Choice and Procrastination
Author(s): Ted O'Donoghue and Matthew Rabin
Reviewed work(s):
Source: *The Quarterly Journal of Economics*, Vol. 116, No. 1 (Feb., 2001), pp. 121-160
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/2696445

CHOICE AND PROCRASTINATION*

TED O'DONOGHUE AND MATTHEW RABIN

Recent models of procrastination due to self-control problems assume that a procrastinator considers just one option and is unaware of her self-control problems. We develop a model where a person chooses from a menu of options and is partially aware of her self-control problems. This menu model replicates earlier results and generates new ones. A person might forgo completing an attractive option because she plans to complete a more attractive but never-to-be-completed option. Hence, providing a nonprocrastinator additional options can induce procrastination, and a person may procrastinate worse pursuing important goals than unimportant ones.

"The better is the enemy of the good."
—Voltaire

I. INTRODUCTION

Most of us procrastinate. We delay doing unpleasant tasks that we wish we would do sooner. Such procrastination can be very costly. We skip enjoyable events in mid-April because we procrastinate in completing our taxes; we die young because we procrastinate in quitting smoking, starting a diet, or scheduling a medical checkup; and we are denied tenure because of our own, coauthors', or journal referees' procrastination.

There is a growing literature in economics that assumes people have self-control problems, conceived of as a time-inconsistent taste for immediate gratification. An often discussed implication of such preferences is procrastination.[1] These models of procrastination assume that a potential procrastinator has only one task under consideration, and hence the only concern is when the person completes the task. In most situations, however, a person must decide not only *when* to complete a task, but also

* We thank Erik Eyster, Edward Glaeser, an anonymous referee, and seminar participants at Harvard University, London School of Economics, University of Rochester, University of California at Berkeley, and Princeton University for helpful comments, and Christopher Carpenter and Erik Eyster for valuable research assistance. For help in interpreting Voltaire, we thank Roland Bénabou, Alan Durell, Erik Eyster, Geraldine Legendre, Julian LeGrand, Drazen Prelec, and Daniel Rabin. For financial support, we thank the National Science Foundation (Awards SBR-9709485 and SES-0078796), and Rabin thanks the Russell Sage, Alfred P. Sloan, and MacArthur Foundations. This research was started while Rabin was a Fellow at the Center for Advanced Study in the Behavioral Sciences, supported by NSF Grant SBR-960123. He is grateful for the Center's hospitality and the NSF's support.

1. See, for instance, Prelec [1989], Akerlof [1991], Fischer [1997], and O'Donoghue and Rabin [1999a, 1999b, 1999c].

The Quarterly Journal of Economics, February 2001

which task to complete, or how much effort to apply to a chosen task. If a person must revise a paper for resubmission, she can either respond minimally to the editor's suggestions or expend more effort to respond thoroughly. If she is choosing how to invest some money, she can either thoughtlessly follow the advice of a friend, or thoroughly investigate investment strategies. If she is putting together a montage of Johnny Depp photos, she can either haphazardly throw together a few press clippings or work devoutly to construct the shrine that he deserves.

In this paper we develop a model of procrastination in which a person must choose not only when to do a task, but also which task to do. The model makes a number of realistic predictions incompatible with the conventional assumption of time-consistent preferences. These include the possibilities that providing a person with an attractive new option can cause her to switch from doing something beneficial to doing nothing at all, and that a person may procrastinate more severely when pursuing important goals than unimportant ones.

We also develop a formal model of *partial naivete,* where a person is aware that she will have future self-control problems, but underestimates their magnitude. The literature on self-control problems has focused entirely on two assumptions regarding a person's beliefs about her future self-control problems: that she is *sophisticated*—fully aware of her future self-control problems—or that she is *naive*—fully unaware of her future self-control problems. We believe that introducing a model of partial naivete to the growing literature on time-inconsistent preferences is an important ancillary contribution of this paper. Economists have been predisposed to focus on complete sophistication; but since our results show that *any* degree of naivete can yield different predictions than complete sophistication, our analysis suggests that restricting attention to complete sophistication could be a methodological and empirical mistake even if people are mostly sophisticated.

In Section II we describe a formalization of time-inconsistent preferences originally developed by Phelps and Pollak [1968] in the context of intergenerational altruism and later employed by Laibson [1994] to capture self-control problems within an individual: in addition to time-consistent discounting, a person always gives extra weight to current well-being over future well-being. These "present-biased preferences" imply that each period a person tends to pursue immediate gratification more than she would have preferred if asked in any prior period.

In Section III we present our model of task choice. We suppose that a person faces a menu of possible tasks. Each period she must either complete one of these tasks or do nothing, without being able to commit to future behavior. Completing a task requires that the person incur an immediate cost, but generates an infinite stream of delayed benefits; tasks may differ in both their costs and their benefits. We assume that the person behaves optimally given her taste for immediate gratification and given her beliefs as to how she will behave in the future, where her beliefs reflect her (sophisticated, naive, or partially naive) perceptions of her future self-control problems.

