Revealing Sophistication and Naïveté from Procrastination

David Freeman*

March 1, 2016

Abstract

Consider a person who must complete a task, and is given a set of options for completing the task at different times. The person cannot commit her future behaviour except by completing the task. This paper shows that comparing the person's completion time across different sets of completion opportunities can reveal her sophistication or naïveté about her dynamic inconsistency. I show that adding an extra opportunity to complete the task can lead a naïve (but not a sophisticated) person to complete it even later, and can lead a sophisticated (but not a naïve) person to complete the task even earlier, even if the extra opportunity is not used. This result can be obtained with or without parametric assumptions about utility. Additional results completely characterize models of naïve and sophisticated individuals in this environment. These results provide the framework for revealing the preference and sophistication types studied in O'Donoghue and Rabin (1999) from behaviour in a generalization of their environment.

JEL Codes: D03, D84, D90.

Keywords: sophistication, naïveté, procrastination, task completion, present bias, time inconsistency.

*Department of Economics, Simon Fraser University. E-mail: david_freeman@sfu.ca.. I'd like to thank to my colleagues at SFU and to Terri Kneeland, Guy Mayraz, and Philipp Sadowski for helpful comments and encouragement, and especially thank David Ahn and Todd Sarver for comments which greatly improved the paper. I also thank Matthew Pentecost and Garrett Petersen for excellent research assistance; special thanks to Garrett for suggesting the example in the second paragraph of the introduction.

1 Introduction

Behavioural economic models of intertemporal choice following Strotz (1955) incorporate two assumptions. First, a person may be dynamically inconsistent, that is, her current preferences over future actions may differ from the future preferences she will use to when she chooses. Second, a person may not be perfectly self-aware of her own dynamic inconsistency when she forms expectations of her own future behaviour. Strotz proposed two ways a dynamically inconsistent person might form expectations. She can be naïve and expect her future selves to behave according to her current preferences, or she can be sophisticated and hold correct expectations about her future behaviour. However, a person's self-awareness is not directly observed. This makes it difficult to understand which assumption is more descriptively appropriate for applications. This paper asks how economic choices can reveal an individual's sophistication or naïveté about her dynamic inconsistency.

Imagine a professor extending the deadline for a major assignment from the start of an exam period to the end of that exam period. A naïve student might (falsely) expect herself to use the extra time wisely and start the assignment later. On the other hand, a more sophisticated student might recognize her own tendency to procrastinate, realize that delaying would lead her to do the work before important exam (when she has a higher opportunity cost) and might finish the assignment sooner than she would have without the extension. This study demonstrates that decisions like this can reveal an individual's sophistication or naïveté.

This paper studies a individual's choice of when to complete a task that must be done exactly once, as in O'Donoghue and Rabin (1999). I suppose that we observe when the individual would complete the task given the set of completion opportunities (actions) available to her. I show that how she responds to adding additional actions can reveal her sophistication or naïveté, even if we do not observe the person's preferences directly. I show that sophistication and naïveté have sufficient force, even in choice problems in which a person's higher-order beliefs about her future behaviour are relevant, to give directly testable and economically intuitive predictions in the domain of task completion.

The main results of this paper show that choice reversals with a "doing-it-earlier" flavour are a hallmark of sophistication, while choice reversals with a "doing-it-later" flavour are a hallmark of naïveté. That is, adding an unused completion opportunity can lead to delay for a naïf but not a sophisticate, while adding an unused completion opportunity can lead to earlier completion for a sophisticate but not a naïf. The intuition is that adding an additional action makes waiting appear weakly more attractive at earlier periods for a naïf who does not appreciate her dynamic inconsistency. Because a sophisticate anticipates her future inconsistency, she may respond to an added action by doing the task earlier than the added action when she forsees that the added action will affect her future behaviour against an earlier self's wishes. However, since a sophisticate correctly anticipates her future behaviour, she would not delay her completion to a time after newly added action unless the added action does not affect the action she completes.

The paper proceeds as follows. Section 2 introduces the task completion environment considered, O'Donoghue and Rabin's perception perfect equilibrium concept for sophisticated, naïve, and partially naïve Strotzian models of behaviour, and how they relate to decision-theoretic notions of sophistication and naïveté (Proposition 1). Section 3 introduces "doing-it-later" and "doing-it-earlier" reversals and shows that a naif commits only the former while a sophisticate commits only the latter (Propositions 2 and 3). Section 4 (Proposition 4) shows that, when looking at choice sets with at most three actions, doing-it-earlier and doing-it-later reversals provide a basis for comparing the sophistication of two people with the same preferences but possibly different beliefs. Section 5 establishes the implications of these results for revealing partial sophistication and naïveté in the partially naïve quasi-hyperbolic discounting model from doing-it-earlier and doing-it-later reversals and dominance violations. Section 6 shows that the results on sophistication in Proposition 3 extend to the case where a person experiences intrinsic self-control costs (as in Gul and Pesendorfer (2001)).

1.1 Preview of results

To motivate the analysis that follows, consider a quasi-hyperbolic discounter who chooses when to complete a task. Each action has the same benefits that are realized at the same time in the distant future, but actions may differ in when they are available and their immediate costs or benefits. Represent each action as a pair $(x,t) \in \mathbb{R}_+ \times \mathbb{N}$, where x is the cost of doing the task and t is the time at which that task can be completed. At time τ, our person evaluates an action (x,t) with $t \geq \tau$ according to

$$U_\tau((x,t)) = \begin{cases} u(x) & \text{if } \tau = t \\ \beta\delta^{\tau-t}u(x) & \text{if } \tau > t \end{cases}$$

where u is a continuous and strictly decreasing utility function. Thus given the options $\{(x,t),(y,t+1)\}$, she does (x,t) if and only if $u(x) \geq \beta\delta u(y)$ (assuming that she does the earlier action when indifferent).

If $\beta < 1$, the person is more willing to have herself incur a higher cost to do the action later when the earlier action involves a cost in the immediate present than if she were to compare the same two options before either action could actually be taken. Because of this, there will exist situations in which the person would like herself to take a particular action in a future period that she will not actually take when that future period becomes the present. For example, if there are x, y, z for which $\delta u(z) > \beta\delta^2 u(y) > u(x) > \beta\delta u(z)$, then $(x,t) = c(\{(x,t),(z,t+1)\})$, $(z,t+1) = c(\{(z,t+1),(y,t+2)\})$, $(y,t+2) = c(\{(x,t),(y,t+2)\})$. This apparent choice cycle arises from the individual's changing preferences over actions. If the person faces the action set $c(\{(x,t),(z,t+1),(y,t+2)\})$, the resolution of this conflict will be determined by how the person forecasts her future behaviour. If she is sophisticated and correctly anticipates her future behaviour, she would take the earliest action (x,t) since she knows that her $t+1$ self would do $(z,t+1)$; notice that she acts earlier than when facing $\{(x,t),(y,t+2)\}$ even though the added option relative to that set isn't used. If she is naïve and incorrectly thinks that her $t+1$ self will share her current ranking of actions, then she will wait at time t but then do $(z,t+1)$; notice that she acts later than in $\{(x,t),(z,t+1)\}$ even though the added option relative to that set isn't used.

When a person has more than three different opportunities to complete a task that are available at more than three different times, a current self's decision will depend on not just how on she forecasts her future selves' preferences. It will also depend on how she forecasts future selves will forecast their future selves' preferences (and so on for higher orders). The definitions of doing-it-later and doing-it-earlier reversals and results linking them to naïveté and sophistication show how the intuition from this example extends generally, both within the quasi-hyperbolic discounting model and without parametric assumptions on utility.

1.2 Related literature

The behavioural economics literature has extensively applied models of time inconsistency following Strotz (1955), in particular to capture present-bias (Laibson 1997, O'Donoghue and Rabin 1999, 2001). A person's degree of sophistication is an important variable in many applications of present-biased preferences (e.g. DellaVigna and Malmendier (2004); Heidhues and Kőszegi (2010)). This paper follows O'Donoghue and Rabin (1999; 2001) in its choice of a doing-it-once environment. The analysis of sophistication and naïveté offered here extend O'Donoghue and Rabin's "smoking guns" for dynamic inconsistency in the quasi-hyperbolic discounting model to provide separate hallmarks for sophistication versus naïveté without assuming quasi-hyperbolic discounting. O'Donoghue and Rabin (2001) propose that we learn little when we observe a person who faces one menu of tasks – their "Weak" and "Strong" Axioms of Revealed Procrastination. This paper shows that we can draw meaningful inferences about her naïveté/sophistication under weak assumptions if we can observe an individual in two different but appropriately comparable environments; I show this is the case even when preferences and the individual's sophistication or naïveté must both be elicited from choices, and when preferences are not restricted to quasi-hyperbolic discounting.[1]

[1] In related work on wide class of stopping problems, Quah and Strulovici (2013, Proposition 10) extend O'Donoghue and Rabin's (1999) Proposition 2 to show that, in a particular class of problems, a more sophisticated

Existing work has shown how discount and utility functions of dynamically inconsistent models can be characterized and measured from preferences over consumption streams (Hayashi, 2003; Olea and Strzalecki, 2014), dated outcomes (Ok and Masatlioglu, 2007; Attema et al., 2010), dated outcome streams (Attema et al., 2016), and convex budget sets of pairs of dated outcomes (Echenique et al., 2015). Other work has studied the commitment preferences of sophisticated decision-makers and have provided axiomatic characterizations of sophistication (Gul and Pesendorfer, 2005; Noor, 2011). With the recent exception of Ahn et al. (2015), this literature has not explored the analogous characterizations for naïveté. Lab and field experiments studying such commitment preferences typically find that only a minority of the study population demands commitment devices when they are free (e.g. 28% in Ashraf et al. (2006), 35% in Kaur et al. (2010)); attempts to estimate a degree of sophistication have been limited and require strong assumptions (Fang and Wang, 2015; Augenblick and Rabin, 2015).[2] Moreover, Blow et al. (2015) show that present-bias with naïveté cannot be distinguished from time consistent preferences using consumption-savings behaviour. In my view, this past work leaves open the question of what restricted domain can be used to tests, and do there exist equally-intuitive tests for naïve as for sophisticated present-bias? Past experiments have studied task completion (Ariely and Wertenbroch, 2002; Burger et al., 2011; Bisin and Hyndman, 2014), but have not used task completion data to distinguish naïveté from sophistication.

2 Modelling sophistication and naïveté

2.1 Environment

I consider an environment in which a person faces a set of actions, A, and must do exactly one action a in A. Each action $a \in A$ is associated with a time, t_a. At any given time τ, the person can take a action available at time τ - that is, an $a \in A$ which satisfies $t_a = \tau$ - or she can decide not to take an action and let herself act in the future. The person cannot commit the actions of her future selves except by completing an action now and thereby preventing her future selves from acting. However, the person is constrained to take an action in any given action set, thus if no actions are available in future periods, the person always takes one of the currently available actions.

agent will stop before a less sophisticated agent with the same preferences.

[2] Specifically, the exclusion restriction in Fang and Wang (2015) requires that a host of demographic variables that affect the likelihood of future health states are uncorrelated with preferences and degree of sophistication. Experiments that elicit beliefs about future behaviour at the individual level provide provide one way around this issue, and Augenblick and Rabin (2015) pursue this direction. But because beliefs are about one's own later behaviour, such studies are subject to possibility that belief elicitation might affect one's own subsequent behaviour either due to its incentives or a psychological taste for consistency. Such a limitation is above and beyond methodological issues standard to belief elicitation, which are reviewed in Schotter and Trevino (2014).

