bar{\beta}\,\bar{\delta}^T\right)^{\frac{\alpha-1}{\alpha}}, \tag{7}$$

and the final progress achieved $x_T = \theta$.

(b) *If* (β, δ) is *TAI, then the optimal goal is*

$$\theta = \left(\frac{R}{q_{T-1}}\right)^{\frac{1}{\alpha}} = R^{\frac{1}{\alpha}} \beta^{\frac{1}{\alpha}} \delta^{\frac{1}{\alpha}} \prod_{t=1}^{T-1} \left(1 + \frac{\bar{\beta}\,\bar{\delta}^{T-t+1}}{\sum_{i=1}^{T-t} \bar{\delta}^i}\right), \tag{8}$$

and again $x_T = \theta$.

Theorem 4 (Exploitative Case)**.** *Suppose exploitative rewards* are *permitted.*

(a) *If* (β, δ) *is* not *TAI, then the optimal goal remains*

$$\theta = \left(\frac{R}{q_0}\right)^{\frac{1}{\alpha}} = R^{\frac{1}{\alpha}} \left(\sum_{i=1}^{T-1} \bar{\delta}^i + \bar{\beta}\,\bar{\delta}^T\right)^{\frac{\alpha-1}{\alpha}}, \tag{9}$$

as in the non-exploitative case, and $x_T = \theta$.

(b) *If* (β, δ) is *TAI, then the optimal goal is*

$$\theta = \left(\frac{R}{\max\{q_0,\, q_{t^*-1}\}}\right)^{\frac{1}{\alpha}},$$

where

$$t^* := \operatorname*{argmax}_{t \in \{\tilde{t},\ldots,T\}} u_t, \tag{10}$$

$$u_t := \left(\frac{R}{\max\{q_0,\, q_{t-1}\}}\right)^{\frac{1}{\alpha}} \left(1 - \prod_{i=1}^{t} p_i\right),$$

and $\tilde{t}$ *is the smallest* t *such that* $q_t > q_0$*. In this case,* $x_T = u_{t^*}$.

The right-hand sides of (7), (8), and (9) are strictly increasing in δ. Hence, in these cases, a smaller δ calls for a smaller optimal goal θ, resulting in a smaller final progress.

In the case where the parameters (β, δ) are TAI and exploitative rewards are permitted (Theorem 4(b)), the optimal goal is determined by the optimum solution of (10), making it challenging to derive qualitative properties directly from the formula. To explore these properties empirically, we plot the values of u_t for several representative cases in Figure 5.

From Figure 5, we observe the following:

(a) As δ decreases, u_{t^*} (where t^* is defined in (10)) decreases; that is, the maximum final progress achievable becomes smaller.

(b) As δ increases, the gap between the maximum u_{t^*} and u_T widens. Since u_{t^*} is the maximum progress under exploitative rewards and u_T is the maximum progress when exploitative rewards are disallowed, a larger δ amplifies the exploitative-reward effect.

To examine point (b) in more detail, we define $\tau := \frac{u_{t^*}}{u_T}$, and for fixed $\alpha = 2, 5$ and $T = 100$, we display τ over the (β, δ) plane as a heat map in Figure 6. Since τ measures the ratio of maximum progress with exploitative rewards to that without, it indicates the effectiveness of exploitative rewards. Larger τ implies a stronger exploitative effect. Figure 6 suggests that the exploitative effect is most pronounced when β is small and δ is large.

These observations indicate that β and δ exert opposing influences on the effectiveness of exploitative rewards. We can interpret this phenomenon as follows. For exploitative rewards to be effective, the agent must be motivated by anticipating a reward that is both difficult to obtain and temporally distant. A small β (strong present bias) makes the agent prone to procrastination, leading to sustained partial progress toward an unattainable goal and thus accumulating considerable progress for exploitative reward. In contrast, a small δ (steep long-term discount) diminishes the agent's valuation of distant future rewards, reducing its incentive to accumulate progress and weakening the exploitative effect.