Naivete about future self-control problems leads a person to be overoptimistic about how soon she would complete a task if she were to delay now, and hence is an important determinant of procrastination. Akerlof [1991] emphasizes the role of naivete in putting off unpleasant tasks, and O'Donoghue and Rabin [1999a] show that even mild self-control problems can cause severe procrastination for a completely naive person, but not for a completely sophisticated person.[2] Section III fleshes out the logic behind these earlier results, and generalizes them by allowing for both a menu of tasks and partial naivete. We show that for any specific environment there is a lower bound on the degree of naivete needed to generate severe procrastination. But we also show that for a person with any degree of naivete, no matter how little, there exist environments where that person procrastinates severely.

In Section IV we turn to the core new results of this paper—those regarding the role of choice for procrastination. The implications of choice for procrastination derive from the fact that the two aspects of a person's decision—which task to do and when to do it—are determined by two different criteria. A person plans to

2. Prelec [1989] discusses how time-inconsistent preferences can lead a person to avoid doing an unpleasant task. Because he does not look at a dynamic model, sophistication is not relevant. Fischer [1997] considers procrastination of a task that may take a while to complete. She assumes sophistication, although because she explores long-term projects she finds that substantial procrastination is still possible. Akerlof [1991] does not frame his analysis of procrastination in terms of time-inconsistent preferences, but his model implicitly corresponds to a model of present-biased preferences, and he highlights the role of naive beliefs in generating severe procrastination. O'Donoghue and Rabin [1999a] explicitly compare the naive to the sophisticated model; O'Donoghue and Rabin [1999b, 1999c] explore procrastination with naive beliefs.

do the task which, taking into account her taste for immediate gratification, yields her the highest long-run net benefit. But whether the person ever completes that task depends on a comparison of its immediate cost to the benefits forgone by brief delay, and has very little to do with either its long-run benefit or the features of other tasks available.

The disjunction between these two criteria can produce some realistic behavior patterns inconsistent with conventional economic models. Our first main finding is that providing a person with additional options can induce procrastination. If a new option has a sufficiently high long-run net benefit, the person will plan to do this new option rather than what she would have otherwise done; and if this new option has a sufficiently large cost relative to its immediate benefit, the person now procrastinates. For example, a person might immediately invest her savings in her company's 401(k) plan if there were a single investment option available, but might procrastinate if she must choose from a menu of different investment options because she constantly plans to figure out her best option in the near future. As Voltaire should have meant by the opening quote (but did not), a person may never complete a good task because of persistent but unfulfilled aspirations to do a better job.[3]

Our second main finding is that people may procrastinate more in pursuit of important goals than unimportant ones, or equivalently that increasing importance can exacerbate procrastination. The more important are a person's goals, the more ambitious are her plans. But the more ambitious are her plans—i.e., the higher is the effort she intends to incur—the more likely she is to procrastinate in executing those plans. We formalize this intuition by supposing that the long-run net benefit of all tasks are increased either by making the person more patient or by increasing per-period benefits, and identify classes of situations where a sufficiently large increase in the long-run benefits of all tasks induces a person to procrastinate.

Our model does not imply that people *always* procrastinate the most when pursuing their most important goals. Indeed, this possibility requires the combination of self-control problems, naivete, and multiple options. If any of the three factors is missing,

3. Although we, like many people, interpreted Voltaire to be referring to procrastination, a proper reading of Voltaire's [1878] statement in the original Italian makes clear that he meant something more akin to, "If it ain't broke, don't fix it."

increasing the long-run net benefits of all tasks makes the person more likely to do a task. Even with all three factors present, increased importance can sometimes reduce procrastination. But our model shows that any presumption that people do not procrastinate on important tasks should be dismissed.[4]

We view it as neither a flaw nor a virtue that some of our results are paradoxical from the perspective of traditional economic analysis. Rather, we are interested in their economic relevance. In O'Donoghue and Rabin [1999c], for instance, we argue with some calibration exercises that such issues can be an important determinant of whether and how a person invests her savings for retirement. Investing for retirement is perhaps the single most important economic decision that people (should) make. Our theoretical model matches what seems to be empirically true: in spite of—or perhaps *because* of—its immense importance, many people never get around to carefully planning their investment for retirement. We conclude the paper in Section V with a brief discussion of the results in that paper, as well as a discussion of how the intuitions in this paper might play out in extensions of our model, such as supposing that a person must allocate time among more than one task, or can improve upon what she has done in the past.

II. Present-Biased Preferences and Beliefs

The standard economics model assumes that intertemporal 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. But there is a mass of evidence that intertemporal preferences take on a specific form of *time inconsistency:* a person's relative preference for well-being at an earlier date over a later date gets stronger as the earlier date gets closer.[5] In other words, people have self-control problems caused by a

4. And as such, our model is another example where careful analysis does not bear out the commonplace conjecture that harmfully irrational behavior is eliminated by ("sufficiently") large stakes.