I use the following notation in the analysis that follows. Let $\bar{A}$ be a set of actions, and let $\mathscr{A}$ denote the set of all non-empty and finite subsets of $\bar{A}$. Refer to any $A \in \mathscr{A}$ as a action set or a choice problem. For each action $a \in \bar{A}$ let $t_a \in \mathbb{N}$ specify the time when each action a is available. Given a set $A \in \mathscr{A}$, let $A_{=\tau} = \{a \in A : t_a = \tau\}$ and let $A_{\geq\tau} = \{a \in A : t_a \geq \tau\}$; similarly let $\mathscr{A}_{\geq\tau}$ denote the set of non-empty finite subsets of $\bar{A}_{\geq\tau}$. Given any $A \in \mathscr{A}$, let $\mathscr{T}_A = \{\tau : \exists a \in A \text{ s.th. } t_a = \tau\}$ and assume $1 = \min[\tau \in \mathscr{T}_{\bar{A}}] < \max[\tau \in \mathscr{T}_{\bar{A}}] < \infty$.

Let $c : \mathscr{A} \to \bar{A}$ denote a choice function. The analysis that follows assumes that c is observed for a single person. This assumption is reasonable if the same person's actions are observed from directly comparable but different action sets. Alternatively, it may possible to observe many people who face different choice sets.

Given a choice function c, we can summarize behaviour in a given action set A at a given point in time by a set of functions $\{c_\tau\}_{\tau\in\mathscr{T}_{\bar{A}}}$ where $c_\tau : \mathscr{A}_{\geq\tau} \to \bar{A}_{=\tau} \cup \mathscr{A}_{\geq\tau}$ gives either the person's chosen action if she does it at τ or a continuation menu if she doesn't. Here, $a = c_\tau(A)$ denotes that the person does a at time τ when she facing set A, and $A_{>\tau} = c_\tau(A)$ denotes that the person takes no action at τ either because she chooses to wait, no action is currently available, or she has already done the task.

Below, I provide examples of types of actions that can fit into this framework.

Example 1. A statistics assignment must be done in a computer lab which is only open on a set number of week days, announced in advanced. The student completes a version of the assignment that is specific to the day on which she completes it, and the assignment version can be easy (e), medium (m), or hard (h) in difficulty and this is observed by students in advance. Our grand space of set of possible completion opportunities for this example is given by $\bar{A} = \{M, Tu, W, Th, F\} \times \{e, m, h\}$. With abuse of notation, let $\mathscr{T}_{\bar{A}} = \{M, Tu, W, Th, F\}$ when I refer back to this example.

Example 2. Let $\mathscr{S} \subset \mathbb{R}^T$ for $1 \leq T \leq \infty$ denote a subsets of (possibly infinite dimensional) streams, let $\mathscr{T}_{\bar{A}} \subseteq \mathbb{N}$, and let $\bar{A} \subseteq \mathscr{S} \times \mathscr{T}_{\bar{A}}$. For a given (x,t), x_τ denotes the real-valued number (e.g. a consumption level or effort requirement) at time $\tau - t$. [3]

2.2 Preferences

Consider the following model of Strotzian preferences. Each person has a set of time-dependent utility functions. For each $\tau \in \mathscr{T}_{\bar{A}}$, let $U_\tau : \bar{A}_{\geq\tau} \to \mathbb{R}$ denote her time-τ utility function, that is, the utility function she uses at time-τ when she evaluates actions. Let $\mathscr{U}$ denote a collection of time-τ utility functions with one U_τ for each $\tau \in \mathscr{T}_{\bar{A}}$.

[3] Note I work with a choice function c, and my analysis will assume that a person always breaks ties the same way. This tie-breaking is only relevant if $\mathscr{C}$ is sufficiently rich that a person is sometimes indifferent.

Note that the structure of the action space and preferences in the general Strotzian model are minimally restricted. This allows for arbitrary attitudes to timing, as well as arbitrarily time-variant preferences that are fully subjective, due to, say, the presence of opening night for a Johnny Dep movie on Wednesday night, or a Stanley Cup finals game on Tuesday night.

Quasi-hyperbolic model. The quasi-hyperbolic model is a special case of Strotzian preferences. In the setting of Example 2, the quasi-hyperbolic discounting model readily applies under the assumption that not doing an action in a given period yields an instantaneous utility of zero. The quasi-hyperbolic utility function is given by:

$$U_\tau((x,t)) = \begin{cases} u(x_0) + \beta \sum_{k=1}^{T} \delta^k u(x_k) & \text{if } t = \tau \\ \beta \delta^{t-\tau} \sum_{k=0}^{\infty} \delta^k u(x_k) & \text{if } t > \tau \end{cases} \tag{1}$$

In this model (with $1 \leq T \leq \infty$), at time τ a reward or cost in a future period time $t > \tau$ is discounted by $\beta\delta^{t-\tau}$. The parameter $\beta < 1$ captures a person's present bias and implies that the person discounts more steeply between the current and next period than between any two future periods. The parameter $\delta \in (0,1]$ captures standard discounting, and gives the discount factor that applies between any two periods that lie strictly in the future. The period-utility function is given by u, which I assume is monotonically increasing, and this formulation implicitly normalizes the utility of not doing any action at a point in time to zero. O'Donoghue and Rabin's (1999) analysis is based on this formulation, but with each $\mathscr{X}$ restricted to acts that only allow $u(x_k) \neq 0$ at $k = 0$ and one other point in time..

2.3 Time consistency

I adopt the following definitions of time consistency and inconsistency for a collection $\mathscr{U}$ as a restriction that all utility functions in $\mathscr{U}$ assign the same ranking to future actions. This maps directly to time consistency as a property of choice function.

Definition. The collection $\mathscr{U}$ is <u>time consistent</u> if, for any $\tau, \tau' \in \mathscr{T}_{\bar{A}}$, U_τ and $U_{\tau'}$ are ordinally equivalent on $A_{\geq \max[\tau,\tau']}$. Otherwise, the collection $\mathscr{U}$ is <u>time inconsistent</u>.

In the simple task completion environment I consider, time consistency of $\mathscr{U}$ requires that preferences at each time never disagree ordinally – which is behaviourally equivalent to rationalizability by a single utility function. The Independence of Irrelevant Alternatives (IIA) property of a choice function characterizes the rational choice model on a finite domain.

IIA. If $a = c(A)$ and $a \in A' \subset A$, then $a = c(A')$.

In the enriched choice domain considered here, IIA has additional bite since we interpret the person as unable to commit to her future actions except by doing a current action. In particular, time consistent choices satisfy IIA, and the IIA property completely characterizes choice functions with a time consistent representation under the technical assumption that $\bar{A}$ is finite.

Lemma 1. *If the choice function c has a time consistent representation then it satisfies IIA. If $\bar{A}$ is finite and c satisfies IIA, then c has a time consistent representation.*

It is well known that the above version of IIA is a necessary and sufficient condition for a choice function on a finite domain to be rationalizable by a single utility function, and the proof in the appendix translates a standard proof to the problem here.

This gives us a basis for viewing time inconsistency as a property of a choice function in our setting. Whether a person is naïve or sophisticated is not directly observable. I will ask whether the time inconsistent behaviour of naïve and sophisticated people are qualitatively different in ways that can reveal a person's naïveté or sophistication by relating each model to properties that restrict the scope of how IIA can be violated.

2.4 Beliefs and behaviour under time inconsistency

The behaviour of a time-inconsistent person will depend both on her preferences ($\mathscr{U}$) and also on her expectations about her future behaviour. The term "sophisticated" refers to a person who correctly forecasts her future behaviour, and and term "naïve" refers to a person who incorrectly believes her future behaviour will align with her current tastes. In the sophisticated and naïve cases, a Strotzian's behaviour in any action set can be derived from her preferences.

Beliefs and behaviour under sophistication and naïveté The perception perfect equilibrium concept from O'Donoghue and Rabin provide how a person would behave in task completion environment given her preferences and expectations. Their sophisticated perception perfect equilibrium concept captures how a person with sophisticated expectations about her future behaviour would behave in each period given an action set A.

Definition. A <u>sophisticated perception perfect equilibrium</u> for the choice problem A given the collection of utility functions $\mathscr{U}$ consists of a strategy $s_\tau^{\mathscr{U},s} : \mathscr{A} \to \bar{A} \cup \mathscr{A}$ for each τ such that

(i) $A_{>\tau} = \emptyset$ implies $s_\tau^{\mathscr{U},s}(A) = \arg\max_{a \in A_{=\tau}} U_\tau(a)$,

(ii) $$s_\tau^{\mathscr{U},s} = \begin{cases} \arg\max_{a \in A_{=\tau}} U_\tau(a) & \text{if } \max_{a \in A_{=\tau}} U_\tau(a) \geq U_\tau(s_{\tau'}^{\mathscr{U},s}(A_{\geq \tau'})) \\ A_{>\tau} & \text{otherwise} \end{cases}$$

for $\tau' = \min\left[\check{\tau} > \tau : s_{\check{\tau}}^{\mathscr{U},s}(A_{\geq\check{\tau}}) \neq A_{>\check{\tau}}\right]$.

The naïve perception perfect equilibrium concept provides an analogous way of solving for behaviour under naïveté.

Definition. A naïve perception perfect equilibrium for the choice problem A given the collection of utility functions $\mathscr{U}$ consists of a strategy $s_\tau^{\mathscr{U},s} : \mathscr{A} \to \bar{A} \cup \mathscr{A}$ for each τ such that

(i) $A_{>\tau} = \emptyset$ implies $s_\tau^{\mathscr{U},s}(A) = \arg\max_{a \in A_{=\tau}} U_\tau(a)$,

(ii) $s_\tau^{\mathscr{U},s} = \begin{cases} \arg\max_{a \in A_{=\tau}} U_\tau(a) & \text{if } \max_{a \in A_{=\tau}} U_\tau(a) \geq \max_{a \in A_{>\tau}} U_\tau(a) \\ A_{>\tau} & \text{otherwise} \end{cases}$

Notice that to avoid the possibility of anomalous behaviour that arises from tie-breaking, I assume that for a given individual, if there are two actions a and a' with $t_a = t_{a'}$ and $\hat{U}_{t_a|\tau}(a) = \hat{U}_{t_a|\tau}(a')$, the individual at τ always perceives her time t_a self as breaking ties in the same way when forming an expectation about a'_τ. Similarly, I assume that for any two such actions with $t_a = t_{a'}$ and $U_{t_a}(a) = U_{t_a}(a')$, the individual breaks the tie the same way at t_a regardless of her choice set.

Beliefs and behaviour in the partially-naïve Strotz model I introduce the partially naïve Strotz model, which generalizes the partially naïve quasi-hyperbolic model of O'Donoghue and Rabin (2001), to capture the possibility of a dynamically inconsistent person who is aware of but does not correctly forecast her dynamic inconsistency. In this model, a person at time τ to perceives that her time-τ' self will use the utility function $\hat{U}_{\tau'|\tau}$, which need not represent the same ranking of actions as either U_τ or $U_{\tau'}$. Let $\hat{\mathscr{U}}$ denote a set of perceived utility functions, with one function $\hat{U}_{\tau'|\tau} : \bar{A}_{\geq\tau'} \to \mathbb{R}$ for each pair $\tau, \tau' \in \mathscr{T}_{\bar{A}}$ with $\tau' > \tau$, and let $\hat{\mathscr{U}}_{\cdot|\tau}$ denote the set of utility functions $\hat{U}_{\tau'|\tau}$ in $\hat{\mathscr{U}}$. O'Donoghue and Rabin's (1999; 2001) perception perfect equilibrium concept applies in this case as well, and is defined below.

