Doing It Now or Later

Ted O'Donoghue
The Center for Mathematical Studies in Economics and Management Sciences
Northwestern University

and

Matthew Rabin
Department of Economics
University of California, Berkeley

First Draft: May 16, 1996
This Draft: December 9, 1996

Abstract

Though economists assume that intertemporal preferences are time-consistent, evidence suggests that a person's relative preference for well-being at an earlier moment over a later moment increases as the earlier moment gets closer. We explore the behavioral and welfare implications of such time-inconsistent preferences in a simple model where a person must engage in an activity exactly once during some duration. We focus on two sets of distinctions. First, do choices involve salient costs — where the costs of an action are immediate but any rewards are delayed — or do they involve salient rewards — where the rewards of an action are immediate but any costs are delayed? Second, are people sophisticated — they foresee future self-control problems — or are they naive — they do not anticipate these self-control problems? Naive people procrastinate activities with salient costs and preproperate — do too soon — activities with salient rewards. If costs are salient, sophistication mitigates procrastination, and can even lead sophisticated people to do the activity sooner than if they had no self-control problem. If rewards are salient, sophistication exacerbates preproperation. These behavioral results have corresponding welfare implications: With salient costs, mild self-control problems can severely damage a person only if she is naive, while with salient rewards mild self-control problems can severely damage a person only if she is sophisticated. We also consider a multiple-activity version of the model, and discuss how our results might apply to savings, addiction, and other behaviors.

Keywords: Doing It, Hyperbolic Discounting, Preproperation, Procrastination, Time Inconsistency.

JEL Classifications: A12, B49, C70, D11, D60, D74, D91, E21

Acknowledgments: We thank Steven Blatt, Erik Eyster, and Clara Wang for useful research assistance, and Steven Blatt, Erik Eyster, David Laibson, and seminar participants at UC Berkeley, Northwestern, Russell Sage Foundation, Columbia, Yale, and Harvard/MIT for helpful comments. O'Donoghue thanks the Alfred P. Sloan Foundation and Rabin thanks the Alfred P. Sloan and Russell Sage Foundations for financial support.

Mail: Ted O'Donoghue / The Center for Mathematical Studies in Economics and Management Sciences / Northwestern University / 2001 Sheridan Road, Room 371 / Evanston, IL 60208-2014. Matthew Rabin / Department of Economics / 549 Evans Hall #3880 / University of California, Berkeley / Berkeley, CA 94720-3880.

Email: eodonogh@skew2.kellogg.nwu.edu and rabin@econ.berkeley.edu.

1. Introduction

People are impatient ˘ they like to experience rewards soon and to delay costs until later. Economists almost always capture impatience by assuming that people discount streams of utility over time exponentially. Such preferences are time-consistent: A person's relative preference for well-being at an earlier date over a later date is the same no matter when she is asked.

Casual observation, introspection, and psychological research all suggest that the assumption of time-consistency is importantly wrong.[1] It ignores the human tendency to grab immediate rewards and to avoid immediate costs in a way that our "long-run selves" do not appreciate. For example, when presented a choice between doing 7 hours of an unpleasant task on April 1 versus 8 hours on April 15, if asked on February 1 virtually everyone would prefer the 7 hours on April 1. But come April 1, given the same choice, most of us are apt to put off the work until April 15. We call such tendencies presently preferences: When considering trade-offs between two future moments, presently preferences give stronger relative weight to the earlier moment as it gets closer.[2]

In this paper, we explore the behavioral and welfare implications of presently preferences in a simple model where a person must engage in an activity exactly once during some length of time. Our analysis emphasizes two sets of distinctions. The first distinction is whether choices involve salient costs ˘ where the costs of an action are immediate but any rewards are delayed ˘ or salient rewards ˘ where the benefits of an action are immediate but any costs are delayed. By exploring these two different settings under the rubric of presently preferences, we unify the investigation of phenomena (e.g., procrastination and overeating) that have often been explored separately, but which clearly come from the same underlying propensity for immediate gratification.

The second distinction is whether people are sophisticated, and foresee that they will have self-control problems in the future, or people are naive, and don't foresee these self-control problems. By explicitly comparing these competing assumptions ˘ each of which has received attention in

---

1 Loewenstein (1992) reviews how the economics profession evolved from perceiving exponential discounting as a useful, ad hoc approximation of intertemporal-choice behavior, to perceiving it as a fundamental axiom of (rational) human behavior.

2 Many researchers have studied such preferences. Several researchers have posited a specific functional form, hyperbolic discounting, to account for such preferences. See Ainslie (1991, 1992), Ainslie and Haslam (1992b), Ainslie and Herrnstein (1981), Chung and Herrnstein (1967), and Loewenstein and Prelec (1992). We have contrived the term "presently preferences" as a more descriptive term for the underlying human characteristic that hyperbolic discounting represents. A small set of economists have over the years proposed formal models of time-inconsistent preferences more generally. See Goldman (1979, 1980), Koopmans (1960), Phelps and Pollak (1968), Pollak (1968), and Strotz (1955). For papers modeling time-variant tastes more generally, see Glazer and Weiss (1992), Hammond (1976a,b), Koopmans (1960), Peleg and Yaari (1973), Pessemier (1978), Pollak (1970, 1978), and Yaari (1977).

the economics literature ˘ we hope to delineate which predictions come from presently preferences per se, and which come from these assumptions about foresight.[3]

In Section 2, we further motivate and formally define presently preferences and other general features of our framework. We then describe a simplified form of presently preferences (originally proposed by Phelps and Pollak (1968) and later employed by Laibson (1994)) that we study in this paper: Relative to time-consistent preferences, a person always gives extra weight to well-being now over any future moment, but weighs all future moments equally. In Section 3, we set up our model of a one-time activity, and translate the general issues from Section 2 into the context of our specific model. We suppose that a person must engage in an activity exactly once during some length of time. Importantly, at each moment the person can choose only whether or not to do it now, and cannot choose when later she will do it. Within this scenario, we consider a general class of reward and cost schedules for completing the activity.

Section 4 explores the behavioral implications of presently preferences in our model. We present two simple results characterizing how behavior depends on whether rewards or costs are salient, and on whether people are sophisticated or naive. The presently effect characterizes the direct implications of presently preferences: You procrastinate ˘ wait when you should do it ˘ if actions involve salient costs (writing a paper), and preproperate ˘ do it when you should wait ˘ if actions involve salient rewards (seeing a movie). Naive people are influenced solely by the presently effect. The sophistication effect characterizes the direct implications of sophistication versus naivete: A sophisticated person does the activity sooner than does a naive person with the same preferences, irrespective of whether rewards or costs are salient. Intuitively, a sophisticated person is correctly pessimistic about her future behavior ˘ a naive person believes she will behave herself in the future while a sophisticated person knows she may not. As a result, waiting always seems less attractive for a sophisticated person. Although the direction is the same, the sophistication effect has very different connotations for salient costs versus salient rewards. When costs are salient, sophistication mitigates the tendency to procrastinate. (And in fact, the sophistication effect can outweigh the presently effect so that a sophisticated person may perform an onerous task before she would if

[3] Strotz (1955) and Pollak (1968), two of the seminal papers on time-inconsistent preferences, carefully lay out these two assumptions, but do not much consider the implications of one versus the other. More recent papers have assumed either one or the other, without attempting to justify the choice on behavioral grounds. For instance, Akerlof (1991) assumes naive beliefs, while Laibson (1994,1995) assumes sophisticated beliefs. Each paper states its assumption about beliefs used (and Akerlof (1991) posits that his main welfare finding depends on his assumption of naive beliefs), but conspicuously neither paper argues why its assumption is correct.

she had no self-control problem.) When rewards are salient, on the other hand, sophistication exacerbates the tendency to preproperate.

While our main concern is predicting behavior given time inconsistency, in Section 5 we show how some of the behavior we predict can arise only from time inconsistency. The literature often cites the use of self-limiting "commitment devices" (e.g., Christmas clubs, fat farms) as proof that people are time-inconsistent, since no time-consistent agent would ever limit future choice sets. Section 5 presents some specific examples where people with presently preferences engage in behavior that time-consistent agents would never exhibit, no matter their impatience. Our examples show that such "smoking guns" can exist even within our simple model, where there are no explicit commitment devices. Furthermore, although the literature has focused on smoking guns for sophisticated people, we identify such smoking guns for naive people as well.

In Section 6, we turn to the welfare results.[4] Again, the two distinctions ˘ salient costs vs salient rewards and sophistication vs naivete ˘ are crucial. When costs are salient, a person is always better off with sophisticated beliefs than with naive beliefs. Naivete can lead you to repeatedly procrastinate an unpleasant task under the incorrect belief that you will do it tomorrow, while sophistication means you know exactly how costly delay would be. In fact, even with an arbitrarily small amount of impatience, for salient costs naive people can experience severe welfare losses, while the welfare loss from a small amount of impatience is small if you are sophisticated.[5] When rewards are salient, however, a person can be better off with naive beliefs. In this case, people with presently preferences tend to do the activity when they should wait. Naivete helps motivate you to wait because you overestimate the benefits of waiting. Sophistication makes you (properly) skeptical of future behavior, so you are more tempted to grab today's salient reward. This can lead to "unwinding" similar to that in the finitely repeated prisoners' dilemma: In the end, you will give in to temptation and grab a reward too soon; because you realize this, near the end you will cave in a little sooner than if you thought you would resist temptation in the end; realizing this, you will cave in a little sooner, etc. As a result, for salient rewards it is sophisticated people who can

[4] Welfare comparisons for people with time-inconsistent preferences are in principle problematic; the very premise of the model is that a person's preferences disagree at different times, so that a change in behavior may make some selves better off while making other selves worse off. We feel the natural perspective in most situations is the "long-run perspective" ˘ what you would wish now (if you were fully informed) about your profile of future behavior. However, few of our comparisons rely on this perspective, and most of our welfare comparisons can be roughly conceived of as "Pareto comparisons," where one outcome is better than another from all of a person's vantage points.

[5] In other words, naivete can lead to very bad outcomes, and sophistication limits how badly you can do. This result formalizes and generalizes a similar claim made in Akerlof (1991), and implied in Strotz (1955).

experience severe welfare losses with an arbitrarily small amount of impatience, while the welfare loss from a small amount of impatience is small if you are naive.

One goal of this paper is to develop a simple model that will begin to delineate some general implications of time-inconsistent preferences and self-control problems for formal economic analysis. We feel our simple setting encompasses an important class of situations, and enables us to lay bare some basic principles. But important contexts in which self-control problems are very important (e.g., savings and addiction) cannot be satisfactorily analyzed in our one-time-activity model. To begin the process of generalizing our model, we present in Section 7 an extension where, rather than being performed exactly once, the activity must be performed more than once during some length of time. We show that for naive people the lessons of the one-time-activity model more or less carry over. For sophisticated people, however, things can change dramatically: Behavior can be very sensitive to how many times the activity must be performed, and the effects of sophistication vs naivete become more subtle.

Finally, in Section 8 we discuss how the various lessons from our model might extend to particular areas of application. Our analysis points to some surprising results in savings behavior and in addiction. For instance, sophisticated people with presently preferences may save more than people with time-consistent preferences who are more patient, and they may respond to a permanent increase in the price of an addictive good by increasing short-term consumption. Along with results from earlier in the paper, these examples suggest that many types of surprising and counter-intuitive behaviors arise not from presently preferences per se, but from presently preferences combined with sophistication effects.

2. Presently Preferences

In this section, we formally define presently preferences and other elements of our model. Throughout this paper, we will use a simplified form of presently preferences, (β, δ)-preferences, which we define below. To provide some context for these preferences, however, we briefly present a more general model of intertemporal preferences. Let u_t be somebody's instantaneous utility in period t.[6] A person in period t cares not only about her present instantaneous utility, but also about her future instantaneous utilities. We let $U^t(u_t, u_{t+1}, ..., u_T)$ represent a person's intertemporal preferences from the perspective of period t, where U^t is continuous and increasing in all components.

If a person makes decisions in periods $0, 1, ..., T$, we must consider the relationship among her intertemporal utility functions for each vista, $U^0, U^1, ..., U^T$. The standard economist's assumption about this relationship is time-consistency:

Definition 1 Intertemporal preferences $\{U^t\}_{t=0}^T$ are time-consistent if for all t, for all $t' \in \{1, ..., T\}$ such that $t < t'$, for all $\{u_k\}_{k=t}^T$, and for all $\{\tilde{u}_k\}_{k=t'}^T$,

$$U^t(u_t, ..., u_{t'}, ...u_T) > U^t(u_t, ...u_{t'-1}, \tilde{u}_{t'}, ..., \tilde{u}_T) \quad \text{if and only if} \quad U^{t'}(u_{t'}, ...u_T) > U^{t'}(\tilde{u}_{t'}, ..., \tilde{u}_T)$$

Definition 1 says that for any fixed sequence of instantaneous utilities prior to period t', U^t and $U^{t'}$ yield the same preference ordering on instantaneous utilities beginning in period t'. But there is evidence that intertemporal preferences are not time-consistent. We formally define a specific form of time-inconsistent preferences: When considering trade-offs between two future moments, presently preferences give stronger relative weight to the earlier moment as it gets closer.[7]

Definition 2 Intertemporal preferences $\{U^t\}_{t=0}^T$ are presently if

a) for all $t < t' \leq \tau < \tau'$, and for all $\{u_k\}_{k=t}^T$ and $(\tilde{u}_\tau, \tilde{u}_{\tau'})$ such that $u_\tau > \tilde{u}_\tau$:

$$U^t(u_t, ..., u_\tau, ..., u_{\tau'}, ...u_T) = U^t(u_t, ..., \tilde{u}_\tau, ..., \tilde{u}_{\tau'}, ...u_T) \quad \text{implies}$$
$$U^{t'}(u_{t'}, ..., u_\tau, ..., u_{\tau'}, ...u_T) \geq U^{t'}(u_{t'}, ..., \tilde{u}_\tau, ..., \tilde{u}_{\tau'}, ...u_T) \text{ ; and}$$

b) there exists $t < t' \leq \tau < \tau'$, $\{u_k\}_{k=t}^T$ and $(\tilde{u}_\tau, \tilde{u}_{\tau'})$ such that $u_\tau > \tilde{u}_\tau$,

$$U^t(u_t, ..., u_\tau, ..., u_{\tau'}, ...u_T) = U^t(u_t, ..., \tilde{u}_\tau, ..., \tilde{u}_{\tau'}, ...u_T) \text{ ,} \quad \text{and}$$
$$U^{t'}(u_{t'}, ..., u_\tau, ..., u_{\tau'}, ...u_T) > U^{t'}(u_{t'}, ..., \tilde{u}_\tau, ..., \tilde{u}_{\tau'}, ...u_T)$$

If the period-t person is indifferent between the utility profiles $(u_\tau, u_{\tau'})$ and $(\tilde{u}_\tau, \tilde{u}_{\tau'})$, then if

[6] While it is common to assume that a person's instantaneous utility u_t depends only on her consumption bundle in period t, many researchers have outlined how past consumption and other reference points can affect a person's current preferences. (See Kahneman, Knetsch, and Thaler (1991) and Becker, Grossman, Murphy (1991).) Loewenstein and Prelec (1992) and Hoch and Loewenstein (1991) emphasize the interaction between the issues explored in this paper and loss aversion and other reference-level effects. Our formulation assumes that such effects can be represented as arguments within the instantaneous utility functions, $u_t(\cdot)$.