5. See, for instance, Ainslie [1975, 1991, 1992], Ainslie and Haslam [1992a, 1992b], Loewenstein and Prelec [1992], Thaler [1991], and Thaler and Loewenstein [1992]. While the rubric of "hyperbolic discounting" is often used to describe such preferences, the qualitative feature of the time inconsistency is more general, and more generally supported by empirical evidence, than the specific hyperbolic functional form.

tendency to pursue immediate gratification in a way that their "long-run selves" do not appreciate.

In this paper we apply a simple form of such *present-biased preferences,* using a model originally developed by Phelps and Pollak [1968] in the context of intergenerational altruism and later used by Laibson [1994] to model time inconsistency within an individual.[6] Let u_t be the instantaneous utility a person gets in period t. Then her intertemporal preferences at time t, U^t, can be represented by the following utility function:

$$U^t(u_t,u_{t+1}, \ldots ,u_T) \equiv \delta^t u_t + \beta \sum_{\tau=t+1}^{T} \delta^\tau u_\tau.$$

This two-parameter model is a simple modification of the standard one-parameter, exponential-discounting model. The parameter δ represents standard "time-consistent" impatience, whereas the parameter β represents a time-inconsistent preference for immediate gratification. For $\beta = 1$, these preferences are time-consistent. But for $\beta < 1$, at any given moment the person has an extra bias for now over the future.

To examine intertemporal choice given time-inconsistent preferences, one must ask what a person believes about her own future behavior. Two extreme assumptions have appeared in the literature: *sophisticated* people are fully aware of their future self-control problems and therefore correctly predict how their future selves will behave, and *naive* people are fully *un*aware of their future self-control problems and therefore believe their future selves will behave exactly as they currently would like them to behave.[7]

While our main goal in this paper is to analyze the role of choice for procrastination, an ancillary goal is to extend the analysis of time-inconsistent preferences beyond the extreme assumptions of sophistication and naivete. Hence, we also examine behavior for a person who is *partially naive*—she is aware that she has future self-control problems, but she underestimates

6. This model has since been used by Laibson [1996, 1997], Laibson, Repetto, and Tobacman [1998], O'Donoghue and Rabin [1999a, 1999b, 1999c], Fischer [1997], and others.

7. Strotz [1956] and Pollak [1968] carefully lay out these two assumptions (and develop the labels), but do not much consider the implications of assuming one versus the other. Fischer [1997] and Laibson [1994, 1996, 1997] assume sophisticated beliefs. O'Donoghue and Rabin [1999a] consider both, and explicitly contrast the two.

their magnitude. To formalize this notion, let $\hat{\beta}$ be a person's beliefs about her future self-control problems—her beliefs about what her taste for immediate gratification, β, will be in all future periods. A sophisticated person knows exactly her future self-control problems, and therefore has perceptions $\hat{\beta} = \beta$. A naive person believes she will not have future self-control problems, and therefore has perceptions $\hat{\beta} = 1$. A partially naive person has perceptions $\hat{\beta} \in (\beta, 1)$. In the next section we shall define within our specific model a formal solution concept that applies to sophisticates, naifs, partial naifs, and (by setting $\beta = 1$) time-consistent agents. We then show in the context of our model how and when partial naivete leads to procrastination.[8]

III. The Model and Some Results

Suppose that there are an infinite number of periods in which a person can complete a task, and each period the person chooses from the same menu of tasks, $X \subset \mathbb{R}^2_+$. While we permit X to be finite or infinite, we assume that it is closed. Task $x \in X$ can be represented by the pair (c,v), where if a person completes task x in period τ she incurs cost $c \geq 0$ in period τ and initiates a stream of benefits $v \geq 0$ in each period from period $\tau + 1$ onward. While we discuss more realistic alternatives in the conclusion, throughout our analysis we assume that the tasks are mutually exclusive and final: the person can complete at most one task, and can complete that task at most once.

The set of actions available each period is $A \equiv X \cup \{\emptyset\}$. Action $x \in X$ means "complete task x," and action $\emptyset$ means "do nothing." We describe behavior by a *strategy* $\mathbf{s} \equiv (a_1, a_2, \ldots)$

8. For simplicity, we abstract away from some complications that might arise with partial naivete. First, we assume that a person is absolutely positive—although wrong when $\hat{\beta} > \beta$—about her future self-control problems. We doubt that our qualitative results would change much if the person had probabilistic beliefs whose mean underestimated the actual self-control problem. But it is central to our analysis that a person not fully learn over time her true self-control problem, or, if she does come to recognize her general self-control problem, she still continues to underestimate it on a case-by-case basis. Second, we assume that all higher-order beliefs—e.g., beliefs about future beliefs—are also equal to $\hat{\beta}$. Hence, a person has what might be called "complete naivete about her naivete." A partially naive person thinks she will be entirely aware in the future of what she now believes is the extent of her future self-control problems (since otherwise she predicts she will forget what she currently knows). While alternatives are not without merit—we suspect that people do sometimes realize that they are too often overoptimistic—we think our modeling choice here is the most realistic and most tractable.