Definition. A partially naïve perception perfect equilibrium for the choice problem A given the collection of utility functions $\mathscr{U}$ and the collection of perceived utility functions $\hat{\mathscr{U}}$ consists of a strategy $s_\tau^{\mathscr{U},\hat{\mathscr{U}}} : \mathscr{A} \to \bar{A} \cup \mathscr{A}$ for each τ such that

(i) $A_{>\tau} = \emptyset$ implies $s_\tau^{\mathscr{U},\hat{\mathscr{U}}}(A) = \arg\max_{a \in A_{=\tau}} U_\tau(a)$,

(ii) $s_\tau^{\mathscr{U},\hat{\mathscr{U}}} = \begin{cases} \arg\max_{a \in A_{=\tau}} U_\tau(a) & \text{if } \max_{a \in A_{=\tau}} U_\tau(a) \geq U_\tau(s_{\tau'}^{\hat{\mathscr{U}}_{\cdot|\tau},S}(A_{\geq\tau'})) \\ A_{>\tau} & \text{otherwise} \end{cases}$

for $\tau' = \min\left[\check{\tau} > \tau : s_{\check{\tau}}^{\hat{\mathscr{U}}_{\cdot|\tau},s}(A_{\geq\check{\tau}}) \neq A_{>\check{\tau}}\right]$.

In a perception perfect equilibrium, the person believes at time τ that her future selves will use utility functions from $\hat{\mathscr{U}}_{\cdot|\tau}$, and forecasts her future actions by computing the sophisticated strategy that corresponds to $\hat{\mathscr{U}}_{\cdot|\tau}$. She then uses U_τ to decide whether to wait or act, if to select a time τ available action in the latter case. The partially naïve Strotz model captures all naïveté by "first-order" beliefs, since it does allows a time τ self to be naïve about her future utility functions, but it allows a time τ self to be naïve in forecasting future behaviour, but assumes that she (naïvely) forecast all future selves will behave as sophisticated relative to these forecasts. This rules out the possibility of a person who, at time τ, correctly forecasts her future preferences, but underestimates her own future naïveté.

Given these definitions, we can give a definitions of sophisticated, partially naïve Strotzian, and fully naïve choice.

Definition. The choice function c is a <u>sophisticated choice function</u> if there exists a $\mathscr{U}$ such that for each $A \in \mathscr{A}$, $s_\tau^{\mathscr{U},S}(A) = c_\tau(A_{\geq\tau})$. A choice function c is called a <u>naïve choice function</u> if c is a partially naïve choice function corresponding to $\{s_\tau^{\mathscr{U},\hat{\mathscr{U}}}\}_\tau$ in which each $\hat{U}_{\tau'|\tau} \in \hat{\mathscr{U}}_{\cdot|\tau}$ is ordinally equivalent to U_τ on $\bar{A}_{>\tau}$ for each τ and each $\tau' > \tau$. The choice function c is a <u>partially naïve Strotzian choice function</u> if there exist $\mathscr{U}$ and $\hat{\mathscr{U}}$ such that for each $A \in \mathscr{A}$ and each τ, $s_\tau^{\mathscr{U},\hat{\mathscr{U}}}(A) = c_\tau(A_{\geq\tau})$.

Observation. For any partially naïve choice function c, beliefs about future behaviour ($\hat{\mathscr{U}}$) are irrelevant in choice sets with two options. In such sets, choosing $a = c(\{a,a'\})$ directly reveals that $U_\tau(a) \geq U_\tau(a')$ for $\tau = \min[t_a, t_{a'}]$.

Partially naïve quasi-hyperbolic discounting model. Consider a quasi-hyperbolic discounter facing the domain of Example 2. Their $\mathscr{U}$ is given by (1), and changes in her ranking of actions due to the passage of time arise only due to the role of $\beta < 1$. A parsimonious model of partial naïvete about one's present bias has the person underestimate the extent of present bias she will exhibit in the future while correctly forecasting u and δ. In O'Donoghue and Rabin's (2001) partially naïve quasi-hyperbolic discounting model, a person perceives that the present bias parameter that will apply in her future decisions is given by $\hat{\beta} \in [\beta, 1]$. This gives a convenient specification for $\hat{\mathscr{U}}$:

$$\hat{U}_{\tau'|\tau}((x,t)) = \begin{cases} u(x_0) + \hat{\beta} \sum\limits_{k=1}^{T} \delta^k u(x_k) & \text{if } t = \tau' \\ \hat{\beta}\delta^{t-\tau} \sum\limits_{k=0}^{\infty} \delta^k u(x_k) & \text{if } t > \tau' \end{cases} \tag{2}$$

The special case with $\hat{\beta} = 1$ corresponds to a naïve representation, while the case with $\hat{\beta} = \beta$ corresponds to a sophisticated representation.

2.5 Comparison with decision-theoretic definitions of sophistication

The notions of sophistication and naïveté I use are based on a person's perception of her future utility functions in a Strotzian model, following O'Donoghue and Rabin. An alternative approach, used in the decision theory literature on multi-period choice problems, would be to define sophistication in terms of behaviour directly as in Gul and Pesendorfer (2005) and Noor (2011). In these papers, each period a person chooses a pair that specifies a current level of consumption and a continuation choice set. The domain of task completion naturally fits within this framework with the restriction that choice in each period τ, as given by c_τ, is restricted to either choose a current task and a continuation menu that gives a singleton menu in all subsequent periods, or to choose to not act ($\emptyset$) which is associated with a single continuation menu. Since I place limited restrictions on $\bar{A}$ and on $\mathscr{U}$, this framework is still rich enough to capture actions that specify a stream of payoffs in future periods.

Given any binary relation $\succ$ and set A, let m be the function that picks $\succ$- undominated elements of A, defined by $m(A,\succ) = \{a \in A : \nexists a' \in A \text{ s.th. } a' \succ a \text{ but not } a \succ a'\}$.

First, let $\succsim_\tau$ denote the binary relation on $\bar{A}_{=\tau} \cup \mathscr{A}_{>\tau}$ that captures how a person ranks the desirability of actions available at time τ against each other and against possible continuation menus; let $\succ_\tau$ its asymmetric subrelation. Below, I impose that $\succsim_\tau$ is "standard" in the sense that it is consistent with pairwise choices, and that it is complete and transitive on its full domain.

S. (i) $a = c_\tau(A_{=\tau} \cup A_{>\tau})$ implies $a \succsim_\tau A_{>\tau}$ and $a \succ_\tau a'$ for all $a' \in A_{=\tau}$, $A_{>\tau} = c_\tau(A_{=\tau} \cup A_{>\tau})$ implies $A_{>\tau} \succ_\tau a$ for all $a \in A_{=\tau}$. (ii) $\succsim_\tau$ is complete and transitive on $\bar{A}_{\geq\tau} \cup \mathscr{A}_{>\tau}$.

Notice that S(i) imposes how $\succsim_\tau$ is determined by c_τ. Given the relationship between $\succsim_\tau$ and c_τ from S(i), axiom S(ii) restricts that each c_τ is determined by maximizing $\succsim_\tau$ (with S(i) imposing that ties are broken in favour of acting in the present).

Gul and Pesendorfer (2005) define sophisticated choices by an axiom, Independence of Redundant Alternatives, on the pair $\{c_\tau\}_\tau$ and $\{\succsim_\tau\}_\tau$ that extends the logic of IIA by restricting that adding a completion opportunity to an action set only affects earlier behaviour if this opportunity if it would be used were it to be reached. My adaptation of their axiom is below.

IRA. Given τ and $\tau' > \tau$, $A_{\geq\tau'} \sim_\tau c_{\tau'}(A_{\geq\tau'})$.

While the decision theory literature has not done so, it is possible to give a completely analogous definition of naïveté based on the idea that a naïve person thinks that her next decision will be consistent with her current tastes. I call this axiom Independence of Worse Alternatives (IWA). For notational convenience, I use $m_{\tau'}(\cdot,\succsim_\tau)$ as an analogous object to c_τ that instead makes the $\succsim_\tau$- preferred "act or wait" decision available in any set $A_{\geq\tau'}$, that is, $m_{\tau'}(A_{\geq\tau'},\succsim_\tau) = m(A_{=\tau'} \cup \{A_{>\tau'}\},\succsim_\tau)$.

IWA. Given τ and $\tau' > \tau$, $A_{\geq\tau'} \sim_\tau m_{\tau'}(A_{\geq\tau'}, \succsim_\tau)$.

While the axiomatic approach to defining sophistication and naïveté appears very different from the definitions of sophisticated and naïve choice functions based on PPE, they are observationally identical, as summarized in Proposition 1, which builds on results for sophistication in Gul and Pesendorfer (2005).

Proposition 1. *Suppose $\bar{A}$ is finite.*

(i) If c and $\{\succsim_\tau\}_\tau$ satisfy S and IRA, then c is a sophisticated choice function. Conversely, if c is a sophisticated choice function, then there exists $\{\succsim_\tau\}_\tau$ such that c and $\{\succsim_\tau\}_\tau$ satisfy S and IRA.

(ii) If c and $\{\succsim_\tau\}_\tau$ satisfy S and IWA, then c is a naïve choice function. Conversely, if c is a naïve choice function, then there exists $\{\succsim_\tau\}_\tau$ such that c and $\{\succsim_\tau\}_\tau$ satisfy S and IWA.

Notice that IRA only implies "first-order" sophistication – that is, that a person is sophisticated about the next decision she makes. It does not impose anything directly about higher-order beliefs. Nonetheless, a folding back argument establishes that S and IRA imply higher-order sophistication as well. Similarly, IWA only implies "first-order" naïveté, but an analogous folding-back argument using S and IWA establishes higher-order naïveté. This is a key step in each proof.

In the reverse direction, it remains to construct each $\succsim_\tau$ to extend the order consistent with U_τ. A similar folding back through the PPE argument in each case verifies that, for the appopriate choice of $\{\succsim_\tau\}_\tau$, we have that $\{\succsim_\tau\}_\tau$ and c satisfy S and IRA under sophistication and IWA under naïveté. I leave the detailed proof for the appendix.

3 Choice reverals under naïveté and sophistication

Subsection 1.1 shows that reversals that violate IIA are possible under quasi-hyperbolic discounting, either under sophistication or naïveté. We classify two types of reversals a person might exhibit.

Definition. Say that c exhibits a reversal if there exists an $A \subseteq \bar{A}$ and an $a' \in \bar{A}$ such that $a = c(A)$, and $a, a' \neq a'' = c(A \cup \{a'\})$. We call a given reversal a doing-it-later reversal if $t_a < t_{a'}, t_{a''}$. We call a given reversal a doing-it-earlier reversal if $t_{a''} < t_a$.

Below, I show that naïveté, as captured by the IWA axiom and as satisfied by the naïve PPE model, allows for doing-it-later reversals but not doing it earlier reversals. In contrast, sophistication allows doing-it-earlier but not doing-it-later reversals.

Intuitively, a naïve person can be induced to delay when she fails to anticipate that her future self will take an action that differs from the action she would currently prefer. Adding new options to a time inconsistent person's action set might tempt her at some point, but a naïve person will not anticipate any inconsistency with her current tastes when deciding whether or not to take a currently available action. Putting these pieces of intuition together, it seems possible that adding an action to an action set might lead a naïve person to delay when her earlier selves do not anticipate that her later selves will be tempted to delay by the added action, even if the added action is not itself taken. This type of doing-it-later behaviour cannot be accommodated under time consistency since it violates IIA. The "Irrelevant Alternatives Delay" property allows for doing-it-later reversals, but rules out doing-it-earlier reversals (in a way that the statement of the property makes precise).

Irrelevant Alternatives Delay. $a = c(A)$ and $a'' = c(A \cup \{a'\}) \neq a, a'$ implies that $t_{a'}, t_{a''} > t_a$.

The Irrelevant Alternatives Delay property can be tested by observing only two choices. Example 3 gives an example of a choice function that violates IIA but is consistent with the Irrelevant Alternatives Delay property.

Example 3. Revisit Example 1. Suppose we observe that $(M,e) = c(\{(M,e),(F,h)\})$ and $(F,h) = c(\{(M,e),(W,m),(F,h)\})$. These choices are jointly inconsistent with IIA, and hence imply that preferences are time inconsistent. However, these choices are consistent with the Irrelevant Alternatives Delay property, since adding the irrelevant option (W,m), which is available on Wednesday (which is later than Monday) leads the person to delay until Friday (which is later than Monday and Wednesday).