[7] We coined the term "presently preferences" to connote that people's preferences have a bias for the "present" over the "future" (where the "present" is constantly changing). But this is a new term for an array of older models that went under different names. In fact, the (β, δ)-preferences that we will use in this paper are identical to the preferences studied by Laibson (1994) and Akerlof (1991), though Akerlof frames his discussion very differently. Laibson uses the term "hyperbolic discounting" because he appeals to psychological evidence that discount functions are approximately hyperbolic. Akerlof uses the term "salience" because he appeals to psychological evidence that costs incurred today are more salient than costs incurred in the future.

$u_\tau > \tilde{u}_\tau$ it must be that $u_{\tau'} < \tilde{u}_{\tau'}$. Presently preferences imply that closer to time τ, the person will then prefer the pair $(u_\tau, u_{\tau'})$, which does better sooner than $(\tilde{u}_\tau, \tilde{u}_{\tau'})$.

The difference between time-consistent and presently preferences is typically captured with two functional forms for discounting. Time-consistent preferences are generally modeled with (time-constant) exponential discounting: For all t, $U^t(u_t, u_{t+1}, ..., u_T) \equiv \sum_{\tau=t}^{T} \delta^\tau u_\tau$, where $\delta \in (0, 1]$ is a "discount factor." Those researchers studying what we have dubbed presently preferences often assume hyperbolic discounting: For all t, $U^t(u_t, u_{t+1}, ..., u_T) \equiv \sum_{\tau=t}^{T} \frac{1}{(1+\tau-t)} u_\tau$.[8]

In this paper, we adopt an elegant simplification for presently preferences developed by Phelps and Pollak (1968), and later employed by Laibson (1994). They capture the most basic form of presently preferences ˘ a bias for the "present" over the "future" ˘ with a simple two-parameter model that modifies exponential discounting:

Definition 3 (β, δ)-preferences are preferences that can be represented by:

$$\text{For all } t, \quad U^t(u_t, u_{t+1}, ..., u_T) \equiv \delta^t u_t + \beta \sum_{\tau=t+1}^{T} \delta^\tau u_\tau \qquad \text{where } 0 < \beta,\ \delta \leq 1$$

In this model, δ represents long-run, time-consistent discounting. The parameter β, on the other hand, represents "short-term impatience" ˘ how you favor now versus later. If $\beta = 1$, then (β, δ)-preferences are simply exponential discounting. But $\beta < 1$ implies presently preferences: The person gives more relative weight to period τ in period τ than she did in any period prior to period τ.

Researchers have converged on a simple strategy for modeling time-inconsistent preferences: The person at each point in time is modeled as a separate "agent" who is choosing her current behavior to maximize current preferences, where her future selves will control her future behavior. In such a model, we must ask what a person believes about her future selves' preferences. Strotz (1955) and Pollak (1968) carefully lay out two extreme assumptions. A person could be sophisticated and know exactly what her future selves' preferences will be. Or, a person could be naive

[8] As noted by Loewenstein and Prelec (1993), a more general form of hyperbolic discounting uses discount factor $\frac{1}{1 + a(\tau-t)}$, where $a > 0$ is a measure of impatience (i.e., a larger a implies more impatience). Ainslie (1992) observes that hyperbolic discounting is connected to formulations of a "matching law" of intertemporal choice (see also Chung and Herrnstein (1967) and Ainslie and Herrnstein (1981)).

That exponential discounting implies time-consistency while hyperbolic discounting implies presently preferences relies on an assumption that the discount function is the same for all vistas, $U^0, U^1, ..., U^T$. If the discount function changes, the reverse relationship can hold. Time-dependent exponential discounting can imply presently preferences: For all t, $U^t(u_t, u_{t+1}, ..., u_T) \equiv \sum_{\tau=t}^{T} \delta_t^\tau u_\tau$, where for all t $\delta_t > \delta_{t+1}$. Date-specific hyperbolic discounting can imply time-consistent preferences: For all t, $U^t(u_t, u_{t+1}, ..., u_T) \equiv \sum_{\tau=t}^{T} \frac{1}{(1+\tau)} u_\tau$.

and believe her future selves' preferences will be identical to her current self's, not realizing that as she gets closer to executing decisions her tastes will have changed.

With (β, δ)-preferences, β completely captures a person's "presentliness". The implications of sophistication versus naivete depend on a person's perceptions of future presentliness, which we denote by $\hat{\beta}$. In any period $t' < t$ the person believes her period-t utility function will be $U^t(u_t, u_{t+1}, ..., u_T) \equiv \delta^t u_t + \hat{\beta} \sum_{\tau=t+1}^{T} \delta^\tau u_\tau$. This formulation makes four simplifying assumptions. First, a person uses a "point estimate" of her future presentliness: Rather than having probabilistic beliefs, she is completely certain (though possibly wrong) that her presentliness is $\hat{\beta}$. Second, a person uses the same estimate for all future selves. Third, there is "consensus" ˘ all selves prior to the period-t self have the same perception of the period-t self's presentliness. Finally, a person believes it is common knowledge among all future selves that all future selves use the perception $\hat{\beta}$.

With this formulation, people are sophisticated if $\hat{\beta} = \beta$, and people are naive if $\hat{\beta} = 1$. We could, of course, imagine more intermediate assumptions where $\beta < \hat{\beta} < 1$: You are aware that your future selves will have presently preferences, but you underestimate the degree of presentliness. Except for some brief comments in Section 8, we focus in this paper entirely on the two extreme assumptions.

Are people sophisticated or naive?[9] The use of self-commitment devices, such as alcohol clinics, Christmas clubs, or fat farms, provides evidence of sophistication.[10] Only sophisticated people would want to commit themselves to smaller choice sets: If you were naive, you would never worry that your tomorrow self might choose an option you don't like today. Despite the existence of some sophistication, however, it does appear that people underestimate the degree to which their future behavior won't match their current preferences over future behavior. For example, people may repeatedly not have the "will power" to forgo tempting foods or to quit smoking, while predicting that tomorrow they will have this will power. We think there are elements of both sophistication

[9] Most economists modeling time-inconsistent preferences assume sophistication. Indeed, sophistication implies that people have "rational expectations" about future behavior, so it is a natural assumption for economists. Akerlof (1991) uses a variant of the naivete assumption.

[10] The very term "self control" implies that people are aware that it may be prudent to control their future selves. For analyses of self control in people, see Ainslie (1974,1975,1987,1992), Ainslie and Haslam (1992a), Funder and Block (1989), Glazer and Weiss (1992), Hoch and Lowenstein (1991), Laibson (1994, 1995), Schelling (1978,1984,1992a), Shefrin and Thaler (1992), Thaler (1980), Thaler and Shefrin (1981), and Wertenbroch (1993). Ainslie (1974) explores similar issues with pigeons. As many have emphasized, especially Ainslie (1992) and Watterson (1993, pp. 83-88), a sort of intrapersonal "bargaining" can arise because of the basic disagreements we have with ourselves about when we should do something.

and naivete in the way people anticipate their own future preferences. In any event, our goal is to clarify the logic of each, and in the process we delineate which predictions come purely from presently preferences, and which come from the "sophistication effects" of people being aware of their own time inconsistency.[11]

3. Doing It Once

Suppose there is an activity that a person must perform exactly once, and there are $T < \infty$ periods in which she can do it. Let $\mathbf{v} \equiv (v_1, v_2, ..., v_T)$ be the reward schedule, and let $\mathbf{c} \equiv (c_1, c_2, ..., c_T)$ be the cost schedule, where $v_t \geq 0$ and $c_t \geq 0$ for each $t \in \{1, 2, ..., T\}$. In period $t \in \{1, 2, ..., T-1\}$, the person must choose either to do it or to wait. If she does the activity, she receives reward v_t but incurs cost c_t, and makes no further choices. If she waits, she then will face the same choice in period $t+1$. Importantly, she cannot choose when later to do it if she waits. If the person waits until period T, she must do it then.

The reward schedule $\mathbf{v}$ and the cost schedule $\mathbf{c}$ represent rewards and costs as a function of when the person does the activity. However, rewards and costs are not necessarily immediate upon completion of the activity. Indeed, we differentiate cases precisely by when rewards and costs are experienced. Some activities, such as writing a paper or mowing the lawn, are unpleasant to perform, but create future benefits. The costs of performing the activity are salient, whereas the rewards are dispersed over future dates, and far less salient. This is our first case: salient costs. Contrarily, other activities, such as seeing a movie or taking a vacation, are pleasurable to perform, but may create future costs (e.g., forgone future consumption). Here, rewards are salient, costs are not. This is our second case: salient rewards. While we concentrate primarily on these two cases, we occasionally make reference to a third case, that both rewards and costs are salient.[12]

We analyze these cases using the (β, δ)-preferences outlined in Section 2. For simplicity, we

[11] Not all aspects of impulsive choice are captured by the framework of time-inconsistent discounting. See Loewenstein (1996). For instance, while we are emphasizing that people tend to overweight rewards and costs that are in close temporal proximity, people also tend to overweight rewards and costs that are in close spatial proximity, and more generally are attentive to rewards and costs that are salient. That is, we are focusing in this paper only on temporal salience, not on all elements of salience. Nonetheless, we feel that time-inconsistency represents a simple important step that improves the behavioral realism of models of intertemporal choice.

[12] The fourth case ˘ neither rewards nor costs are salient ˘ is not of interest because it is equivalent to the case of time consistency, which we study.

assume $\delta = 1$; i.e., we assume that there is no "long-term" discounting.[13] Given $\delta = 1$, without loss of generality we can interpret delayed rewards or costs as being experienced in period $T + 1$. Then, the only periods in which a person has non-zero instantaneous utility are the period in which she performs the activity and period $T + 1$. We can thus translate our cases into the language of Section 2 as follows:[14]

1) Salient Costs: If a person does the activity in period τ, then instantaneous utilities are

$$u_\tau = -c_\tau \qquad u_{T+1} = v_\tau \qquad u_t = 0 \quad \text{for all}\, t \notin \{\,\tau\,,\ T+1\,\}$$

Hence, for each $t \leq \tau$, intertemporal utility is

$$U^t(\tau) \equiv \begin{cases} \beta\, v_\tau - c_\tau & \text{if } \tau = t \\ \beta\, v_\tau - \beta\, c_\tau & \text{if } \tau > t \end{cases}$$

2) Salient Rewards: If a person does the activity in period τ, then instantaneous utilities are

$$u_\tau = v_\tau \qquad u_{T+1} = -c_\tau \qquad u_t = 0 \quad \text{for all}\, t \notin \{\,\tau\,,\ T+1\,\}$$

Hence, for each $t \leq \tau$, intertemporal utility is

$$U^t(\tau) \equiv \begin{cases} v_\tau - \beta\, c_\tau & \text{if } \tau = t \\ \beta\, v_\tau - \beta\, c_\tau & \text{if } \tau > t \end{cases}$$

A strategy $s \equiv (s_1, s_2, ..., s_T) \in \{Y, N\}^T \equiv S$ specifies for each period $t \in \{1, 2, ..., T\}$ whether or not to do it in period t given the person has not yet done it. The strategy s specifies doing it in period t if $s_t = Y$ and waiting if $s_t = N$. Since the person must do it in period T if she has not yet done it, without loss of generality we require $s_T = Y$. A strategy is a complete description of how a person will behave in this environment. In addition to specifying when the person will do it, a strategy also specifies what the person "would" do in periods after she has already done it; e.g., if $s_t = Y$, we still specify $s_{t'}$ for all $t' > t$.[15]

We will define below two functions, U^t_Y and U^t_N. In this notation, the superscript on U denotes the vista from which utility is considered, and the subscript on U denotes the action choice for the period denoted in the superscript. U^t_Y is the utility to the period-t self from doing it in period t, and U^t_N is the perceived utility to the period-t self from waiting in period t. To determine the action

[13] The results are easily generalized to $\delta < 1$. Suppose the "true" reward schedule is $\boldsymbol{\pi} \equiv (\pi_1, \pi_2, ..., \pi_T)$, the "true" cost schedule is $\boldsymbol{\phi} \equiv (\phi_1, \phi_2, ..., \phi_T)$, and $\delta < 1$. If, for instance, costs are salient and rewards are received in period $T+1$, then if we let $v_t \equiv \delta^{T+1}\pi_t$ and $c_t \equiv \delta^t \phi_t$ for each t, doing the analysis with $\mathbf{v}$, $\mathbf{c}$, and no discounting is identical to doing the analysis with $\boldsymbol{\pi}$, $\boldsymbol{\phi}$, and δ.

[14] Our assumption that a person's "base" instantaneous utility is zero is purely for convenience. For instance, suppose costs are salient and rewards are received in period $T + 1$. For any $\overline{u}$ and $\widehat{u}$, the following formulations are identical to ours: If a person does the activity in period τ, then instantaneous utilities are,

(i) $u_{T+1} = \widehat{u} + v_\tau$, $u_\tau = \overline{u} - c_\tau$, and $u_t = \overline{u}$ for all $t \notin \{\tau, T + 1\}$, or

(ii) $u_{T+1} = \widehat{u} + v_\tau$, $u_\tau = \overline{u}$, and $u_t = \overline{u} + c_t$ for all $t \notin \{\tau, T + 1\}$.