which specifies an action $a_t \in A$ for each period t.[9] In this environment there are two relevant questions about a person's behavior: (1) when, if at all, does she complete a task? and (2) which task does she complete? Given a strategy $\mathbf{s} \equiv (a_1, a_2, \ldots)$, let $\tau(\mathbf{s})$ denote the period in which the person completes a task, and let $x(\mathbf{s})$ denote the specific task that the person completes. Formally, $\tau(\mathbf{s}) = \min\{t | a_t \neq \varnothing\}$ and $x(\mathbf{s}) = a_{\tau(\mathbf{s})}$, with $\tau(\mathbf{s}) = \infty$ and $x(\mathbf{s}) = \varnothing$ if $a_t = \varnothing$ for all t. While the question of which task the person completes and when she completes that task are of obvious interest, we shall often focus only on whether the person *ever* completes *any* task. Hence, the strategy $\mathbf{s}^{\varnothing} \equiv (\varnothing, \varnothing, \ldots, \varnothing, \ldots)$ plays a prominent role in our analysis.

Our solution concept, "perception-perfect strategies," requires that at all times a person have reasonable beliefs about how she would behave in the future following any possible current action, and that she choose her current action to maximize her current preferences given these beliefs. Let $\hat{\mathbf{s}}^t \equiv (\hat{a}^t_{t+1}, \hat{a}^t_{t+2}, \ldots)$ represent the person's period-t beliefs about future behavior, where $\hat{a}^t_\tau$ represents the person's belief in period t for what action she would choose in period τ if she were to enter period τ not yet having completed a task. Given the person's beliefs $\hat{\mathbf{s}}^t$, let $V^t(a_t, \hat{\mathbf{s}}^t, \beta, \delta)$ represent the person's period-t preferences over current actions conditional on following strategy $\hat{\mathbf{s}}^t$ beginning in period $t+1$.[10] Then,

$$V^t(a_t, \hat{\mathbf{s}}^t, \beta, \delta) \equiv \begin{cases} -c + \beta\delta v/(1-\delta) & \text{if } a_t = (c,v); \\ \beta\delta^\tau[-c + \delta v/(1-\delta)] & \text{if } a_t = \varnothing, \\ & \quad \tau \equiv \min\{d > 0 | \hat{a}^t_{t+d} \neq \varnothing\} \\ & \quad \text{exists,} \\ & \quad \text{and } \hat{a}^t_{t+\tau} = (c,v); \\ 0 & \text{if } a_t = \varnothing \text{ and } \hat{a}^t_{t+d} = \varnothing \\ & \quad \text{for all } d > 0. \end{cases}$$

The three cases in this equation correspond to three different

9. Because a person's choice in period τ matters only following the history with $a_t = \varnothing$ for all $t < \tau$, defining strategies to be independent of history is not restrictive in our model. Our definitions below also rule out mixed strategies; it is perhaps best to interpret our analysis as applying to equilibrium strategies for an infinite horizon that correspond to some equilibrium strategy for a long, finite horizon, which (generically) does not involve mixed strategies.

10. Formally, V^t represents preferences conditional on having chosen $a_\tau = \varnothing$ for all $\tau < t$.

possibilities of when, relative to period t, the person completes the task. In the first case, the person completes task (c,v) now, and therefore she does not discount the immediate cost c by β, but does discount the delayed reward $\delta v/(1-\delta)$ by β. In the second case, the person does nothing now and expects to complete task (c,v) in τ periods, and therefore discounts both the cost and reward by β. In the third case, she does nothing now and expects never to complete any task, and therefore her payoff is zero.

With this notation, a person in period t chooses her current action a_t to maximize her current preferences V^t given her beliefs $\hat{\mathbf{s}}^t$. To predict behavior in our model, however, we do not allow arbitrary beliefs. Rather, a person's beliefs should be a function of her perception of her future self-control problems, $\hat{\beta}$, in conjunction with some coherent theory of how she will behave given such self-control problems. We require beliefs to be dynamically consistent:

DEFINITION 1. Given $\hat{\beta} \leq 1$ and δ, a set of beliefs $\{\hat{\mathbf{s}}^1,\hat{\mathbf{s}}^2, \ldots\}$ is **dynamically consistent** if
(i) for all $\hat{\mathbf{s}}^t$, $\hat{a}^t_\tau = \arg\max_{a\in A} V^\tau(a,\hat{\mathbf{s}}^t,\hat{\beta},\delta)$ for all τ, and
(ii) for all $\hat{\mathbf{s}}^t$ and $\hat{\mathbf{s}}^{t'}$ with $t < t'$, $\hat{a}^t_\tau = \hat{a}^{t'}_\tau$ for all $\tau > t'$.