These insights are crucial to design interventions: we must distinguish between β and δ, measure them accurately, and tailor interventions accordingly; otherwise, interventions may be ineffective or unintentionally exploitative.

6 Reward Scheduling Optimization

Next, we consider an optimization problem that aims to maximize the agent's final progress by splitting the given total reward and appropriately presenting it to the agent. This problem is an extension of the goal-optimization problem in Section 5, allowing for more flexible interventions. We call this the *optimal reward scheduling problem*.

We mathematically formulate this problem. We assume that exploitative rewards are disallowed. First, the intervention designer is given a total period $T \in \mathbb{Z}_{>0}$ and a total reward $R \in \mathbb{R}_{\geq 0}$. The intervention designer splits these into $N \in \mathbb{Z}_{>0}$ parts, with subperiods $T_1, \ldots, T_N \in \mathbb{Z}_{>0}$ and subrewards $R_1, \ldots, R_N \in \mathbb{R}_{\geq 0}$. In the i-th period, the agent accumulates progress under the condition "increasing

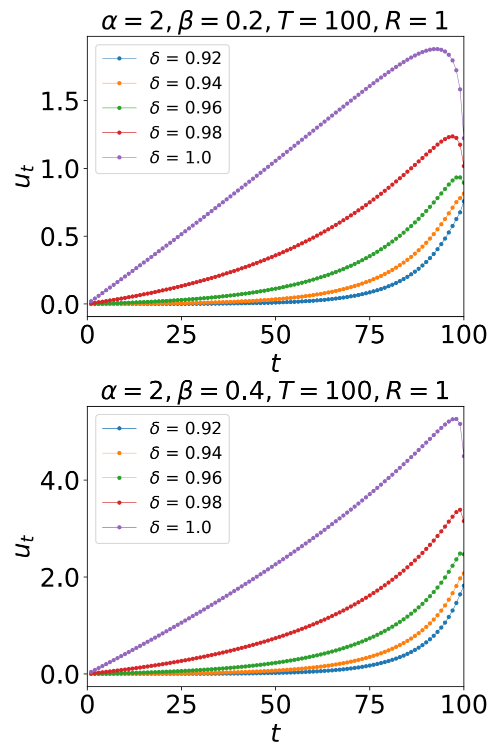

Figure 5: Relationship between t and u_t for varying β, δ with $\alpha = 2$, $T = 100$, and $R = 1$. Top: $\beta = 0.2$; bottom: $\beta = 0.4$.

progress by $\theta_i \in \mathbb{R}_{\geq 0}$ within period T_i yields reward R_i". The intervention designer's objective is to determine N, $(T_i)_{i=1}^N$, $(R_i)_{i=1}^N$, and $(\theta_i)_{i=1}^N$ so as to maximize the sum of progress across all periods. Note that rewards are presented sequentially: while working in period i, the agent does not observe any future rewards or goals for periods $i+1, \ldots, N$. Figure 7 illustrates an example of reward scheduling with $N = 4$.

Theorem 5. *For given (β, δ), define*

$$F(x) := \begin{cases} \bar{\beta}\bar{\delta} \prod_{t=1}^{x-1} \left(1 + \frac{\bar{\beta}\,\bar{\delta}^{x-t+1}}{\sum_{i=1}^{x-t} \bar{\delta}^i}\right)^{\frac{\alpha}{\alpha-1}} & (\beta, \delta) \text{ is TAI,} \\ \sum_{t=1}^{x-1} \bar{\delta}^t + \bar{\beta}\,\bar{\delta}^x & \text{otherwise.} \end{cases}$$

Let $N, (T_i)_{i=1}^N$ be the optimal solution to

$$\max_{N, (T_i)_{i=1}^N} \sum_{i=1}^N F(T_i), \text{ s.t. } \sum_{i=1}^N T_i = T,\ T_i \in \mathbb{Z}_{>0}, \quad (11)$$

and denote its optimum by $N^*, (T_i^*)_{i=1}^{N^*}$. Then the optimal reward schedule is

$$N = N^*,\ T_i = T_i^*,\ R_i \propto F(T_i),\ \theta_i = R_i^{\frac{1}{\alpha}} F(T_i)^{\frac{\alpha-1}{\alpha}}.$$