[15] This specification of strategy corresponds to the standard definition of strategies used in game theory.

choice s_t, we must compare U_Y^t to U_N^t.

The utility to the period-t self from doing it in period t is straightforward.

$$U_Y^t(\beta) \equiv \begin{cases} \beta v_t - c_t & \text{if costs are salient} \\ v_t - \beta c_t & \text{if rewards are salient} \end{cases}$$

U_N^t is more complicated since it depends on beliefs about future behavior. If the period-t self believes she will follow continuation strategy s, then $U_N^t(s, \beta) \equiv \beta(v_\tau - c_\tau)$ where $\tau = min_{\hat{t}>t}\{\ \hat{t} \mid s_{\hat{t}} = Y\}$. Because we assume all period selves know about and agree on v and c, the crucial determinant of the strategy that the period-t self believes she will follow in future periods is her perceptions of future selves' presentliness $\hat{\beta}$. Although we restrict attention in this paper to perceptions $\hat{\beta} = \beta$ or $\hat{\beta} = 1$ (the two cases discussed in Section 2), we define our solution concept for any perception $\hat{\beta} \in (0, 1]$.

Given $\hat{\beta}$, the perceived continuation strategy is $\sigma(\hat{\beta}) \equiv \left(\sigma_2(\hat{\beta}), \sigma_3(\hat{\beta}), ..., \sigma_T(\hat{\beta})\right)$ satisfying $\sigma_T(\hat{\beta}) = Y$ and for all $t < T$ $\sigma_t(\hat{\beta}) = Y$ if and only if $U_Y^t(\hat{\beta}) \geq U_N^t(\sigma(\hat{\beta}), \hat{\beta})$. This apparatus allows us to define our solution-concept, perception-perfect strategy.

Definition 4 Given presentliness $\beta \in (0, 1]$ and perceptions of future presentliness $\hat{\beta} \in (0, 1]$, a strategy $s \equiv (s_1, s_2, ..., s_T)$ is perception-perfect if $s_T = Y$ and for all $t < T$ $s_t = Y$ if and only if $U_Y^t(\beta) \geq U_N^t(\sigma(\hat{\beta}), \beta)$.[16]

Intuitively, a perception-perfect strategy means each period self decides whether to do it using backwards-induction logic in her perceived continuation game. In the period-t self's perceived continuation game, the period-t self's payoffs reflect her true presentliness β while her future selves' payoffs reflect her perception of future presentliness $\hat{\beta}$.[17]

The concept of a perception-perfect strategy is a translation of the game-theoretic notion of subgame-perfection into our setting. Like subgame perfection in game theory, perception-perfect strategies will rule out seemingly irrational behavior "off the equilibrium path," such as "incredible

[16] We have assumed that people do it when indifferent, and that people believe their future selves will do it when indifferent, and that people believe that their future selves will believe that their future selves will do it when indifferent, and so on. For generic values of v, c, and β, nobody will ever be indifferent, so these assumptions are irrelevant. In non-generic games, a more general definition of perfection-perfect could lead to additional equilibria.

[17] For $\hat{\beta} = \beta$, perception-perfect strategy is an identical solution concept to those used by Strotz (1955), Pollak (1968), Laibson (1994, 1995), and others. In addition, for $\hat{\beta} = 1$, it is essentially the same solution concept to those used by Akerlof (1991) and Pollak (1968). It may seem strange to assume that people with incorrect perceptions (i.e., $\widehat{\beta} \neq \beta$) can do correct backwards induction given these perceptions. However, we feel it is equally strange to believe that having incorrect perceptions should imply an inability to do backwards induction. Even so, we note that for $\beta < \widehat{\beta} = 1$ (the case of incorrect perceptions we study) a perception-perfect strategy is equivalent to a strategy determined by the agent choosing in period t the action consistent with her optimal lifetime strategy for the period-t continuation game.

threats" by future selves to behave very badly (e.g., to do it in a period with an exorbitant cost). In fact, when $\hat{\beta} = \beta$, perception-perfect strategy is identical to subgame-perfect equilibrium in the game where each period self is modeled as a separate agent. When $\hat{\beta} \neq \beta$, perception-perfect strategies impose "perfection" on each perceived continuation game. The following Lemma is comparable to the well-known result that in generic games of complete and perfect information, there is a unique subgame-perfect equilibrium:

Lemma 1 For any $\beta \in (0,1]$ and $\hat{\beta} \in (0,1]$, there is a unique perception-perfect strategy.[18]

As discussed in Section 2, we focus on two extreme types of agents with presently preferences. We call people with sophisticated perceptions ($\hat{\beta} = \beta$) sophisticates. Sophisticates know exactly how presently their future selves will be, and therefore have correct expectations about future behavior. We call people with naive perceptions ($\hat{\beta} = 1$) naifs. In each period, naifs overvalue current payoffs (and know they do), but they believe they will never do so in the future. When $\beta = \hat{\beta} = 1$, preferences are standard exponential preferences and therefore time-consistent. We refer to people with time-consistent preferences as TCs.

We conclude this section by defining some notation that we will use in the next three sections. Let s^{tc}, s^s and s^n be the unique perception-perfect strategies for TCs, sophisticates, and naifs respectively. Given s^{tc}, s^s and s^n, let τ_{tc}, τ_s, τ_n be the periods in which each of the three types of agents do the activity. (I.e. given $a \in \{tc, s, n\}$, $\tau_a \equiv min_t\{t \mid s_t^a = Y\}$.)

4. Behavior

In this section, we examine the behavioral implications of presently preferences, and of sophistication versus naivete. By assuming $\delta = 1$ for all agents, naifs, sophisticates, and TCs have identical long-run preferences. Hence, a comparison of naifs or sophisticates to TCs reflects how people with presently preferences behave relative to how they would like to behave from a long-run perspective. Furthermore, a comparison of naifs to sophisticates reflects the implications of sophistication versus naivete. We begin by analyzing in some detail a pair of related examples, to demonstrate the types of behavior that can arise.

Suppose you usually go to the movies on Saturdays. The schedule at the local cinema consists

18 All proofs are in the Appendix.

of a mediocre movie this week, a good movie next week, a great movie in two weeks, and (best of all) a Johnny Depp movie in three weeks. Now suppose you must complete a report for work within four weeks, and to do so you must skip the movie on one of the next four Saturdays. When do you complete the report?

The activity you must do exactly once is the report. The reward from doing the report is received at work in the future, and hence is not salient. We'll assume the reward is independent of when you complete the report, and denote it by $\bar{v}$. The cost of doing the report on a given Saturday ˘ not seeing the movie shown that day ˘ is experienced immediately, and hence is salient. Letting valuations of the mediocre, good, great and Depp movies be 3, 5, 8, and 13, we formalize this situation in the following example, where we present both the parameters of the example and the perception-perfect strategy for each type of agent.

Example 1 Suppose costs are salient, $T = 4$, and $\beta = \frac{1}{2}$. Let $\mathbf{v} = (\bar{v}, \bar{v}, \bar{v}, \bar{v})$ and $\mathbf{c} = (3, 5, 8, 13)$.
$s^{tc} = (Y, Y, Y, Y)$, so TCs do the report in period $\tau_{tc} = 1$.
$s^{n} = (N, N, N, Y)$, so naifs do the report in period $\tau_{n} = 4$.
$s^{s} = (N, Y, N, Y)$, so sophisticates do the report in period $\tau_{s} = 2$.

TCs do the report on the first Saturday, skipping the mediocre movie. In general, TCs do the activity in the period t that maximizes $v_t - c_t$. Since Example 1 has a stationary reward schedule, TCs do the report in the period with the minimum cost.

Naifs procrastinate until the last Saturday, forcing themselves to skip the Depp movie. On the first Saturday, naifs give in to their self-control problem and see the mediocre movie because they believe they will skip the good movie in week 2 and still be able to see the great movie and the Depp movie. The period-1 naif prefers incurring a cost of 5 next week as opposed to a cost of 3 now. However, when the second Saturday arrives, naifs again give in to their self-control problem and see the good movie, now believing they will skip the great movie in week 3 and still get to see the Depp movie. Finally, when the third Saturday arrives, naifs have self-control problems for a third time and see the great movie, forcing themselves to miss the Depp movie. If either the period-1 naif or period-2 naif knew that he would have future self-control problems, then he would have done the report immediately. This example demonstrates a typical problem for naifs when costs are salient: They incorrectly predict that they will not procrastinate in the future, and consequently underestimate the cost of procrastinating now.

Sophisticates procrastinate one week, but they do the report on the second Saturday, skipping the good movie and enabling themselves to see the great movie and the Depp movie. The period-1 sophisticate correctly predicts that he would have self-control problems on the third Saturday and see the great movie. However, the period-1 sophisticate also correctly predicts that knowing about period-3 self-control problems will induce him to do the report on the second Saturday. Hence, the period-1 sophisticate can safely procrastinate and see the mediocre movie, and indeed he then does the report on the second Saturday. Example 1 illustrates typical behavior for sophisticates when costs are salient. Although sophisticates have a tendency to procrastinate (they don't write the report right away, which their long-run selves prefer), perfect foresight can mitigate this problem because sophisticates will do it now when they (correctly) foresee costly procrastination in the future.

Example 1 illustrates a belief in the literature that sophistication is "good" because it helps overcome self-control problems. As in Akerlof's (1991) procrastination example, naifs repeatedly put off a task because they believe they will do it tomorrow. Akerlof intuits that sophistication could overcome this problem, and Example 1 demonstrates this intuition.

However, this intuition may not hold when rewards are salient. Consider a similar scenario: Suppose you have a coupon to see one movie over the next four Saturdays, and your allowance is such that you cannot afford to pay for a movie. The schedule at the local cinema is the same as for the above example ˘ a mediocre movie this week, a good movie next week, a great movie in two weeks, and (best of all) a Johnny Depp movie in three weeks. Which movie do you see?

Now, the activity you must do exactly once is going to a movie, and the reward, seeing the movie, is experienced immediately.[19] Using the same payoffs for seeing a movie as in Example 1, we have the following formalization.

Example 2 Suppose rewards are salient, $T = 4$, and $\beta = \frac{1}{2}$. Let $\mathbf{v} = (3, 5, 8, 13)$ and $\mathbf{c} = (0, 0, 0, 0)$.
$s^{tc} = (N, N, N, Y)$, so TCs see the movie in period $\tau_{tc} = 4$.
$s^{n} = (N, N, Y, Y)$, so naifs see the movie in period $\tau_{n} = 3$.
$s^{s} = (Y, Y, Y, Y)$, so sophisticates see the movie in period $\tau_{s} = 1$.

TCs wait and see the Depp movie since it yields the highest reward. Naifs see the great movie. On the first two Saturdays, naifs skip the mediocre and good movies incorrectly believing they will wait to see the Depp movie. However, on the third Saturday, they give in to self-control problems

[19] That seeing a movie is a "cost" in Example 1 and a "reward" in Example 2 reflects that the rewards and costs are defined with respect to the activity being done once.

and see the great movie. This example illustrates that when rewards are salient, naive perceptions can help naifs overcome their self-control problem because it makes them overoptimistic about future behavior. If naifs knew they would end up seeing the great movie and not the Depp movie, they would have given in to self-control problems and seen the good movie.

Sophisticates are not so fortunate. They see the mediocre movie because of an unwinding similar to that in the finitely-repeated prisoners' dilemma. The period-2 sophisticate would choose to see the good movie because he correctly predicts that he would give in to self-control problems on the third Saturday, and see merely the great movie rather than the Depp movie. The period-1 sophisticate correctly predicts this reasoning and behavior by his period-2 self. Hence, the period-1 sophisticate realizes that he will see merely the good movie if he waits, so he concludes he might as well see the mediocre movie now. This example demonstrates a typical problem for sophisticates when rewards are salient: Knowing about future self-control problems can lead you to give in to them today, because you realize you will give in to them tomorrow. The example also shows why sophisticates would like ways to "commit" the behavior of their future selves, as discussed by many researchers: The period-1 sophisticate would be better off if he could commit himself to seeing the Depp movie, or even the great movie.

We now present some propositions that characterize presently behavior more generally. We refer to the most basic intuition concerning how presently preferences affect behavior as the presently effect : When costs are salient people with presently preferences tend to procrastinate ˘ wait when they should do it ˘ while when rewards are salient they tend to preproperate ˘ do it when they should wait.[20] With salient costs, in periods where they should do it they do not because they want to avoid the salient cost. With salient rewards, in periods where they should not do it they do because they want the salient reward now. Proposition 1 captures that naifs are influenced solely by the presently effect.

Proposition 1 (1) If costs are salient, then $\tau_n \geq \tau_{tc}$.
(2) If rewards are salient, then $\tau_n \leq \tau_{tc}$.[21]

Proposition 1 is as simple as it seems ˘ naifs believe they will behave like TCs in the future but are more impatient now. Hence, comparisons of naive behavior and TC behavior are straightforward

[20] Throughout this paper, "procrastination" means that an agent chooses to wait when her long-run self (i.e., a TC) would choose to do it, and "preproperation" means that an agent chooses to do it when her long-run self would choose to wait.

[21] All propositions are stated with weak inequalities; but in each case, examples exist where the inequalities are strict.

and intuitive.

Comparisons of sophisticated behavior and TC behavior are more complicated because there is a second effect influencing sophisticated behavior. The sophistication effect captures that sophisticates are fully aware of any self-control problems they might have in the future, and this awareness can influence behavior now. In our one-task model, however, the sophistication effect is straightforward: Because sophisticates are (correctly) pessimistic that they will behave themselves in the future, they are more inclined than naifs to do it now, irrespective of whether it is costs, rewards, or both that are salient.[22]

Proposition 2 For all cases, $\tau_s \leq \tau_n$.

Even though the sophistication effect is behaviorally the same for both salient costs and salient rewards (i.e., $\tau_s \leq \tau_n$), it has two different interpretations. For salient costs, that sophisticates do it before naifs reflects that sophistication helps overcome the tendency to procrastinate, as discussed in Example 1. For salient rewards, that sophisticates do it before naifs reflects that sophistication can exacerbate the tendency to preproperate, as discussed in Example 2. These alternative interpretations will have important welfare implications, as we discuss in Section 6.

