OSIPP Discussion Paper: DP-2020-E-001

Procrastination and Learning about Self-Control

March 4, 2020

Else Gry Bro Christensen*

RBB Economics, Düsseldorf

Takeshi Murooka†

Osaka School of International Public Policy (OSIPP), Osaka University

Abstract: We study a model of task completion with the opportunity to learn about own self-control problems over time. While the agent is initially uncertain about her future self-control, in each period she can choose to learn about it by paying a non-negative learning cost and spending one period. If the agent has time-consistent preferences, she always chooses to learn whenever the learning is beneficial. If the agent has time-inconsistent preferences, however, she may procrastinate such a learning opportunity. Further, if her time preferences exhibit inter-temporal conflicts between future selves (e.g., hyperbolic discounting), the procrastination of learning can occur even when the learning cost is zero. Such procrastination also leads to non-completion of the task. Our results help explain why people pursue implausible dreams and never start any task instead of taking better alternatives. When the agent has multiple initially uncertain attributes (e.g., own future self-control and own ability for the task), the agent's endogenous learning decisions may be misdirected – she chooses to learn what she should not learn from her initial perspective, and she chooses not to learn what she should.

JEL Codes: C70, D83, D90, D91

Keywords: procrastination, self-control, naivete, hyperbolic discounting, misdirected learning

* Email: e.g.b.christensen@gmail.com

† Email: murooka@osipp.osaka-u.ac.jp

Procrastination and Learning about Self-Control*

Else Gry Bro Christensen† Takeshi Murooka‡

March 4, 2020

Abstract

We study a model of task completion with the opportunity to learn about own self-control problems over time. While the agent is initially uncertain about her future self-control, in each period she can choose to learn about it by paying a non-negative learning cost and spending one period. If the agent has time-consistent preferences, she always chooses to learn whenever the learning is beneficial. If the agent has time-inconsistent preferences, however, she may procrastinate such a learning opportunity. Further, if her time preferences exhibit inter-temporal conflicts between future selves (e.g., hyperbolic discounting), the procrastination of learning can occur even when the learning cost is zero. Such procrastination also leads to non-completion of the task. Our results help explain why people pursue implausible dreams and never start any task instead of taking better alternatives. When the agent has multiple initially uncertain attributes (e.g., own future self-control and own ability for the task), the agent's endogenous learning decisions may be misdirected — she chooses to learn what she should not learn from her initial perspective, and she chooses not to learn what she should.

JEL Codes: C70, D83, D90, D91

Keywords: procrastination, self-control, naivete, hyperbolic discounting, misdirected learning

---

*We thank Florian Englmaier, Matthias Fahn, Dominik Fischer, Fabian Herweg, Daisuke Hirata, Alexander Koch, Botond Kőszegi, David Laibson, Yves Le Yaouanq, Johannes Maier, Matthew Rabin, Klaus Schmidt, Peter Schwardmann, Marco Schwarz, Adam Szeidl, and seminar and conference audiences for helpful comments. Financial support from the Deutsche Forschungsgemeinschaft through CRC/TRR 190 (B01) and JSPS KAKENHI (JP16K21740) are gratefully acknowledged.

†RBB Economics, Kasernenstr. 1, 40213 Duesseldorf, Germany (Email: e.g.b.christensen@gmail.com).

‡Corresponding author. Osaka School of International Public Policy, Osaka University, Machikaneyama 1-31, Toyonaka, Osaka, 560-0043, Japan (Email: murooka@osipp.osaka-u.ac.jp).

1 Introduction

People procrastinate (i.e., make a delay that was not fully anticipated ex-ante) in a range of tasks, including filing tax returns, doing reports, setting up financial portfolios, or making educational decisions.[1] In some situations, people keep procrastinating the same or similar tasks again and again. Such "naivete" about own future self-control appears prevalent in many economic environments, and its implications have been investigated in the literature. Most of the literature, however, assumes (either explicitly or implicitly) that people do not learn about their self-control problems over time. This opens up a natural question: under what conditions do people fail to learn about their self-control problems?

Building upon the work on naivete about self-control by O'Donoghue and Rabin (1999a, 2001), we investigate models in which an agent faces a task to complete and potentially has time-inconsistent preferences.[2] Section 2 introduces our model and discusses its key assumptions. In our basic model, an agent is initially not sure whether her future selves will be time-consistent or time-inconsistent, and she may initially underestimate her probability of being time-inconsistent. Her time-preference type is persistent over time. In each period, she voluntarily chooses whether or not to update her beliefs about own self-control problem (which may affect her subsequent actions on task completion) by incurring a non-negative cost of learning. We highlight the case in which the cost of learning is zero and the agent can (in principle) perfectly learn about her self-control. In contrast to Ali (2011), in our model the agent cannot commit to any future actions.

Section 3 analyzes an illustrative model of task completion. Our key mechanism is that the agent may procrastinate the opportunity to learn about her future self-control type. The intuition consists of two steps. First, because learning itself changes own future actions, a time-inconsistent agent may indirectly incur a cost from learning if changing future actions

[1] For empirical and experimental evidence, see, for example, Ariely and Wertenbroch (2002), DellaVigna and Malmendier (2006), Augenblick and Rabin (2019), and Le Yaouanq and Schwardmann (2019). Following O'Donoghue and Rabin (2001) and DellaVigna (2009), we say that an agent "procrastinates" if she plans to do a task with some probability but actually does it with lower probability, whereas an agent "delays" if she plans not to do a task with some probability and actually does not do it with the same probability.

[2] For papers analyzing task completion and procrastination, see, for example, O'Donoghue and Rabin (1999a, 1999b, 2001, 2008), Ariely and Wertenbroch (2002), Herweg and Müller (2011), and Fischer (2018).

involves both future benefit (e.g., returns from setting up a financial portfolio) and future cost (e.g., effort cost to set up a financial portfolio). We highlight that this indirect cost occurs when an agent's time preferences exhibit intra-personal conflicts between future selves (e.g., hyperbolic discounting), whereas there is no such indirect cost when she has intra-personal conflicts only between her current self and future selves (e.g., quasi-hyperbolic discounting). Second, because the time-consistent type of agent initially overestimates the probability that she will be time-consistent, she (erroneously) believes that she will choose to learn and complete a task with higher probability than in reality. Because of this overconfidence about her own future behavior, she may prefer to postpone the learning opportunity in order to complete the task later rather than sooner.[3] As a result, the agent may not use the opportunity to learn about her own self-control problem, even when the direct cost of learning is zero. We also derive conditions in which procrastination of learning and non-completion of a task is a *unique* equilibrium outcome. Our results help explain why people procrastinate even when they face similar situations repeatedly.

This non-learning result is in contrast to the one found by Ali (2011), who shows that time-inconsistent people will learn perfectly about their self-control under flexible commitment devices. We discuss assumptions leading to our non-learning result: absence of commitment devices, voluntary learning, and conflict between future selves. Our result could help bridge the gap between Ali's (2011) theoretical results and the empirical relevance of naivete about self-control. Perhaps surprisingly, our result does not depend on anticipatory utility, ego utility, private image, or social image (e.g., Kőszegi 2006, Gottlieb 2014, 2016). Furthermore, even when people are inherently aware of their own self-control problems, when facing a new task, they may have an incentive to neglect this awareness and take action as an ignorant self. In this sense, we provide a rationale of O'Donoghue and Rabin's (2001, page 127) assumption: "it is central to our analysis that a person not fully learn over time her

[3] Note that the second logic looks akin to that in O'Donoghue and Rabin (1999a, 2001). However, we identify assumptions and conditions for when and how procrastination can occur in the domain of learning about own self-control problems, whereas O'Donoghue and Rabin (1999a, 2001) assume that an agent cannot learn about this over time. We further show that procrastination of learning can occur even when the direct cost of learning is zero under general time-inconsistent preferences, whereas it does not occur under quasi-hyperbolic discounting as employed by O'Donoghue and Rabin (1999a, 2001).

true self-control problem, or, if she does come to recognize her general self-control problem, she still continues to underestimate it on a case-by-case basis."

Section 4 analyzes a model of task choice in which an agent first chooses a task to complete and then engages in the task. To describe the situation, suppose that an agent chooses to either pursue a university degree, do an apprenticeship, or postpone the decision. If she knew she were time-consistent, she would go to university now. If she knew she were time-inconsistent, the high upfront cost of the university degree would be too much for her, so she would be better off choosing the apprenticeship now. But if she is time-inconsistent and does not know her future self-control type, she may believe that her future selves will be time-consistent with a high probability. Given this belief and her time-inconsistent preferences, she may prefer to go to university next year and not now. This is why she does not start the apprenticeship now, and once next year comes she may postpone going to university again; in this way, she may never complete any further education. In addition, she may also not choose to learn about her own naivete, because she may have an incentive to procrastinate doing so, as illustrated above. We derive conditions under which a time-inconsistent type of agent neither chooses a task nor learns about her own future self-control, which leads to non-completion of any task. Interestingly, lower learning costs can make time-inconsistent agents more likely to procrastinate on the equilibrium path.

As an important extension, Section 5 investigates a case in which an agent initially has multiple uncertain attributes and can choose to learn about either (or both) of them. For example, the agent may be initially uncertain about both her own future self-control and her own ability for the task. As analyzed by Carrillo and Mariotti (2000) and Bénabou and Tirole (2002), under some conditions the agent strictly prefers to avoid learning about her own ability to motivate her future self to complete the task. This strategic ignorance can be beneficial from the agent's initial perspective. Under other conditions, however, we show that the agent's endogenous learning decisions are *misdirected* — she chooses to learn what, from the initial perspective, she should not learn (in our example, about her own self-control), and she chooses not to learn what she should learn from the initial perspective (in our example, about her own ability).

Section 6 briefly discusses other extensions. Section 7 concludes. Proofs and additional analysis are provided in the Appendix.

Related literature. There are several strands of literature that investigate why people do not learn over time. The most closely related literature to our paper is on strategic ignorance: Carrillo and Mariotti (2000) and Bénabou and Tirole (2002) show how, under the presence of self-control problems, an agent may strategically abstain from learning about her own ability for the task, to keep motivating her own future selves to work harder. In these papers, the strategic ignorance is beneficial from the agent's initial perspective. Relatedly, Bénabou and Tirole (2004) study how an agent may commit to a personal rule to maintain self-reputation. The implications of our results are different from those of the previous studies mentioned here, however: in our model, non-learning is harmful from the agent's initial perspective because it leads to non-completion of a task.[4] Furthermore, we demonstrate that if the agent is uncertain about both her own self-control and her own ability, then under some conditions the agent's learning decisions will be misdirected and harmful.

Within the literature on self-esteem, Kőszegi (2006), Gottlieb (2014), and Gottlieb (2016) investigate models in which an agent may avoid learning because of the presence of ego utility, private image, or anticipatory utility. In our model, such psychological costs are captured as a direct cost of learning; we show that non-learning can occur even when there is no direct (physical or psychological) cost of learning.

The literature on selective attention, such as Schwartzstein (2014) and Gagnon-Bartsch, Rabin and Schwartzstein (2018), analyzes situations in which an agent systematically does not encode a certain type of signal. In contrast to these studies, we focus on the situation in which the cost of learning can be zero and an agent can perfectly learn about her own self-control problems but may voluntarily choose not to learn about it.[5]

[4] There is also a difference regarding the mechanisms: an agent in the cited papers chooses not to learn as a means of internal commitment (and to improve her own future payoffs), whereas an agent in our paper does so because she overestimates the probability that she will behave time-consistently.

[5] In a similar way, our results also differ from the literature on non-Bayesian updating rules, in which an agent cannot update her own beliefs according to Bayes's rule. Further, most of the mechanisms proposed

Finally, our results on multiple uncertain attributes are related to the recent literature on learning under misspecified models (Fudenberg, Romanyuk and Strack 2017, Heidhues, Kőszegi and Strack 2018, Hestermann and Le Yaouanq 2020). Building upon and extending this literature, our mechanism based on procrastination explains why an agent's learning decisions can be misdirected even when she can choose to perfectly learn about all initially uncertain attributes.

2 Model

This section introduces our basic model. Section 2.1 sets up the model. Section 2.2 discusses our key assumptions.