Example 4 gives an example of a choice function that violates both IIA and the Irrelevant Alternatives Delay property.

Example 4. Revisit Example 1. Suppose we observe that $(W,m) = c(\{(M,e),(W,m)\})$ and $(M,e) = c(\{(M,e),(W,m),(F,h)\})$. These choices are jointly inconsistent with IIA, and hence with time consistent preferences. But they are also inconsistent with the Irrelevant Alternatives Delay property, since adding the unused alternative (F,h) leads the person to complete the assignment earlier than Wednesday.

Example 3 shows that there exist choices that are consistent with the Irrelevant Alternatives Delay property but inconsistent with IIA, while Example 4 establishes that Irrelevant Alternatives Delay places a testable restriction on behaviour.

Proposition 2 shows that not only is it possible that an added unused alternative leads to delay for a naïf, but that this is a general property of behaviour satisfied by any naïve person. An implication of this result is that the choices in Example 4 are inconsistent with naïve decision-making, while those in Example 3 violate time consistency but are consistent with naïve decision-making.

Proposition 2. *If c is a naïve choice function, then c satisfies the Independent Alternatives Delay property.*

Proof. If c is a naïve choice function, then by Proposition 1, there is a $\{\succsim_\tau\}_\tau$ for which $\{\succsim_\tau\}_\tau$ and c satisfy S and IWA. But then, by the proof of Proposition 1, $A_{>\tau} \sim_\tau m(A_{>\tau}, \succsim_\tau)$ for each A, τ with $A_{>\tau} \neq \emptyset$.

Then, for any τ, A, and a' with $A_{>\tau} \neq \emptyset$ and $t_{a'} \geq \tau'$, we have by S(i) that $c_\tau(A_{\geq\tau}) = A_{>\tau}$ holds if and only if $A_{>\tau} \succ_\tau a\ \forall a \in A_{=\tau}$. But since $A_{>\tau} \cup \{a'\} \sim_\tau m(A_{>\tau} \cup \{a'\}, \succsim_\tau) \succsim_\tau m(A_{>\tau}, \succsim_\tau) \sim_\tau A_{>\tau}$, this implies by S(ii) that $A_{>\tau} \cup \{a'\} \succ_\tau a\ \forall a \in A_{=\tau}$, which implies by S(i) that $c_\tau(A_{\geq\tau} \cup \{a'\}) = A_{>\tau} \cup \{a'\}$. Thus $c_\tau(A_{\geq\tau}) = A_{>\tau}$ and $t_{a'} > \tau$ implies $c_\tau(A_{\geq\tau} \cup \{a'\}) = A_{>\tau} \cup \{a'\}$. It follows that $t_{c(A)} \leq t_{c(A\cup\{a'\})}$.

Let $a = c(A)$. If $t_{c(A\cup\{a'\})} = t_a$, then since (by S) $\succsim_\tau$ is transitive, either $a = c(A \cup \{a'\})$ or $t_{a'} = t_a$ and $a' = c(A \cup \{a'\})$. Thus $c(A \cup \{a'\}) \neq a, a'$ implies $t_{c(A\cup\{a'\})} > t_a$.

If $t_{a'} \leq t_a$, then since $a \succsim_\tau A_{>\tau}$, we must have $c_\tau(A_{\geq\tau} \cup \{a'\}) = a$ or $= a'$. Thus $t_{c(A\cup\{a'\})} \leq t_a$ and either $c(A \cup \{a'\}) = a$ or $= a'$, a contradiction. It follows that $t_{a'} > t_a$. □

Intuitively, sophisticates act earlier because they anticipate their future self-control problems. Thus, they might do an earlier action to avoid the temptation that they anticipate their future selves will succumb to against her current preferences, exercising only type of commitment they have access to in this choice environment. The type of "doing it earlier" behaviour that cannot be accommodated under time consistency is when adding a new action leads to an earlier but previously available action being chosen. This is allowed for under sophistication. The "Irrelevant Alternatives Expedite" property gives a precise restriction that allows doing-it-earlier reversals, but rules out doing-it-later reversals.

Irrelevant Alternatives Expedite. If $a = c(A)$ and $a'' = c(A \cup \{a'\})$ with $a'' \neq a, a'$, then $t_{a''} < t_{a'}$.

To see how Irrelevant Alternatives Expedite accommodates behaviour that violates Independence of Irrelevant Alternatives while placing a meaningful restriction on behaviour, revisit Examples 3 and 4. Choices in Example 3 are consistent with the Irrelevant Alternatives Delay property, but inconsistent with the Irrelevant Alternatives Expedite property, since adding the option to do the task on Wednesday leads her to do instead on Friday (later than Wednesday). The choices in Example 4 are consistent with Irrelevant Alternatives Expedite but inconsistent with Irrelevant Alternatives Delay, since adding the option to do the task on Friday leads her to do it on Monday (earlier than Friday).

Proposition 3 states any sophisticated choice function satisfies the Irrelevant Alternatives Expedite property.

Proposition 3. *If c is a sophisticated choice function, then c satisfies the Independent Alternatives Expedite property.*

Proof. If c is a sophisticated choice function, then by Proposition 1, there is a $\{\succsim_\tau\}_\tau$ for which $\{\succsim_\tau\}_\tau$ and c satisfy S and IRA. But then, by the proof of Proposition 1, $A_{>\tau} \sim_\tau c(A_{>\tau})$ for each A, τ with $A_{>\tau} \neq \emptyset$.

First, suppose $a' = c_{t_{a'}}(A_{\geq t_{a'}} \cup \{a'\})$. Then, either $c(A \cup \{a'\}) = a'$ or $t_{c(A\cup\{a'\})} < t_{a'}$.

Second, suppose $a' \neq c_{t_{a'}}(A_{\geq t_{a'}} \cup \{a'\})$. Then by S, $c_{t_{a'}}(A_{\geq t_{a'}} \cup \{a'\}) = c_{t_{a'}}(A_{\geq t_{a'}})$. Thus by applying IRA sequentially backwards from $t_{a'}-1, t_{a'}-2, \ldots$ it follows that $c_\tau(A_{\geq\tau} \cup \{a'\}) = c_\tau(A_{\geq\tau})$ for each $\tau < t_{a'}$. Since $A_{\geq\tau} = (A \cup \{a'\})_{\geq\tau}$ whenever $\tau > t_{a'}$, we have $c_\tau((A \cup \{a'\})_{\geq\tau}) = c_\tau(A_{\geq\tau})$ in this case as well. Thus $c(A) = c(A \cup \{a'\})$ in this case.

It follows that if $c(A \cup \{a'\}) \neq c(A), a'$ then $t_{c(A\cup\{a'\})} < t_{a'}$. □

4 Comparing sophistication and naïveté with task completion data

This next result shows that in the partially naïve Strotz model, all naïveté in time-τ beliefs can be revealed from DiL and DiE reversals, up to limitations arising from coarseness of the space of actions available at time τ. This in turn provides a natural notion of comparative naïveté for this model. Consider two partially naïve Strotz model, each with collection of utility functions $\mathscr{U}$ but with different forecasts of future utility functions at each period, captured by $\hat{\mathscr{U}}$ and $\hat{\mathscr{U}}'$. Intuitively, we can define $\hat{\mathscr{U}}$ is more sophisticated than $\hat{\mathscr{U}}'$ if, whenever the perceived utility function in $\hat{\mathscr{U}}'$ makes a correct forecast about future behaviour, so does the corresponding perceived utility function in $\hat{\mathscr{U}}$.

Definition. Given two partially naïve Strotz choice functions, c and c' with the same $\mathscr{U}$ and with $\hat{\mathscr{U}}$ and $\hat{\mathscr{U}}'$ respectively, c is <u>more sophisticated than</u> c' if $U_{\tau'}(a) > U_{\tau'}(a')$ and $\hat{U}'_{\tau'|\tau}(a) > \hat{U}'_{\tau'|\tau}(a')$ imply $\hat{U}_{\tau'|\tau}(a) > \hat{U}_{\tau'|\tau}(a')$.

One might ask whether doing-it-earlier and doing-it-later reversals can be used to compare the naïveté of two choice functions. Example 5 shows that a partially sophisticated person can exhibit a doing-it-earlier reversal not exhibited by a fully sophisticated person with the same preferences. This intuition being more optimistic about future behaviour can sometimes work counterintuitively, since in the partially naïve Strotz model if you're currently more optimistic about your behaviour in the more distant future, you also expect your less-distant-future-selves to share that optimism, which can make you expect them to delay against your current wishes.

Example 5. Consider four actions a_0, a_1, a_2, a_3 with $t_{a_\tau} = \tau$, and $\mathscr{U}$ such that $U_\tau(a_\tau) < U_\tau(a_{\tau+1})$ for $\tau = 0, 1, 2$ and $U_\tau(a_\tau) > U_\tau(a_{\tau'})$ for $\tau = 0, 1$ and each $\tau' > \tau + 1$. A sophisticate has $a_1 = c(\{a_0, a_1, a_3\}) = c(\{a_0, a_1, a_2, a_3\})$. However, a partially naïve person who, at $\tau = 0$ incorrectly projects that $\hat{U}_{2|0}(a_2) > \hat{U}_{2|0}(a_3)$ but make the correct prediction of $\hat{U}_{1|0}(a_1) < \hat{U}_{1|0}(a_2)$ will expect herself to do it at $\tau = 2$ and delay at $\tau = 1$, and would do it at $\tau = 0$, leading to a doing-it-earlier reversal since $a_1 = c(\{a_0, a_1, a_3\})$ but $a_0 = c(\{a_0, a_1, a_2, a_3\})$. Thus $\{a_0, a_1, a_3\}$ and a_2 generate a doing-it-earlier reversal for the partially sophisticated person, but not the fully sophisticated one. This example also illustrates that a more sophisticated person need not complete an action earlier than a less sophisticated person with the same preferences.

However, if we restrict attention to reversals involving only three actions, then there is no "beliefs-about-beliefs" channel by which a change in a sophistication might affect behaviour. To show that more doing-it-earlier reversals are equivalent to more sophistication we need to introduce a restriction on $\bar{A}$ to ensure that beliefs about future behaviour can actually be revealed from choice.

Definition. $\bar{A}$ is <u>rich</u> if, for each $\tau \in \mathscr{T}_{\bar{A}}$ and each $a', a'' \in \bar{A}_{>\tau}$ with $U_\tau(a'') > U_\tau(a')$, there exists $a \in \bar{A}_\tau$ such that $U_\tau(a'') > U_\tau(a) > U_\tau(a')$.

Richness need not imply that $\bar{A}$ is infinite - but it does require that $\bar{A}$ has a sufficiently "large" number of alternatives at each period relative to later periods.

Proposition 4. *Then for two partially naïve Strotz choice functions, c and c', with the same $\mathscr{U}$, if c is more sophisticated than c', then for any three actions a, a', a'' with $t_a < t_{a'}, t_{a''}$ and $a = c(\{a, a'\})$, $a'' = c(\{a, a''\})$ and $a' = c(\{a', a''\})$, $a = c'(\{a, a', a''\})$ implies $a = c(\{a, a', a''\})$. If $\bar{A}$ is rich and c and c' make the same choices when two options are available, then the converse also holds.*

Proposition 4 says that in the partially naïve Strotz model, when restricting to the three action case, a more sophisticated choice function exhibits more doing-it-earlier reversals. This is equivalent to the (analogously-phrased) statement that, when restricting to the three action case, a more sophisticated choice function exhibits fewer doing-it-later reversals; this equivalence comes from the fact that when three actions involve a choice cycle, behaviour when all three are available necessarily involves either a doing-it-earlier or a doing-it-later reversal, but not both.