What can we say about sophisticates versus TCs? When rewards are salient, the sophistication effect says sophisticates do it before naifs, and the presently effect says naifs do it before TCs, so the clear implication is that sophisticates do it before TCs:

Corollary 1 When rewards are salient, $\tau_s \leq \tau_{tc}$.

When costs are salient, however, we cannot make a general statement because the sophistication effect can sometimes outweigh the presently effect. Suppose you must write a paper this weekend, on Friday night, Saturday, or Sunday. You know the paper will be better if written on either Saturday or Sunday (when you have an entire day). However, it is a mid-November weekend with plenty of sports on TV ˘ pro basketball on Friday night, college football on Saturday, and pro football on Sunday. You prefer watching pro football to college football, and prefer college football to pro basketball. Which sports event do you miss to write the paper? We can represent this scenario with the following example, where the activity to be done once is writing the paper and the costs correspond to the attractiveness of the sports event missed.

[22] This strong result depends very much on the one-task model as opposed to having to do the activity more than once. Sophisticates are correctly pessimistic about future total utility. In general, the relevant comparison is the utility from doing it versus the marginal future utility from delaying one task. In the one-task model, future total utility is identical to the marginal future utility from delaying one task.

Example 3 Suppose costs are salient, $T = 3$, and $\beta = \frac{1}{2}$. Let $\mathbf{v} = \{12, 18, 18\}$ and $\mathbf{c} = \{3, 8, 13\}$. Then $\tau_s = 1$ and $\tau_{tc} = 2$ (and $\tau_n = 3$).

TCs write the paper on Saturday because the marginal benefit of a better paper outweighs the marginal cost of giving up college football for pro basketball. Since the example involves salient costs, the presently effect suggests that sophisticates should procrastinate. However, the sophistication effect leads sophisticates to write the paper on Friday night, before TCs. On Friday, sophisticates correctly predict that they will end up writing the paper on Sunday if they do not do it now. Hence, although sophisticates would prefer to write the paper on Saturday, they do it on Friday to prevent themselves from waiting until Sunday.

Example 3 demonstrates how sophisticates might preproperate even when costs are salient as a means to preempt costly procrastination. Nonetheless, we can find sufficient conditions where such "preemptive over-control" does not occur. When costs are salient, sophisticates perform the activity before the optimal date (i.e., optimal from a long-run perspective) only if doing so prevents procrastination beyond the optimal date. In other words, the period-τ_s sophisticate does it because the period-τ_{tc} sophisticate would wait. Given τ_{tc} yields higher utility than τ_s from a long-run perspective, the period-τ_s sophisticate can be more willing to do it than the period-τ_{tc} sophisticate only if he faces a smaller salient cost, or $c_{\tau_s} < c_{\tau_{tc}}$. But $c_{\tau_s} < c_{\tau_{tc}}$ implies TCs will wait until period τ_{tc} only if $v_{\tau_s} < v_{\tau_{tc}}$. Hence, sophisticates preproperate when costs are salient only if $c_{\tau_s} < c_{\tau_{tc}}$ and $v_{\tau_s} < v_{\tau_{tc}}$, which cannot hold if either the reward schedule or the cost schedule is weakly decreasing.

Proposition 3 If costs are salient, and either the reward schedule or the cost schedule is weakly decreasing, then $\tau_{tc} \leq \tau_s$.

Although Proposition 3 specifies sufficient conditions for sophisticates not to preproperate when costs are salient, preproperation does not require particularly odd reward or cost schedules. In addition to simple examples like Example 3, one class of salient-cost scenarios where we might expect to see sophisticates doing it before TCs is with an increasing concave reward schedule and an increasing convex cost schedule. Suppose, for example, you must decide when to start a project that has a fixed deadline. The project cost is increasing and convex in the start date because you have fewer days to put in a fixed number of hours. However, information that will increase the value of the project becomes available over time. If you are a sophisticate, you might start the project before the optimal information/cost trade-off is reached, because you fear procrastination beyond that time.

While the above propositions represent our main behavioral results, we briefly examine a surprising result for sophisticates: The sophistication effect can lead sophisticates to have a "cyclical" strategy in a non-cyclical environment. Consider the following example.

Example 4 Suppose costs are salient. Let the reward schedule $\mathbf{v} \in \mathbb{R}^T$ satisfy $v_t = v_0 - t\,\gamma$ for each t, where $\gamma > 0$. Let the cost schedule $\mathbf{c} \in \mathbb{R}^T$ satisfy $c_t = \bar{c}$ for each t. Then:

1) $s^{tc} = (Y, Y, ..., Y)$.

2) If $\beta \geq \frac{\bar{c}}{\gamma+\bar{c}}$ then $s^n = (Y, Y, ..., Y)$, and if $\beta < \frac{\bar{c}}{\gamma+\bar{c}}$ then $s^n = (N, N, ..., Y)$.

3) If $\beta \in \left[\frac{\bar{c}}{w\gamma+\bar{c}}, \frac{\bar{c}}{(w-1)\gamma+\bar{c}}\right)$ where $w \in \{1, 2, ..., T\}$, then s^s satisfies $s^s_t = Y$ if and only if $t \in \{T - w, T - 2w, ...\}$. If $\beta < \frac{\bar{c}}{T\gamma+\bar{c}}$ then $s^s = (N, N, ..., Y)$.

In the example, TCs have strategy $(Y, Y, ..., Y)$, and naifs have either strategy $(Y, Y, ..., Y)$ or $(N, N, ..., Y)$. However, when naifs have strategy $(N, N, ..., Y)$, sophisticates can have a cyclical strategy (e.g., $(...N, N, Y, N, N, Y)$). That is, they plan to do it on the last day, the fourth-to-last day, the seventh-to-last day, and so on. In this "stationary" environment, TCs and naifs have a stationary perception of when they will do it if they wait, and hence perceive the same problem every period. Sophisticates, on the other hand, can have cyclical perceptions of when they will do it if they wait. For example, suppose a sophisticate will do it in period $t+1$. In period t, he perceives that he will do it next period if he waits. If this leads him not to do it in period t, then in period $t-1$ he perceives he will do it in two periods if he waits. Hence, the period-t and period-$(t-1)$ selves can make different choices. Furthermore, if the period-$(t-1)$ self will do it, the period-$(t-2)$ and period-$(t-3)$ selves will be in identical positions to the period-t and period-$(t-1)$ selves respectively, so the period-$(t-2)$ self will wait and the period-$(t-3)$ self will do it; hence the cyclicality.[23]

The point of this example is that sophistication can lead to strange and complicated behavior. We have already seen that sophistication can lead to preproperation when presentliness suggests people should procrastinate. Here, we see that sophistication can lead to cyclical strategies in a stationary environment. We will return to this theme in Section 7 where we show that doing the activity more than once can drastically alter sophisticated behavior.

[23] The cyclical strategy of sophisticates is, of course, not going to be observed in behavior since all we would observe is the first period in which they do it.

5. Is It Just Impatience?

In our model, sophisticates and naifs are both impatient and time-inconsistent, while TCs are neither impatient nor time-inconsistent. A natural question to ask is whether differences between presently behavior and TC behavior are driven entirely by impatience, or is time inconsistency playing an important role?

It should be clear that time inconsistency is playing a crucial role. Example 3, for instance, shows that sophisticates can preproperate when impatience suggests they should procrastinate. In this section, we consider this theme more generally by presenting some specific examples where people with presently preferences behave in ways that time-consistent agents would never behave.

Many researchers studying time-inconsistent preferences emphasize the use of "commitment" devices that limit future choice sets, in part because such devices provide smoking guns that prove time consistency wrong. However, most have emphasized the use of external commitment devices (e.g., Christmas clubs, fat farms) that involve explicit attempts by a person to restrict her future choice sets. A second motivation for this section is to demonstrate the existence of smoking guns in the simple one-time-activity model, where no such commitment devices are available.[24]

The following example demonstrates that sophisticates can engage in behavior that time-consistent agents would never exhibit, even time-consistent agents that discount differently from period to period.

Example 5 Suppose rewards are salient, $T = 3$, $\mathbf{v} = (0, 5, 1)$, and $\mathbf{c} = (1, 8, 0)$.
(i) Sophisticates with $\delta = 1$ and $\beta = \frac{1}{2}$ will do it in period 1.
(ii) A time-consistent agent will never do it in period 1. For any discounting, the utility from doing it in period 1 is negative, and the utility from doing it in period 3 is positive.

In Example 5, time-consistent agents will never do it in period 1 when they have the option of period 3; period 1 has a cost and no reward, whereas period 3 has a reward and no cost. Sophisticates do it in period 1 not because it is their most preferred time to do it, but instead because they want to avoid doing it in period 2. Like time-consistent agents, the period-1 sophisticate would prefer doing it

[24] Despite our emphasis in this section, we feel that the search for "smoking guns" in the form of behavior that is incompatible with any time-consistent discounting is too narrow an exercise. We should also consider "calibrated smoking guns" ˘ behavior that could be compatible with time-consistent discounting, but only for implausible discount rates. For instance, our example of somebody avoiding 7 hours of work today at the expense of having to do 8 hours two weeks from now could be explained by time-consistent discounting. But if we merely assumed that people do not have decreasing cost of work, this would imply an annual discount factor of $\frac{31}{1000}$ ˘ that a person cares 32 times more about her well-being 2 years from now than 3 years from now. A "discount factor" of $\frac{7}{8}$ per every two weeks is simply not plausible.

in period 3 to period 1. Unfortunately, the period-2 sophisticate gets to choose between doing it in period 2 versus period 3, and he will choose period 2. Because of his time inconsistency, period 3 is not an option for the period-1 sophisticate.

That smoking guns exist for sophisticates in our setting extends a theme from previous literature on commitment devices. Perhaps more interesting is that smoking guns can exist for naifs.

Example 6 Suppose costs are salient, $T = 3$, $\mathbf{v} = (1, 6, 0)$, and $\mathbf{c} = (0, 4, 1)$.
(i) Naifs with $\delta = 1$ and $\beta = \frac{1}{2}$ will do it in period 3.
(ii) A time-consistent agent will never do it in period 3. For any discounting, the utility from doing it in period 3 is negative, and the utility from doing it in period 1 is positive.

In Example 6, time-consistent agents will never do it in period 3 when they have the option of period 1. Naifs end up doing it in period 3, even though the period-1 naif prefers doing it in period 1 to period 3. The period-1 naif waits incorrectly believing he will do it in period 2, but the period-2 naif in fact prefers doing it in period 3. Because of his time inconsistency, waiting in period 1 leads naifs to do it in period 3, but because of his naivete, the period-1 naif doesn't realize this.

Another type of smoking gun arises in cross-model comparisons. Both sophisticates and naifs can have preference reversals: Eliminating an option that was not chosen can alter behavior. In other words, eliminating an "irrelevant alternative" can reverse a person's choice over existing possibilities. Under time-consistent preferences, eliminating an opportunity that is not chosen will not affect behavior.

We can observe a preference reversal for sophisticates in Example 1, where you must decide when to skip a movie and do a report for work. The following example shows the implications of eliminating the option of doing the report on the fourth Saturday.

Example 7 Suppose costs are salient, and consider a sophisticate with $\beta = \frac{1}{2}$.
(i) If $T = 4$, $\mathbf{v} = (\bar{v}, \bar{v}, \bar{v}, \bar{v})$, and $\mathbf{c} = (3, 5, 8, 13)$, then he does it in period 2.
(ii) If $T = 3$, $\mathbf{v} = (\bar{v}, \bar{v}, \bar{v})$, and $\mathbf{c} = (3, 5, 8)$, then he does it in period 1.

The preference reversal for sophisticates is driven by the sophistication effect. Part (i) is identical to Example 1. As discussed there, the period-2 sophisticate will do the report because he knows he would procrastinate in period 3 and be forced to miss the Depp movie in period 4. Since the period-2 sophisticate will not procrastinate, the period-1 sophisticate can safely procrastinate. In part (ii), however, missing the Depp movie is no longer a possibility since the report must be completed before then. Now, the period-2 sophisticate can safely procrastinate knowing he will be forced to do the report in period 3. Given the period-2 sophisticate would now procrastinate, the period-1 sophisticate prefers doing the report as opposed to procrastinating until period 3. In both (i) and (ii), sophisticates do the report before the Depp movie; but whether sophisticates are allowed to do the report that Saturday changes behavior.

We can observe a preference reversal for naifs in Example 2, where you must decide when to see a movie with your coupon. The following example shows the implications for naifs of whether the coupon is valid for the Depp movie.

Example 8 Suppose rewards are salient, and consider a naif with $\beta = \frac{1}{2}$.
(i) If $T = 4$, $\mathbf{v} = (3, 5, 8, 13)$, and $\mathbf{c} = (0, 0, 0, 0)$, then he does it in period 3.
(ii) If $T = 3$, $\mathbf{v} = (3, 5, 8)$, and $\mathbf{c} = (0, 0, 0)$, then he does it in period 2.

The preference reversal for naifs is driven by optimism about future behavior. Part (i) is identical to Example 2. As discussed there, naifs wait in periods 1 and 2 incorrectly believing they will see the Depp movie. In period 3, they end up giving in to self-control problems and seeing the great movie. In part (ii), the coupon is not valid for the Depp movie. Now, the best movie naifs can expect to see is the great movie, and as a result they give in to self-control problems earlier than in (i). The Depp movie is not chosen in (i), and yet eliminating that option changes naive behavior.

6. Welfare

In this section, we examine the welfare implications of presently preferences, particularly the implications of sophistication versus naivete. We begin with some qualitative welfare comparisons of sophisticates to naifs. Then we ask when moderate self-control problems can cause severe welfare losses.

Welfare comparisons for people with time-inconsistent preferences are in principle problematic; the very premise of the model is that a person's preferences at different times disagree, so that a change in behavior may make some selves better off while making other selves worse off. The consumption/savings literature addresses this issue by introducing a Pareto-efficiency criterion (e.g., Goldman (1979,1980), Laibson (1994)), asking when all period selves weakly prefer one strategy to another. If a strategy is Pareto-superior to another, then it is clearly better. However, we feel a Pareto-efficiency criterion is too strong (i.e., there can be situations where two strategies cannot be Pareto-ranked but one seems clearly better than the other). Since presently preferences are often meant to capture self-control problems, we feel the natural perspective in most situations is the "long-run perspective" ˘ what would you wish now about your profile of future behavior. (See Schelling (1984) for an thoughtful discussion of some of these issues.) To formalize the long-run perspective, we suppose there is a (fictitious) period 0 where the person has no decision to make and weights all future periods equally. We can then denote a person's long-run utility from doing it in period τ by $U^0(\tau) \equiv v_\tau - c_\tau$.[25]

In Section 4, we implied that sophistication is good when costs are salient because it mitigates the tendency to procrastinate. Indeed, in Example 1 $U^0(\tau_s) > U^0(\tau_n)$. Proposition 4 establishes that this intuition generalizes: If costs are salient, sophisticates always do at least as well as naifs.