2.1 Setup

A risk-neutral agent can choose a task from a menu of tasks in periods $t = 1, 2, \cdots, T$, where $T \geq 2$ is either finite or infinite. Once she chooses a task, she can complete the task from the next period on.[6] Let $X = \{x_1, \ldots, x_N\}$ be a menu of tasks. Each task is represented by $x_n = (-c_n, b_n)$: completing the task in period t gives a cost $c_n \geq 0$ in period t and brings a benefit $b_n \geq 0$ in period $t+1$. There is no (physical or psychological) cost of choosing a task. The agent can choose (and complete) at most one task during the game.

Let u_t denote the agent's period-t instantaneous utility. There are two types of agents: time-consistent and time-inconsistent ones. The type of each agent is persistent throughout the game. For the time-consistent type, her total utility in period t is $\sum_{\tau=t}^{\infty} \delta^{\tau-t} u_\tau$ with $\delta \in (0,1)$. For the time-inconsistent type, her total utility in period t is $\sum_{\tau=t}^{\infty} D(\tau - t) u_\tau$, where $D(\tau - t) \in (0,1]$ represents her time discounting in $\tau - t$ periods.[7] We assume

in previous studies (e.g., conservatisms) would predict that learning is merely slowed down but would occur eventually (see Benjamin, Rabin and Raymond (2016) for a notable exception).

[6] Motivated by real-world examples such as making educational decisions (and then studying) or setting up a financial portfolio (and then using it to accumulate savings), we focus on situations in which an agent takes multiple steps to complete a task in Section 4.

[7] See, for example, Echenique, Imai and Saito (2020) for the analysis under general time preferences. Relatedly, Schweighofer-Kodritsch (2018) analyzes Rubinstein's alternative bargaining problem with general

that $D(0) = 1$, $D(1) < \delta$, $D(t)$ is strictly decreasing in t with $\lim_{t\to\infty} D(t) = 0$, and $D(1)/D(0) \neq D(2)/D(1)$.

Both types of agent are initially uncertain about their own future self-control. We assume that all types share the same initial belief about their own future time consistency, denoted by $\hat{D}(t)$, and that it is $\hat{D}(t) = D(t)$ with probability $1-q$ and $\hat{D}(t) = \delta^t$ with probability q.[8] In line with the literature on limited cognition and beliefs as assets (Bénabou 2015, Bénabou and Tirole 2016), we assume that updating the initial belief about own self-control requires the agent to voluntarily choose to learn; formally, in each period the agent can acquire a signal about her own self-control by incurring a cost $m \geq 0$, while she sticks to her initial beliefs if she has not acquired the signal. To focus on our main mechanism, we assume that learning is perfect and perpetual: the signal is perfectly informative and the agent becomes completely sophisticated about their own self-control problems for the rest of the game. It is worth emphasizing that learning is made by the agent's voluntary choice, as we discuss in detail in Section 2.2.

Let $k \in \{C, I\}$ be the agent's true discounting type, where C represents a time-consistent type and I represents a time-inconsistent type. Let $U_t^C(x_n; \tau - t) = -\delta^{\tau-t}c_n + \delta^{\tau-t+1}b_n$ and $U_t^I(x_n; \tau - t) = -D(\tau - t)c_n + D(\tau - t + 1)b_n$ respectively denote type-C's and type-I's total utility evaluated in period t (not taking into account the learning cost m) when the agent completes task x_n in period τ. Let $s_t \in \{0, 1\}$ denote the agent's decision on acquiring a signal about own self-control (i.e., her learning decision) in period t. If the agent has already acquired the signal (i.e., $s_{t_0} = 1$ for some $t_0 < t$), the cost of learning $m \geq 0$ has been paid, and hence the agent's utility only depends on when she will complete the task. Note that, at the beginning of period τ, the agent has not learned about her self-control if and only if $s_t = 0$ for all $t < \tau$. When the agent acquires a signal about her type in period $t_0 \geq 0$, each type of agent's subjective expectation of total utility evaluated in period t is

time preferences.

[8] In some applications, it could make more sense to assume that a time-consistent type knows own future self-control, whereas a time-inconsistent type is initially (stochastically) overconfident about own future self-control. As we will discuss later, none of our results would qualitatively change with this alternative specification.

$\hat{U}_t^k(s_{t_0} = 1) = qU_t^k(x(C); \tau(C) - t) + (1 - q)U_t^k(x(I); \tau(I) - t)$, where $x(k)$ and $\tau(k)$ indicate the chosen tasks and completion periods for an agent of type k after learning her type in period t_0. We abbreviate subscript t whenever it is clear. Note that the utility is evaluated according to each type's own preferences: the only reason why each type of agent cares about own future preferences is to predict future outcomes, and each period-t self maximizes the agent's own period-t utility given her (possibly erroneous) expectations.

The timing of the game is as follows. If the agent has not chosen a task, in each period she takes one of the following actions: learning about her own self-control problem (if she has not yet done so), choosing a task, or not doing anything. If the agent has chosen a task, in each period the agent takes one of the following actions: learning about her own self-control problem (if she has not yet done so), completing the task, or not doing anything.

As an equilibrium concept, we adopt a natural extension of the perception-perfect equilibria (O'Donoghue and Rabin 2001) in which (i) in each continuation game an agent chooses her best response given her current belief and preference and (ii) an agent keeps using the initial prior about her own self-control problem if she has not acquired a signal about her own self-control.[9] We focus on pure strategies. While all of our theoretical results are descriptive, whenever we discuss potential welfare implications, we evaluate each type of agent's welfare based on the initial-self's preference.

2.2 Discussion of Key Assumptions

Voluntary learning. A crucial assumption of our model is that the agent has to voluntarily choose whether to learn about her own self-control problem: whenever updating her beliefs, she has to "encode" a signal (e.g., she needs to introspect from her own past experience).[10] Hence, our model is different from a classical one in which Bayesian updating is automatically applied. It is also different from a self-confirming equilibrium, as the agent in

[9] The second part is a natural extension of O'Donoghue and Rabin (2001), because a (partially) naive present-biased agent in their original model does not update beliefs about own future self-control over time.

[10] In period $t = 1$, we interpret the learning opportunity as constituting the option for the agent to look back to similar past experiences to infer her self-control problem. In period $t \geq 2$, the agent can try to learn about her own self-control problem from her own past actions in the game.

our model updates her beliefs only when she chooses to do so.[11] In our model, learning is thus analyzed through the lens of the agent's cost-benefit analysis.

Unless the agent chooses to learn, she does not revise her beliefs, and keeps using her prior belief. Hence, unlike Bénabou and Tirole (2004), we rule out self-signaling over time. We also assume away the possibility that the agent becomes aware of her own self-control by inferring from her own current preferences and beliefs (i.e., we assume away inferences like "because my preferences are time-inconsistent today, I must be a time-inconsistent type in the future").[12] These assumptions are consistent with our assumption that the agent has to voluntarily choose to update her belief, and are also in line with the assumptions in O'Donoghue and Rabin (1999a, 2001). In contrast to these two papers, however, our model does not assume that the possibility of learning is excluded. Indeed, our results complement O'Donoghue and Rabin (1999a, 2001) in the sense that, given the conditions provided in this paper, their original insight and results are robust to introducing the opportunity to learn about own self-control.

(At most) one action per period. Another assumption is that the agent cannot choose learning and other actions at the same time (i.e., an agent has to spend one period to update her beliefs). While this assumption is restrictive, it captures the notion that we need to spend time thinking about ourselves whenever changing our self-perception; this "time cost" is key to generating an intra-personal conflict between future selves (and hence procrastination on learning). In other words, an agent in our model can be Bayesian, but the Bayesian inference is not free: the agent needs to pay the cost of learning $m \geq 0$ and take some time to think.[13]

[11] Esponda and Pouzo (2016) provide a framework to analyze agents with misspecified models. Note that our model is also different from theirs, because in our model the agent voluntarily chooses whether or not to update her own initial belief.

[12] Akin to the discussion in Footnote 8 of O'Donoghue and Rabin (2001), a time-inconsistent type in our model does not make such an inference. That is, a time-inconsistent type cannot revise her beliefs unless she encodes a signal. By contrast, all of our results are qualitatively robust (and often become simpler) if a time-consistent type becomes aware of her own future self-control by inferring from the fact that she is currently time-consistent.

[13] Note that this assumption may not be necessary if future selves can misuse the acquired information from the perspective of today's self. To both simplify the analysis and highlight the procrastination mechanism, however, we shut down such a possibility in our model.

Although the direct cost of learning can be zero, because of the delay introduced by voluntary learning (as described above), the agent may also incur a cost from non-learning. One of the contributions of this paper is to highlight how an agent may not learn about her own future self-control even when the direct learning cost is zero (i.e., $m = 0$) and how this can be a unique equilibrium outcome. Our model also allows a positive learning cost (i.e., $m > 0$), which represents, for example, a physical recollection cost or a psychological cost of deteriorating self-esteem (Kőszegi 2006; Gottlieb 2014, 2016). Incorporating such a learning cost is plausible in many real-world situations, and comparative statics with respect to m hence generate additional implications for such settings.

No commitment. In the model, the agent cannot commit to any particular future behavior, including future learning decisions. In Section 6, we discuss the case in which an agent can take up a commitment device. If the agent can commit to specific future actions, under zero learning cost, non-learning will not occur in our model. In this respect, our non-learning result complements the study by Ali (2011), who shows that time-inconsistent people will learn perfectly about their self-control if they can take up a flexible commitment device and do Bayesian updating automatically.

3 Task Completion

This section analyzes an illustrative model of task completion; we focus on the case in which the agent has already selected a task $x = (c, b)$.[14] We can then restrict attention to the agent's learning and task completion choices. To shed light on our main mechanism in the simplest manner, we further assume throughout this section that the agent can work on the task only after acquiring a signal about her own self-control.[15] In Section 4, we analyze a full model without these assumptions.

[14] Alternatively, this can be a situation in which the agent is facing a menu of only one task.

[15] As we formally investigate in Section 4, this assumption simplifies a situation in which working on a task without learning is costly (e.g., the agent cannot choose an optimal task without knowing her own type), and hence the agent does not want to engage in any task before learning about her own self-control.

3.1 Task Completion: Setup

The timing of the game in this section is as follows. If the agent has not acquired a signal $s_t \in \{0, 1\}$ about her self-control problem, in each period she chooses to either acquire it ($s_t = 1$) or not ($s_t = 0$). If the agent has acquired it, in each period she chooses whether to complete the task or not.

In what follows, we focus on the most interesting case, in which each type of agent prefers to complete the task as soon as possible rather than never:

Assumption 1 (Task is Worthwhile)**.** (i) $U^I(x; 0) > 0$, (ii) $\min\{U^C(x; 1), U^I(x; 1)\} > m$.

Because $U^C(x; 0) \geq U^I(x; 0)$ by assumption $D(1) < \delta$, Assumption 1 (i) means that each type of agent prefers to complete the task right now as opposed to never .[16] Assumption 1 (ii) means that each type of agent prefers to acquire information right now and to complete the task in the next period rather than never.[17] Note that under Assumption 1, both types of agent agree that the task should be completed. Hence, uncertainty and naivete about own future type is the only reason for not acquiring the signal and for non-completion of the task.

3.2 Illustrative Example in a Finite Horizon Model

In this subsection, we illustrate a simple numerical example with a finite time horizon (i.e., $4 \leq T < \infty$) to highlight how procrastination on learning can occur in our model. Because of the finite time horizon, equilibrium can be pinned down by backward induction and is generically unique. See Appendix B for the full analysis of the finite horizon model.

Suppose that $m = 0$, $c = 1$, $b = \frac{3}{2} + \epsilon$ for a sufficiently small $\epsilon > 0$, $\delta = 1$, and $D(t) = \frac{1}{1+rt}$ with $r = \frac{1}{2}$, that is, a time-inconsistent type's preference exhibits hyperbolic discounting. In this case, it can be shown that in every continuation game the time-consistent type always

[16] If $U^I(x; 0) < 0$, a time-inconsistent type never completes the task. In this case, if $U^C(x; 0) < 0$ or if the time-inconsistent type would not want the task to be completed in any future period, then obviously the time-inconsistent type would never encode the signal. Otherwise, the time-inconsistent type may have an incentive to acquire a signal, but procrastination on acquiring the signal can still occur in this case.