The optimization problems (11) can be solved exactly in $O(T^2)$ time via dynamic programming (Algorithm 1), so we can solve the original reward scheduling optimization problem in $O(T^2)$ time. For $T = 100$, we computed the optimal schedule over the grid $\{(\beta, \delta) = (0.01i, 0.01j) \mid$

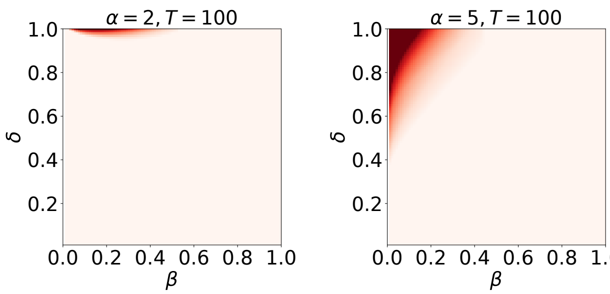

Figure 6: Heatmap of $\tau := \frac{u_{t^*}}{u_T}$ on the (β, δ) plane. Darker colors indicate a larger effect of exploitative reward.

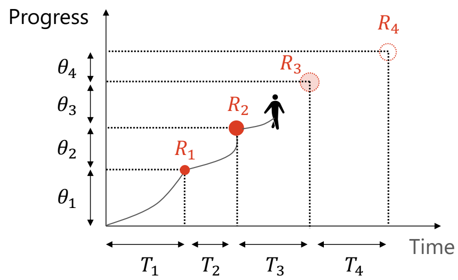

Figure 7: An example of reward scheduling with $N = 4$.

$i, j = 1, \ldots, 100\}$ by dynamic programming and plotted $\max_{i=1}^{N} T_i$ in Figure 8. In our computations, we observed $\max_{i=1}^{N} T_i - \min_{i=1}^{N} T_i \leq 1$ for all (β, δ), so we can regard $\max_{i=1}^{N} T_i$ as the optimal reward-interval length. From Figure 8, we see:

- For fixed β, smaller δ leads to shorter optimal intervals.
- Fixing δ and varying β: for $\alpha = 2$, when $\delta \leq 0.6$, the optimal interval is 1 (i.e. every time step), regardless of β. Otherwise, as β increases from 0, the optimal interval first lengthens, then shortens; the maximal interval occurs near the TAI boundary $\beta_0(\delta)$. A similar pattern holds for $\alpha = 5$.

These results further indicate the importance of δ in intervention design. Even for the same β, variations in δ change the optimal intervention. In particular, in the human's parameter range $\beta \approx$0.5–0.9, $\delta \approx$0.90–0.99 estimated in field studies (Laibson et al. 2024; Cheung, Tymula, and Wang 2021), optimal solution is very sensitive to the value of δ. Therefore, accurate measurement of δ is crucial for effective intervention design.

7 Comparison to Previous Models

This section discusses the relationships and differences among the insights gained from (Akagi, Marumo, and Kurashima 2024), (Akagi, Kim, and Kurashima 2025), and this work. The model of Akagi, Marumo, and Kurashima (2024) corresponds to fixing $\delta = 1$ in our framework, and thus can be viewed as a special case of our model. The model of Akagi, Kim, and Kurashima (2025) is a continuous-time model and employs hyperbolic discounting rather than β–δ discounting. We refer to our model as the *β–δ model*, to Akagi, Marumo, and Kurashima (2024)'s as the *β–1 model*, and to Akagi, Kim, and Kurashima (2025)'s as the *continuous-time hyperbolic model (CTHM)*.