Proposition 4 If costs are salient, then $U^0(\tau_s) \geq U^0(\tau_n)$.

With salient costs, sophisticates and naifs both have a tendency to procrastinate. But sophisticates know when procrastination will be extremely costly, and can preempt costly procrastination by doing the activity now. Hence, with salient costs the conventional wisdom that sophistication helps you is correct.[26]

When rewards are salient, however, sophistication can hurt you. We saw in Section 4 how sophistication exacerbates the tendency to preproperate, and in Example 2 we have $U^0(\tau_s) < U^0(\tau_n)$.[27]

Although sophistication exacerbates the tendency to preproperate, doing so does not necessarily imply lower long-run utility (so there is no analogue to Proposition 4). Consider the following scenario. In May, you receive a special vacation offer: On June 1 you can claim an immediate mediocre vacation at a fairly low price; on July 1 you can claim an immediate spectacular vacation but at an extremely high price; or on August 1 you can claim an immediate good vacation for free. Which vacation do you claim? We can represent this scenario with the following example.

[25] In fact, most of our welfare comparisons can be roughly conceived of as "Pareto comparisons," and we shall note Pareto-efficiency "analogues" as we go. To formalize a Pareto-efficiency criterion in our model, however, we must define period-t utility from doing the activity prior to period t. We suppose preferences over past events are symmetric to preferences over future events, so

$$U^t(\tau) = \begin{cases} \beta(v_t - c_t) & \text{if } \tau \neq t \\ \beta v_\tau - c_\tau & \text{if } \tau = t \text{ and costs are salient} \\ v_\tau - \beta c_\tau & \text{if } \tau = t \text{ and rewards are salient} \end{cases}$$

[26] There is an identical result for Pareto-comparisons: If costs are salient, the sophisticated strategy s^s is (weakly) Pareto-superior to the naive strategy s^n.

[27] The comparison of sophisticates to naifs or TCs when rewards are salient is the main case in our model where there are substantive differences between long-run-utility comparisons and Pareto comparisons. In particular, $U^0(\tau_n) > U^0(\tau_s)$ does not imply that s^n is Pareto-superior to s^s. Nonetheless, it will never be the case that s^s is Pareto-superior to s^n, and there exist examples where s^n is Pareto-superior to s^s (e.g., Example 2). It is also true that $U^0(\tau_{tc}) > U^0(\tau_n)$ does not imply s^{tc} is Pareto-superior to s^n; however, at most one perspective of the naif prefers s^n to s^{tc}, while many other perspectives may be made worse off.

Example 9   Suppose rewards are salient, $T = 3$, and $\beta = \frac{1}{2}$. Let $\mathbf{v} = \{1, 6, 2\}$ and $\mathbf{c} = \{1, 8, 0\}$.
Then $\tau_s = 1$, $\tau_n = 2$, and $\tau_{tc} = 3$, so $U^0(\tau_s) \;=\; 0 \;>\; -2 \;=\; U^0(\tau_n)$.

Example 9 contains a temptation trap: On July 1 people with presently preferences cannot resist the spectacular vacation even though it is too expensive from a long-run perspective. (In fact, the spectacular vacation is the worst choice from a long-run perspective.) Sophisticates avoid this temptation trap by taking the mediocre vacation in June. In contrast, naifs fall into the temptation trap: Naifs believe in June that they will wait until August and take the good vacation for free, and as a result end up taking the spectacular but too costly vacation in July.

With salient rewards, sophisticates can do better than naifs only if naifs fall into a temptation trap that sophisticates avoid. However, since a temptation trap requires increasing the reward and the cost, this cannot occur if either rewards or costs are weakly decreasing:

Proposition 5   If rewards are salient, and either the reward schedule or the cost schedule is weakly decreasing, then $U^0(\tau_n) \geq U^0(\tau_s)$.

Perhaps more interesting than simple comparisons of sophisticates to naifs is the question of when moderate self-control problems can cause severe welfare losses. Since sophisticates, naifs, and TCs have identical long-run utility, we can measure the welfare loss from self-control problems by the deviation from TC long-run utility (i.e., $U^0(\tau_{tc}) - U^0(\tau_s)$ and $U^0(\tau_{tc}) - U^0(\tau_n)$). We first point out that if rewards and costs can be arbitrarily large, then both naifs and sophisticates can suffer arbitrarily severe welfare losses. Consider the following two-period example (so sophistication is irrelevant), where we use $\beta = \frac{1}{2}$ but could construct analogous examples for any $\beta \in (0, 1)$. By letting $X \to \infty$, both naifs and sophisticates can suffer arbitrarily severe welfare losses.

Example 10   Let $T = 2$ and $\beta = \frac{1}{2}$.
(i) Suppose costs are salient. Let $\mathbf{v} = (X + \epsilon, 0)$ and $\mathbf{c} = (X + \epsilon, X)$ for some $X, \epsilon > 0$. Then:
$\tau_{tc} = 1$ and $\tau_s = \tau_n = 2$ so $U^0(\tau_{tc}) - U^0(\tau_s) \;=\; U^0(\tau_{tc}) - U^0(\tau_n) \;=\; X$.
(ii) Suppose rewards are salient. Let $\mathbf{v} = (X + \epsilon, 0)$ and $\mathbf{c} = (2X + \epsilon, 0)$ for some $X, \epsilon > 0$. Then:
$\tau_{tc} = 2$ and $\tau_s = \tau_n = 1$ so $U^0(\tau_{tc}) - U^0(\tau_s) \;=\; U^0(\tau_{tc}) - U^0(\tau_n) \;=\; X$.

The welfare losses in Example 10 are the result of a single decision to give in to self-control: In period 1, both sophisticates and naifs are pleased with their decision to wait in part (i) and to do it in part (ii). For a variety of reasons, we are cautious of the importance of severe welfare losses arising solely from one-time procrastination or preproperation. It relies heavily on the particular formulation of presentliness, and the extreme rewards and costs. Example 10 stretches the credibility of the model by supposing that a person will accept unboundedly severe welfare losses to satisfy a single salient desire.

Focusing on severe welfare losses driven by a single bad decision hides a more interesting result: Even if the welfare loss from any individual decision is small, severe welfare losses can still arise when self-control problems are compounded. As a framework to demonstrate this result, we suppose there is some upper bound on rewards and costs.[28] To formalize this, let $\bar{v}$ and $\bar{c}$ be positive numbers. Let $P(\bar{v},\bar{c}) \equiv \{\ (\mathbf{v},\mathbf{c}) \mid \text{for all } t\ v_t \le \bar{v} \text{ and } c_t \le \bar{c}\ \}$ be the set of all reward/cost schedule combinations satisfying the upper bounds $\bar{v}$ and $\bar{c}$ on rewards and costs. Note that the set $P(\bar{v},\bar{c})$ does not put any restrictions on the number of periods T. Let $M^s(\beta;\bar{v},\bar{c})$ and $M^n(\beta;\bar{v},\bar{c})$ be the upper bounds on how severe the welfare losses from self-control problems can be given β and the bounds $\bar{v}$ and $\bar{c}$:

$$M^s(\beta;\bar{v},\bar{c}) \ \equiv\ sup_{(\mathbf{v},\mathbf{c})\in P(\bar{v},\bar{c})} \left[\, U^0(\tau_{tc}) \ - \ U^0(\tau_s)\, \right]$$
$$M^n(\beta;\bar{v},\bar{c}) \ \equiv\ sup_{(\mathbf{v},\mathbf{c})\in P(\bar{v},\bar{c})} \left[\, U^0(\tau_{tc}) \ - \ U^0(\tau_n)\, \right]$$

For any t, $-\bar{c} \le U^0(t) \le \bar{v}$, so for all β, $M^s(\beta;\bar{v},\bar{c}) \in [\,0\ ,\ \bar{v}+\bar{c}\,]$ and $M^n(\beta;\bar{v},\bar{c}) \in [\,0\ ,\ \bar{v}+\bar{c}\,]$.

The following proposition establishes that with salient costs, however small self-control problems may be, naifs can still suffer severe welfare losses, while for sophisticates welfare losses go to zero as self-control problems go away.

Proposition 6 Suppose costs are salient:
(i) $lim_{\beta\to 1}\,[M^s(\beta;\bar{v},\bar{c})] \ =\ 0$
(ii) For any $\beta < 1$, $M^n(\beta;\bar{v},\bar{c}) \ =\ \bar{v}+\bar{c}$.

When costs are salient, presentliness leads you to procrastinate. Since sophisticates know exactly when they will do it if they wait, delaying from period τ_{tc} to period τ_s is a single decision to procrastinate for sophisticates. Hence, for sophisticates small self-control problems cannot cause severe welfare losses. Naifs, on the other hand, can compound self-control problems by making repeated decisions to procrastinate, each time believing they will do it next period. With each decision to procrastinate, they incur a small welfare loss, but the total welfare loss is the sum of these increments. No matter how small the individual welfare losses, the sum can be large if naifs procrastinate enough times. Proposition 6 supports an intuition that appears in the literature (e.g.,

[28] With such bounds in a two-period model, welfare losses from self-control problems go to zero as self-control problems go away (i.e., $[U^0(\tau_{tc}) - U^0(\tau_s)] \to 0$ and $[U^0(\tau_{tc}) - U^0(\tau_n)] \to 0$ as $\beta \to 1$). Similarly, bounds on rewards and costs when $T > 2$ will limit the welfare loss from any single decision to procrastinate or preproperate.

Strotz (1955) and Akerlof (1991)) that naivete about future preferences can lead you to experience very bad outcomes, and sophistication can save you.

But in the salient-reward case, sophistication can be a problem. Although sophisticates can do better than naifs, naifs are generally better off than sophisticates since sophistication exacerbates the tendency to preproperate. The following proposition supports this intuition: When rewards are salient, naif welfare losses go to zero as self-control problems go away, while sophisticates can suffer severe welfare losses even with small self-control problems.

Proposition 7 Suppose rewards are salient:
(i) $lim_{\beta \to 1} [M^n(\beta; \bar{v}, \bar{c})] = 0$
(ii) For any $\beta < 1$, $M^s(\beta; \bar{v}, \bar{c}) = \bar{v} + \bar{c}$.

When rewards are salient, presentliness leads you to preproperate. Naifs always believe that if they wait they will do it when TCs do it, so doing it in period τ_n as opposed to waiting until period τ_{tc} is a single decision to preproperate for naifs. Hence, for naifs small self-control problems cannot cause severe welfare losses. But sophisticates can compound self-control problems because of an unwinding: In the end, sophisticates will preproperate; because they realize this, near the end they will preproperate; realizing this they preproperate a little sooner, etc. For each step of this unwinding, the welfare loss may be small, but the total welfare loss is the sum of multiple steps. As with naifs and salient costs, no matter how small the individual welfare losses, the sum can be large if the unraveling occurs over enough periods.[29] [30]

[29] Propositions 6 and 7 formalize when small self-control problems can be very costly from a long-run perspective. A less strong formalization can be made using Pareto comparisons: If costs are salient, self-control problems can be costly only if you are naive in the sense that s^s is Pareto-optimal while s^n may not be; and if rewards are salient, self-control problems can be costly only if you are sophisticated in the sense that s^n is Pareto-optimal while s^s may not be.

[30] We feel that our limit results qualitatively capture very real differences in when moderately impatient sophisticates and naifs can suffer severe welfare losses, but there are reasons to be cautious in interpreting them too literally. For instance, since "unwinding" drives severe preproperation for sophisticates, it seems natural to ask whether a small amount of uncertainty could reverse this tendency, much as Kreps et al (1982) showed that a small amount of uncertainty can lead to extensive cooperation in the finitely repeated prisoners' dilemma. We suspect that there is something to this story, but the analogy is problematic on two fronts. First, although players may cooperate for most of a very long horizon, there is still a long duration at the end of the repeated prisoners' dilemma where players are unlikely to cooperate. Such an "endgame" could still create significant welfare losses. Second, in the Kreps et al result a player's current behavior will signal something about her future behavior to other players. Since each "player" in our game plays only once, the comparable signal is that a person in period t infers something about the propensity of her period-$(t+1)$ self to wait from the fact that her period-$(t-1)$ self waited, which requires that the period-t self doesn't know β. While we believe that such self-inference and self-signaling go on, there are many issues to be worked out to understand the strategic logic and psychological reality of such phenomena.

7. Multi-Tasking

In this section, we begin to explore how our results might carry over to more general settings. We present a simple extension of our model where the activity must be performed more than once. The basic structure of the model is exactly as in Section 3, but now the person must do the activity exactly $M \geq 1$ times, and she can do it at most once in any given period. For each period t in which she does it, she receives reward v_t and incurs cost c_t, and these can be experienced immediately or with some delay. Using the interpretations of salient costs and salient rewards from Section 3, preferences take the following form:

1) Salient Costs: If the set Θ is the M periods in which a person does it, then instantaneous utilities are

$$u_t = -c_t \quad \text{for all } t \in \Theta \qquad u_{T+1} = \sum_{t \in \Theta} v_t \qquad u_t = 0 \quad \text{for all } t \notin \Theta \cup \{T+1\}$$

Then intertemporal utility is

$$U^t(\tau) \equiv \begin{cases} -(1-\beta)c_t + \beta\left(\sum_{\tau\in\Theta} v_\tau - \sum_{\tau\in\Theta} c_\tau\right) & \text{if } t \in \Theta \\ \beta\left(\sum_{\tau\in\Theta} v_\tau - \sum_{\tau\in\Theta} c_\tau\right) & \text{if } t \notin \Theta \end{cases}$$

2) Salient Rewards: If the set Θ is the M periods in which a person does it, then instantaneous utilities are

$$u_t = v_t \quad \text{for all } t \in \Theta \qquad u_{T+1} = -\sum_{t \in \Theta} c_t \qquad u_t = 0 \quad \text{for all } t \notin \Theta \cup \{T+1\}$$

Then intertemporal utility is

$$U^t(\tau) \equiv \begin{cases} (1-\beta)v_t + \beta\left(\sum_{\tau\in\Theta} v_\tau - \sum_{\tau\in\Theta} c_\tau\right) & \text{if } t \in \Theta \\ \beta\left(\sum_{\tau\in\Theta} v_\tau - \sum_{\tau\in\Theta} c_\tau\right) & \text{if } t \notin \Theta \end{cases}$$

Given these preferences, we can define perception-perfect strategies analogously to Definition 4. We omit the formal definitions here. To discuss behavior, we define some notation. Let $\tau_a^M(i)$ be such that if an agent of type $a \in \{tc, s, n\}$ must do the activity exactly M times, then in period

A comparable worry about our extreme results for naifs is that they will eventually learn that they have a tendency to procrastinate. Again, we think there is something to this intuition, but we suspect the issue is complicated. The issue of self-inference again arises. Further, people seem to have a powerful ability not to apply general lessons they understand well to specific situations. For instance, we are all familiar with the sensation of being simultaneously aware that we tend to be over-optimistic in completing projects, but still being over-optimistic regarding our current project. (See Kahneman and Lovallo (1993) for evidence on related issues.)