Comparison with Ahn et al. (2015). In a two-period setting, Ahn et al. provide a definition for a comparative of when person 1 is more naïve than person 2. Their two-period setting is unable to capture naïveté about future behaviour when choices are made at more than two periods, since they do not provide any restrictions on continuation menus are evaluated in such a setting. However, in the case where a person only faces actions in at most three periods $t_1 < t_2 \leq t_3$, the person only faces at most two choices so long as there is at most one action available at t_3. Since my comparative notion of sophistication here makes this restriction, it aligns with Ahn et al.'s, although (unlike Ahn et al.) I restrict such comparisons to individuals with the same preferences over actions (i.e. same $\mathscr{U}$).

5 Revealing sophistication and naïveté under quasi-hyperbolic discounting

5.1 Doing-it-earlier and doing-it-later reversals and dominance violations

As in Example 2, let $\mathscr{S} \subseteq \mathbb{R}^T$ for $1 \leq T \leq \infty$, $\mathscr{T}_{\bar{A}} \subseteq \mathbb{Z}_+$, and $\bar{A} \subseteq \mathscr{S} \times \mathscr{T}_{\bar{A}}$. Restrict attention to the partially naïve quasi-hyperbolic discounting model given by (1) and (2). In this model, a person's decision-making is captured by a quadruple, $\left(u, \delta, \beta, \hat{\beta}\right)$.

In the partially naïve quasi-hyperbolic discounting model, the condition that $\beta = 1$ is a sufficient condition for a person's collection of utility functions to be time consistent; this condition will also evidently be a necessary condition for time consistency under mild technical assumptions. Lemma 1 then implies that choices induced by this model will violate IIA if and only if $\beta < 1$. This implication extends O'Donoghue and Rabin's (1999) Proposition 5.2, which covers only the cases of full sophistication ($\hat{\beta} = \beta$) and full naïveté ($\hat{\beta} = 1$). However, such a result about the link between the existence of reversals and time inconsistency of a person tells us nothing about the person's self-awareness.

Proposition 5 shows that IIA violations can provide hallmarks of naïveté and sophistication. It shows that doing-it-later reversals generically indicate some degree of naïvete (i.e. $\hat{\beta} > \beta$), while doing-it-earlier reversals indicate some degree of sophistication (i.e. $\hat{\beta} < 1$) in the partially naïve quasi-hyperbolic discounting model.

Proposition 5. *If c has a partially naïve quasi-hyperbolic representation, $[\underline{m}, \overline{m}] \times \{0\}^\infty \subseteq \mathscr{S}$, $\{0,1,2\} \subseteq \mathscr{T}_{\bar{A}}$, and $u(\underline{m}) < \beta\delta^2 u(\overline{m})$, then: (i) $\hat{\beta} > \beta$ if and only if c exhibits a doing-it-later reversal, and (ii) $1 > \hat{\beta}$ if and only if c exhibits a doing-it-earlier reversal.*

The 'only if' direction of the proof of Proposition 5 uses the model to directly construct such a reversal in a three-period case. Because of the structure imposed by the partially naïve quasi-hyperbolic representation, the 'if' direction of the proof emerges as the contapositive of Propositions 2 and 3, which establish (without restriction to the partially naïve quasi-hyperbolic representation) that naïveté rules out doing-it-earlier reversals while sophistication rules out doing-it-later reversals.[4]

An (arguably) more basic axiom of choice than IIA is that a person never chooses a "dominated" action. I give a behavioural definition of dominance below adapted to the choice environ-

[4] The notion that it is possible to add unused tasks that induce delay when $\hat{\beta} > \beta$ lies behind O'Donoghue and Rabin's (2001) Propositions 2 and 5, though they only allow environments in which the set of tasks in each period is the exact same in terms of payoff structure as in preceding periods, and they only on the extreme case of failing to do any task.

ment of this section under the assumption that u is increasing and $u(0) = 0$.[5]

Definition. A choice function c satisfies dominance if for any two actions $((x_0, x_1, \dots), t)$ and $((x'_0, x'_1, \dots), t')$ in $\bar{A}$:

(i) if $t' > t$, $x_\tau \le 0$ for each $\tau < t' - t$, $x_\tau \le x'_{\tau+t-t'}$ for all $\tau \ge t' - t$, and either $x_\tau < 0$ for some $\tau < t' - t$ or $x_\tau < x'_{\tau+t-t'}$ for some $\tau \ge t' - t$, then c never chooses (x,t) from an action set that includes (x', t'), and

(ii) if $t' < t$, $x'_\tau \ge 0$ for all $\tau < t - t'$, $x_\tau \le x'_{\tau+t'-t}$ for all $\tau > t - t'$, and either $x'_\tau > 0$ for some $\tau < t - t'$ or $x_\tau < x'_{\tau-t'+t}$ for some $\tau \ge t - t'$, then c never chooses (x,t) from an action set that includes (x', t').

If c violates requirement (i), then say that it exhibits a doing-it-later dominance violation, and if it violates requirement (ii), then say it exhibits a doing-it-earlier dominance violation.

In the "only if" proof of Proposition 5 used actions for which there is only one non-zero (utility) payoff. However, actions that generate with non-zero utilities at more than one point in time can be used to show that when $\beta < 1$, subjects violate dominance (as shown in O'Donoghue and Rabin's (1999) Proposition 5.1). The nature of dominance violations will depend crucially on a person's self-awareness of her present bias ($\hat{\beta}$), as Proposition 5 suggests. In particular, a person who is partially naïve ($\hat{\beta} > \beta$) will exhibit behaviour that violates dominance by failing to do an earlier dominating action, while an a person who is partially sophisticated ($\hat{\beta} < 1$) will exhibit behaviour that violates dominance by doing an earlier dominated action.

Proposition 6. *If c has a partially naïve quasi-hyperbolic representation, $T \ge 3$, any $[\underline{m}, \overline{m}] \times [\underline{m}, \overline{m}] \times [\underline{m}, \overline{m}] \times \{0\}^{T-3} \subseteq \mathscr{S}$, $\{0,1,2\} \subseteq \mathscr{T}_{\bar{A}}$, $\underline{m} < 0 < \overline{m}$, and $u(\underline{m}) < \beta\delta^2 u(\overline{m})$, then: (i) if $\hat{\beta} > \beta$, then c exhibits doing-it-later dominance violations and (ii) if $1 > \hat{\beta}$ then c exhibits doing-it-earlier dominance violations.*

A related conjecture is that, when comparing two partially naïve people with the same preferences u, β, δ but with $\beta \le \hat{\beta}_1 < \hat{\beta}_2 \le 1$, that person 1 will act earlier, in any action set, than person 2. O'Donoghue and Rabin (1999), Proposition 2, prove this conjecture holds true when only considering pure sophistication ($\hat{\beta}_1 = \beta$) and naïveté ($\hat{\beta}_2 = 1$). The conjecture that this extends to cases that involve partial naïveté is incorrect, as shown in the example below.

Example 6. Consider two people who have the same preferences: a linear $u(x) = x$, $\delta = 1$, and $\beta = 1$, but person 1 is sophisticated while person 2 is partially naïve with $\hat{\beta}_2 = .8$. Suppose $T = 1$ and each must do a task that entails immediate costs and no benefits at one of $\tau = 0,1,2,3$; the set of available tasks is given by $\{(-40,0), (-60,1), (-70,2), (-100,3)\}$. Facing this set, person 1 will do the task at $\tau = 1$ while person 2 will do the task at $\tau = 0$.

[5] The definition below ignores dominance violations in choosing between actions available at the same date. Such violations will never occur in the partially naïve quasi-hyperbolic discounting model.

5.2 Measuring and comparing self-awareness

Behaviour when two completion opportunities are available can reveal time preference parameters β, δ, and the utility function u. Conditional on time preferences, behaviour when facing three completion opportunities can reveal an individual's perception of her future behaviour, captured by the parameter $\hat{\beta}$. Below, I illustrate how conditional on β, δ, and u, the bounds on the parameter $\hat{\beta}$ can be revealed from behaviour in simple 3-period, 3-task examples.

Suppose that we know β, δ, and u.[6] Consider the three actions that involve only an immediate cost, $((x,0,0,\dots),0)$, $((y,0,0,\dots),1)$, and $((z,0,0,\dots),2)$ with $x,y,z<0$. Pick x,y,z such that $\beta\delta^2 u(z) < u(x) < \beta\delta u(y) < \beta^2\delta^2 u(z)$; behaviourally, this is revealed by:

$$((y,0,0,\dots),1) = c\left(\{((x,0,0,\dots),0),((y,0,0,\dots),1)\}\right),$$

$$((x,0,0,\dots),0) = c\left(\{((x,0,0,\dots),0),((z,0,\dots),2)\}\right),$$

$$((z,0,0,\dots),2) = c\left(\{((y,0,0,\dots),1),((z,0,\dots),2)\}\right).$$

Now consider how this person would behave when she has the three actions available, $((x,0,0,\dots),0),((y,0,0,\dots),1)$, and $((z,0,\dots),2)$. Her choices imply that, if she expects that she would delay at $t=1$, then she would want to do $((x,0,0,\dots),0)$ at $t=0$. In partially naïve quasi-hyperbolic discounting model, this occurs when:

$$u(y) < \hat{\beta}\delta u(z)$$

which occurs if and only if $\hat{\beta} < \frac{u(y)}{\delta u(z)}$. If instead at $t=0$ she thinks that she would do $((y,0,0,\dots),1)$ at $t=1$, then she would want to wait at $t=0$. In the model, this occurs when:

$$u(y) \geq \hat{\beta}\delta u(z)$$

which occurs if and only if $\hat{\beta} \geq \frac{u(y)}{\delta u(z)}$. This example illustrates how simple experiments can be used to measure an individual's $\hat{\beta}$ from task completion behaviour.

[6] Existing papers have shown how the parameters of the quasi-hyperbolic model can be measured from a person's $t=0$ ranking of consumption streams, e.g. Attema et al. (2010); Olea and Strzalecki (2014).

6 Extension to models with intrinsic self-control costs

Intrinsic self-control costs model. Gul and Pesendorfer (2001; 2004) introduce a model in which an individual demands commitment due to her desire to avoid tempting options. Adapting their dynamic model to the setting here, a person's decision-making at period τ maximizes a recursively-defined value function, W_τ, derived from her set of normative utility functions $\{u_{\tau'}\}_{\tau'}$, temptation utility functions $\{v_{\tau'}\}_{\tau'}$, and discount factor δ, given by:

$$W_\tau(A_{\geq\tau}) = \max_{x\in A_{=\tau}\cup\{A_{>\tau}\}} [u_\tau(x) + v_\tau(x) + \delta W_{\tau+1}(x)] - \max_{x\in A_{=\tau}\cup\{A_{>\tau}\}} v_\tau(x) \tag{3}$$

We normalize that $v_\tau(x) = 0\,\forall x \notin \bar{A}_{=\tau}$ (the temptation utility of waiting is zero), in line with the assumption in Gul and Pesendorfer (2004) that temptation utility derives only from current actions.

Since W_τ dislikes tempting options even if they are not chosen, the $\succsim_\tau$ and c consistent with this model may not satisfy IRA. I introduce below a weaker notion of sophistication for models that incorporate instrinsic cost of resisting temptation, DRA (dislike of redundant alternatives), which weakens IRA; this is an adaptation of Lipman and Pesendorfer's (2013) axiomatic definition of sophistication for this model.

DRA. If $t_a > \tau$, $A_{\geq t_a} \cup \{a\} \succ_\tau A_{\geq t_a}$ implies $a = c_{t_a}(A_{\geq t_a} \cup \{a\})$.

The Gul and Pesendorfer model satisfies DRA.