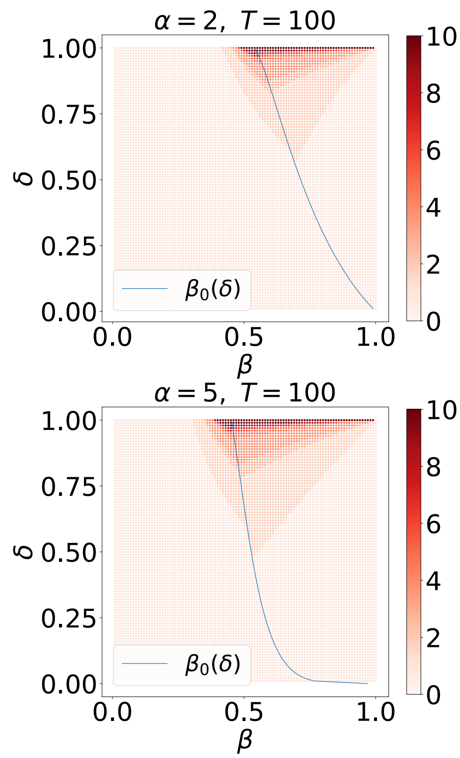

Figure 8: Top: heatmap of $\max_{i=1}^{N} T_i$ for $\alpha = 2, T = 100$. Bottom: for $\alpha = 5, T = 100$. The blue curve denotes the TAI boundary $\beta_0(\delta)$.

Task Abandonment The β–1 model has a threshold β_0 such that the discount parameters are TAI if and only if $\beta < \beta_0$ (Akagi, Marumo, and Kurashima 2024, Theorem 2). In the β–δ model, the analogous generalization (Theorem 2 in this paper) holds: β_0 depends on δ and tends to increase as δ decreases. In the CTHM , steep discounting induces task abandonment (Akagi, Kim, and Kurashima 2025, Proposition 1), displaying the same trend as other models.

Goal-Setting Optimization Under the β–1 model, exploitative rewards become more effective as β decreases. The β–δ model exhibits a similar pattern, but exploitative rewards are effective only when δ is sufficiently large. In contrast, the CTHM demonstrates that exploitative rewards are never beneficial: non-exploitative goals always produce higher final progress. This highlights a fundamental divergence between discrete-time models and the CTHM .

Reward-Scheduling Optimization In the β–1 model, large β favors lump-sum rewards, whereas small β favors reward splitting. The β–δ model also requires parameter-dependent reward splitting, but the mapping from (β, δ) to the optimal splitting is more intricate, as in Figure 8. By contrast, in the CTHM , fine-grained reward splitting is optimal regardless of parameter values. Here, the discrete-time and

Algorithm 1: Dynamic Programming for Problem (11)

Input: Total period T, function F
Output: Optimal segments $(T_1, \ldots, T_k)$

Initialize $v[0] \leftarrow 0$ and $v[t] \leftarrow -\infty$ for $t = 1, \ldots, T$
for $t = 1$ to T **do**
  for $s = 1$ to t **do**
    if $F(s) + v[t-s] > v[t]$ **then**
      $v[t] \leftarrow F(s) + v[t-s]$
      $prev[t] \leftarrow s$
    end if
  end for
end for
$t \leftarrow T$
$segments \leftarrow []$
while $t > 0$ **do**
  $s \leftarrow prev[t]$
  prepend s to $segments$
  $t \leftarrow t - s$
end while
return $segments$

continuous-time frameworks also yield markedly different optimal structures.