[17] If $U^C(x; 1) < m$, then a time-consistent type never encodes a signal.

acquires information about her own type (if she has not acquired it yet) and always works on the task (if she has acquired information). For the time-consistent type, $U^I(x;\tau) > 0$ for all τ, and completing the task in two periods is the most preferred for her, that is, $\max_\tau U^I(x;\tau) = U^I(x;2)$. As the agent has to learn her type before completing the task, the earliest point in time the task can be completed is $t = 2$ (with acquisition of information in $t = 1$).

Next, we analyze the behavior of the time-inconsistent type by backward induction. First, we characterize the behavior of the time-inconsistent type in $t = T$. If she has acquired information, then she completes the task, because $U^I(x;0) = -c + D(1)b = \frac{2}{3}\epsilon > 0$. If she has not acquired information, then she cannot complete the task, because $t = T$ is the last period, and hence, her payoff is zero.

Second, we characterize the behavior of the time-inconsistent type in $t = T - 1$. If she has already acquired information, then she does not complete the task, because $U^I(x;1) = -D(1)c + D(2)b = \frac{1}{12} + \frac{1}{2}\epsilon > U^I(x;0)$. Intuitively, akin to O'Donoghue and Rabin (1999a, 2001) as well as the subsequent literature on behavior of sophisticated time-inconsistent agents, a time-inconsistent type may prefer to delay the completion of the task by one period rather than to complete it now. Note that this result itself is not procrastination, as the agent has correct beliefs about when the task will be completed and her actual decisions follow the beliefs. If she has not yet acquired information, then she acquires it because both types will complete the task in the next period, and hence $\hat{U}^I(s_{T-1} = 1) = U^I(x;1) > 0 = \hat{U}^I(s_{T-1} = 0)$.

Third, consider a continuation game in $t = T - 2$ in which the time-inconsistent type has not acquired information. If she chooses to acquire it now, she anticipates that she will complete the task in $t = T - 1$ with probability q and do so in $t = T$ with probability $1 - q$, and hence $\hat{U}^I(s_{T-2} = 1) = qU^I(x;1) + (1-q)U^I(x;2)$. If she does not acquire information now, she anticipates that she will complete the task in $t = T$ with probability one, and hence $\hat{U}^I(s_{T-2} = 0) = U^I(x;2)$. Because $U^I(x;2) = -D(2)c + D(3)b = \frac{1}{10} + \frac{2}{5}\epsilon > U^I(x;1)$, she does not acquire information in $t = T - 2$. Intuitively, a time-inconsistent type prefers to delay acquiring information in order to delay the completion of the task.

Finally, we show the condition under which the time-inconsistent type *procrastinates* acquiring information. Consider continuation games in $t \leq T-3$, in which the time-inconsistent type has not acquired information. If she acquires it now, she anticipates that she will complete the task in the next period with probability q and do so in some future period with probability $1-q$.[18] Because $\max_\tau U^I(x;\tau) = U^I(x;2)$ in this case, her anticipated expected utility when she acquires information is at most $qU^I(x;1) + (1-q)\max_\tau U^I(x;\tau) = qU^I(x;1) + (1-q)U^I(x;2)$. If she does not acquire information now, she anticipates that she will do so in the next period and will complete the task in the next two periods with probability q. Hence, her anticipated expected utility when she does not acquire information is at least $qU^I(x;2)$. This implies that if $qU^I(x;2) > qU^I(x;1) + (1-q)U^I(x;2)$, or if $q > \frac{6}{7}$ when ϵ approaches 0, the time-inconsistent type will choose not to acquire information in any $t \leq T-3$.[19] Hence, if $q > \frac{6}{7}$, the time-inconsistent type procrastinates acquiring information until $t = T-2$ on the equilibrium path.

Intuitively, because a time-inconsistent type is (probabilistically) overconfident about own future self-control, she erroneously thinks that she will acquire information with probability q in the next period, even when she does not do so now. Hence, as in O'Donoghue and Rabin (1999a, 2001), the time-inconsistent type may procrastinate information acquisition — she thinks that she will acquire information with probability q in the next period, but she actually will not do so.[20] However, in contrast to O'Donoghue and Rabin (1999a, 2001), the agent in our model is never surprised by a zero-probability event on the equilibrium path, even when she keeps procrastinating.

If the time-inconsistent type is sufficiently naive (i.e., q is sufficiently close to 1), she (erroneously) believes that time inconsistency is of little relevance to her. Hence, she be-

[18] Here, for simplicity of exposition, we do not describe exact equilibrium strategies and beliefs of the time-inconsistent type. See Appendix B for the full characterization.

[19] For a sufficiently small but positive $\epsilon > 0$, the condition becomes $q + \frac{7}{24}(2q-1)\epsilon > \frac{6}{7}$.

[20] Note that, for continuation games in $t \leq T-3$ in which the time-inconsistent type has not acquired information, the beliefs for her future learning decisions are *wrong*: in reality, if she does not learn about her own type in period $t = T-3$, she will not do so in period $t = T-2$ either (because her type is persistent over time). However, she is naive about it exactly because she has not learned about it yet (and hence keeps using her initial prior belief).

lieves that if she acquires information now, she is almost certainly going to complete the task tomorrow, and if she does not acquire it now, she believes that she will most likely acquire it in the next period and then complete the task in the next two periods. The naive time-consistent type (erroneously) believes that this would be the main trade-off she is facing: acquiring information now and completing the task tomorrow, or acquiring information tomorrow and completing the task in two periods. This leads to procrastination if the latter option outweighs the former one.

Perhaps surprisingly, and beyond the original logic of O'Donoghue and Rabin (1999a, 2001), procrastination in our model can occur even when $m = 0$. This is because the conflict between future selves exists if $U^I(x; 2) > U^I(x; 1)$, which never arises under quasi-hyperbolic discounting but can occur under hyperbolic discounting (as well as other functional forms of declining impatience over time).

3.3 Analysis in an Infinite Horizon Model

In this subsection, we analyze a model with $T = \infty$ and show under which conditions a time-inconsistent agent will procrastinate learning indefinitely, which implies never completing the task on the equilibrium path. Specifically, we look for an equilibrium in which: I) once the agent learns about her own self-control, both types will complete the task in the next period; II) the agent always acquires the signal when she is time-consistent; and III) the agent will never acquire information when she is time-inconsistent.

We first specify continuation games of type I: the agent's behavior in continuation games in which the agent has already learned her type. If the agent has learned that she is time-consistent, she chooses to complete the task in any period, because Assumption 1 (i) implies $U^C(x; 0) = -c + \delta b > 0$. If the agent has learned that she is time-inconsistent, Assumption 1 (i) ensures that there exists an equilibrium in continuation games in which the agent always completes the task in the next period after learning, as shown in the proof. When showing the existence of a non-learning equilibrium, we focus on such an equilibrium in continuation games. We then derive the condition in which non-learning for time-inconsistent agents is a unique equilibrium outcome.

We next investigate the agent's behavior regarding learning about her own self-control. Consider continuation games of type II, where the agent is time-consistent. In this case, Assumption 1 (ii) ensures that she chooses to learn about her own self-control in our candidate equilibrium.

Now suppose continuation games of type III, where the agent is time-inconsistent. Note that her beliefs about her own future actions depend on her beliefs about her own future self-control. On the one hand, the agent (erroneously) believes that she will be time-consistent and hence will acquire information in the next period, with probability q. On the other hand, she (correctly) anticipates that she will not acquire information in any future period, with probability $1-q$, in our candidate equilibrium. To show the existence of this equilibrium, we first focus on an equilibrium in continuation games in which the time-inconsistent type always completes the task in period $t+1$ if she learns about own self-control in period t. Given these beliefs, if the time-inconsistent type chooses not to learn, her anticipated expected utility is:

$$\hat{U}_t^I(s_t = 0) = q\left[-D(1)m + U^I(x;2)\right] + (1-q)\cdot 0.$$

In contrast, if the time-inconsistent type chooses to learn, her anticipated expected utility is:

$$\hat{U}_t^I(s_t = 1) = -m + U^I(x;1).$$

Hence, there exists an equilibrium in which the time-inconsistent type prefers not to learn in each period if:

$$\begin{aligned} & \hat{U}_t^I(s_t = 0) > \hat{U}_t^I(s_t = 1) \\ \Longleftrightarrow \quad & q\left[U^I(x;2) - U^I(x;1)\right] + m\left[1 - qD(1)\right] > (1-q)U^I(x;1). \end{aligned} \tag{1}$$

Next, we derive the condition under which a time-inconsistent type never chooses to learn about her self-control problems — and hence never completes the task — in any equilibrium outcome. In the proof, we show that in any continuation game, a time-consistent type will always choose to learn and complete the task immediately if

$$(q-\delta)U^C(x;1)+m(1-\delta)\geq 0. \tag{2}$$

Suppose continuation games where a time-inconsistent type has not learned. Denote the maximum and minimum utility of a time-inconsistent type from completing a task by $\overline{U}^I(x)=\max_\tau U^I(x;\tau)$ and $\underline{U}^I(x)=\min_\tau U^I(x;\tau)$, respectively. Note that $\overline{U}^I(x)>0\geq \underline{U}^I(x)$ by Assumption 1 and that $\lim_{\tau\to\infty} U^I(x;\tau)=0$. Then, the lower bound of anticipated expected utility when the time-inconsistent type chooses not to learn is:

$$\underline{U}_t^I(s_t=0)=q\left[-D(1)m+U^I(x;2)\right]+(1-q)\left[-D(1)m+\underline{U}^I(x)\right].$$

To see the intuition here, note that a time-inconsistent type knows how a time-consistent type will behave. With probability q, she believes that she will be time-consistent in the future, acquire the signal, and complete the task in two periods. In contrast, however, with probability $1-q$ she believes that she will be time-inconsistent in the future. In this case, the lowest possible payoff is to pay the information acquisition cost as soon as possible but to complete the task in the least desired period. Similarly, an upper bound on anticipated expected utility when the time-inconsistent type chooses to learn is:

$$\overline{U}_t^I(s_t=1)=-m+qU^I(x;1)+(1-q)\overline{U}^I(x).$$

Hence, in any equilibrium outcome, the agent prefers not to learn if:

$$\begin{aligned}
&\underline{U}_t^I(s_t=0)>\overline{U}_t^I(s_t=1)\\
\iff\ & q\left[U^I(x;2)-U^I(x;1)\right]+m\left[1-D(1)\right]>(1-q)\left[\overline{U}^I(x)-\underline{U}^I(x)\right].
\end{aligned} \tag{3}$$

The result is summarized as follows.

Proposition 1. Suppose $T=\infty$ and Assumption 1 holds.

(i) If Inequality (1) holds, there exists an equilibrium in which a time-inconsistent type never learns about her own future self-control (and hence never completes the task).

(ii) If Inequalities (2) and (3) hold, a time-inconsistent type never learns about her own future self-control (and hence never completes the task) in any equilibrium outcome.

The intuition of the result is two-fold. First, if the agent is time-inconsistent, she may prefer to complete a task later rather than sooner. Specifically, if the time-inconsistent type prefers to complete the task in two periods rather than in one period (i.e., if $U^I(x;2) > U^I(x;1)$), then she has an incentive to delay acquiring the signal in order to delay task completion. Second, because the time-inconsistent type is (probabilistically) overconfident about her own future self-control, she overestimates the likelihood that she will acquire the signal in the future. Specifically, if the degree of naivete q is sufficiently large, the time-inconsistent type (erroneously) believes that she will most likely acquire the signal in the next period; hence, she prefers not acquiring it now, to delay task completion. Because the time-inconsistent type fails to infer from her own actions that she cannot be time-consistent in the future unless acquiring a signal about her own type, she may procrastinate indefinitely.

It is worth mentioning that procrastination on learning about own self-control can occur even when the time-inconsistent type prefers to complete the task in the next period rather than never (i.e., $U^I(x;1) > m$). Intuitively, because the time-inconsistent type is overconfident about her own future self-control, she underestimates the possibility that she will procrastinate on learning. As a result, she may prefer to postpone a valuable learning opportunity.