Definition 1 incorporates two aspects of dynamic consistency. First, each period's beliefs must be *internally consistent*: the beliefs must consist of a behavior path such that each period's action is optimal given that the person will stick to that behavior path in the future. Internal consistency implies that the person perceives in all future periods she will have "rational expectations" about her own behavior even further in the future. Second, the set of beliefs must be *externally consistent*: a person's beliefs must be consistent across periods, which means that her belief of what she will do in period τ must be the same in all $t < \tau$. This restriction rules out procrastination arising from a form of irrational expectations that goes beyond merely mispredicting self-control. For example, if in period 1 a person decides to delay based on a belief that she will complete a task in period 2 in order to avoid procrastination in period 3, then we do not allow this person to delay in period 2 based on a new belief that she will complete a task in period 3.[11]

11. The restrictions imposed by external consistency essentially correspond to the additional restrictions that subgame-perfect equilibrium imposes beyond nonequilibrium backwards induction. By the same token, these restrictions would

Once we impose external consistency on beliefs, we can simplify our notation: given $\hat{\beta}$ and δ, any set of dynamically consistent beliefs can be represented by a single vector of period-1 beliefs $\hat{\mathbf{s}}(\hat{\beta},\delta) = (\hat{a}_2(\hat{\beta},\delta),\ \hat{a}_3(\hat{\beta},\delta),\ \ldots)$, because for all $t > 1$ external consistency requires that period-t beliefs be $\hat{\mathbf{s}}^t(\hat{\beta},\delta) = (\hat{a}_{t+1}(\hat{\beta},\delta),\ \hat{a}_{t+2}(\hat{\beta},\delta),\ \ldots)$.

A perception-perfect strategy is a set of plans where in each period the person chooses an action to maximize her current preferences given dynamically consistent beliefs about future behavior:

DEFINITION 2. A **perception-perfect strategy** for a $(\beta,\hat{\beta},\delta)$ agent is $\mathbf{s}^{pp}(\beta,\hat{\beta},\delta) \equiv (a_1(\beta,\hat{\beta},\delta),\ a_2(\beta,\hat{\beta},\delta),\ \ldots)$ such that there exists dynamically consistent beliefs $\hat{\mathbf{s}}(\hat{\beta},\delta)$ such that $a_t(\beta,\hat{\beta},\delta) = \arg\max_a V^t(a,\ \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta)$ for all t.

This definition includes as special cases the three cases previously studied in the literature: time consistency, sophisticated time inconsistency, and naive time inconsistency. A person with time-consistent preferences is characterized by $\hat{\beta} = \beta = 1$. A completely sophisticated person is characterized by $\hat{\beta} = \beta < 1$; for such a person a perception-perfect strategy is identical to its corresponding dynamically consistent beliefs. A completely naive person is characterized by $\hat{\beta} = 1 > \beta$; such a person believes she will behave like a time-consistent person in the future. Definition 2 generalizes the previous literature by allowing for partial naivete.

In our model, there are two reasons a person might never do any task: because no task is worth doing, and because she "procrastinates." The following terminology will prove useful in distinguishing these cases:

DEFINITION 3. Given β and δ, a task (c,v) is **β-worthwhile** if $\beta\delta v/(1 - \delta) - c \geq 0$; and given X, the **β-best task** in X is $x^*(\beta,\delta,X) \equiv \arg\max_{(c,v)\in X} [\beta\delta v/(1 - \delta) - c]$.[12]

be unnecessary in generic, finite-period situations where "perceptual" backwards induction would yield a unique prediction. Previous analyses of time-inconsistent preferences, whether examining sophistication or complete naivete, have implicitly assumed that people have beliefs that are both internally and externally consistent. Without the assumption of external consistency, even a person who knows exactly her future self-control problems could fail to exhibit rational expectations.

12. Because the set $\mathbf{B} \equiv \arg\max_{(c,v)\in X} [\beta\delta v/(1 - \delta) - c]$ need not be a singleton, $x^*(\beta,\delta,X)$ is not necessarily well-defined. If $\mathbf{B}$ is not a singleton, we define $x^*(\beta,\delta,X)$ to be the task $(c^*,v^*) \in \mathbf{B}$ such that $v^* = \max\ \{v|(c,v) \in$

A task is β-worthwhile if a person prefers doing it now to never doing anything given her taste for immediate gratification; we do not classify as "procrastination" never doing a task merely because no task is β-worthwhile. When there exist β-worthwhile tasks, two types of behavior will emerge in our model: early completion of the β-best task (perhaps with a short delay) or infinite delay. Because it turns out that early completion is for the most part not inefficient, our focus shall be infinite delay. Hence, for our formal analysis we define procrastination to mean never completing a task when there exists some task that is β-worthwhile.[13]

DEFINITION 4. A person **procrastinates** if she follows strategy $\mathbf{s}^{\varnothing}$ when there exists $x \in X$ that is β-worthwhile.