In summary, while discrete-time models (both β–1 and β–δ) and the CTHM share similar behavior regarding task abandonment, they diverge completely in their optimal solutions for goal-setting and reward scheduling. The precise reason for these differences remains unclear. One possible explanation lies in the differing shapes of the discount functions. In hyperbolic discounting used in the CTHM , the discount function is given by $D(t) = \frac{1}{1+kt}$, where the single positive parameter k determines both the short-term and long-term discount rates, by contrast, in the β–δ model, β governs short-term discounting and δ governs long-term discounting independently, thereby allowing a wider variety of discounting behaviors than hyperbolic discounting. This greater flexibility may introduce more complex effects on the optimal intervention strategy under β–δ preferences. More detailed theoretical analyses may clarify this issue in future work. Additionally, because Akagi, Kim, and Kurashima (2025) has not undertaken empirical validation, assessing experimentally which model more accurately captures real-world human behavior and yields superior intervention strategies is important. Various techniques to estimate parameters of the β–δ model have been established (Laibson et al. 2024; Cheung, Tymula, and Wang 2021), and by comparing the estimated parameters with observed human behavior and responses to interventions, such validation can be carried out.

8 Conclusion

This study extends the time-inconsistent agent behavior model proposed by Akagi, Marumo, and Kurashima (2024) to the case $\delta \neq 1$, successfully deriving a closed-form mathematical description of the agent's behavior. Based on this expression, we characterize the conditions of task abandonment, derive optimal intervention algorithms, and analyze the relationship between δ and the optimal interventions. Our results demonstrate that δ plays a critical role in agent behaviors and the structure of optimal interventions, suggesting that, unlike prior work which fixes $\delta = 1$, this assumption may constitute an undue simplification of real-world decision processes.

9 Proofs

Due to the space limitation, we only provide proof of the main results (Theorems 1 and 2). For other proofs, please refer to the full version (Akagi and Kurashima 2025).

9.1 Proof of Theorem 1

Proof. For simplicity, let $S_t := \sum_{i=1}^{t} \delta^{\frac{i}{\alpha-1}} = \sum_{i=1}^{t} \bar{\delta}^i$ and let $\bar{\beta} := \beta^{\frac{1}{\alpha-1}}$ and $\bar{\delta} := \delta^{\frac{1}{\alpha-1}}$ as defined in Theorem 1.

Consider the inner minimization in (2). If $y_T < \theta$, the minimum is clearly attained at $y_t = y_{t+1} = \cdots = y_T$, yielding an objective value of 0. Otherwise, the minimum is attained at $y_T = \theta$. Hence,

$$\begin{aligned}&\min_{y_{t+1},\ldots,y_T \in \mathbb{R}} \mathcal{C}_t(y_t, \ldots, y_T)\\ &\quad= \min\left\{0, \min_{y_{t+1},\ldots,y_{T-1} \in \mathbb{R}} \mathcal{C}_t(y_t, \ldots, y_{T-1}, \theta)\right\}.\end{aligned}$$

The second term of $\mathcal{C}_t(y_t, \ldots, y_{T-1}, \theta)$ can be bounded by Jensen's inequality as follows:

$$\sum_{i=t+1}^{T} \delta^{i-t}(y_i - y_{i-1})^\alpha \geq \delta^{T-t+1} S_{T-t}^{-(\alpha-1)}(\theta - y_t)^\alpha,$$

where equality holds if and only if

$$\frac{y_i - y_{i-1}}{\delta^{-\frac{i-t}{\alpha-1}}} = \text{constant}, \quad i = t+1, \ldots, T.$$

Using this, we obtain

$$\begin{aligned}&\mathcal{C}_t(y_t, \ldots, y_{T-1}, \theta)\\ &\geq (y_t - x_{t-1})^\alpha + \beta\delta^{T-t+1} S_{T-t}^{-(\alpha-1)}(\theta - y_t)^\alpha.\end{aligned}$$

To further bound the right-hand side, we apply Hölder's inequality.

Lemma 1 (Hölder's inequality)**.** *Let $p, q > 0$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, and let $(a_i)_{i=1}^n$, $(b_i)_{i=1}^n$ be real sequences. Then*

$$\left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |b_i|^q\right)^{\frac{1}{q}} \geq \sum_{i=1}^{n} |a_i b_i|,$$

with equality if there exists a constant C such that $|a_i|^p = C|b_i|^q$ for all i.