Furthermore, if the time-inconsistent type is sufficiently naive and prefers a delay, procrastination of learning is a unique equilibrium outcome. It extends the result in Section 3.2 in which non-learning can occur even when $m = 0$. Further, even when $m = 0$, non-learning can be a unique equilibrium outcome:

Corollary 1. Suppose $T = \infty$, Assumption 1 holds, and $U^I(x;2) > U^I(x;1)$. Then, there exists $\bar{q} \in (0,1)$ such that for any $q > \bar{q}$ a time-inconsistent type never learns about own future self-control (and hence never completes the task) in any equilibrium outcome.

4 Task Choice

In Section 3, we assumed that the agent already faces the need to complete a certain task. This section analyzes situations in which an agent first chooses a task to complete and then engages in the task; for example, one could choose to either pursue a university degree, do an apprenticeship, or postpone the decision. Formally, the agent in this section has multiple tasks, but she can choose only one task throughout the game, and the optimal choice depends on her type. In this section, we allow the agent to complete a task without acquiring a signal about her own type. As an illustration, we first focus on a case where the agent is facing two possible tasks to choose from, and then provide a more general analysis.

4.1 Analysis with a Two-Task Menu

Suppose $T = \infty$ and the agent can choose one of the following two tasks: $x = (-c, b)$ and $x' = (0, b')$ with $-c + \delta b > \delta b' > 0 > -c + D(1)b$. We also assume that $\min_{\tau \geq 0} U^I(x, \tau) = U^I(x, 0) = -c + D(1)b$. As described in Section 2.1, we assume that the task cannot be completed in the same period it is chosen. We derive the conditions for an equilibrium in which the time-inconsistent type never learns about her own self-control problems or chooses a task, whereas the time-consistent type chooses to learn about her own self-control problems and then complete task x on the equilibrium path.

First, we characterize the agent's task-completion behavior after learning her own type. It is straightforward to show that if the agent has chosen task x', she will complete that task in every period irrespective of her type. Also, if the agent has chosen task x, she will complete that task in every period if she is time-consistent and will never complete the task if she is time-inconsistent.

Second, suppose the agent has learned her own type but has not yet chosen a task. In this case, it is straightforward to show that the agent chooses x in any period if she is time-consistent and chooses x' in any period if she is time-inconsistent.

Third, suppose the agent has chosen a task but has not learned her own type. If the agent has chosen task x', she will complete the task in every period irrespective of her type.

Also, if the agent has chosen task x, she will complete the task in every period if she is time-consistent and will never complete the task if she is time-inconsistent.

Fourth, suppose that the agent has neither learned about her own type nor chosen a task and that the agent is time-consistent. Given the behavior shown in the above continuation games, the perceived expected utility of the time-consistent type for each action in our candidate equilibrium is as follows:

- Choosing x without learning own type: $qU^C(x;1) + (1-q) \cdot 0$.
- Choosing x' without learning own type: $U^C(x';1)$.
- Learning own type: $-m + qU^C(x;2) + (1-q)U^C(x';2)$.
- Not doing anything: $q(-m\delta + U^C(x;3)) + (1-q) \cdot 0$.

Note that for the time-consistent type, not doing anything is dominated by choosing x without learning own type. Hence, if the agent is time-consistent, she will choose to learn if and only if:

$$-m + qU^C(x;2) + (1-q)U^C(x';2) > \max\{qU^C(x;1), U^C(x';1)\}$$

or

$$\frac{(1-q)\delta}{1-\delta}U^C(x';0) - \frac{m}{\delta(1-\delta)} > qU^C(x;0) > \frac{1-(1-q)\delta}{\delta}U^C(x';0) + \frac{m}{\delta^2} \tag{4}$$

Note that Inequality (4) always holds for $m = 0$ and δ close to one. In what follows , we focus on the case in which Inequality (4) holds.

Finally, suppose that the agent has neither learned about her own type nor chosen a task and that she is time-inconsistent. If she chooses to learn, she expects that she will choose task x in the next period and then complete it in two periods with probability q, whereas she will choose task x' in the next period and then complete it in two periods with probability $1-q$. Hence, the perceived expected utility of the time-inconsistent type for each action in our candidate equilibrium is as follows:

- Choosing x without learning own type: $qU^I(x;1) + (1-q)\cdot 0$.

- Choosing x' without learning own type: $U^I(x';1)$.

- Learning own type: $-m + qU^I(x;2) + (1-q)U^I(x';2)$.

- Not doing anything: $-q(mD(1) + U^I(x;3)) + (1-q)\cdot 0$.

Thus, if the agent is time-inconsistent, she will choose not to do anything if

$$U^I(x;3) - mD(1) > \max\left\{U^I(x;1), \frac{1}{q}U^I(x';1),\ U^I(x;2) + \frac{1-q}{q}U^I(x';2) - \frac{m}{q}\right\} \quad (5)$$

and Inequality (4) holds.

The intuition is as follows. First, the time-inconsistent type thinks that she will acquire a signal about her own self-control if she is time-consistent. Second, because she is time-inconsistent, she may prefer to complete a task later rather than sooner. For example, if $m = 0$ and b' is sufficiently close to zero, then the time-inconsistent type has an incentive to postpone acquiring the signal to delay task completion if $U^I(x;3) > \max\{U^I(x;1), U^I(x;2)\}$. Third, because the agent is (probabilistically) overconfident about her own future self-control, she underestimates the likelihood that she will not acquire the signal in the future. The next proposition summarizes the result:

Proposition 2. Suppose $T = \infty$ and the agent faces a menu of $x = (-c, b)$ and $x' = (0, b')$. If Inequalities (4) and (5) hold, there exists an equilibrium in which a time-inconsistent type never learns about her own future self-control or chooses any task on the equilibrium path.

Proposition 2 highlights a perverse welfare effect of non-learning in our model: If the time-inconsistent type were to know that she is time-inconsistent, then she would choose task x', which would give her strictly positive utility. But because she (erroneously) will that she is time-consistent with a high probability, she believes that she would choose task x in the future with a high probability. Given this (erroneous) belief, she prefers to do nothing in order to delay completing task x. Because her type is time-inconsistent and is persistent over time, however, she forever keeps doing nothing on the equilibrium path.

Proposition 2 may help explain why people pursue implausible dreams and never actually start anything instead of taking better alternatives . As a real-world example, suppose that an agent can choose whether to pursue a university degree or do an apprenticeship. If she knew she were time-consistent, she would go to the university this year. If she knew she were time-inconsistent, she would realize that the high upfront cost of the university degree is too much for her, so she would choose the apprenticeship this year. But if she is time-inconsistent and does not know her future type, she may believe that her future selves would be time-consistent with a high probability, and hence, may plan to go to university next year. This is why she does not start the apprenticeship this year, and as a result, ends up never completing any further education.

Example 4.1. Consider the case in which $D(t) = \frac{1}{1+rt}$, $r = \frac{1}{2}$, and $\delta \simeq 1$. Then, the assumption $-c + \delta b > 0 > -c + D(1)b$ holds if and only if $\frac{b}{c} \in (1, \frac{3}{2})$.

(i) If $m = 0$, Inequality (4) always holds and Inequality (5) is equivalent to the following condition:

$$q > \max \left\{ \frac{25\frac{b'}{c}}{1 + 20\frac{b'}{c}}, \frac{15}{4}\frac{b'}{c} \right\}.$$

Hence, for $\frac{b'}{c} < \frac{1}{5}$ there exists a $\bar{q} < 1$ such that for all $q > \bar{q}$ the time-consistent type acquires information and completes task x, while a time-inconsistent type never does anything on the equilibrium path.

(ii) If $m = \frac{1}{50}c$ and $\frac{b'}{c} = \frac{1}{10}$, Inequality (4) and (5) become:

$$\frac{4}{5} > q > \max \left\{ \frac{1}{50\frac{b}{c} - 55}, \frac{15}{44} \right\}.$$

Hence, for $\frac{b}{c} \in (\frac{9}{8}, \frac{3}{2})$, there is a range for q such that the time-consistent type acquires information and completes task x, while a time-inconsistent type never does anything on the equilibrium path.

Example 4.1 (i) shows that a sufficiently naive time-inconsistent type (i.e., q is close to one) procrastinates information acquisition and task choice indefinitely if b' is small. The intuition is as follows: if q is large, the time-inconsistent type believes that x is likely to be the optimal task; therefore, choosing task x' becomes unattractive. Also, because choosing task x now will lead to the completion of the task in the next period with probability q, the time-inconsistent type might prefer to postpone choosing the task. Therefore, not doing anything is the only way to commit to postponing the completion of task x — which leads to procrastination.

Note that if b' approaches zero in Example 4.1 (i), the degree of naivete q ensures procrastination goes to zero. This means that when the value of knowing that she will be time-inconsistent is sufficiently small, the agent becomes more likely to procrastinate and stick to suboptimal decisions from her initial perspective. Similarly, if c increases while $\frac{b}{c}$ is fixed, the time-inconsistent type is more likely to procrastinate.

Example 4.1 (ii) illustrates a case in which the time-consistent type may choose a task without having acquired information under a positive m, posing additional restrictions on the degree of naivete q. Intuitively, if the time-consistent type believes that she will be a particular type with sufficiently high probability (q is close to either zero or one), the information has low perceived value and she will choose a task without learning her own type. This consideration imposes an upper bound of q for the existence of our candidate equilibrium. Interestingly, the range of q in the candidate equilibrium can be smaller in Example 4.1 (ii) than in Example 4.1 (i); that is, lower learning costs can make time-inconsistent agents more likely to procrastinate on the equilibrium path.

4.2 Analysis with a General Menu

This subsection analyzes a generalized version of the above infinite-horizon model, including more than two tasks in a menu. We look for an equilibrium in which the time-inconsistent type never learns about own type nor chooses a task on the equilibrium path.

Let $X = \{x_1, \ldots, x_N\}$ be a menu of tasks of the form $x_n = (-c_n, b_n)$ where $c_n, b_n \geq 0$ for each n. Without loss of generality, we assume that $U^C(x_n; 0) > 0$ for all i.[21] Given this, in any continuation games in which the time-consistent type of agent has already chosen

[21] Any task x_n that gives the time-consistent type negative utility would never be considered or completed by either type because of the assumption $D(1) < \delta$.

x_n, she will always choose to complete the task. We also assume that each type of agent has strict preferences for any task completed in any period: $U^k(x;t) \neq U^k(x';t')$ for any $x, x' \neq x$, $k \in \{C, I\}$, and $t, t' \geq 0$. Note that this assumption holds generically.

Let $x_0^k = \text{argmax}_{x_n} U^k(x_n; 0)$ denote the most preferred task by type k when she completes a task. Note that $x_0^C = \text{argmax}_{x_n} U^C(x_n;t)$ for any $t \geq 0$. Let $x_1^k = \text{argmax}_{x_n \text{ s.t. } U^I(x_n;0)>0} U^k(x_n;1)$ denote the most preferred task by type k that is feasible to complete for both types, given that the agent will complete a task in the next period. Because $D(1) < \delta$, $U^I(x_n;0) > 0$ implies $U^C(x_n;t) > 0$ for any t. Denote the maximum utility level of the time-inconsistent type by $\overline{U}^I = \max_{(x_n,t)} U^I(x_n;t)$. To focus on interesting cases, we assume the following:

Assumption 2 (Feasibility of Task)**.** (i) $U^C(x_0^C;0) > 0 > U^I(x_0^C;0)$, (ii) $U^I(x_0^I, 1) > 0$.

Assumption 2 implies that the time-inconsistent type will never complete the time-consistent type's preferred task and that the optimal task depends on her type.[22] Note that Assumption 2 implies $x_0^C \neq x_1^C$.