Proposition 7. *If there exists a δ and sets $\{u_\tau\}_\tau, \{v_\tau\}_\tau, \{W_\tau\}_\tau$ related by (3), then the binary relation $\succsim_\tau$ on $\bar{A}_{\geq\tau} \cup \mathscr{A}_{>\tau}$ represented by W_τ is complete and transitive for each τ and if each W_τ strictly ranks $\bar{A}_{=\tau}$, then $\{\succsim_\tau\}_\tau$ induces a unique choice function c that satisfies S. This pair $\{\succsim_\tau\}_\tau$, c satisfies DRA.*

The Gul-Pesendorfer model allows for a preference for commitment without dynamic inconsistency. Intuitively, this intrinsic preference for commitment due to the desire to avoid future costly temptation gives an additional motive to do it sooner. Proposition 8 shows that whenever the pair $\{\succsim_\tau\}_\tau$ and c satisfies S and DRA, then c also satisfies the Independent Alternatives Expedite property.

Proposition 8. *If c satisfies S and DRA, then c satisfies the Independent Alternatives Expedite property.*

The proof of Proposition 8 closely follows 3, and is provided in the Appendix.

7 Discussion

This paper makes precise how adding unused options can lead to delay but cannot expedite task completion for a naïve but not a sophisticated person, establishing this behaviour as a hallmark of naïveté. Similarly, adding unused options can expedite but not delay for a sophisticated person. These results establish intuitive empirical tests for sophistication and naïveté. Each model has additional content, and the examples here show how content of these models that are not implied by the Independent Alternatives Delay or Irrelevant Alternatives Expedite properties suggest intuitive tests. The examples and results here suggests how empirical work on task completion can be used to measure naïveté versus sophistication from behaviour, and how related but model-specific results can be obtained for models of partial naïveté. The results of such tests would provide a means of evaluating the appropriateness of alternative assumptions about sophsitication and naïveté in applications of models of time inconsistent preferences.

One implication of the analysis here is that firms that have data on task completion can learn the sophistication or naïveté of their clients or employees, especially if they can experiment. For example, a financial institution that observes when a client pays her bills can use this information to target her with financial products that exploit her degree of naïveté, without having to offer her a menu of contracts that screen for this subject to an incentive compatibility constraint. Similarly, a manager who observes when an employee completes work assignments can use this information to infer the degree of sophistication of her employee, and can use this to better tailor his work responsibilities and deadlines.

Appendix: Proofs.

Proof of Lemma 1

Proof. Suppose c has a time consistent representation, $a = c(A)$, and $a \in A' \subset A$. Then $U_1(a) > U_1(a') \forall a \in A$, thus $U_1(a) > U_1(a') \forall a \in A'$. It follows that $a = c(A')$, and thus IIA holds.

Now suppose c satisfies IIA and $\bar{A}$ is finite. Let $a_1 = c(\bar{A})$, $a_2 = c(\bar{A} \backslash \{a_1\})$, ..., $a_k = c(\bar{A} \backslash \{a_1, \dots, a_{k-1}\})$, stop when$\{a_1, \dots, a_{k-1}\} = \bar{A}$. By IIA, $c(A) = a_{\min[k:\, a_k \in A]}$. Define $V(a_k) = |\bar{A}| - k$. Now define $U_\tau(a_k) = V(a_k) \forall a_k \in \bar{A}_{\geq \tau}$. By construction, $\mathscr{U} = \{U_\tau\}_{\tau \in \mathscr{T}_{\bar{A}}}$ is time consistent, and $c(A) = \underset{a \in A}{\arg\max}\, U_1(a)$. □

Proof of Proposition 1. The proof for part (ii) has four steps. Lemma 2 shows that naïveté has an equivalent, simpler representation which I then make use of in subsequent Lemmas. Lemma 3 shows that given S, IWA holds if and only if $\succsim_\tau$ is indifferent between any continuation menu and its favourite item on that menu, i.e. $A_{\geq \tau'} \sim_\tau m(A_{\geq \tau'}, \succsim_\tau)$. Lemma 4. Thus Lemmas 3 and 4 imply sufficiency of the axioms. Lemma 5 shows us that given any naïve representation we can construct $\succsim_\tau$'s to satisfy S; then applying Lemmas 4 and 3 implies that IWA holds as well, establishing necessity of the axioms.

□

Lemma 2. *c is a naïve choice function represented by $\mathscr{U}$, if and only if* $c_\tau(A_{\geq\tau}) = \begin{cases} \arg\max\limits_{a\in A_{=\tau}} U_\tau(a) & \text{if } \max\limits_{a\in A_{=\tau}} U_\tau(a) = \max\limits_{a\in A_{\geq\tau}} U_\tau(a) \\ A_{>\tau} & \text{otherwise} \end{cases}$.

Proof. Notice that c is a naïve choice function if and only if $c_\tau(A_{\geq\tau}) = s_\tau^{\mathscr{U},\hat{\mathscr{U}}}(A)$ for all A. Constructing the PPE given $\mathscr{U}$ and $\hat{\mathscr{U}}$, we obtain $s_\tau^{\mathscr{U},\hat{\mathscr{U}}}(A_{=\tau}) = \arg\max\limits_{a\in A_{=\tau}} U_\tau(a) = c_\tau(A_{=\tau})$ for each A,τ follows by our construction of U_τ. Then, noting that $\min\left[\check{\tau} > \tau : s_{\check{\tau}}^{\hat{\mathscr{U}}_{\cdot|\tau,s}}(A_{\geq\check{\tau}}) \neq A_{>\check{\tau}}\right] = \min\left[\check{\tau} > \tau : \max\limits_{a\in A_{=\check{\tau}}} U_\tau(a) \geq \max\limits_{a'\in A_{>\check{\tau}}} U_\tau(a')\right]$, we can equivalently write: $s_\tau^{\mathscr{U},\hat{\mathscr{U}}}(A) = \begin{cases} \arg\max\limits_{a\in A_{=\tau}} U_\tau(a) & \text{if } \max\limits_{a\in A_{=\tau}} U_\tau(a) \geq \max\limits_{a'\in A_{>\tau}} U_\tau(a') \\ A_{>\tau} & \text{otherwise} \end{cases}$. Thus, $c_\tau(A_{\geq\tau}) = a$ if and only if $s_\tau^{\mathscr{U},\hat{\mathscr{U}}}(A) = a$ if and only if $U_\tau(a) \geq U_\tau(a')$ for all $a' \in A_{\geq\tau}$ and with strict inequality for all $a' \in A_{=\tau}$. □

Lemma 3. *If c and $\{\succsim_\tau\}_\tau$ satisfy S, then they also satisfy IWA if and only if $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$ for any $\tau, \tau' > \tau$, and A.*

Proof. With notational sloppiness, I will sometimes view a' and $\{a'\}$ as interchangable from the perspective of $\succsim_\tau$ for $\tau < t_{a'}$, using the statement that $a \succsim_\tau a'$ to denote that $a \succsim_\tau \{a'\}$ when $t_{a'} > \tau$ and $m(A_{\geq\tau}, \succsim_\tau)$ to denote the $\succsim_\tau$-maximal action in $A_{\geq\tau}$.

Suppose c and $\{\succsim_\tau\}_\tau$ satisfy S.

Consider the following relaxation of IWA, I call k-IWA, where k is an integer. IWA holds if k-IWA holds for arbitrary k.

***k*-IWA.** For any τ and $\tau' > \tau$, $A_{\geq\tau'} \sim_\tau m_\tau(A_{\geq\tau'}, \succsim_\tau)$ so long as $A_{\geq\tau'}$ has actions available in at most k periods.

Since $m_{\tau'}(A_{=\tau'}, \succsim_\tau) = m(A_{=\tau'}, \succsim_\tau)$, 1-IWA holds if and only if $A_{=\tau'} \sim_\tau m(A_{=\tau'}, \succsim_\tau)$ for any τ, $\tau' > \tau$, and $A_{=\tau}$.

Now we proceed by induction to show that $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$. Suppose that k-IWA holds, and that whenever $A_{\geq\tau'}$ has actions that are available at most at k distinct times, that we have $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$. Now consider an $A_{\geq\tau'}$ with actions available at $k+1$ distinct times, and the earliest action available at τ'. Then, $A_{>\tau'} \sim_\tau m(A_{>\tau'}, \succsim_\tau)$. Then $k+1$-IWA holds if and only if $A_{\geq\tau'} \sim_\tau m_{\tau'}(A_{\geq\tau'}, \succsim_\tau)$. If for some $a \in A_{=\tau'}$ we have $a \succsim_\tau A_{>\tau'}$, then $A_{\geq\tau'} \sim_\tau m_{\tau'}(A_{\geq\tau'}, \succsim_\tau$

$) = m(A_{=\tau'}, \succsim_\tau) \succsim m(A_{>\tau'}, \succsim_\tau)$ (by k-IWA) and thus $m_{\tau'}(A_{\geq\tau'}, \succsim_\tau) = m(A_{\geq\tau'}, \succsim_\tau)$. Otherwise, $A_{\geq\tau'} \sim_\tau m_{\tau'}(A_{\geq\tau'}, \succsim_\tau) = A_{>\tau'} \sim_\tau m(A_{>\tau'}, \succsim_\tau)$ and $A_{>\tau'} \succ_\tau a$ for all $a \in A_{=\tau'}$, thus $m(A_{>\tau'}, \succsim_\tau) = m(A_{\geq\tau'}, \succsim_\tau)$. Thus, $k+1$-IWA holds if and only $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$ for each $A_{\geq\tau'}$ with actions available at $k+1$ or fewer periods.

By induction, it follows that $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$ for any $A_{\geq\tau'}$with $\tau' > \tau$ holds if and only if k-IWA holds for all k, or equivalently, IWA holds. $\square$

Lemma 4. *Let each U_τ represent $\succsim_\tau$ restricted to $\bar{A}_{\geq\tau}$. Then if S is satisfied, then $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$ for any $\tau, \tau' > \tau$, and A if and only c has a naïve representation.*

Proof. Since S implies that each $\succsim_\tau$ is complete and transitive, if $\bar{A}$ is finite, then each $\succsim_\tau$ restricted to $\bar{A}_{\geq\tau}$ has a utility representation, U_τ, with $U_\tau(a) \geq U_\tau(a')$ if and only if $a \succsim_\tau a'$ (or the analogous ranking with one or both being singleton continuation menus). Then, $m(A_{\geq\tau'}, \succsim_\tau) = \underset{a \in A_{\geq\tau'}}{\arg\max}\, U_\tau(a)$. Then applying the representation of $\succsim_\tau$, if S(i) holds, this is equivalent to choice being given by $c_\tau(A_{\geq\tau}) = \begin{cases} \underset{a\in A_{=\tau}}{\arg\max} U_\tau(a) & \text{if } \underset{a\in A_{=\tau}}{\max} U_\tau(a) = \underset{a\in A_{\geq\tau}}{\max} U_\tau(a) \\ A_{>\tau} & \text{otherwise} \end{cases}$. By Lemma 2, this is equivalent to c having a naïve representation. $\square$

Lemma 5. *4. If c is a naïve choice function, we can construct $\succsim_\tau$ from $\mathscr{U}$ by $a \succ_\tau a'$ iff $U_\tau(a) > U_\tau(a')$, and $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$ for each $A, \tau, \tau' > \tau$. This $\{\succsim_\tau\}_\tau$ and c will satisfy S.*

Proof. Let c be a naïve choice function. Construct such a $\succsim_\tau$; by construction $\succsim_\tau$ is complete and transitive.

If $a = c_\tau(A_{=\tau} \cup A_{>\tau})$ then, by the equivalent naïve representation in Lemma 2, $U_\tau(a) \geq U_\tau(a')$ for all $a' \in A_{>\tau}$ and also strictly for all $a' \in A_{=\tau}$. Thus, $a \succsim_\tau a'$ for all $a' \in A_{>\tau}$ and $a \succ_\tau a'$ for all $a' \in A_{=\tau}$. Since we picked $\succsim_\tau$ to satisfy $A_{\geq\tau'} \sim_\tau m(A_{\geq\tau'}, \succsim_\tau)$, it follows that $a \succsim_\tau A_{>\tau}$.