Setting $n = 2$ and

$$\begin{aligned}&a_1 = y_t - x_{t-1}, \quad a_2 = \beta^{\frac{1}{\alpha}} \delta^{\frac{T-t+1}{\alpha}} S_{T-t}^{-\frac{\alpha-1}{\alpha}}(\theta - y_t),\\ &b_1 = 1, \quad b_2 = \beta^{-\frac{1}{\alpha}} \delta^{-\frac{T-t+1}{\alpha}} S_{T-t}^{\frac{\alpha-1}{\alpha}},\\ &p = \alpha, \quad q = \tfrac{\alpha}{\alpha-1},\end{aligned}$$

we obtain

$$\Big((y_t - x_{t-1})^\alpha + \beta\delta^{T-t+1} S_{T-t}^{-(\alpha-1)}(\theta - y_t)^\alpha\Big)^{\frac{1}{\alpha}} \cdot \Big(1 + \beta^{-\frac{1}{\alpha-1}} \delta^{\frac{T-t+1}{\alpha-1}} S_{T-t}\Big)^{\frac{\alpha-1}{\alpha}} \geq \theta - x_{t-1},$$

hence

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

Equality holds if and only if

$$y_t = \frac{S_{T-t} x_{t-1} + \bar{\beta}\bar{\delta}^{T-t+1}\theta}{S_{T-t} + \bar{\beta}\bar{\delta}^{T-t+1}}. \tag{12}$$

Consequently, when

$$\begin{aligned}&\Big(1 + \beta^{-\frac{1}{\alpha-1}} \delta^{\frac{T-t+1}{\alpha-1}} S_{T-t}\Big)^{-(\alpha-1)} (\theta - x_{t-1})^\alpha \\ &\qquad - \beta\delta^{T-t+1} R \leq 0 \\ &\Leftrightarrow x_{t-1} \geq \theta - R^{\frac{1}{\alpha}} \big(S_{T-t} + \bar{\beta}\bar{\delta}^{T-t+1}\big)^{\frac{\alpha-1}{\alpha}}\end{aligned} \tag{13}$$

holds, the minimum in (2) is attained by (12), and otherwise the minimum is attained at $y_t = x_{t-1}$. Therefore, for $t = 1, \ldots, T$,

$$x_t = \begin{cases} \frac{S_{T-t} x_{t-1} + \bar{\beta}\bar{\delta}^{T-t+1}\theta}{S_{T-t} + \bar{\beta}\bar{\delta}^{T-t+1}} & \text{if (13) holds,} \\ x_{t-1} & \text{otherwise.} \end{cases}$$

A simple calculation shows the right-hand side of (13) is strictly increasing in t, so there exists $t^* \in \{0, \ldots, T\}$ such that

$$x_t = \begin{cases} \frac{S_{T-t} x_{t-1} + \bar{\beta}\bar{\delta}^{T-t+1}\theta}{S_{T-t} + \bar{\beta}\bar{\delta}^{T-t+1}} & t \leq t^*, \\ x_{t-1} & \text{otherwise,} \end{cases}$$

and by unrolling this recursion we obtain (3). Finally, t^* is the smallest $t \in \{0, \ldots, T-1\}$ satisfying

$$\theta\Big(1 - \prod_{i=1}^{t} p_i\Big) < \theta - R^{\frac{1}{\alpha}} \big(S_{T-t-1} + \bar{\beta}\bar{\delta}^{T-t}\big)^{\frac{\alpha-1}{\alpha}},$$

which is easily seen to be equivalent to (4). □

9.2 Proof of Theorem 2

Proof. First, we prove the following lemma.