In what follows, we derive sufficient conditions for the non-learning outcome, in which the time-inconsistent type does not do anything while the time-consistent type learns own type in $t = 1$, chooses x_0^C in $t = 2$, and completes it in $t = 3$. Note again that because of the assumption $U^C(x_n;0) > 0$ for all $x_n \in X$, in any continuation games in which the time-consistent type of agent has already chosen x_n, she will always choose to complete the task. We focus on the following candidate equilibrium: for continuation games of type I, in which the agent has first learned own type and then has chosen a task x_n, the time-consistent type always completes it while the time-inconsistent type completes it immediately after choosing it if $U^I(x_n, 0) > 0$; for continuation games of type II, in which the agent has not chosen a task but has learned own type, the time-consistent type always chooses x_0^C while the time-inconsistent type always chooses task x_1^I; and for continuation games of type III, in which the agent has not done anything yet, the time-consistent type always learns own type while

[22] If Assumption 2 (i) were not satisfied, the time-consistent agent would always be better of picking x_0^C immediately without learning her type. If Assumption 2 (ii) were not satisfied, it might be the best for the agent to not complete any task.

the time-inconsistent type does not do anything.[23]

In the proof, we show that there is an equilibrium in which each type of agent takes a best response in continuation games of types I and II. In what follows, we investigate continuation games of type III.

The time-consistent type of agent prefers to learn own type immediately rather than to choose a task immediately without learning (or not doing anything) if

$$\begin{aligned} & -m + qU^C(x_0^C;2) + (1-q)U^C(x_0^I;2) \\ & > \max\left\{qU^C(x_0^C;1),\ qU^C(x_1^C;1) + (1-q)U^C(x_1^C;1)\right\}, \end{aligned} \tag{6}$$

where $qU^C(x_0^C;1)$ is the upper bound of the agent's expected payoff when she chooses a task that is not feasible for the time-inconsistent type to complete, and $qU^C(x_1^C;1) + (1-q)U^C(x_1^C;1)$ is the upper bound of the payoff when she chooses a task that is feasible for both types to complete.

Because the time-inconsistent type of agent believes that once she chooses a task she will complete it in the next period with probability q, she prefers to do nothing rather than to choose a task immediately without learning if

$$q(-mD(1) + U^I(x_0^C;3)) > qU^I(x_1^I;1) + (1-q)\overline{U}^I. \tag{7}$$

Similarly, because the time-inconsistent type of agent believes that once she learns her own type she will choose x_0^C in the next period with probability q, she prefers to do nothing rather than to acquire information if

$$q(-mD(1) + U^I(x_0^C;3)) > -m + qU^I(x_0^C;2) + (1-q)\overline{U}^I. \tag{8}$$

The result is summarized as follows:

Proposition 3. Suppose $T = \infty$ and Assumption 2 holds. If Inequalities (6), (7), and (8) also hold, there exists an equilibrium in which the time-consistent type acquires information immediately and completes her preferred task, whereas the time-inconsistent type procrastinates indefinitely.

[23] As we derive only sufficient conditions for the non-learning outcome using the upper bound $\overline{U}^I$, we do not need to explicitly specify continuation games in which the time-inconsistent type of agent has chosen a task x_n but has not learned own type yet.

For illustration, suppose that $m = 0$ and δ is close to 1. In this case, for a sufficiently large q, Inequality (6) holds: the time-consistent type always acquires information. Also, for $m = 0$ and sufficiently large q, Inequalities (7) and (8) holds if:

$$U^I(x_0^C; 3) > \max\{U^I(x_1^I; 1),\ U^I(x_0^C; 2)\}.$$

Intuitively, the time-inconsistent type thinks that she will almost surely acquire information in the next period, choose x_0^C in two periods, and complete it in three periods. If her perceived utility of completing x_0^C in three periods is higher than completing any task within two periods due to her time inconsistency, then she prefers to do nothing — leading to non-completion of any task.

5 Procrastination and Misdirected Learning

This section investigates an extension to our basic model by including learning about an additional payoff-relevant attribute, which we call ability. Suppose that there are two initially uncertain attributes: self-control and ability. Though the agent is initially uncertain about both of them, she can acquire perfectly informative signals about each attribute.

As we now have two attributes, we introduce further notation. Let $q_d \in (0, 1)$ be the probability with which the agent believes her discounting is exponential, that is, the probability that she is time-consistent, and let $(1 - q_d)$ be the probability the agent attaches to having a discount function $D(t)$. Additionally, the agent now has beliefs about her ability. We assume that there are only two ability types: high ability $a_H > 0$, and low ability $a_L < a_H$. The agent has beliefs about her ability: she believes with probability $q_a \in (0, 1)$ that she is of high ability and with probability $(1 - q_a)$ that she is of low ability.

To highlight our key mechanism and results in a simple manner, we examine a model in which the agent has an opportunity to costlessly and perfectly learn about her own ability in $t = 0$ and then plays the game described in Section 3 with $m = 0$. That is, the agent faces a task x with payoff $(-c, b_i)$ where $b_i = a_i b$. The agent's expected ability is $\hat{a} = q_a a_H + (1 - q_a) a_L$, and the expected benefit of the task is $\hat{b} = \hat{a} b$. Slightly abusing the notation, let $U_t^k(b_i;\tau-t)=-D(\tau-t)c+D(\tau-t+1)b_i$ denote a type-k agent's total utility evaluated in period t when she completes a task in period τ. We abbreviate subscript t whenever it is clear.

To focus on interesting cases, throughout this section we consider the case in which both types would acquire information about their own self-control and work on the task as soon as possible if they do not learn about their own ability. Formally, we assume that $\max_\tau U^I(\hat{b};\tau)=U^I(\hat{b};1)>0$ and $U^I(\hat{b};0)>0$. The latter condition implies $U^C(\hat{b};0)=-c+\delta\hat{b}>0$, and hence both types would acquire information about own self-control in $t=1$ and work on the task in $t=2$ if they would not learn about their own ability in $t=0$. It also implies that if each type would learn about own ability in $t=0$ and realize that own ability would be b_H, then both types would acquire information about their own self-control and work on the task as soon as possible.

We investigate two cases. First, consider the case in which $U^C(b_L;0)>0$ and $U^I(b_L;2)>0>U^I(b_L;0)$. In this case, a time-inconsistent agent never works on the task once she learns that her ability is a_L. However, from the $t=0$ perspective, both types strictly prefer to complete the task in $t=2$. Also, because by the assumption, $\max_\tau U^I(\hat{b};\tau)=U^I(\hat{b};1)>0$, both types would immediately learn about their own self-control and then complete the task if they do not learn about their own ability. That is, the agent would strictly prefer to avoid learning about her own ability to motivate her future self to complete the task, as analyzed by Carrillo and Mariotti (2000) and Bénabou and Tirole (2002).

Second, consider the case in which $U^C(b_L;0)<0$, $U^I(b_L;2)<0$, and $U^I(b_H;2)>U^I(b_H;1)$. The first condition implies that $U^I(b_L;0)<0$. In this case, an agent of either type never works on the task once she learns that her ability is a_L. We show that if $U^k(b_L;2)<-\frac{q_a(1-q_d)}{1-q_a}U^k(b_H;2)<0$ holds for each type $k\in\{C,I\}$, then both time-consistent and time-inconsistent types strictly prefer to acquire information about their own ability in $t=0$. As shown in Section 3, if $U^I(b_H;2)>U^I(b_H;1)$, a time-inconsistent type may procrastinate learning about own self-control. That is, learning about own ability can lead to non-completion of the task even when ability is b_H.

The following proposition summarizes the results:

Proposition 4. Suppose $T = \infty$, $m = 0$, $\max_\tau U^I(\hat{b};\tau) = U^I(\hat{b};1) > 0$, and $U^I(\hat{b};0) > 0$.

(i) Assume $U^C(b_L;0) > 0$ and $U^I(b_L;2) > 0 > U^I(b_L;0)$. There exists an equilibrium in which a time-inconsistent type of agent strictly prefers to not learn about own ability and both types of agent complete the task in $t = 2$.

(ii) Assume $U^C(b_L;0) < 0$, $U^I(b_H;2) > U^I(b_H;1)$, and $U^k(b_L;2) < -\frac{q_a(1-q_d)}{1-q_a}U^k(b_H;2)$ for type $k \in \{C, I\}$. There exists an equilibrium in which a time-inconsistent type of agent strictly prefers to learn about own ability but never learns about own future self-control or completes the task.

On the one hand, the result in Proposition 4 (i) is in line with the literature on strategic ignorance (Carrillo and Mariotti 2000, Bénabou and Tirole 2002). In this literature, an agent chooses not to learn about her own ability as a means of internal commitment to improve her own future payoffs. In our model, the agent chooses not to learn about own ability to commit to completing the task in $t = 2$.

On the other hand, Proposition 4 (ii) highlights that a time-inconsistent type may not learn about own future self-control. Non-learning about own self-control in this case is harmful from the period-0 perspective, because it leads to a non-completion of a task. This is different from the literature on strategic ignorance, as it focuses on learning about ability when the agent perfectly knows her own future self-control problems.

Interestingly, in contrast to Carrillo and Mariotti (2000) and Bénabou and Tirole (2002), the time-inconsistent type of agent endogenously chooses to learn about own ability, which is harmful from the initial perspective. Without learning about her own ability, an agent of this type will complete the task in $t = 2$, and her expected utility from her initial perspective will be higher than in the case when she procrastinates and hence never completes the task. Because an agent of the time-inconsistent type is initially overconfident about her own future self-control, however, she takes the same action as the time-consistent type does, which leads to procrastination in the end. In particular, the result and its welfare consequences in Bénabou and Tirole (2002) — that sophisticated time-inconsistent agents optimally choose not to learn about their own ability — can be reversed when we investigate partially naive agents.

Most importantly, Proposition 4 (ii) highlights how the agent's endogenous learning decisions can be misdirected under naivete about her own future self-control — she chooses to learn what she should not learn from her initial perspective (i.e., about her ability) and chooses not to learn what she should (her self-control). Complementing recent papers such as Heidhues et al. (2018) and Hestermann and Le Yaouanq (2020), our mechanism based on procrastination explains why an agent's learning decisions can be self-defeating even when she can voluntarily choose to learn about all uncertain attributes.

6 Further Extensions and Discussion

Leisure goods. So far, we have only considered situations in which the agent faces investment goods: tasks or goods with an immediate cost and a delayed benefit (DellaVigna and Malmendier 2004). A natural extension is to analyze the agent's learning incentives if they face leisure goods: a task or good with an immediate benefit and a delayed cost. To describe the implications of such a case in a simple manner, consider an infinite horizon model in Section 3 with $m = 0$ and a modification under which the agent now receives a benefit $b > 0$ immediately and a delayed cost $c > 0$ in the next period upon the completion of a task.

In this case, if $U^I(x; 1) < 0 < U^I(x; 0)$, an agent of the time-inconsistent type may not learn about her own future self-control to prevent herself from completing the task. Intuitively, she chooses not to learn about her own future self-control as a means of internal commitment, to avoid temptation. This result is akin to Proposition 4 (i) as well as the literature on strategic ignorance (Carrillo and Mariotti 2000, Bénabou and Tirole 2002), while the agent here would not learn about her own future self-control rather than her own ability.

Commitment to future actions. We discuss the predictions of our model when we allow for commitment technologies. First, in the infinite horizon model in Section 3, if the agent were given the opportunity to commit to acquiring information in any future period, then both types of agent would commit to doing so. Intuitively, as a time-inconsistent type would prefer to complete a task in some future period rather than never by Assumption 1, she would commit to acquiring the information in the future if such a commitment were possible. By a similar logic, if the agent were given the opportunity to commit to acquiring information in any future period, the non-learning results in Section 4 would not occur. Therefore, the assumption of non-commitment is crucial to derive our non-learning result.

Negative learning cost. While we have highlighted how an agent may not learn about her own future self-control even when the learning cost is zero (i.e., $m = 0$), this result implies that non-learning can occur for some small $m < 0$. In other words, time-inconsistent people may be willing to pay a positive cost to stay ignorant. Notably, as our underlying mechanism is procrastination regarding learning opportunities, this willingness to pay to stay ignorant can occur even in the absence of any type of self-esteem concern (Kőszegi 2006; Gottlieb 2014, 2016).

Partial naivete and sophistication. In the literature, partial naivete is often modeled as a point belief about own future self-control, specifically as $\hat{\beta} \in (\beta, 1)$ for quasi-hyperbolic discounting. To allow for the possibility of learning, we assume that an agent's belief about her own future self-control is a bi-modal distribution where the true value is in its support.