$\tau_a^M(i)$ she will do it for the i^{th} time (according to her perception-perfect strategy). Let $\tau_a^M \equiv \{\ \tau_a^M(1)\ ,\ \tau_a^M(2)\ ,\ ...\ ,\ \tau_a^M(M)\ \}$.

We first consider TCs and naifs. Proposition 8 addresses how behavior depends on M: If TCs or naifs must do the activity an extra time, they do it in all periods they used to do it, and some additional period.

Proposition 8 For all cases, given a reward schedule $\mathbf{v}$ and a cost schedule $\mathbf{c}$, for each $M \in \{1, 2, ..., T-1\}$:
$\tau_{tc}^M \subset \tau_{tc}^{M+1}$ and $\tau_n^M \subset \tau_n^{M+1}$.

If they have k activities to do in future periods, both TCs (correctly) and naifs (incorrectly) perceive that they will choose the k best remaining periods. As a result, the marginal value of having an additional activity in the future is decreasing in k, so the perceived payoff from waiting is decreasing in k. Consequently, if in period t TCs or naifs have more tasks remaining to be done, then they are more likely to perform the activity in period t. Hence, for TCs and naifs, changing M essentially does not alter behavior.

Moreover, the following proposition states that the relationship between TC behavior and naive behavior for $M > 1$ is essentially identical to that for $M = 1$.

Proposition 9 (1) If costs are salient, then for all $i \in \{1, 2, ..., M\}$, $\tau_n^M(i) \geq \tau_{tc}^M(i)$.
(2) If rewards are salient, then for all $i \in \{1, 2, ..., M\}$, $\tau_n^M(i) \leq \tau_{tc}^M(i)$.

Proposition 9 says that naive behavior in the multi-activity model is exactly analogous to naive behavior in the one-time activity model. If costs are salient, naifs procrastinate: They are always behind TCs in terms of tasks completed so far. If rewards are salient, naifs preproperate: They are always ahead of TCs in terms of tasks completed so far. Hence, the presently effect extends directly to the multi-activity setting; and again naifs exhibit the pure effects of presently preferences.

Propositions 8 and 9 show that even if a person must do an activity more than once, TCs and naifs are "normal" and intuitive. Sophisticates, on the other hand, are "weird." For example, changes in M can dramatically alter behavior. Recall that the sophistication effect reflects how awareness of future self-control problems will affect behavior now. In fact, there are two embodiments of the sophistication effect. First, you are correctly pessimistic about future behavior. Second, you look for "commitment devices" that limit future choice sets. In the multi-activity case, these motives can lead to dramatic behavioral changes for sophisticates. Example 11 illustrates the role of pessimism.

Example 11 Suppose costs are salient, and $\beta = \frac{1}{2}$. Let $\mathbf{v} = (0, 0, 0)$ and $\mathbf{c} = (2, 3, 5)$.
If $M = 1$, then $\tau_s = 1$, $\tau_n = 3$, and $\tau_{tc} = 1$.
If $M = 2$, then $\tau_s^2 = \{2, 3\}$, $\tau_n^2 = \{1, 3\}$, and $\tau_{tc}^2 = \{1, 2\}$.

In Example 11, two activities leads sophisticates to procrastinate, while they do not procrastinate with one activity. With salient costs, sophisticates have a tendency to procrastinate, but they do not procrastinate if doing so is costly. In Example 11, when $M = 1$ the period-1 sophisticate knows that procrastination implies delaying until period 3, so it is worthwhile to incur the immediate cost. When $M = 2$, however, the relevant comparison for the period-1 sophisticate is the marginal cost of delaying one task. Since the period-1 sophisticate realizes he will do one task in period 3 no matter what, period-1 procrastination merely implies delaying the second task until period 2, so it is not worth incurring the immediate cost. Sophisticates' correct pessimism leads to procrastination when $N = 2$ because sophisticates realize the marginal cost of delaying one task is not too large.

In the one-activity model, the only form of "commitment" was to do the activity now. Example 12 illustrates the way delay can be used as a commitment device in the multiple-activity model.

Example 12 Suppose rewards are salient, $T = 3$,and $\beta = \frac{1}{2}$. Let $\mathbf{v} = (6, 11, 21)$ and $\mathbf{c} = (0, 0, 0)$.
If $M = 1$, then $\tau_s = 1$, $\tau_n = 2$, and $\tau_{tc} = 3$.
If $M = 2$, then $\tau_s^2 = \{2, 3\}$, $\tau_n^2\{1, 2\}$, and $\tau_{tc}^2 = \{2, 3\}$.

In Example 12, one activity leads sophisticates to preproperate, while they do not preproperate with two activities. With salient rewards, sophisticates have a tendency to preproperate, but having multiple tasks provides a commitment device for sophisticates to overcome this tendency. If there is only one activity, there is no way to commit future selves not to preproperate. If there is a second activity, however, a commitment device becomes available: Waiting now prevents you from doing the activity for the second time tomorrow; you can only do it for the first time tomorrow. Thus, forgoing the reward today makes you delay at least until period 3. In Example 12, when $M = 1$ the period-1 sophisticate does the activity because he (correctly) predicts that he will just do it in period 2 if he waits. When $M = 2$ the period-1 sophisticate knows he will do the second activity in period 2 if he does the first now, but he can force himself to do it in periods 2 and 3 if he waits now.

In the one-activity model, we are able to make a strong statement comparing naifs to sophisticates: The sophistication effect always leads sophisticates to do it before naifs. Examples 11 and 12 show that this result does not hold if you must do the activity more than once. For salient costs, we showed in Example 11 how (correct) pessimism can make sophisticates procrastinate with multiple tasks because sophisticates realize procrastination may not be too costly. The period-1 sophisticate correctly predicts that he will do one task in period 3 no matter what, so period-1 procrastination merely implies delaying the first task until period 2. In contrast, the period-1 naif incorrectly believes he will do one task in period 2 no matter what, so he views period-1 procrastination as delaying the second task until period 3, which is more costly. For salient rewards, we showed in Example 12 how multiple tasks provide a commitment device to help sophisticates overcome the tendency to preproperate, and therefore do things after naifs.

Furthermore, when rewards are salient and $M > 1$, the commitment component of the sophistication effect can lead to surprising sophisticated behavior analogous to that in Example 3. In Example 3, sophisticates do it before TCs even though costs are salient. In the following example, sophisticates do things after TCs even though rewards are salient:

Example 13 Suppose rewards are salient, and $\beta = \frac{1}{2}$. Let $\mathbf{v} = (12, 6, 11, 21)$ and $\mathbf{c} = (0, 0, 0, 0)$.
If $M = 2$, then $\tau_{tc}^2 = \{1, 4\}$, $\tau_n^2 = \{1, 3\}$, and $\tau_s^2 = \{3, 4\}$.
If $M = 1$, then $\tau_{tc} = 4$, $\tau_n = 1$, and $\tau_s = 1$.

In Example 13, the situation beginning in period 2 is identical to Example 12, and the intuition for why sophisticates do it later than TCs is related to the intuition of Example 12. The period-1 sophisticate knows that if he has one activity left in period 2, he will do it in period 2, while if he has two activities left in period 2, he can commit himself to waiting until periods 3 and 4. Hence, even though the period-1 sophisticate's most preferred periods for doing it are periods 1 and 4, he realizes he won't do it in period 4 if he does it in period 1. The choice for the period-1 sophisticate is between doing it in periods 1 and 2 versus doing it in periods 3 and 4.

The upshot of these examples is that sophistication can lead to strange behavior, a theme we mentioned at the end of Section 4. Sophistication makes behavior very sensitive to how many times you must perform the activity. Furthermore, even if you have presently preferences, sophistication can lead you to behave exactly opposite of what presentliness would suggest.[31]

[31] Given the unpredictableness of sophisticated behavior, welfare comparisons (in terms of long-run utility) become more complicated. The discussion of the sophistication effect above suggests that when rewards are salient sophisticates are better off with more activities (as in Example 12), and when costs are salient sophisticates are better off with fewer activities (as in Example 11). Nonetheless, there are results analogous to those in Proposition 6 and Proposition 7. However, we must be careful: Results analogous to Proposition 6 and Proposition 7 hold if we fix an absolute M and let T become large; but if instead we fix M as a proportion of T (i.e., $M = \alpha T$) and let T become large, the results do not hold.

8. Discussion and Conclusion

Many realms where presently preferences are clearly important, such as savings and addiction, cannot generally be put into the framework of this paper. We examine only the restrictive setting where "doing an activity" is a discrete event, whereas in realms such as savings, marginal changes could be quite important. Furthermore, we assume that the reward and cost of doing the activity in any given period are unaffected by whether and when the agent previously did the activity. Particularly for addiction, but also for savings, past behavior may influence payoffs from current behavior.

Nonetheless, we feel that the lessons from the previous sections provide some insight into such realms. Presently preferences are relevant for consumption/savings decisions because consuming now yields salient payoffs, whereas the increased future consumption that saving allows is not salient. Does the analysis of our paper help shed any additional light on this topic?[32] In fact, the multi-tasking model can readily be translated into the context of savings. We can give the following savings interpretation to a multi-task model with $\mathbf{c} = (0, 0, ..., 0)$: People have time-variant instantaneous utility functions, where in any period t the marginal utility of consuming the first dollar is v_t, and the marginal utility for any consumption beyond the first dollar is negligible. Then given wealth $M \in \{\$1, \$2, ..., \$T\}$, you must decide in which periods to consume.

With this savings interpretation, the examples in Section 7 yield surprising results. For instance, Example 12 shows that sophisticates can have a negative marginal propensity to consume (MPC) over some ranges of income. With wealth \$1 sophisticates consume \$1 in period 1, while with wealth \$2 sophisticates consume \$0 in period 1. Intuitively, sophisticates can have negative MPC when additional income makes it worthwhile to save enough to overcome self-control problems. While this example corresponds to time-variant utility functions, sophisticates can have a negative MPC even with time-constant concave utility functions.[33] (The concavity of the utility function is

[32] There has been a lot of previous research on time inconsistency in savings models; see, for instance, Laibson (1994, 1995), Phelps and Pollak (1968), Pollak (1968), Shefrin and Thaler (1988, 1992), Strotz (1955), Thaler (1994), and Thaler and Shefrin (1981).

[33] To illustrate, suppose that in all periods the marginal utility is 1 for the first dollar consumed, $\beta + \epsilon$ for the second dollar consumed, $(\beta+\epsilon)^2$ for the third dollar consumed, etc, where $\epsilon > 0$ is very small. Now suppose you live for $T = 3$ periods, with $\delta = 1$. Assume no liquidity constraints and zero interest on savings. Given total lifetime wealth, W, let $c_t(W)$ be the amount consumed in period t. Then for sophisticates $c_1(4) = 3$, but $c_1(5) = 2$. The intuition is that a sophisticate knows that if he saves less than \$2 to period 2, he will consume it all in period 2, but if he saves \$3 to period 2, he will save \$1 to period 3. With \$4, your period-1 self compares $(c_1, c_2, c_3) = (3, 1, 0)$ to $(c_1, c_2, c_3) = (2, 2, 0)$, and chooses the former. With \$5, your period-1 self compares $(2, 2, 1)$ and $(3, 2, 0)$, and chooses $(2, 2, 1)$. Laibson (1994) gives a similar example with CRRA utility functions that invokes liquidity constraints. Although they appear necessary for a negative MPC with CRRA utilities, this example shows that liquidity constraints are not required with general concave utility functions.

sufficient to guarantee that TCs and naifs don't have a negative MPC.)

A translation of Example 13 yields an even more surprising result: sophisticates may sometimes save more than TCs. With wealth \$2, TCs consume \$1 in year 1 and save \$1 (which is consumed in year 4), while sophisticates consume \$0 in year 1 and save \$2 (which is consumed in years 3 and 4). The odd behavior of sophisticates is driven by the commitment component of the sophistication effect. Sophisticates know they will have future self-control problems, and that additional savings may help to overcome them. Sophisticates save more than TCs because they need the additional savings to overcome future self-control problems.[34] (Proposition 9 implies naifs always save less than TCs in the context of our model, no matter the utility functions.)

Presently preferences would also seem to have important implications in the realm of addiction: Addiction is all about giving in to some salient desire today that has costs in the future. Recently, economists have proposed models of "rational addiction," where people with time-consistent intertemporal preferences get addicted because they choose to do so.[35] These models insightfully formalize the essence of (bad) addictive goods: Consuming more of the good today decreases overall utility but increases marginal utility for consumption of the same good tomorrow. However, these models a priori rule out the time-inconsistency and self-control issues modeled in this paper, and which many observers consider important in addiction.