Lemma 2. *Fix α, δ, and T. Then the followings hold:*

(a) *If $\beta \leq h_1^{-1}(\delta)$, then $q_0 < q_1 < \cdots < q_{T-1}$.*
(b) *If $h_1^{-1}(\delta) < \beta < h_2^{-1}(\delta)$, then there exists $t \in \{0, \ldots, T-1\}$ such that $q_0 \geq q_1 \geq \cdots \geq q_t < q_{t+1} < \cdots < q_{T-1}$.*
(c) *If $\beta \geq h_2^{-1}(\delta)$, then $q_0 > q_1 > \cdots > q_{T-1}$.*

Proof. By the definition (5) of q_t, we have

$$\log \frac{q_t}{q_{t-1}} = -(\alpha - 1)\log\Big(1 - \frac{1-\bar{\beta}}{x}\Big) - \log\Big(1 + \frac{\bar{\beta}\bar{\delta}}{x}\Big),$$

where $x := \frac{S_{T-t}}{\bar{\delta}^{T-t}} = 1 + \frac{1}{\bar{\delta}} + \frac{1}{\bar{\delta}^2} + \cdots + \frac{1}{\bar{\delta}^{T-t-1}}$ decreases monotonically in t and always lies in $[1, \infty)$ for $t = 1, \ldots, T-1$. We investigate the function

$$f(x) := -(\alpha - 1)\log\Big(1 - \tfrac{1-\bar{\beta}}{x}\Big) - \log\Big(1 + \tfrac{\bar{\beta}\bar{\delta}}{x}\Big).$$

We have

$$\lim_{x \searrow 1-\bar{\beta}} f(x) = +\infty, \quad \lim_{x \to +\infty} f(x) = 0,$$
$$f'(x) = \frac{\big(\bar{\beta}\bar{\delta} - (\alpha-1)(1-\bar{\beta})\big)x - \alpha\bar{\beta}\bar{\delta}(1-\bar{\beta})}{x(x - (1-\bar{\beta}))(x + \bar{\beta}\bar{\delta})}$$

hold.

Case (a): $\beta \leq h_1^{-1}(\delta)$. In this case, $\bar{\beta}\bar{\delta} - (\alpha-1)(1-\bar{\beta}) \leq 0$ and $\alpha\bar{\beta}\bar{\delta}(1-\bar{\beta}) \geq 0$, so $f'(x) \leq 0$ for all $x > 1 - \bar{\beta}$. Because f is decreasing on $(1-\bar{\beta}, \infty)$, we get $f(x) > 0$ for all $x \geq 1$. Therefore $\log(q_t/q_{t-1}) > 0$ for $t = 1, \ldots, T-1$, proving (a).

Case (b): $h_1^{-1}(\delta) < \beta < h_2^{-1}(\delta)$. In this case, $\bar{\beta}\bar{\delta} - (\alpha-1)(1-\bar{\beta}) > 0$, and we can check $f'(1) < 0$. Thus $f'(x) \leq 0$ for $x \leq \frac{\alpha\bar{\beta}\bar{\delta}(1-\bar{\beta})}{\bar{\beta}\bar{\delta} - (\alpha-1)(1-\bar{\beta})}$ and $f'(x) > 0$ for $x > \frac{\alpha\bar{\beta}\bar{\delta}(1-\bar{\beta})}{\bar{\beta}\bar{\delta} - (\alpha-1)(1-\bar{\beta})}$. So, there is a unique threshold $\tilde{x} < 1 - \bar{\beta}$ and $f(x) > 0$ for $x < \tilde{x}$ and $f(x) \leq 0$ for $x \geq \tilde{x}$. Since $S_{T-t}/\bar{\delta}^{T-t}$ decreases in t, this yields the claimed pattern of monotonicity in (b).

Case (c): $\beta \geq h_2^{-1}(\delta)$. In this case $\bar{\beta}\bar{\delta} - (\alpha-1)(1-\bar{\beta}) \geq 0$ and $f'(1) \geq 0$, so $f'(x) \geq 0$ for all $x \geq 1$. Hence f is strictly increasing on $[1, \infty)$, we have $f(x) < 0$ for $x \geq 1$. Therefore $\log(q_t/q_{t-1}) < 0$ for all $t = 1, \ldots, T-1$, proving (c). □

Next, through a more detailed analysis of the case $h_1^{-1}(\delta) < \beta < h_2^{-1}(\delta)$, we obtain the following lemma.