First, we discuss how we can extend our results to general non-degenerate distributions for own future self-control. They can be extended straightforwardly, as in Section 3, because the agent faces only one task and the only reason an agent cares about her own future self-control is to predict how likely it is the task will be completed and when. Specifically, for a given general non-degenerate distribution on own future self-control, an agent of either type can compute the probability of completing the task in t periods for each $t \geq 1$ when she does not learn now. Given these probabilities, each type of agent maximizes own perceived utility. By contrast, the analysis in Section 4 under general non-degenerate distributions on own future self-control can be much more complicated, as different types can choose different tasks at different points in time. The analysis in such cases is left for future research.

Second, there are other ways of modeling partial naivete even within a bi-modal distribution. For example, instead of an agent's type being constant over time, agents can have stochastic preferences in each period. Specifically, the true probability of being time-consistent in each period is $q_L < 1$ for the time-inconsistent type and $q_H = 1$ for the time-consistent type. The realization is i.i.d. across time, but the time-inconsistent type is uncertain about the probability of being time-consistent in a given period. An agent of this type believes with probability $q \in (0, 1)$ that she will be time-consistent with probability 1, and with probability $1 - q$ that she will be time-consistent with probability q_L. In a given period, she believes she will be time-consistent in future periods with probability $\hat{q} \equiv q \cdot q_H + (1 - q)q_L = q_L + (1 - q_L)q > q_L$. Defining partial naivete in this way is akin to the approach in models of cue-based consumption, where agents are facing stochastic temptation and are overestimating how well they will withstand it. In the limit case where $q = 0$, the two types are identical to the ones in our model, in both actions and beliefs. If $q_L > 0$, that is, both types of agent are time-consistent in any given period with a positive probability, the beliefs are distinct from those in our setting. Our main results would not be qualitatively changed by this alternative specification, however.

Last but not least, it could make more sense in many situations to assume that a time-consistent type is certain about own future self-control from the beginning of the game, whereas a time-inconsistent type is initially (stochastically) overconfident about own future self-control. Given the assumption that a time-inconsistent type cannot infer own type without acquiring information, our results would qualitatively be the same under this specification, because the main force of our results is that an agent of the time-inconsistent type wants to procrastinate on her learning opportunity, to delay completing the task.

7 Concluding Remarks

This paper provides a new mechanism explaining why people do not learn about their self-control problems over time. We find that individuals may procrastinate on a free learning opportunity indefinitely, even though learning about her own self-control problems would make the agent better off. Notably, our results do not rely on agents having image concerns or on the use of overconfidence as commitment. In our model, naivete about own self-control problem is often harmful.

One potential future direction of research is applying our mechanism to situations that involve strategic interactions between players. For example, when firms offer contracts which specify a base contract and a set of options, they often have an incentive to make consumers procrastinate on cancelling or switching the option when a fraction of consumers are naive (DellaVigna and Malmendier 2004, Heidhues and Kőszegi 2010, Murooka and Schwarz 2018). Our results imply that firms may have an incentive to let consumers endogenously procrastinate on learning about their own naivete by offering such "exploitative" contracts. How procrastination on learning about own self-control problems may interact with strategic concerns of other parties is left for future research.

References

Ali, Nageeb S., "Learning Self-Control," *Quarterly Journal of Economics*, 2011, *126* (2), 857–893.

Ariely, Dan and Klaus Wertenbroch, "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science*, 2002, *13* (3), 219–224.

Augenblick, Ned and Matthew Rabin, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," *Review of Economic Studies*, 2019, *86* (3), 941–975.

Bénabou, Roland, "The Economics of Motivated Beliefs," *Revue d'Economie Politique*, 2015, *125* (5), 665–685.

____ **and Jean Tirole**, "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

____ **and** ____ , "Willpower and Personal Rules," *Journal of Political Economy*, 2004, *112* (4), 848–886.

____ **and** ____ , "Mindful Economics: The Production, Consumption, and Value of Beliefs," *Journal of Economic Perspectives*, 2016, *30* (3), 141–164.

Benjamin, Daniel J., Matthew Rabin, and Collin Raymond, "A Model of Nonbelief in the Law of Large Numbers," *Journal of the European Economic Association*, 2016, *14* (2), 515–544.

Carrillo, Juan D. and Thomas Mariotti, "Strategic Ignorance as a Self-Disciplining Device," *Review of Economic Studies*, 2000, *67* (3), 529–544.

DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 2009, *47* (2), 315–372.

___ **and Ulrike Malmendier**, "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

___ **and** ___ , "Paying Not to Go to the Gym," *American Economic Review*, 2006, *96* (3), 694–719.

Echenique, Federico, Taisuke Imai, and Kota Saito, "Testable Implications of Models of Intertemporal Choice: Exponential Discounting and Its Generalizations," *American Economic Journal: Microeconomics*, 2020. Forthcoming.

Esponda, Ignacio and Demian Pouzo, "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica*, 2016, *84* (3), 1093–1130.

Fischer, Dominik, "Motivation by Naivete," 2018. Working Paper.

Fudenberg, Drew, Gleb Romanyuk, and Philipp Strack, "Active Learning with a Misspecified Prior," *Theoretical Economics*, 2017, *12* (3), 1155–1189.

Gagnon-Bartsch, Tristan, Matthew Rabin, and Joshua Schwartzstein, "Channeled Attention and Stable Errors," 2018. Working Paper.

Gottlieb, Daniel, "Imperfect Memory and Choice under Risk," *Games and Economic Behavior*, 2014, *85*, 127–158.

___ , "Will You Never Learn? Self Deception and Biases in Information Processing," 2016. Working Paper.

Heidhues, Paul and Botond Kőszegi, "Exploiting Naivete about Self-Control in the Credit Market," *American Economic Review*, 2010, *100* (5), 2279–2303.

___ **,** ___ **, and Philipp Strack**, "Unrealistic Expectations and Misguided Learning," *Econometrica*, 2018, *86* (4), 1159–1214.

Herweg, Fabian and Daniel Müller, "Performance of Procrastinators: On the Value of Deadlines," *Theory and Decision*, 2011, *70* (3), 329–366.

Hestermann, Nina and Yves Le Yaouanq, "Experimentation with Self-Serving Attribution Biases," *American Economic Journal: Microeconomics*, 2020. Forthcoming.

Kőszegi, Botond, "Ego Utility, Overconfidence, and Task Choice," *Journal of the European Economic Association*, 2006, *4* (4), 673–707.

Murooka, Takeshi and Marco A. Schwarz, "The Timing of Choice-Enhancing Policies," *Journal of Public Economics*, 2018, *157*, 27–40.

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

___ **and** ___ , "Incentives for Procrastinators," *Quarterly Journal of Economics*, 1999, *114* (3), 769–816.

___ **and** ___ , "Choice and Procrastination," *Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

___ **and** ___ , "Procrastination on Long-Term Projects," *Journal of Economic Behavior & Organization*, 2008, *66* (2), 161–175.

Schwartzstein, Joshua, "Selective Attention and Learning," *Journal of the European Economic Association*, 2014, *12* (6), 1423–1452.

Schweighofer-Kodritsch, Sebastian, "Time Preferences and Bargaining," *Econometrica*, 2018, *86* (1), 173–217.

Yaouanq, Yves Le and Peter Schwardmann, "Learning about One's Self," 2019. Working Paper.

A Appendix: Proofs

A.1 Proof of Proposition 1.

(i) We derive conditions for an equilibrium in which the agent always acquires information about her own type when she is time-consistent, never when she is time-inconsistent, and agents of both types will complete the task in the next period once they know their own type.

First, we investigate continuation games, in which the agent has already learned her type. If the agent has learned that she is time-consistent, she always chooses to complete the task, because of Assumption 1 (i), which implies $U^C(x;0) = -c + \delta b > 0$. Suppose, in contrast, that the agent has learned that she is time-inconsistent. Because $D(t)$ is decreasing in t, $-cD(t) + bD(t+1)$ approaches zero as t goes to $+\infty$, Assumption 1 (i) ensures that there exists a $\tau \in \mathbb{N}$ such that $U^I(x;0) = -c + D(1)b \geq -cD(\tau) + D(\tau+1)b = U^I(x;\tau)$. Let denote $\underline{\tau} = \min\{\tau | U^I(x;0) \geq U^I(x;\tau)\}$. Consider an equilibrium candidate in continuation games in which, if the agent has learned in period τ that her type is time-inconsistent, she will complete the task in periods $t = \tau + 1 + n\underline{\tau}$, where $n \in \mathbb{N}_0$. Then, akin to O'Donoghue and Rabin (1999a, 2001), by the definition of $\underline{\tau}$, the agent strictly prefers not to complete the task in $t \neq \tau + 1 + n\underline{\tau}$ and weakly prefers to complete it in $t = \tau + 1 + n\underline{\tau}$. Hence, there exists an equilibrium in continuation games in which the time-inconsistent type always completes the task in the next period after learning.

Second, we investigate continuation games in which the agent is of the time-consistent type and in which she has not acquired the signal. In this case, the expected payoff of her acquiring information immediately in our candidate equilibrium is:

$$\hat{U}^C(s_0 = 1) = -m + U^C(x;1).$$

If she does not do so, in our candidate equilibrium she believes that she will acquire information in the next period if she is a time-consistent type and that she will never do so if she is a time-inconsistent type. Hence, her expected payoff is:

$$\hat{U}^C(s_0 = 0) = q(-m\delta + U^C(x;2)) + (1-q) \cdot 0 = q\delta(-m + U^C(x;1)).$$

Under Assumption 1 (ii), such an agent always prefers to acquire information immediately in our candidate equilibrium.

Continuation games in which the agent is a time-inconsistent type and in which she has not acquired the signal are characterized in the main text. □

(ii) Note that in any continuation game in which the agent is a time-consistent type and in which she has acquired the signal, she always chooses to complete the task, under Assumption 1 (i).

We here prove that in continuation games in which an agent of a time-consistent type has not learned her type, she always chooses to learn. Given the above behavior, note that the expected payoff of her acquiring information immediately is at least $-m + qU^C(x;1) + (1-q)U^I(x;\underline{\tau}) > -m + qU^C(x;1)$, whereas the expected payoff of her not acquiring information is at most $-\delta m + U^C(x;2) = -m\delta + \delta U^C(x;1)$. Hence, in any such continuation games, the time-consistent type always chooses to learn if $(q-\delta)U^C(x;1) + m(1-\delta) \geq 0$.

Continuation games in which the agent is a time-inconsistent type and in which she has not acquired the signal are characterized in the main text. □

A.2 Proof of Proposition 2.

Provided in the main text. □

A.3 Proof of Proposition 3.

We derive conditions for the equilibrium outcome described in the main text. Specifically, we investigate the following candidate equilibrium: for continuation games of type I, in which the agent has first learned her own type and then has chosen a task x_n, the time-consistent type always completes the task while the time-inconsistent type completes it in the next period after choosing the task if $U^I(x_n, 0) > 0$ and never completes it if $U^I(x_n, 0) \leq 0$; for continuation games of type II, in which the agent has not chosen a task but has learned her own type, the time-consistent type always chooses x_0^C while the time-inconsistent type always chooses task x_1^I; and for continuation games of type III, in which the agent has not done anything yet, the time-consistent type always learns own type while the time-inconsistent type does not do anything.

It is straightforward to check that the time-consistent type takes the best response in each continuation game of types I and II. Suppose a continuation game of type I with the time-inconsistent type of agent. When $U^I(x_n, 0) \leq 0$, there is an equilibrium in the continuation game in which the agent never completes the task. When $U^I(x_n, 0) > 0$, by the same logic as O'Donoghue and Rabin (1999a, 2001) and the above analysis, she will complete the task in every $\underline{\tau}_n$ periods where $\underline{\tau}_n = \min\{\tau | U^I(x_n; 0) \geq U^I(x_n; \tau)\}$. Hence, there exists an equilibrium in continuation games in which the time-inconsistent type always completes the task in the next period after choosing a task if she has already learned her own type. Given this, if the time-inconsistent type has learned her own type but has not chosen a task, her best response is to choose x_1^I.