We have not analyzed the general implications of presently preferences in the context of addictive goods, but several preliminary examples indicate that some of the lessons of this paper provide a useful framework for doing so.[36] Along the lines of Proposition 6, we suspect that even moderate self-control problems can lead naive people to slip into more and more severe addiction. Intuitively, naifs always believe they will reverse addictions tomorrow, so they indulge today. Sophistication should enable sophisticates to avoid this problem. Against this intuition, however, we have found examples along the lines of Proposition 7, where sophistication can lead to damaging addiction

[34] Examples 12 and 13, and the example from the previous footnote, all use rather special utility functions, and we don't know whether there are such examples with more realistic specifications of utility. However, the examples do assume concave and increasing utility functions, and could easily be made differentiable. Hence, any sufficient conditions that rule out odd behavior must go well beyond assuming diminishing marginal utility for consumption. We suspect, but have not proven, that sophisticates will never save more than TCs if utility functions are constant over time.

[35] See Becker, Grossman and Murphy (1991, 1994) and Becker and Murphy (1988).

[36] For related papers on addiction, see Hernnstein and Prelec (1992), Heyman (1994), Orford (1985), Schelling (1992b), and Winston (1980). Addictive goods are but one context where there are "negative internalities"–consumption of a good today causes harm tomorrow (see Hernnstein and Prelec (1992)). Many implications of presently preferences are likely to be similar for non-addictive goods that exhibit negative internalities (such as fatty food) and harmful but nonaddictive drugs.

that naive people would avoid. To illustrate, suppose a person is very sure she will live to be 100 years old, and sure that she will become addicted to sleeping pills in the last year of her life. The unwinding logic in the salient-rewards case can lead sophisticates to get addicted immediately and ruin their lives, whereas naifs don't get addicted until close to the end of their lives.

In addiction, as in other realms, sophistication can also lead to odd behavior that is qualitatively distinct from the types of behavior that either time-consistent people or naive people would have. For example, addictive goods can be Giffen goods for sophisticates ˘ non-addicts may buy more of a good in response to a permanent price increase, because high prices act as a sort of commitment device not to become addicted in the future. The example of Giffen goods invites one last reiteration of a theme we've returned to several times throughout the paper: Behavioral implications of presently preferences are generally going to be pretty intuitive and "clean" if we assume naivete, but often complicated and counter-intuitive with sophistication.

Finally, we note that our focus throughout on the extreme cases ˘ that people are either fully aware of or fully ignorant of future self-control problems ˘ is not likely to be completely innocuous in general settings. In fact, even when sophisticates and naifs behave similarly, a sophisticate/naif hybrid can exhibit completely different behavior due to a miscalibrated scheme of self-control. Consider a hybrid who knows she will have future self-control problems, but underestimates them. She can buy chocolate in either a small package at a high per-unit cost, or in a large package at a low per-unit cost.[37] Aware of her tendency to overeat chocolate, she chooses the small package to prevent herself from eating too much chocolate later. However, she miscalibrates: Her craving is quite strong, so she returns to the store to buy more chocolate (in small packages). The net result may be that she buys just as much chocolate at a much higher price than if she were either fully aware or fully ignorant of her self-control problem. A naif would buy the large package, thinking he wouldn't eat it all immediately upon getting home. A sophisticate will realize the inevitability of eating too much and buy the large packages so as to economize. The hybrid behaves qualitatively differently, and is worse off than either of the extremes.[38]

[37] The role of self-control in purchasing decisions is well-known and much studied by marketing experts. Naughty goods are sold in small packages, because people tend not to buy large packages of goods they know they have a regrettable tendency to overconsume. Marketing consumer goods is also a realm where the type of non-temporal saliency and impulsivity ignored in this paper (e.g., having goods enticingly packaged and prominently displayed at checkout counters) play an immense role. For discussions of self control in the context of consumer choice, see Thaler (1980), Rook (1987), Hoch and Loewenstein (1991), and Wertenbroch (1993).

[38] For other parameter values, naifs too would buy multiple small packages rather than one large package. Hence, while we have emphasized that (in contrast to sophisticates) naifs behave relatively normally in the model we develop in this paper, in more general settings naifs may also exhibit extremely strange behavior.

Appendix: Proofs

Proof of Lemma 1: For each $t \in \{1, 2, ..., T-1\}$, the choice s_t is determined by backwards induction in a finite complete- and perfect-information game. Since we have assumed that people do it when indifferent, believe this about their future selves, believe their future selves believe this about their future selves, etc., there is a unique choice. $\square$

Proof of Proposition 1: (1) Suppose $\tau_n < \tau_{tc}$. Since τ_{tc} is the smallest t that maximizes $v_t - c_t$, $v_{\tau_{tc}} - c_{\tau_{tc}} > v_{\tau_n} - c_{\tau_n}$. When $\hat{\beta} = 1$, $U_N^t(\sigma(1), \beta) = \beta \; \left(max_{\hat{t} \in \{t+1, t+2, ..., T\}}[v_{\hat{t}} - c_{\hat{t}}]\right)$. Therefore, in any period $t < \tau_{tc}$, $U_N^t(\sigma(1), \beta) = \beta \; (v_{\tau_{tc}} - c_{\tau_{tc}})$. For any $t < \tau_{tc}$, $s_t^n = Y$ if and only if $U_Y^t(\beta) \geq U_N^t(\sigma(1), \beta)$ or $\beta v_{\tau_n} - c_{\tau_n} \geq \beta(v_{\tau_{tc}} - c_{\tau_{tc}})$. However, this implies $v_{\tau_n} - c_{\tau_n} \geq v_{\tau_{tc}} - c_{\tau_{tc}}$, a contradiction. Hence, we must have $\tau_n \geq \tau_{tc}$.

(2) Suppose $\tau_n > \tau_{tc}$. Since τ_{tc} maximizes $v_t - c_t$, $U_N^{\tau_{tc}}(\sigma(1), \beta) \leq \beta \; (v_{\tau_{tc}} - c_{\tau_{tc}})$. $s_{\tau_{tc}}^n = N$ only if $v_{\tau_{tc}} - \beta c_{\tau_{tc}} < U_N^{\tau_{tc}}(\sigma(1), \beta)$. However, $(v_{\tau_{tc}} - \beta \; c_{\tau_{tc}}) \geq \beta \; (v_{\tau_{tc}} - c_{\tau_{tc}})$, a contradiction. Hence, we must have $\tau_n \leq \tau_{tc}$. $\square$

Proof of Proposition 2: We show that for each $t \in \{1, 2, ..., T-1\}$, if naifs do it in period t then sophisticates do it. In period t, naifs do it only if $U_Y^t(\beta) \geq U_N^t(\sigma(1), \beta)$, and sophisticates do it if $U_Y^t(\beta) \geq U_N^t(\sigma(\beta), \beta)$. Since $U_N^t(\sigma(1), \beta) = \beta \; \left(max_{\hat{t} \in \{t+1, t+2, ..., T\}}[v_{\hat{t}} - c_{\hat{t}}]\right)$, we must have $U_N^t(\sigma(1), \beta) \geq U_N^t(\sigma(\beta), \beta)$ for all t. The result follows. $\square$

Proof of Corollary 1: The result follows directly from combining part 2 of Proposition 1 with Proposition 2. $\square$

Proof of Proposition 3: For any $t \in \{1, 2, ..., T-1\}$, let $\hat{t}$ be the smallest $t' > t$ such that $s_{t'}^s = Y$. When costs are salient, $s_t^s = Y$ only if $\beta v_t - c_t \geq \beta(v_{\hat{t}} - c_{\hat{t}})$, which implies $v_t - c_t \geq v_{\hat{t}} - c_{\hat{t}}$. Since this can be iterated, for any $t < \bar{t}$ where $s_t^s = s_{\bar{t}}^s = Y$, we have $v_t - c_t \geq v_{\bar{t}} - c_{\bar{t}}$.

Now suppose $\tau_s < \tau_{tc}$. We must have $s_{\tau_{tc}}^s = N$ because otherwise the above logic implies $v_{\tau_s} - c_{\tau_s} \geq v_{\tau_{tc}} - c_{\tau_{tc}}$, a contradiction. Thus, $\tau_{tc} < T$. Using the logic above, $s_{\tau_s}^s = Y$ implies $\beta v_{\hat{\tau}_s} - c_{\hat{\tau}_s} \geq \beta(v_{\hat{t}} - c_{\hat{t}})$. $s_{\tau_{tc}}^s = N$ only if $\beta v_{\tau_{tc}} - c_{\tau_{tc}} < \beta(v_{\hat{\tau}_{tc}} - c_{\hat{\tau}_{tc}})$, so $\tau_s < \tau_{tc}$ only if $\beta v_{\tau_s} - c_{\tau_s} > \beta v_{\tau_{tc}} - c_{\tau_{tc}}$. Since τ_{tc} is the smallest t that maximizes $v_t - c_t$, $v_{\tau_{tc}} - c_{\tau_{tc}} > v_{\tau_s} - c_{\tau_s}$. $\beta v_{\tau_s} - c_{\tau_s} > \beta v_{\tau_{tc}} - c_{\tau_{tc}}$ and $v_{\tau_{tc}} - c_{\tau_{tc}} > v_{\tau_s} - c_{\tau_s}$ can both hold only if $v_{\tau_{tc}} > v_{\tau_s}$ and $c_{\tau_{tc}} > c_{\tau_s}$, which cannot hold if either v or c is weakly decreasing. Hence, we must have $\tau_{tc} \leq \tau_s$. $\square$

Proof of Proposition 4: Proposition 2 says $\tau_s \leq \tau_n$, and from the proof of Proposition 2, we must have $s^s_{\tau_n} = Y$. By definition, $s^s_{\tau_s} = Y$. As argued in the proof of Proposition 3, if $\tau_s < \tau_n$ and $s^s_{\tau_s} = s^s_{\tau_n} = Y$, then $U^0(\tau_s) \geq U^0(\tau_n)$. □

Proof of Proposition 5: Proposition 2 shows that $\tau_s \leq \tau_n$. If $\tau_s = \tau_n$, the result is trivially true. Consider $\tau_s < \tau_n$. Proposition 1(2) shows $\tau_n \leq \tau_{tc}$. If $\tau_n = \tau_{tc}$, the result follows because $U^0(\tau_n) = U^0(\tau_{tc}) > U^0(\tau_s)$. Since $U^t_N(\sigma(1), \beta) = \beta(v_{\tau_{tc}} - c_{\tau_{tc}})$ for all $t < \tau_{tc}$, when $\tau_n < \tau_{tc}$ then $s^n_{\tau_s} = N$ only if $U^{\tau_s}_Y(\beta) < U^{\tau_n}_Y(\beta)$ or $v_{\tau_s} - \beta\, c_{\tau_s} < v_{\tau_n} - \beta\, c_{\tau_n}$. Given $v_{\tau_s} - \beta\, c_{\tau_s} < v_{\tau_n} - \beta\, c_{\tau_n}$, $U^0(\tau_s) > U^0(\tau_n)$ only if $v_{\tau_s} < v_{\tau_n}$ and $c_{\tau_s} < c_{\tau_n}$, which cannot hold if either v or c is weakly decreasing. □

Proof of Proposition 6: (i) When costs are salient, $U^0(\tau_s) < U^0(\tau_{tc})$ only if $s^s_{\tau_{tc}} = N$. Let $\hat{t} \equiv min_{t > \tau_{tc}}\{\ t \mid s^s_t = Y\ \}$, so $\hat{t}$ is when sophisticates would do it if they waited in all $t \leq \tau_{tc}$. Either $\tau_s = \hat{t}$ or $\tau_s < \tau_{tc}$, but using the logic from the proof of Proposition 3, in either case $U^0(\tau_{tc}) - U^0(\tau_s) \leq U^0(\tau_{tc}) - U^0(\hat{t})$.

Given $\hat{t}$, $s^s_{\tau_{tc}} = N$ only if $\beta v_{\tau_{tc}} - c_{\tau_{tc}} < \beta U^0(\hat{t})$ or $-\frac{1-\beta}{\beta} c_{\tau_{tc}} + U^0(\tau_{tc}) < U^0(\hat{t})$. Given the upper bound on costs $\bar{c}$, we have $U^0(\tau_{tc}) - U^0(\hat{t}) < \frac{1-\beta}{\beta}\bar{c}$, so $M^s(\beta; \bar{v}, \bar{c}) < \frac{1-\beta}{\beta}\bar{c}$ and $lim_{\beta \to 1} M^s(\beta; \bar{v}, \bar{c}) = 0$.

(ii) Fix $\beta < 1$. We will show that for any $\epsilon \in (0, \bar{c})$ there exist reward/cost schedule combinations such that $U^0(\tau_{tc}) - U^0(\tau_n) = \bar{v} + \bar{c} - \epsilon$, from which the result follows. Choose $\gamma > 0$ such that $\beta + \gamma < 1$. Let x satisfy $\frac{\epsilon}{(\beta+\gamma)^x} < \bar{c} \leq \frac{\epsilon}{(\beta+\gamma)^{x+1}}$, and let y satisfy $\bar{v} - y\left(\frac{1-\beta}{\beta+\gamma}\bar{c}\right) > 0 \geq \bar{v} - (y+1)\left(\frac{1-\beta}{\beta+\gamma}\bar{c}\right)$. Consider the following reward and cost schedules where $T = x + y + 3$ is finite:

$$\mathbf{v} = (\ \bar{v}\ ,\ \bar{v}\ ,\ \ldots\ ,\ \bar{v}\ ,\ \bar{v}\ ,\ \bar{v} - \left(\tfrac{1-\beta}{\beta+\gamma}\bar{c}\right)\ ,\ \ldots\ ,\ \bar{v} - y\left(\tfrac{1-\beta}{\beta+\gamma}\bar{c}\right)\ ,\ 0)$$
$$\mathbf{c} = (\ \epsilon\ ,\ \tfrac{\epsilon}{(\beta+\gamma)}\ ,\ \ldots\ ,\ \tfrac{\epsilon}{(\beta+\gamma)^x}\ ,\ \bar{c}\ ,\ \bar{c}\ ,\ \ldots\ ,\ \bar{c}\ ,\ \bar{c})$$

Under $\mathbf{v}$ and $\mathbf{c}$, $\tau_{tc} = 1$ so $U^0(\tau_{tc}) = \bar{v} - \epsilon$, and $\tau_n = T$ so $U^0(\tau_n) = -\bar{c}$. Hence, $U^0(\tau_{tc}) - U^0(\tau_n) = \bar{v} + \bar{c} - \epsilon$. □

Proof of Proposition 7: (i) When rewards are salient, by Proposition 1 $\tau_n \leq \tau_{tc}$. For any $t < \tau_{tc}$, naifs believe they will do it in period τ_{tc} if they wait. Hence, $v_{\tau_n} - \beta c_{\tau_n} \geq \beta U^0(\tau_{tc})$, which we can rewrite as $\frac{1-\beta}{\beta} v_{\tau_n} + U^0(\tau_n) \geq U^0(\tau_{tc})$. Given the upper bound on rewards $\bar{v}$, we have $U^0(\tau_{tc}) - U^0(\tau_n) \leq \frac{1-\beta}{\beta}\bar{v}$, so $M^n(\beta; \bar{v}, \bar{c}) \leq \frac{1-\beta}{\beta}\bar{v}$ and $lim_{\beta \to 1} M^n(\beta; \bar{v}, \bar{c}) = 0$.