Lemma 3. *Fix α, δ, and T. There exists a unique*

$$\beta_0 \in \big(h_1^{-1}(\delta), h_2^{-1}(\delta)\big)$$

such that:

(a) *If $h_1^{-1}(\delta) < \beta < \beta_0$, then $q_0 < q_{T-1}$.*
(b) *If $\beta = \beta_0$, then $q_0 = q_{T-1}$.*
(c) *If $\beta_0 < \beta < h_2^{-1}(\delta)$, then $q_0 > q_{T-1}$.*

Proof. Observe that

$$\frac{q_{T-1}}{q_0} = \Big(\tfrac{S_{T-1}}{\bar{\beta}\bar{\delta}} + \bar{\delta}^{T-1}\Big)^{\alpha-1} \prod_{t=1}^{T-1} \Big(\tfrac{S_{T-t}}{S_{T-t} + \bar{\beta}\bar{\delta}^{T-t+1}}\Big)^{\alpha}$$

is strictly decreasing in β. By Lemma 2, at $\beta = h_1^{-1}(\delta)$ we have $q_{T-1}/q_0 > 1$, and at $\beta = h_2^{-1}(\delta)$ we have $q_{T-1}/q_0 < 1$. The intermediate value theorem yields a unique β_0 satisfying $q_{T-1}/q_0 = 1$, and the proof completes. □

Finally, to prove Theorem 2, recall that (β, δ) is TAI if and only if

$$\max_{t=0,\ldots,T-1} q_t \neq q_0$$

by (6). Combining Lemma 2 and Lemma 3 completes the proof. □

References

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

Akagi, Y.; Kim, H.; and Kurashima, T. 2025. A Continuous-time Tractable Model for Present-biased Agents. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 13510–13519.

Akagi, Y.; and Kurashima, T. 2025. Delta Matters: An Analytically Tractable Model for β–δ Discounting Agents. *arXiv preprint arXiv:2509.13637*.

Akagi, Y.; Marumo, N.; and Kurashima, T. 2024. Analytically tractable models for decision making under present bias. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 9441–9450.

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 (TEAC)*, 9(3): 1–18.

Belova, T.; Dementiev, Y.; Fomin, F.; Golovach, P.; and Ignatiev, A. 2024. How to guide a present-biased agent through prescribed tasks? In *ECAI 2024*, 3461–3468. IOS Press.

Burks, S.; Carpenter, J.; Götte, L.; and Rustichini, A. 2012. Which measures of time preference best predict outcomes: Evidence from a large-scale field experiment. *Journal of Economic Behavior & Organization*, 84(1): 308–320.

Cheung, S. L.; Tymula, A.; and Wang, X. 2021. Quasi-hyperbolic present bias: A meta-analysis. *Life Course Centre Working Paper*, (2021-15).

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 2016 ACM Conference on Economics and Computation*, 361–370.

Halpern, J. Y.; and Saraf, A. 2023. Chunking tasks for present-biased agents. In *Proceedings of the 24th ACM Conference on Economics and Computation*, 853–884.

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.

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

Laibson, D.; Lee, S. C.; Maxted, P.; Repetto, A.; and Tobacman, J. 2024. Estimating discount functions with consumption choices over the lifecycle. *The Review of Financial Studies*.

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

Meier, S.; and Sprenger, C. 2008. Impatience and credit behavior: Using choice experiments to explain borrowing and defaulting. *Center for Behavioral Economics and Decision Making, Federal Reserve Bank of Boston*, 600.

O'Donoghue, T.; and Rabin, M. 2006. Incentives and self-control. *Econometric Society Monographs*, 42: 215.

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.

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

Tang, P.; Teng, Y.; Wang, Z.; Xiao, S.; and Xu, Y. 2017. Computational issues in time-inconsistent planning. In *Thirty-First AAAI Conference on Artificial Intelligence*.