In continuation games in which the time-consistent type of agent has chosen task x_n but has not learned her own type, by $U^C(x_n, 0) > 0$ for all $x_n \in X$ and she can complete the task without learning, she always chooses to complete it. In continuation games in which an agent of the time-inconsistent type has chosen task x_n but has not learned her own type, although equilibrium behavior depends on parameters, by construction, her payoff is at most $\overline{U}^I$ in any equilibrium and in any period.

The conditions for continuation games of type III are characterized in the main text. □

A.4 Proof of Proposition 4.

(i) As in the main text, the assumption $U^I(b_L; 0) < 0$ ensures that an agent of a time-inconsistent type never works on the task once she learns that her ability is a_L. However, by the assumption $U^I(b_L; 2) > 0$, from the $t = 0$ perspective, agents of both types prefer to complete the task in two periods rather than never. Also, the assumption $U^I(\hat{b}; 0) = -c + D(1)\hat{b} > 0$ implies that $-c + \delta\hat{b} > 0$ and $\max_\tau U^C(\hat{b}; \tau) = U^C(\hat{b}; 0)$. Hence, if the agent does not learn about her own ability in $t = 0$, then agents of both types would immediately learn about their own self-control in $t = 1$ and complete the task in $t = 2$. Therefore, such an equilibrium exists. $\square$

(ii) We look for an equilibrium in which an agent of a time-consistent type acquires information about own ability and then works on the task if and only if she learns that her ability is a_H while an agent of a time-inconsistent type acquires information about her own ability and never works on the task on the equilibrium path, and both types will work on the task as soon as possible if they do not acquire information about their own ability in $t = 0$.

Note first that agents of both types never work on the task once they learn that their ability is a_L. We here show that an agent of a time-consistent type chooses to learn about her own ability in $t = 0$ and then completes the task if and only if her ability is a_H. Given the above strategies, her anticipated expected utility in $t = 0$ when she chooses to learn about her own ability is:

$$q_a \left[q_d \cdot U^C(b_H; 2) + (1 - q_d) \cdot 0 \right] + (1 - q_a) \cdot 0 = q_a q_d U^C(b_H; 2) > 0.$$

Because of the assumptions, agents of both types complete the task in $t = 2$ if they choose not to learn about their own ability in $t = 0$. Their anticipated expected utility in $t = 0$ when they choose not to learn about their own ability is:

$$q_a \cdot U^C(b_H; 2) + (1 - q_a) \cdot U^C(b_L; 2).$$

Hence, an agent of the time-consistent type strictly prefers to learn about her own ability in $t = 0$ if $U^C(b_L; 2) < -\frac{q_a(1-q_d)}{1-q_a} U^C(b_H; 2)$.

We now show that the time-inconsistent type strictly prefers to acquire information about own ability in $t = 0$. In $t = 0$, given the above strategies, an agent of this type's anticipated expected utility when she chooses to learn about her own ability is:

$$q_a \left[q_d \cdot U^I(b_H; 2) + (1 - q_d) \cdot 0 \right] + (1 - q_a) \cdot 0 = q_a q_d \cdot U^I(b_H; 2).$$

Because of the assumption $\max_\tau U^I(\hat{b}; \tau) = U^I(\hat{b}; 1) > 0$, her anticipated expected utility in

$t = 0$ when she chooses not to learn about her own ability is:

$$q_a \cdot U^I(b_H; 2) + (1 - q_a) \cdot U^I(b_L; 2).$$

Hence, the time-inconsistent type strictly prefers to learn about her own ability in $t = 0$ if $U^I(b_L; 2) < -\frac{q_a(1-q_d)}{1-q_a} U^I(b_H; 2)$.

From $t = 1$ on, notice that if Inequality (1) with $b = b_H$ and $m = 0$ holds, or equivalently if $U^I(b_H; 2) > U^I(b_H; 1) > 0$, then the equilibrium behavior for each type in $t \geq 1$ is exactly the same as in Section 3. □

B Task Completion: Analysis in a Finite Horizon Model

This section analyzes when procrastination occurs in a task completion setting under a finite time horizon (i.e., $T < \infty$). It corresponds to environments in which there is a deadline for task completion (e.g., filing a tax return).

Decisions after learning. By Assumption 1, the task is worthwhile for both types of agents. Because of the properties of finite horizon decision problems, if the agents have learned their type, they will complete the task at latest in the last period. Let $\underline{\tau} = \min\{\tau | U^I(x; 0) \geq U^I(x; \tau)\}$ denote the earliest period in which the agent prefers to complete the task immediately rather than in $\underline{\tau}$ periods; by Assumption 1, both types of agents then complete the task in $t = T$ if they have not yet done so. When the task will be completed is summarized in the following lemma:

Lemma B.1. Suppose $T < \infty$ and Assumption 1 holds. Consider a continuation game in period $\tau \geq 1$ in which the agent has already learned her type. Then,

(i) If the agent is time-consistent, she completes the task immediately.

(ii) If the agent is time-inconsistent, she completes the task in the first period $t \geq \tau$ that satisfies $t = T - n\underline{\tau}$, where $n \in \mathbb{N}_0$.

Proof. For (i), note that $U^C(x;k) = \delta^k U^C(x;0)$ for the time-consistent type. Hence, in any period t, the agent will always complete the task immediately if she is time-consistent and the task has not yet been completed.

We now prove (ii). In $t = T$, the payoff from completing the task for a time-inconsistent agent is $U^I(x;0)$. If she does not complete the task in $t = T$, she receives zero payoffs. By Assumption 1, the agent prefers to complete the task.

In $t = T - 1$, the agent knows that the task will be completed in $t = T$ if it is not completed in the current period. Hence, she completes the task if $U^I(x;0) \geq U^I(x;1)$. By induction applied in each $t \leq T - 2$, the agent will always complete the task immediately when $U^I(x;0) \geq U^I(x;1)$.

If $U^I(x;0) < U^I(x;1)$, the sophisticated time-inconsistent type would not complete the task in $t = T - 1$, instead postponing task completion until $t = T$. If this is the case, then in $t = T - 2$ the agent knows that not completing the task now means that the task will be completed two periods from now; hence, the agent will complete the task in $t = T - 2$ if and only if $U^I(x;0) \leq U^I(x;2)$. If this inequality does not hold, then the agent keeps postponing task completion until period $t = T - \tau$ that satisfies $U^I(x;0) < U^I(x;\tau)$. If there does not exist such a $\tau \leq T - 1$, then no matter when she has learned about own self-control, she postpones completing the task until $t = T$. If such a $\tau \leq T - 1$ does exist, then she postpones at most τ periods.[24] Because $\underline{\tau} = \min\{\tau | U^I(x;0) \geq U^I(x;\tau)\}$, the agent knows that she will complete the task in $t = T$ and in $t = T - \underline{\tau}$ if $\underline{\tau} < T$.

In $t = T - \underline{\tau} - 1$, the agent knows that if she postpones the task, it will be completed in the next period, and she will therefore postpone to complete the task now if and only if $\underline{\tau} > 1$. As described above, she keeps postponing for $\underline{\tau}$ periods. By induction, the agent will complete the task in period $t \geq 1$ if and only if $t \in \{t \in \mathbb{N} | t = T - n\underline{\tau},\, n \in \mathbb{N}_0\}$. This completes Lemma B.1. □

Lemma B.1 (ii) implies that a time-inconsistent type completes the task in every $\underline{\tau}$ periods. The intuition underlying it is as follows. Because the agent is perfectly sophisticated

[24] From Assumption 1, we know that if T is large enough, there exists $\tau < T$ such that the agent completes the task immediately rather than in τ periods.

about her future self-control after learning, she knows she will complete the task in the last period and therefore not complete it in the second-last period if $U^I(x;0) < U^I(x;1)$. In the third-last period, she will complete the task if she prefers completing it immediately over completing it in two periods (i.e., $U^I(x;0) \geq U^I(x;2)$); otherwise, she will not. By induction, the agent completes the task in any period $t = T - n\underline{\tau}$, and there is a cycle of length $\underline{\tau}$. Note that this result does not constitute procrastination, as the agent has a correct belief about when the task will be completed, and her actual decisions coincide with the beliefs. This result is akin to that of O'Donoghue and Rabin (1999a, 2001) as well as subsequent literature on the behavior of sophisticated time-inconsistent agents .

Learning decisions. Next, we investigate conditions under which each type of agent acquires the signal. We focus on deriving the conditions for our equilibrium of interest: one in which the time-inconsistent type procrastinates on information acquisition while the time-consistent type always acquires information.

We first analyze the actions of the time-consistent type. Note that in this case she is underconfident about her own future self-control: she thinks she will be time-inconsistent with probability $1-q$. Nevertheless, if $U^I(x;1) \leq U^I(x;0)$, she will complete the task in the next period with probability one. Otherwise, she takes into account the delay in case her future self will be time-inconsistent, and she will complete the task only in period $t = T - n\underline{\tau}$. In this case, the most stringent condition on acquiring information is that the agent expects a delay of $\underline{\tau}$ periods if her future self is time-inconsistent. To be precise, if $U^I(x;1) > U^I(x;0)$, the time-consistent type acquires information if the following inequality holds:

$$q(1-\delta)U^C(x;1) \geq (1 - \delta q - (1-q)\delta^{\underline{\tau}-1})m. \tag{9}$$

The result is summarized in the following lemma:

Lemma B.2. Suppose $T < \infty$ and Assumption 1 holds. If an agent is time-consistent, she will acquire information in any continuation game if either of the following holds:

(i) $U^I(x;1) \leq U^I(x;0)$,

(ii) $U^I(x;1) > U^I(x;0)$ and Inequality (9) holds.

Proof. (i) If $U^I(x;1) \leq U^I(x;0)$, both time-inconsistent and time-consistent types will complete the task immediately upon learning. Therefore, the payoff from acquiring information for the time-consistent agent is $\hat{U}_t^C(s_t = 1) = -m + U^C(x;1) > 0$, where the last inequality is from Assumption 1.

We show that, if the time-consistent agent postpones information acquisition, she will get a lower payoff no matter what her beliefs are about her future information acquisition. If she believes a time-consistent type will acquire information in k periods and a time-inconsistent type in h periods, her payoff will be:

$$\begin{aligned}\hat{U}_t^C(s_t = 0) &= q(-\delta^k m + U^C(x;k+1)) + (1-q)(-\delta^h m + U^C(x;h+1)) \\ &= (q\delta^k + (1-q)\delta^h)(-m + U^C(x;1)) < -m + U^C(x;1).\end{aligned}$$

Therefore, if $U^I(x;1) \leq U^I(x;0)$, then the time-consistent agent always completes the task immediately.

(ii) If $U^I(x;1) > U^I(x;0)$, the time-inconsistent type may postpone completing the task. From Lemma B.1, the time-inconsistent type will postpone the task for cycles of $\underline{\tau}$ periods. In $t = T-1$, the agent knows the task will be completed in the next period. In $t = T-2$, the time-inconsistent type believes that the task will be completed in $t = T-1$ with probability q. Therefore, she will acquire information if:

$$\begin{aligned}&\hat{U}_{T-2}^C(s_{T-2} = 1) = -m + qU^C(x;1) + (1-q)U^C(x;2) \geq -\delta m + U^C(x;2) = \hat{U}_{T-2}^C(s_{T-2} = 0) \\ \Leftrightarrow \quad & U^C(x;1) \geq \frac{m}{q}.\end{aligned}$$

By the same logic, in $t = T - \tau$ where $\tau \leq \underline{\tau}$, the agent knows that the time-inconsistent type will postpone until $t = T$ to complete the task and until $t = T-1$ to acquire information. Hence, the time-consistent type prefers acquiring information immediately if

$$\begin{aligned}&-m + qU^C(x;1) + (1-q)U^C(x;\tau) \geq q(-\delta m + U^C(x;2)) + (1-q)(-\delta^{\tau-1} m + U^C(x;\tau)) \\ \Leftrightarrow \quad & U^C(x;1) \geq \frac{[1 - q\delta - (1-q)\delta^{\tau-1}]m}{(1-\delta)q}. \qquad (10)\end{aligned}$$

The right-hand side of Inequality (10) is increasing in τ. Because $\underline{\tau}$ is the longest period in which the time-inconsistent agent will delay completing the task upon learning her type, the time-consistent agent will always complete the task immediately if:

$$U^C(x;1) \geq \frac{[1 - q\delta - (1-q)\delta^{\underline{\tau}-1}]}{(1-\delta)q} m.$$

This completes Lemma B.2. □

Condition (i) in Lemma B.2 ensures that both types will complete the task immediately upon learning their type, and hence that the time-consistent agent will acquire information immediately to maximize her expected payoff. Condition (ii) in Lemma B.2 covers the case where the time-inconsistent agent postpones task completion in some periods. In this case, the time-consistent agent is facing a trade-off: acquiring information immediately leads to incurring cost m with certainty, but to completing the task in the next period only with probability q. If Condition (ii) in Lemma B.2 is satisfied, however, the time-consistent agent will not postpone acquiring information in any period. Note that Lemma B.2 holds when m is close to zero and/or q is close to one.