(ii) Fix $\beta < 1$. We will show that for any $\epsilon \in (0, \bar{v})$ there exist reward/cost schedule combinations such that $U^0(\tau_{tc}) - U^0(\tau_s) = \bar{v} + \bar{c} - \epsilon$, from which the result follows. Let x satisfy $\frac{\epsilon}{\beta^x} < \bar{v} \leq \frac{\epsilon}{\beta^{x+1}}$, and let y satisfy $\bar{c} - y\left(\frac{1-\beta}{\beta}\bar{v}\right) > 0 \geq \bar{c} - (y+1)\left(\frac{1-\beta}{\beta}\bar{v}\right)$. Consider the following reward and cost schedules where $T = x + y + 3$ is finite:

$$\mathbf{v} = (\ \epsilon\ ,\ \tfrac{\epsilon}{\beta}\ ,\ \ldots\ ,\ \tfrac{\epsilon}{\beta^x}\ ,\ \bar{v}\ ,\ \bar{v}\ ,\ \ldots\ ,\ \bar{v}\ ,\ \bar{v})$$
$$\mathbf{c} = (\ \bar{c}\ ,\ \bar{c}\ ,\ \ldots\ ,\ \bar{c}\ ,\ \bar{c}\ ,\ \bar{c} - \left(\tfrac{1-\beta}{\beta}\bar{v}\right)\ ,\ \ldots\ ,\ \bar{c} - y\left(\tfrac{1-\beta}{\beta}\bar{v}\right)\ ,\ 0)$$

Under $\mathbf{v}$ and $\mathbf{c}$, $\tau_{tc} = T$ so $U^0(\tau_{tc}) = \bar{v}$, and $\tau_s = 1$ so $U^0(\tau_s) = \epsilon - \bar{c}$. Hence, $U^0(\tau_{tc}) - U^0(\tau_s) = \bar{v} + \bar{c} - \epsilon$. □

Proof of Proposition 8: With M tasks, TCs choose the M periods with the highest values for $v_t - c_t$. With $M+1$ tasks, TCs choose the $M+1$ periods with the highest values for $v_t - c_t$. Since $\mathbf{v}$ and $\mathbf{c}$ are fixed, the result is obvious.

Consider naifs. In any period t, if naifs have k tasks remaining to be done, they do it if and only if $U_Y^t \geq \beta \cdot (k^{\text{th}}$ best future $v_t - c_t)$. Hence, for any $k^* > k$, if naifs do it with k tasks remaining, then they do it with k^* tasks remaining. Letting $K_n^M(t)$ be the number of tasks remaining (under the perception-perfect strategy) for a naif in period t when there are M total tasks, the above logic implies that for any $t \in \tau_n^M$, if $K_n^{M+1}(t) \geq K_n^M(t)$ then $t \in \tau_n^{M+1}$.

It remains to argue $K_n^{M+1}(t) < K_n^M(t)$ is not possible. This follows because for any t^* the action choices s_t^n for $t \geq t^*$ depend only on the number of tasks remaining to be done at t^*. In other words, if $K_n^{M+1}(t) = K_n^M(t)$, then for either M or $M+1$ total tasks, naifs behave identically in periods t through T. Hence, given $K_n^{M+1}(1) > K_n^M(1)$, $K_n^{M+1}(t) \geq K_n^M(t)$ for all t, and the result follows. □

Proof of Proposition 9: (1) Consider some period t^* where naifs and TCs each have k tasks remaining. TCs wait only if $v_{t^*} - c_{t^*} < [k^{\text{th}}$ best future $v_t - c_t]$. Naifs do it only if $\beta v_{t^*} - c_{t^*} \geq \beta[k^{\text{th}}$ best future $v_t - c_t]$ or $-\frac{1-\beta}{\beta}c_{t^*} + v_{t^*} - c_{t^*} \geq [k^{\text{th}}$ best future $v_t - c_t]$. Hence, if TCs and naifs each have k tasks remaining in period t^*, naifs cannot do it if TCs wait. Then, if $\tau_n^M(i) \geq \tau_{tc}^M(i)$ then $\tau_n^M(i+1) \geq \tau_{tc}^M(i+1)$, because otherwise naifs must do it when TCs wait when each has $M - i$ tasks remaining. $\tau_n^M(1) \geq \tau_{tc}^M(1)$ because otherwise naifs do it while TCs wait when each has M tasks remaining, and the result follows.

(2) Consider some period t^* where naifs and TCs each have k tasks remaining. TCs do it only if $v_{t^*} - c_{t^*} \geq [k^{\text{th}}$ best future $v_t - c_t]$. Naifs wait only if $v_{t^*} - \beta c_{t^*} < \beta[k^{\text{th}}$ best future $v_t - c_t]$ or $\frac{1-\beta}{\beta}v_{t^*} + v_{t^*} - c_{t^*} < [k^{\text{th}}$ best future $v_t - c_t]$. Hence, if TCs and naifs each have k tasks remaining in period t^*, naifs cannot wait if TCs do it. Then, if $\tau_n^M(i) \leq \tau_{tc}^M(i)$ then $\tau_n^M(i+1) \leq \tau_{tc}^M(i+1)$,

because otherwise naifs must wait when TCs do it when each has $M - i$ tasks remaining. $\tau_n^M(1) \leq \tau_{tc}^M(1)$ because otherwise naifs wait while TCs do it when each has M tasks remaining, and the result follows. □

References

Ainslie, G. (1974). "Impulse control in pigeons." Journal of the Experimental Analysis of Behavior 21(3): 485-489.

Ainslie, G. (1975). "Specious reward: A behavioral theory of impulsiveness and impulse control." Psychological Bulletin 82(4): 463-496.

Ainslie, G. (1987). "Self-reported tactics of impulse control." International Journal of the Addictions 22(2): 167-179.

Ainslie, G. (1991). "Derivation of 'Rational' Economic Behavior from Hyperbolic Discount Curves." American Economic Review 81(2): 334-340.

Ainslie, G. (1992). Picoeconomics: The strategic interaction of successive motivational states within the person. Cambridge University Press; New York, NY, US.

Ainslie, G. and N. Haslam (1992a). "Self-control." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 177-209.

Ainslie, G. and N. Haslam (1992b). "Hyperbolic discounting." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 57-92.

Ainslie, G. and R. J. Herrnstein (1981). "Preference reversal and delayed reinforcement." Animal Learning & Behavior 9(4): 476-482.

Akerlof, G. A. (1991). "Procrastination and Obedience." American Economic Review 81(2): 1-19.

Becker, G. S., M. Grossman and K. M. Murphy (1991). "Rational Addiction and the Effect of Price on Consumption." American Economic Review 81(2): 237-41.

Becker, G. S., M. Grossman and K. M. Murphy (1994). "An Empirical Analysis Of Cigarette Addiction." American Economic Review 84(3): 396-418.

Becker, G. S. and K. M. Murphy (1988). "A Theory of Rational Addiction." Journal of Political Economy 96(4): 675-700.

Bernheim, B. D. (1994). "Personal Saving, Information, and Economic Literacy: New Directions for Public Policy." Tax Policy for Economic Growth in the 1990's, American Council for Capital Formation: 53-78.

Chung, S-H., and Herrnstein (1967). "Choice and Delay of Reinforcement." Journal of the Experimental Analysis of Behavior 10: 67-74.

Funder, D. C. and J. Block (1989). "The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence." Journal of Personality & Social Psychology 57(6): 1041-1050.

Glazer, J. and A. Weiss (1992). "Conflicting Preferences and Voluntary Restrictions on Choices," Tel Aviv University, Boston University.

Goldman, S. M. (1979). "Intertemporally Inconsistent Preferences and the Rate of Consumption." Econometrica 47(3): 621-626.

Goldman, S. M. (1980). "Consistent Plans." Review of Economic Studies 47: 533-537.

Hammond, P. J. (1976a). "Endogenous Tastes and Stable Long-Run Choice." Journal of Economic Theory 13: 329-340.

Hammond, P. J. (1976b). "Changing Tastes and Coherent Dynamic Choice." Review of Economic Studies 43: 159-173.

Herrnstein, R. J. and D. Prelec (1992). "A Theory of Addiction," in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 331-360.

Heyman, G. M. (1994). "Addiction, Choice or Disease? An Empirical Review and Theoretical Synthesis," Harvard University.

Hoch, S. and G. Loewenstein (1991). "Time-inconsistent Preferences and Consumer Self-Control." Journal of Consumer Research 17: 492-507.

Kahneman, Knetsch, and Thaler (1991). "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." Journal of Economic Perspectives 5(Winter): 193-206.

Kahneman, D. and D. Lovallo (1993). "Timid choices and bold forecasts: A cognitive perspective on risk taking." Management Science 39(1):17-31.

Koopmans, T. (1960). "Stationary Ordinal Utility and Impatience." Econometrica 28: 287-309.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982). "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." Journal of Economic Theory 27: 245-252.

Laibson, D. (1994). "Essays in Hyperbolic Discounting." Economics, MIT.

Laibson, D. (1995). "Hyperbolic Discount Functions, Undersaving, and Savings Policy," Harvard University.

Loewenstein, G. (1988). "Frames of Mind in Intertemporal Choice." Management Science 34(2): 200-214.

Loewenstein, G. (1991). "Intertemporal Choice in Economics." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US.

Loewenstein, G. (1992). "The fall and rise of psychological explanations in the economics of intertemporal choice." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 3-34.

Loewenstein, G. (1996). "Out of Control: Visceral Influences on Behavior." Organizational Behavior and Human Decision Processes 65(2): 001-0021.

Loewenstein, G. and D. Prelec (1991). "Negative Time Preference." American Economic Review 81(2): 347-352.

Loewenstein, G. and D. Prelec (1992). "Anomalies in Intertemporal Choice: Evidence and an Interpretation." Quarterly Journal of Economics 107(2): 573-597.

Loewenstein, G. and D. Prelec (1993). "Preferences for sequences of outcomes." Psychological Review 100(1): 91-108.

Mischel, W., Y. Shoda and M. L. Rodriguez (1992). "Delay of Gratification in children." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 147-164.

Orford, J. (1985). Excessive Appetites: A Psychological View of Addictions, John Wiley & Sons.

Peleg, B. and M. E. Yaari (1973). "On the Existence of a Consistent Course of Action when Tastes are Changing." Review of Economic Studies 40: 391-401.

Pessemier, E. A. (1978). "Stochastic Properties of Changing Preferences." American Economic Review 68(2 May): 381-385.

Phelps, E. S. and R. A. Pollak (1968). "On Second-best National Saving and Game-equilibrium Growth." Review of Economic Studies 35: 185-199.

Pollak, R. A. (1968). "Consistent Planning." Review of Economic Studies 35: 201-208.

Pollak, R. A. (1970). "Habit Formation and Dynamic Demand Functions." Journal of Political Economy 78(4 July/August): 745-763.

Pollak, R. A. (1978). "Endogenous Tastes in Demand and Welfare Analysis." American Economic Review 68(2 May): 374-379.

Prelec, D. (1990). "Decreasing Impatience: Definition and Consequences," WP 90-015, Harvard Business School.

Rachlin, H. and A. Raineri (1992). "Irrationality, impulsiveness, and selfishness as discount reversal effects." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 93-118.

Rook, D. W. (1987). "The Buying Impulse." Journal of Consumer Research 14: 189-199.

Schelling, T. C. (1978). "Egonomics, or the Art of Self-Management." American Economic Review 68(2): 290-294.

Schelling, T. C. (1984). "Self-Command in Practice, in Policy, and in a Theory of Rational Choice." American Economic Review 74(2): 1-11.

Schelling, T. C. (1992a). "Self-command: a new discipline." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 167-176.

Schelling, T. C. (1992b). "Addictive Drugs: The Cigarette Experience." Science 255: 430-433.

Shefrin, H. and R. Thaler (1988). "The Behavioral Life-Cycle Hypothesis." Economic Inquiry 26 (October): 609-643.

Shefrin, H. M. and R. H. Thaler (1992). "Mental accounting, saving, and self-control." in Choice over time, G. Loewenstein and J. Elster, eds., Russell Sage Foundation, New York, NY, US: 287-329.

Strotz, R. H. (1955). "Myopia and Inconsistency in Dynamic Utility Maximization." Review of Economic Studies 23(3): 165-180.

Thaler, R. H. (1980). "Toward a Positive Theory of Consumer Choice." Journal of Economic Behavior and Organization 1:39-60.

Thaler, R. H. (1991). "Some Empirical Evidence on Dynamic Inconsistency." Quasi Rational Economics, Russell Sage Foundation: 127-133.

Thaler, R. H. (1994). "Psychology and Savings Policies." American Economic Review 84(2): 186-192.

Thaler, R. and G. Loewenstein (1992). "Intertemporal Choice." in The Winners Curse: Paradoxes & Anomalies of Economic Life, R. Thaler, ed., New York, Free Press: 92-106.

Thaler, R. H. and H. M. Shefrin (1981). "An Economic Theory of Self-Control." Journal of Political Economy 89(2): 392-406.

Watterson, B. (1993). The Days are Just Packed: A Calvin and Hobbes Collection, Scholastic.

Wertenbroch, K. (1993). "Consumer Self-Control: How Much Vice and Virtue Do We Buy?," Graduate School of Business, University of Chicago.

Winston, G. (1980). "Addiction and Backsliding: A Theory of Compulsive Consumption." Journal of Economic Behavior and Organization 1: 295-324.

Yaari, M. E. (1977). "Consistent Utilization of an Exhaustible Resource, Or, How to Eat an Appetite-Arousing Cake," Center for Research in Mathematical Economics and Game Theory Research memorandum No. 26.