If $A_{>\tau} = c_\tau(A_{=\tau} \cup A_{>\tau})$, then by the equivalent naïve representation in Lemma 2, there exists an $a' \in A_{>\tau}$ for which $U_\tau(a') > U_\tau(a)$ for all $a \in A_{=\tau}$. By our choice of $\succsim_\tau$, $a' \succ_\tau a$ for all $a \in A_{=\tau}$. Then, since $A_{>\tau} \sim_\tau m(A_{>\tau}, \succsim_\tau) \sim_\tau a'$ by construction, transitivity of $\succsim_\tau$ implies that $A_{>\tau} \succ_\tau a$ for all $a \in A_{=\tau}$.

Thus S holds for this choice of $\{\succsim_\tau\}_\tau$. $\square$

The proof for the sophisticated case follows an analogous outline.

Lemma 6. *c is a sophisticated choice function represented by $\mathscr{U}$, if and only if* $c_\tau(A_{\geq\tau}) = \begin{cases} \underset{a\in A_{=\tau}}{\arg\max} U_\tau(a) & \textit{if } \underset{a\in A_{=\tau}}{\max} U_\tau(a) \geq U_\tau(c(A_{>\tau})) \\ A_{>\tau} & \textit{otherwise} \end{cases}$.

Proof. Notice that c is a sophisticated choice function if and only if $c_\tau(A_{\geq\tau}) = s_\tau^{\mathscr{U},s}(A)$ for all A. Constructing the PPE given $\mathscr{U}$, we obtain $s_\tau^{\mathscr{U},s}(A_{=\tau}) = \arg\max_{a\in A_{=\tau}} U_\tau(a) = c_\tau(A_{=\tau})$ for each A,τ follows by our construction of U_τ. Then, noting that $\min\left[\check{\tau} > \tau : s_{\check{\tau}}^{\mathscr{U},s}(A_{\geq\check{\tau}}) \neq A_{>\check{\tau}}\right] = \min[\check{\tau} > \tau : c_{\check{\tau}}(A_{\geq\check{\tau}}) \neq A_{>\check{\tau}}]$, we can equivalently write:

$$c_\tau(A_{\geq\tau}) = \begin{cases} \arg\max_{a\in A_{=\tau}} U_\tau(a) & \text{if } \max_{a\in A_{=\tau}} U_\tau(a) \geq U_\tau(c(A_{>\tau})) \\ A_{>\tau} & \text{otherwise} \end{cases}. \qquad \square$$

Lemma 7. *If c and $\{\succsim_\tau\}_\tau$ satisfy S, then they also satisfy IRA if and only if $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'})$ for any $\tau, \tau' > \tau$, and A.*

Proof. Suppose c and $\{\succsim_\tau\}_\tau$ satisfy S.

Consider the following relaxation of IRA, I call k-IRA, where k is an integer. IRA holds if and only if k-IRA holds for arbitrary k.

***k*-IRA.** For any τ and $\tau' > \tau$, $A_{\geq\tau'} \sim_\tau c_{\tau'}(A_{\geq\tau'})$ so long as $A_{\geq\tau'}$ has actions available in at most k periods.

Since $c_{\tau'}(A_{=\tau'}) = c(A_{=\tau'})$ for any $A_{=\tau'}$, 1-IRA holds if and only if $A_{=\tau'} \sim_\tau c(A_{=\tau'})$ for any τ, $\tau' > \tau$, and $A_{=\tau}$.

Now we proceed by induction to show that $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'})$. Suppose that k-IRA holds, and that whenever $A_{\geq\tau'}$ has actions that are available at most at k distinct times, that we have $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'})$. Now consider an $A_{\geq\tau'}$ with actions available at $k+1$ distinct times, and the earliest action available at τ'. Then, $A_{>\tau'} \sim_\tau c(A_{>\tau'})$. Then $k+1$-IRA holds if and only if $A_{\geq\tau'} \sim_\tau c_{\tau'}(A_{\geq\tau'})$, I now show this is equivalent to the condition that $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'})$. If for some $a \in A_{=\tau'}$ we have $a \succsim_{\tau'} A_{>\tau'}$, then (by S), $A_{\geq\tau'} \sim_\tau c_{\tau'}(A_{\geq\tau'}, \succsim_\tau) = c(A_{\geq\tau'})$. Otherwise (by S and then k-IRA), $A_{\geq\tau'} \sim_\tau c_{\tau'}(A_{\geq\tau'}) = A_{>\tau'} \sim_\tau c(A_{>\tau'}) = c(A_{\geq\tau})$. Thus, $k+1$-IRA holds if and only $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'})$ for each $A_{\geq\tau'}$ with actions available at $k+1$ or fewer periods.

By induction, it follows that $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'})$ for any $A_{\geq\tau'}$with $\tau' > \tau$ holds if and only if k-IRA holds for all k, or equivalently, IRA holds. $\square$

Lemma 8. *Let each U_τ represent $\succsim_\tau$ restricted to $\bar{A}_{\geq\tau}$. Then if S is satisfied, then $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'}, \succsim_\tau)$ for any $\tau, \tau' > \tau$, and A if and only c has a sophisticated representation.$a = c'(\{a, a', a''\})$*

Proof. Since S implies that each $\succsim_\tau$ is complete and transitive, if $\bar{A}$ is finite, then each $\succsim_\tau$ restricted to $\bar{A}_{\geq\tau}$ has a utility representation, U_τ, with $U_\tau(a) \geq U_\tau(a')$ if and only if $a \succsim_\tau a'$ (or the analogous ranking with one or both being singleton continuation menus). Then applying the representation of $\succsim_\tau$, if S(i) holds, this is equivalent to choice being given by $c_\tau(A_{\geq\tau}) =$

$\begin{cases} \arg\max_{a \in A_{=\tau}} U_\tau(a) & \text{if } \max_{a \in A_{=\tau}} U_\tau(a) \geq U_\tau(c(A_{>\tau})) \\ A_{>\tau} & \text{otherwise} \end{cases}$. By Lemma 6, this is equivalent to c having a sophisticated representation. □

Lemma 9. *4. If c is a sophisticated choice function, we can construct $\succsim_\tau$ from $\mathscr{U}$ by $a \succ_\tau a'$ iff $U_\tau(a) > U_\tau(a')$, and $A_{\geq\tau'} \sim_\tau c(A_{\geq\tau'})$ for each $A, \tau, \tau' > \tau$. This $\{\succsim_\tau\}_\tau$ and c will satisfy S.*

Proof. Let c be a sophisticated choice function. Construct such a $\succsim_\tau$, and extend it to rank $A \succsim_\tau A'$ whenever there exist a, a' such that $A \sim_\tau a \succsim_\tau a' \sim_\tau A'$; by construction $\succsim_\tau$ is complete and transitive.

If $a = c_\tau(A_{=\tau} \cup A_{>\tau})$ then, by the equivalent sophisticated representation in Lemma 6, $U_\tau(a) \geq U_\tau(c(A_{>\tau}))$ and also $U_\tau(a) > U_\tau(a')$ for all $a' \in A_{=\tau}$. Thus, by our choice of $\succsim_\tau$, $a \succsim_\tau c(A_{>\tau}) \sim_\tau A_{>\tau}$ thus $a \succsim_\tau A_{>\tau}$, and $a \succ_\tau a'$ for all $a' \in A_{=\tau}$.

If $A_{>\tau} = c_\tau(A_{=\tau} \cup A_{>\tau})$, then by the equivalent sophisticated representation in Lemma 6, $U_\tau(c(A_{>\tau})) > U_\tau(a)$ for all $a \in A_{=\tau}$. By our choice of $\succsim_\tau$, $A_{>\tau} \sim_\tau c(A_{>\tau}) \succ_\tau a$ for all $a \in A_{=\tau}$. Then, transitivity of $\succsim_\tau$ implies that $A_{>\tau} \succ_\tau a$ for all $a \in A_{=\tau}$.

Thus S holds for this choice of $\{\succsim_\tau\}_\tau$. □

Proof of Proposition 4.

Proof. WLOG, I assume that each U_τ and $\hat{U}_{\tau'|\tau}$ is strict in order to avoid the caveats when dealing with how choice reveals each.

Suppose c is more sophisticated than c'. Then if we take actions a, a', a'' with $t_a < t_{a'}, t_{a''}$ and $a = c(\{a, a'\})$, $a'' = c(\{a, a''\})$ and $a' = c(\{a', a''\})$, by the definition of "more sophisticated than", c' makes the same pairwise choices. By the representation, we must have $U_{t_a}(a'') > U_{t_a}(a) > U_{t_a}(a')$ but $U_{\min[t_{a'}, t_{a''}]}(a') > U_{\min[t_{a'}, t_{a''}]}(a'')$, and by the partially naïve PPE, $a = c'(\{a, a', a''\})$ holds iff we have $\hat{U}'_{\min[t_{a'}, t_{a''}]|t_a}(a') > \hat{U}'_{\min[t_{a'}, t_{a''}]|t_a}(a'')$. By the definition of more sophisticated than, $\hat{U}_{\min[t_{a'}, t_{a''}]|t_a}(a') > \hat{U}_{\min[t_{a'}, t_{a''}]|t_a}(a'')$ as well.

Now suppose $\bar{A}$ is rich, c and c' make the same pairwise choices, and for any three actions a, a', a'' with $t_a < t_{a'}, t_{a''}$ and $a = c(\{a, a'\})$, $a'' = c(\{a, a''\})$ and $a' = c(\{a', a''\})$, $a = c'(\{a, a', a''\})$ implies $a = c(\{a, a', a''\})$.

Consider any a', a'' and $\tau < \min[t_{a'}, t_{a''}]$ for which $U_\tau(a'') > U_\tau(a')$. Since $\bar{A}$ is rich, there exists an $a \in \bar{A}_{=\tau}$ with $U_\tau(a'') > U_\tau(a) > U_\tau(a')$. Thus, both c and c' will have $a = c(\{a, a'\})$, $a'' = c(\{a, a''\})$, and $a' = c(\{a', a''\})$. By the definition of PPE, $a = c'(\{a, a', a''\})$ holds if and only if $\hat{U}'_{\min[t_{a'}, t_{a''}]|\tau}(a') > \hat{U}'_{\min[t_{a'}, t_{a''}]|\tau}(a'')$. By our assumption, $a = c'(\{a, a', a''\})$ also implies $a = c(\{a, a', a''\})$, which implies $\hat{U}_{\min[t_{a'}, t_{a''}]|\tau}(a') > \hat{U}_{\min[t_{a'}, t_{a''}]|\tau}(a'')$. Thus, c is more sophisticated than c'. □

Proof of Proposition 5.

Proof. "Only if" part.

Without loss of generality, assume $\underline{m} \leq 0 < \overline{m}$ and $u(0) = 0$.

(i)

Suppose $\hat{\beta} > \beta$. Then since u is continuous, by the intermediate value theorem we can take $x, y, z \in [\underline{m}, \overline{m}]$ such that $\hat{\beta}\delta^2 u(z) > \delta u(y) > \beta\delta^2 u(z) > u(x) > \beta\delta u(y)$. Then

$$((x,0,0,\dots),0) = c\left(\{((x,0,0,\dots),0),((y,0,0,\dots),1)\}\right)$$

but

$$((y,0,0,\dots),1) = c\left(\{((x,0,0,\dots),0),((y,0,0,\dots),1),((z,0,0,\dots),2)\}\right).$$

(ii)

Suppose $\hat{\beta} < 1$. Then since u is continuous, by the intermediate value theorem we can take $x, y, z \in X$ such that $\beta\hat{\beta}\delta^2 u(z) < \beta\delta u(y) < u(x) < \beta\delta^2 u(z)$. Since $u(x) < \beta\delta^2 u(z)$, we have

$$((z,0,0,\dots),2) = c\left(\{((x,0,0,\dots),0),((z,0,0,\dots),2)\}\right)$$

but

$$((x,0,0,\dots),0) = c\left(\{((x,0,0,\dots),0),((y,0,0,\dots),1),((z,0,0,\dots),2)\}\right).$$

"If" part.