To provide some intuition for why a person might procrastinate in this sense, consider the case where there is a singleton task menu; that is, $X \equiv \{(c,v)\}$. If the task is β-worthwhile, then the person will want to (eventually) complete it. Moreover, she will have some maximum tolerable delay d^* such that for any $d \geq d^* + 1$ completing the task today is preferred to completing the task in d periods. When the maximum tolerable delay is zero—she is not willing to tolerate even a one-period delay—then the person clearly will not delay. When the maximum tolerable delay is greater than zero, the person might delay depending on her perceptions of when in the future she would complete the task.

Suppose that the person is completely sophisticated—she has beliefs $\hat{\beta} = \beta$—and therefore accurately predicts her future behavior. Since in period t she completes the task if and only if she

$\mathbf{B}\}$—that is, the task in $\mathbf{B}$ with the largest reward (and therefore the largest cost). If either $\mathbf{B}$ is empty or $\max \{v|(c,v) \in \mathbf{B}\}$ does not exist, then we say the β-best task does not exist. For a given $(\beta,\hat{\beta},\delta)$ combination, there exists a perception-perfect strategy if and only if the β-best task and the $\hat{\beta}$-best task both exist. If the menu of tasks X is finite, existence is guaranteed. If X is infinite, then letting $\bar{v}(c) \equiv \max_{c' \leq c}\{v|(c',v) \in X\}$ be the maximal benefit that can be achieved for cost c or lower, there exists a perception-perfect strategy if $\bar{v}(c)$ is defined for all c (i.e., the person cannot achieve an infinite reward for a finite cost) and $\lim_{c \to \infty} [\bar{v}(c)/c] < (1 - \delta)/(\hat{\beta}\delta)$.

13. While our formal results revolve around whether a person delays forever, this extreme form of procrastination is an artifact of our simple model. More generally, there are other unmodeled forces that prevent infinite delay. An obvious one is external deadlines. A more interesting one is learning—that is, after repeatedly planning to do a task in the near future and not carrying out these plans, the person may realize the futility of such plans and instead just do the task now. While it is likely that such learning occurs, we suspect that in real-world situations such learning does not take place very quickly, and does not generalize from one situation to another.

predicts that she would delay more than d^* periods, her perception-perfect strategy must be a "cyclical" strategy: in every period that she plans to do the task, it must be that waiting would lead to a delay of exactly $d^* + 1$ periods. Any strategy with this feature is a perception-perfect strategy. If $d^* = 2$, for instance, there are three perception-perfect strategies—$(x,\varnothing,\varnothing,x,\varnothing,\varnothing,x,\ldots)$, $(\varnothing,x,\varnothing,\varnothing,x,\varnothing,\varnothing,\ldots)$, and $(\varnothing,\varnothing,x,\varnothing,\varnothing,x,\varnothing,\varnothing,\ldots)$—and there are three perception-perfect outcomes—completing the task on the first, second, or third day. This logic clearly implies that while a completely sophisticated person might delay for a short while, she will not delay indefinitely, and hence would never procrastinate as we have defined it.[14]

Now suppose that the person is partially naive—she has beliefs $\hat{\beta} > \beta$—and therefore perceives that she will behave in the future like a completely sophisticated person with a self-control problem of $\hat{\beta}$. If a sophisticate with self-control problem $\hat{\beta}$ would tolerate a delay of at most $\hat{d}^*$ periods, then the partially naive person believes the most she will delay if she does not do the task now is $\hat{d}^* + 1$ periods. But if $\hat{d}^* + 1 \leq d^*$, then in all periods the partially naive person perceives that she will complete the task within a tolerable number of periods even if she delays now. She will therefore delay forever. In other words, a partially naive person delays indefinitely whenever she perceives that her future tolerance for delay will be at least one period less than her current (and actual future) tolerance for delay.

While there is only one task in this example, similar logic determines procrastination when there is a menu of tasks from which to choose. The main complication is that the person must choose which task to consider completing now, and she must

14. For any finite horizon, there is a unique perception-perfect strategy for sophisticates, and this strategy corresponds to one of the three strategies in the text. For an infinite horizon, there can exist an additional "mixed" perception-perfect strategy for sophisticates. Although we have ruled out such strategies, it is worth noting what they look like. Let $x^*(\beta,\delta,X) \equiv (c^*,v^*)$, and let p satisfy

$$\frac{\beta\delta v^*}{1-\delta} - c^* = \beta\delta\left[\sum_{\tau=1}^{\infty}(1-p)^{\tau-1}p\delta^{\tau-1}\left(\frac{\delta v^*}{1-\delta} - c^*\right)\right].$$