We next analyze the actions of the time-inconsistent type. First, we illustrate conditions in which the time-inconsistent type delays acquiring information in $t = T-2$. Because the last period for an agent of this type to acquire information in order to complete the task is in $t = T-1$, Assumption 1 (ii) and Lemma B.1 ensure that the agent will acquire information in $t = T-1$. Given Lemma B.2, acquiring information in $t = T-2$ will lead the time-consistent type to complete the task in $t = T-1$, but the time-inconsistent type will delay the task until $t = T$ if $U^I(x;1) > U^I(x;0)$. Hence, for the time-inconsistent agent in $t = T-2$, the subjective expected payoff from acquiring information immediately is

$$\hat{U}^I_{T-2}(s_{T-2} = 1) = -m + qU^I(x;1) + (1-q)U^I(x;2).$$

If the time-inconsistent agent does not acquire information, she (correctly) anticipates that she will acquire it in the next period and complete the task with certainty in two periods. Hence, the subjective expected payoff from not acquiring information in $t = T-2$ is

$$\hat{U}^I_{T-2}(s_{T-2} = 0) = -mD(1) + U^I(x;2).$$

Therefore, the time-inconsistent agent will prefer not to acquire information in $t = T - 2$ if

$$\begin{aligned} & -mD(1) + U^I(x;2) > -m + qU^I(x;1) + (1-q)U^I(x;2) \\ \Leftrightarrow \quad & q(U^I(x;2) - U^I(x;1)) + m(1 - D(1)) > 0. \end{aligned} \tag{11}$$

Note that she will always prefer to delay if $U^I(x;2) > U^I(x;1)$. Intuitively, a time-inconsistent type indirectly incurs a cost from learning in this case, because learning itself changes the behavior or future selves.

We now show the condition under which the time-inconsistent type *procrastinates* on acquiring information in $t = T - 3$. In period $t = T - 3$, she wrongly believes that she will acquire information in $t = T - 2$ with positive probability. Suppose that $U^I(x;2) > U^I(x;1) > U^I(x;0)$ and that Inequality (9) holds. If so, upon acquiring information in $t = T - 3$, the agent believes that she will complete the task in the next period if and only if her future type will be time-consistent. Hence,

$$\hat{U}^I_{T-3}(s_{T-3} = 1) = -m + qU^I(x;1) + (1-q)U^I(x;3).$$

If she does not acquire information in $t = T - 3$, she (erroneously) expects that she will acquire information in the next period with probability q. Hence,

$$\hat{U}^I_{T-3}(s_{T-3} = 0) = q(-mD(1) + U^I(x;2)) + (1-q)(-mD(2) + U^I(x;3)).$$

Note that the beliefs for this agent's future learning decisions are *wrong*: in reality, if she does not learn her own type in $t = T - 3$, she will not do so in $t = T - 2$, because her type is persistent over time. However, she is naive about this fact precisely because she has not acquired information about her own type yet, and hence, keeps using her initial prior belief. Hence, the agent will not acquire information in $t = T - 3$, while believing that she will acquire information in $t = T - 2$ with probability q if

$$\begin{aligned} & \hat{U}^I_{T-3}(s_{T-3} = 0) > \hat{U}^I_{T-3}(s_{T-3} = 1) \\ \Leftrightarrow \quad & q(U^I(x;2) - U^I(x;1)) + m(1 - qD(1)) > m(1-q)D(2). \end{aligned} \tag{12}$$

Note that the logic of procrastination here — a time-inconsistent type (in our model, probabilistically) overestimates own future self-control and hence takes no action now — is akin to that one in O'Donoghue and Rabin (1999a, 2001).

The logic and derivations are the same in periods $t < T-3$, although there are additional conditions depending on $\underline{\tau}$. The result for procrastination of learning decisions for the time-inconsistent type of agent until $t = T-1$ is summarized as follows:

Proposition B.1. Suppose $T < \infty$ and Assumption 1 holds. If $U^I(x;2) > U^I(x;1) > U^I(x;0)$, Inequality (9), and

$$q(U^I(x;2) - U^I(x;1)) + m(1 - qD(1)) > \max_{n,k}\,(1-q)[mD(n\underline{\tau}) - U^I(x;n\underline{\tau}+k) + U^I(x;k)]$$

where $n \in \mathbb{N}_0$, $k \in \mathbb{N}$, $k < \underline{\tau}$, and $n\underline{\tau} + k \leq T$ also hold, then there exists a unique equilibrium outcome in which a time-inconsistent type will procrastinate learning about own future self-control until $t = T-1$.

Proof. If $U^I(x;1) > U^I(x;0)$, the time-inconsistent type of agent will complete the task in periods $t = T - n\underline{\tau}$ where $n \in \mathbb{N}_0$. Under Assumption 1, the time-inconsistent type will acquire information in $t = T-1$ and complete the task in $t = T$. In $t = T-2$, if an agent of the time-inconsistent type acquires information, she believes that the task will be completed with probability q in $t = T-1$ and with $1-q$ in $t = T$. If she does not acquire information in $t = T-2$, she will acquire information with certainty in $t = T-1$ and complete the task with certainty in $t = T$. Therefore, the time-inconsistent agent will postpone information acquisition if:

$$\begin{aligned} & -m + qU^I(x;1) + (1-q)U^I(x;2) < -mD(1) + U^I(x;2) \\ \Leftrightarrow \quad & q(U^I(x;2) - U^I(x;1)) + m(1-D(1)) > 0. \end{aligned} \tag{13}$$

Given this inequality and $\underline{\tau} > 2$, then in $t = T-3$, acquiring information immediately means the task will be completed in $t = T-2$ with probability q and in $t = T$ with probability $1-q$. If the time-inconsistent agent postpones now, she believes that with probability q the information will be acquired in $t = T-2$ and the task will be completed in $t = T-1$, but with probability $1-q$ the information will be acquired in period $t = T-1$ and the task will be completed in $t = T$. The time-inconsistent type of agent will therefore postpone information acquisition if:

$$\begin{aligned}
& -m + qU^I(x;1) + (1-q)U^I(x;3) < q(-mD(1) + U^I(x;2)) + (1-q)(-mD(2) + U^I(x;3)) \\
\Leftrightarrow \quad & q(U^I(x;2) - U^I(x;1)) + m(1 - qD(1)) > +m(1-q)D(3).
\end{aligned}$$

In general, when $t = T - \tau$ where $2 \leq \tau \leq \underline{\tau}$, the time-inconsistent type of agent will postpone acquiring information if

$$q(U^I(x;2) - U^I(x;1)) + m(1 - qD(1)) > m(1-q)D(\tau - 1).$$

The right-hand side of this inequality is decreasing in τ, so the inequality becomes the most stringent when $\tau = 2$. Hence, if $q(U^I(x;2) - U^I(x;1)) + m(1 - qD(1)) > m(1-q)D(1)$ or equivalently if $q(U^I(x;2) - U^I(x;1)) + m(1 - D(1)) > 0$, the time-inconsistent agent will postpone information acquisition for at least $\underline{\tau}$ periods. A sufficient condition for the time-inconsistent agent to procrastinate for $\underline{\tau}$ periods is $U^I(x;2) > U^I(x;1)$.

In $t = T - \underline{\tau} - 1$, the time-inconsistent agent knows that if she acquires information, the task will be completed with certainty in the next period, and hence her expected payoff is

$$\hat{U}_{T-\underline{\tau}-1}(s_{T-\underline{\tau}-1} = 1) = -m + U^I(x;1).$$

If she chooses not to acquire information, she knows that the time-consistent type will acquire information in the next period, but the time-inconsistent type will only acquire information in $t = T - 1$. Hence, her expected payoff is

$$\hat{U}_{T-\underline{\tau}-1}(s_{T-\underline{\tau}-1} = 0) = q(-mD(1) + U^I(x;2)) + (1-q)(-mD(\underline{\tau}) + U^I(x;\underline{\tau}+1)).$$

Thus, she will prefer not acquiring the information in $t = T - \underline{\tau} - 1$ if

$$q(U^I(x;2) - U^I(x;1)) + m(1 - qD(1)) > (1-q)(mD(\underline{\tau}) - U^I(x;\underline{\tau}+1) + U^I(x;1)).$$

When q approaches 1, this condition reduces to Inequality (13). Whether this condition is generally stronger or weaker than Inequality (13) depends on parameters . If this condition holds, however, the agent will procrastinate for an additional $\underline{\tau}$ periods. This follows from the same argument as above.

By induction, for each $t = T - n\underline{\tau}$ such that $t \geq 1$ and $n \in \mathbb{N}_0$, we need to check whether the agent will prefer to acquire information or not. The agent is always comparing the outcomes of acquiring information now and completing the task with certainty in the next period, on the one hand, and the case where the time-inconsistent type procrastinates until the second-last period and the time-consistent type acquires information in the next period on the other. Therefore, if

$$\hat{U}_{T-n\underline{\tau}-k}(s_{T-n\underline{\tau}-k} = 1) = -m + qU^I(x;1) + (1-q)U^I(x;k)$$
$$< \hat{U}_{T-n\underline{\tau}-k}(s_{T-n\underline{\tau}-k} = 0) = q(-mD(1) + U^I(x;2)) + (1-q)(-mD(n\underline{\tau}) + U^I(x;n\underline{\tau}+k))$$

, which is rewritten as

$$q(U^I(x;2) - U^I(x;1)) + m(1-qD(1)) > (1-q)(mD(n\underline{\tau}) - U^I(x;n\underline{\tau}+k) + U^I(x;k))$$

, for any $n \in \mathbb{N}_0$ and $k \in \mathbb{N}$ where $k < \underline{\tau}$ and $n\underline{\tau} + k \leq T$, then the time-inconsistent agent will procrastinate information acquisition for any $t < T - 1$. Note also that, given this condition, we have derived the time-inconsistent type's learning decision in all continuation games and hence have uniquely pinned down the equilibrium outcome. □

To see the underlying intuition here, suppose that a time-consistent agent always chooses to acquire information and completes the task in every period. Intuitively, because an agent of the time-inconsistent type is (probabilistically) overconfident about her own future self-control, she erroneously thinks that she will acquire information with probability q in the next period. Note that the time-inconsistent agent evaluates her anticipated future outcomes in terms of her current (i.e., time-inconsistent) preferences. Hence, akin to O'Donoghue and Rabin (1999a, 2001), the time-inconsistent agent may procrastinate information acquisition — in any period $t < T - 1$, she thinks that she will acquire information with probability q in the next period, but in reality she will not do so, with probability one.

Perhaps surprisingly, and beyond the original logic of O'Donoghue and Rabin (1999a, 2001), procrastination in our model can occur even when $m = 0$. This is because the time-inconsistent type indirectly incurs a cost from learning if $U^I(x;2) > U^I(x;1)$, which never occurs under quasi-hyperbolic discounting. Corollary B.1 highlights the result:

Corollary B.1. Suppose $T < \infty$ and $U^I(x;2) > U^I(x;1) > U^I(x;0)$. Then, for any $m \geq 0$, there exists a $\bar{q} < 1$ such that for any $q \in [\bar{q}, 1)$ there is a unique equilibrium outcome in which the time-consistent type will acquire information immediately while the time-inconsistent type will procrastinate on acquiring information until $t = T - 1$.