(i).

It is equivalent to prove contrapositive, i.e. show that if $\hat{\beta} = \beta$, then there are no reversals of the doing-it-later variety. This result will follow from Proposition 3.

(ii).

As in the proof of (i), it is equivalent to prove the contrapositive, i.e .show that if $\hat{\beta} = 1$, then there are no reversals of the doing-it-earlier variety. Then the result will follow from Proposition 2. □

Proof of Proposition 6.

Proof. (i) Suppose $\hat{\beta} > \beta$. Take $z \in \text{int}(\underline{m}, \overline{m})$. Since 0 and z are both in $(\underline{m}, \overline{m})$ and u is continuous, there exist strictly positive ε_1 and ε_2 such that $u(-\varepsilon_1) + \hat{\beta}\delta u(z+\varepsilon_2) > \hat{\beta}\delta u(z)$ but $\beta\delta u(z) > u(-\varepsilon_1) + \beta\delta u(z+\varepsilon_2)$. Then at $t = 0$, this person would expect herself to choose $((-\varepsilon_1, z+\varepsilon_2, 0, 0, \dots), 1)$ over $((z,0,0,\dots),2)$ if at $t = 1$ she were to have only these two actions available, even though she would actually delay at $t = 1$ and end up doing $((z,0,0,\dots),2)$.

Since $u(-\varepsilon_1)+\hat{\beta}\delta u(z+\varepsilon_2)>\hat{\beta}\delta u(z)$, we have $u(-\varepsilon_1)+\delta u(z+\varepsilon_2)>\delta u(z)$. Thus by continuity of u and since $0\in \text{int}[\underline{m},\overline{m}]$, there exists an $\varepsilon_0>0$ such that $\beta\delta u(-\varepsilon_1)+\delta^2 u(z+\varepsilon_2)>u(\varepsilon_0)+\beta\delta^2 u(z)$. By construction,

$$((z,0,0,\dots),2)=c\left(\{((\varepsilon_0,0,z,0,0,\dots),0),((-\varepsilon_1,z+\varepsilon_2,0,0,\dots),1),((z,0,0,\dots),2)\}\right)$$

which violates dominance.

(ii) Suppose $\hat{\beta}<1$. Take $z\in(\underline{m},\overline{m})$. Since 0 and z are both in $(\underline{m},\overline{m})$ and u is continuous, there exist strictly positive ε_1 and ε_2 such that $\hat{\beta}\delta u(z)>u(-\varepsilon_1)+\hat{\beta}\delta u(z+\varepsilon_2)$ but $u(-\varepsilon_1)+\delta u(z+\varepsilon_2)>\delta u(z)$. Then at $t=0$, this person would (correctly, since $\beta\leq\hat{\beta}$) expect herself to pass on $((-\varepsilon_1,z+\varepsilon_2,0,0,\dots),1)$ to wait to do $((z,0,0,\dots),2)$ if at $t=1$ she were to have only these two actions available.

Then at $t=0$, since u is continuous and $u(-\varepsilon_1)+\delta u(z+\varepsilon_2)>\delta u(z)$, there exists a $\varepsilon_0>0$ such that $u(-\varepsilon_0)+\beta\delta u(-\varepsilon_1)+\beta\delta^2 u(z+\varepsilon_2)>\beta^2\delta u(z)$. Thus, this person would choose

$$((-\varepsilon_0,-\varepsilon_1,z+\varepsilon_2,0,0,\dots),0)$$

$$=c\left(\{((-\varepsilon_0,-\varepsilon_1,z+\varepsilon_2,0,0,\dots),0),((-\varepsilon_1,z+\varepsilon_2,0,0,\dots),1),((z,0,0,\dots),2)\}\right)$$

which violates dominance. □

Proof of Proposition 7.

Proof. Since each W_τ is a utility function on $A_{=\tau}\cup\mathscr{A}_{>\tau}$, it represents a complete and transitive binary relation and when it strictly ranks all elements in $\bar{A}_{=\tau}$, then S gives the choice function c_τ at each τ that maximizes $\succsim_\tau$ with ties broken in favour of the earlier action.

To see that this pair must satsify DRA, consider any A,a, and $\tau<t_a$ for which $A_{\geq t_a}\cup\{a\}\succ_\tau A_{\geq t_a}$. By the representation, this is equivalent to $W_\tau(A_{\geq t_a}\cup\{a\})>W_\tau(A_{\geq t_a})$. Since $A_{=\tau'}=\emptyset$ for each $\tau'\in\{\tau,\dots,t_a\}$, (3) implies that $W_{\tau'}(A_{\geq t_a}\cup\{a\})>W_{\tau'}(A_{\geq t_a}))$. Since $A_{\geq t_a}\subset A_{\geq t_a}\cup\{a\}$, $\max\limits_{x\in A_{=t_a}\cup\{A_{>t_a}\}} v_{t_a}(x)\leq \max\limits_{x\in A_{=t_a}\cup\{a\}\cup\{A_{>t_a}\}} v_{t_a}(x)$, and thus for each $x\in A_{=\tau}\cup\{A_{>\tau}\}$, $[u_{t_a}(x)+v_{t_a}(x)+\delta W_{t_a+1}(x)]-\max\limits_{x\in A_{=t_a}\cup\{A_{>t_a}\}} v_{t_a}(x)\leq[u_{t_a}(x)+v_{t_a}(x)+\delta W_{t_a+1}(x)]-\max\limits_{x\in A_{=t_a}\cup\{a\}\cup\{A_{>t_a}\}} v_{t_a}(x)$. Thus by the representation in (3), $W_{t_a}(A_{\geq t_a}\cup\{a\})>W_{t_a}(A_{\geq t_a})$ implies that $a=\arg\max\limits_{x\in A_{=t_a}\cup\{a\}\cup\{A_{>t_a}\}}[u_{t_a}(x)+v_{t_a}(x)+\delta W_{t_a+1}(x)]$, thus $a=c_{t_a}(A_{\geq t_a}\cup\{a\})$ by S. □

Proof of Proposition 8.

Proof. Take $\{\succsim_\tau\}_\tau$ and c that satisfy S and DRA.

First, suppose $a' = c_{t_{a'}}(A_{\geq t_{a'}} \cup \{a'\})$. Then, either $c(A \cup \{a'\}) = a'$ or $t_{c(A\cup\{a'\})} < t_{a'}$.

Second, suppose $a' \neq c_{t_{a'}}(A_{\geq t_{a'}} \cup \{a'\})$. Then by S, $c_{t_{a'}}(A_{\geq t_{a'}} \cup \{a'\}) = c_{t_{a'}}(A_{\geq t_{a'}})$. Thus by applying DRA sequentially backwards from $t_{a'} - 1, t_{a'} - 2, \ldots$, since $A_{\geq \tau} \succsim_\tau A_{\geq \tau} \cup \{a'\}$ for each $\tau < t_{a'}$, thus, by S, we have that $a'' = c_\tau(A_{\geq \tau}) \implies a'' = c_\tau((A \cup a)'_{\geq \tau})$ for such τ. Thus $t_{c(A)} < t_{a'}$ implies $t_{c(A\cup\{a'\})} < t_{a'}$. Since $A_{\geq \tau} = (A \cup \{a'\})_{\geq \tau}$ whenever $\tau > t_{a'}$, if $\emptyset = c_\tau(A_{\geq \tau} \cup \{a'\})$ for all $\tau \leq t_{a'}$, then we have $c_\tau((A \cup \{a'\})_{\geq \tau}) = c_\tau(A_{\geq \tau})$ for all $\tau > t_{a'}$ in this case, and thus $c(A) = c(A \cup \{a'\})$.

It follows that if $c(A \cup \{a'\}) \neq c(A), a'$ then $t_{c(A\cup\{a'\})} < t_{a'}$. $\square$

References

Ahn, David, Ryoto Iijima, and Todd Sarver, "Comparative Measures of Naiveté," Technical Report 2015.

Ariely, Dan and Klaus Wertenbroch, "Procrastination, deadlines, and performance: Self-control by precommitment," *Psychological Science*, 2002, *13* (3), 219–224.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines," *Quarterly Journal of Economics*, 2006, pp. 635–672.

Attema, Arthur E, Han Bleichrodt, Kirsten IM Rohde, and Peter P Wakker, "Time-tradeoff sequences for analyzing discounting and time inconsistency," *Management Science*, 2010, *56* (11), 2015–2030.

Attema, Arthur, Han Bleichrodt, Yu Gao, Zhenxing Huang, and Peter Wakker, "Measuring discounting without measuring utility," *American Economic Review*, 2016.

Augenblick, Ned and Matthew Rabin, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," Technical Report 2015.

Bisin, Alberto and Kyle Hyndman, "Procrastination, Self-Imposed Deadlines and Other Commitment Devices," Technical Report 2014.

Blow, Laura, Martin Browning, and Ian Crawford, "Never mind the hyperbolics: nonparametric analysis of time-inconsistent preferences," Technical Report 2015.

Burger, Nicholas, Gary Charness, and John Lynham, "Field and online experiments on self-control," *Journal of Economic Behavior & Organization*, 2011, *77* (3), 393–404.

DellaVigna, Stefano and Ulrike Malmendier, "Contract design and self-control: Theory and evidence," *Quarterly Journal of Economics*, 2004, *119*, 353–402.

Echenique, Federico, Taisuke Imai, and Kota Saito, "Testable Implications of Models of Intertemporal Choice: Exponential Discounting and Its Generalizations," Technical Report 2015.

Fang, Hanming and Yang Wang, "Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions," *International Economic Review*, 2015, *56* (2), 565–596.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and self-control," *Econometrica*, 2001, *69* (6), 1403–1435.

_ **and** _ , "Self-control and the theory of consumption," *Econometrica*, 2004, pp. 119–158.

_ **and** _ , "The revealed preference theory of changing tastes," *Review of Economic Studies*, 2005, *72* (2), 429–448.

Hayashi, Takashi, "Quasi-stationary cardinal utility and present bias," *Journal of Economic Theory*, 2003, *112* (2), 343–352.

Heidhues, Paul and Botond Kőszegi, "Exploiting naivete about self-control in the credit market," *American Economic Review*, 2010, *100*, 2279–2303.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-control and the development of work arrangements," *American Economic Review P&P*, 2010, *100* (2), 624–628.

Laibson, David, "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 1997, pp. 443–477.

Lipman, Bart and Wolfgang Pesendorfer, "Temptation," 2013, pp. 243–288.

Noor, Jawwad, "Temptation and revealed preference," *Econometrica*, 2011, *79* (2), 601–644.

O'Donoghue, Ted and Matthew Rabin, "Doing it now or later," *American Economic Review*, 1999, pp. 103–124.

_ **and** _ , "Choice and procrastination," *Quarterly Journal of Economics*, 2001, pp. 121–160.

Ok, Efe A and Yusufcan Masatlioglu, “A theory of (relative) discounting,” *Journal of Economic Theory*, 2007, *137* (1), 214–245.

Olea, José Luis Montiel and Tomasz Strzalecki, “Axiomatization and measurement of quasi-hyperbolic discounting,” *The Quarterly Journal of Economics*, 2014, *129* (3), 1449–1499.

Quah, John K-H and Bruno Strulovici, “Discounting, values, and decisions,” *Journal of Political Economy*, 2013, *121* (5), 896–939.

Schotter, Andrew and Isabel Trevino, “Belief elicitation in the laboratory,” *Annual Review of Economics*, 2014, *6* (1), 103–128.

Strotz, Robert, “Myopia and inconsistency in dynamic utility maximization,” *Review of Economic Studies*, 1955, pp. 165–180.