Preferring to complete the β-best task tomorrow rather than today implies that there exists a unique $p \in (0,1)$ that satisfies this condition, in which case it is a perception-perfect strategy to complete the β-best task with probability p in all periods. While this strategy can yield some delay, it does not represent severe procrastination in the sense that it is Pareto-efficient and does not cause severe welfare losses (as defined below). We conjecture but have not proved that this is the only mixed perception-perfect strategy for sophisticates.

predict which task she would complete in the future if she waits now. But clearly the person only considers completing the β-best task $x^*(\beta,\delta,X)$ now, and she perceives that in the future she will only consider completing the $\hat{\beta}$-best task $x^*(\hat{\beta},\delta,X)$, and therefore in each period the person debates completing the β-best task now versus the $\hat{\beta}$-best task in the not-too-distant future. Hence, whether a person procrastinates boils down to comparing her current tolerance for delaying the β-best task now in favor of completing the $\hat{\beta}$-best task in the future with her perceived future tolerance for delay of the $\hat{\beta}$-best task.

Let $d(\beta|\hat{\beta})$ be the maximum delay d such that a person with self-control problem β prefers doing the $\hat{\beta}$-best task in d periods rather than doing the β-best task now. Letting $x^*(\beta,\delta,X) = (c^*,v^*)$ and $x^*(\hat{\beta},\delta,X) = (\bar{c},\bar{v})$, this current tolerance for delay is given by

$$d(\beta|\hat{\beta}) \equiv \max \{d \in \{0,1,\ldots\} | -c^* + \beta\delta v^*/(1-\delta) < \beta\delta^d(-\bar{c} + \delta\bar{v}/(1-\delta))\}.$$

With this notation, a person with beliefs $\hat{\beta}$ perceives that her future tolerance for delay is $d(\hat{\beta}|\hat{\beta})$, and applying the logic above, the person procrastinates if $d(\hat{\beta}|\hat{\beta}) + 1 \le d(\beta|\hat{\beta})$.

Lemma 1 formally characterizes the set of dynamically consistent beliefs.[15]

LEMMA 1. For all $\hat{\beta}$, δ, and X, any dynamically consistent beliefs $\hat{\mathbf{s}}(\hat{\beta},\delta) \equiv (\hat{a}_2(\hat{\beta},\delta), \hat{a}_3(\hat{\beta},\delta), \ldots)$ must satisfy

(1) For all t either $\hat{a}_t(\hat{\beta},\delta) = \varnothing$ or $\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)$, and

(2) If $x^*(\hat{\beta},\delta,X)$ is not $\hat{\beta}$-worthwhile, then $\hat{a}_t(\hat{\beta},\delta) = \varnothing$ for all t. Otherwise there exists $\tau \in \{2,3,\ldots,d(\hat{\beta}|\hat{\beta}) + 2\}$ such that $\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)$ if and only if $t \in \{\tau,\tau + (d(\hat{\beta}|\hat{\beta}) + 1), \tau + 2(d(\hat{\beta}|\hat{\beta}) + 1), \ldots\}$.

Lemma 1 states that the only task a person would ever expect to complete in the future is the $\hat{\beta}$-best task, and moreover, the person will expect to complete the $\hat{\beta}$-best task every $d(\hat{\beta}|\hat{\beta}) + 1$ periods. In other words, any dynamically consistent beliefs must be cyclical, and whenever some task is $\hat{\beta}$-worthwhile the length of the cycle is finite. Notice, however, that whenever $d(\hat{\beta}|\hat{\beta}) > 0$, the

15. All proofs are in the Appendix.

first date of completion is indeterminate, and therefore there can be multiple dynamically consistent beliefs.

Given that there can be multiple dynamically consistent beliefs, there can be multiple perception-perfect strategies.[16] Many of our results state properties of the entire set of perception-perfect strategies, which we denote by $S^{pp}(\beta,\hat{\beta},\delta,X)$. Lemma 2 characterizes $S^{pp}(\beta,\hat{\beta},\delta,X)$.

LEMMA 2. For all β, $\hat{\beta}$, δ, and X, either $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$, or for every $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$, $x(\mathbf{s}) = x^*(\beta,\delta,X)$, $\tau(\mathbf{s}) \leq d(\hat{\beta}|\hat{\beta}) + 1$, and if $\tau(\mathbf{s}) > 1$ then $\tau(\mathbf{s}) = \tau(\hat{\mathbf{s}})$ where $\hat{\mathbf{s}}$ is the corresponding set of dynamically consistent beliefs.

Lemma 2 establishes that either there is a unique perception-perfect strategy under which the person never completes a task, or in every perception-perfect strategy the person eventually completes task $x^*(\beta,\delta,X)$. That is, for given parameter values there can be indeterminacy solely in *when* the person completes a task, and not in either whether she completes a task or which task she completes. The intuition for determinacy in whether a person will (eventually) complete a task should be clear from our earlier discussion. If $d(\hat{\beta}|\hat{\beta}) + 1 \leq d(\beta|\hat{\beta})$, then in all periods the person prefers what she perceives to be her maximum future delay to doing the task now, and hence she delays in all periods. If $d(\hat{\beta}|\hat{\beta}) = d(\beta|\hat{\beta})$, by contrast, then in some period the person must perceive an intolerable delay, and she will therefore complete the task in that period. In the latter case, a multiplicity of perception-perfect strategies can arise because the period of completion depends on the specific dynamically consistent beliefs the person holds, which determine the first period in which she perceives an intolerable delay from waiting. The final part of Lemma 2 establishes that if the person delays but eventually completes a task, then she correctly predicts the period in which she will

16. There is a unique perception-perfect strategy associated with each set of beliefs $\hat{\mathbf{s}}(\hat{\beta},\delta)$, but different beliefs can yield different perception-perfect strategies. We emphasize that the multiplicity does not arise because of reward-and-punishment supergame strategies. The strategies in the infinite-period model correspond to the set of strategies that are the limit of the strategies in the finite-period model as the number of periods becomes arbitrarily large, where (generically) each finite-period situation will have a unique perception-perfect strategy. The multiplicity in the limit comes from the fact that each perception-perfect strategy is "cyclical," so that (say) a person will plan to do the task the last period if not before, and the fourth-to-last if not before, the seventh-to-last if not before, etc., and not do the task in other periods. Such strategies will therefore predict that in a 1008-period model the person does the task in period 2, but in a 1007-period model she does it in period 1.

complete the task (although she incorrectly predicts which task she will complete in that period whenever the $\hat{\beta}$-best task differs from the β-best task).

Lemma 2 establishes that the parameters of the model fully determine whether or not the person does a task. Proposition 1 uses Lemma 2 to characterize more explicitly how the degree of sophistication $\hat{\beta}$ determines whether or not the person completes the task.

PROPOSITION 1. For all β, δ, and X:

(1) if no $x \in X$ is β-worthwhile, then $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\emptyset}\}$ for all $\hat{\beta}$; and

(2) if there exists $x \in X$ that is β-worthwhile, then either a) $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\emptyset}\}$ for all $\hat{\beta}$, or b) "generically" there exist β^* and β^{**} satisfying $\beta < \beta^* \leq \beta^{**} < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\emptyset}\}$ for any $\hat{\beta} < \beta^*$ and $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\emptyset}\}$ for any $\hat{\beta} > \beta^{**}$.[17]

Part 1 merely states that if no task is β-worthwhile then the person does not complete a task regardless of her perceptions. Part 2 considers the role that naivete plays in procrastination in the more interesting case where some task is β-worthwhile. If a person is sophisticated or nearly sophisticated, then she does not procrastinate. Intuitively, if a task is β-worthwhile, then in all periods the person prefers completing that task immediately to never completing any task. Since a sophisticated person correctly predicts future behavior, she cannot delay indefinitely because if she perceives that she will do so, then she completes the task immediately. If a person is nearly sophisticated, then her beliefs $\hat{\mathbf{s}}(\hat{\beta},\delta)$ are nearly identical to a sophisticate's beliefs $\hat{\mathbf{s}}(\beta,\delta)$, and therefore she completes the task when the sophisticate does so. If a person is more naive, on the other hand, she may well be persistently optimistic enough that her taste for immediate gratification always induces her to delay.[18]

In a slightly different framework, O'Donoghue and Rabin

17. The caveat "generically" is required to guarantee that $\beta^* > \beta$, which holds if we rule out knife-edge parameters where $x^*(\beta,\delta,X) \equiv (c,v)$ and $\delta v/(1-\delta) - c/\beta = \delta^{d(\beta|\beta)+1}[\delta v/(1-\delta) - c]$. In such cases, it could be that $\beta^* = \beta$.

18. Implicit in Proposition 1 is that procrastination can be nonmonotonic in $\hat{\beta}$—i.e., there may exist an interval $[\beta^*,\beta^{**}]$ on which increasing $\hat{\beta}$ leads the person to cycle between completing a task and procrastinating. This nonmonotonicity is driven by changes in the $\hat{\beta}$-best task, which for $\hat{\beta} < 1$ can cause discrete shifts in $d(\beta|\hat{\beta})$ and $d(\hat{\beta}|\hat{\beta})$. There is in fact a second type of nonmonotonicity, driven by the discreteness of $d(\beta|\hat{\beta})$ and $d(\hat{\beta}|\hat{\beta})$, that shows up in later propositions. Nonmonotonicities are a pervasive feature of our model, but none of our

[1999a] show that when there is a single task available, a pure sophisticate (i.e., $\hat{\beta} = \beta$) cannot procrastinate whereas a pure naif (i.e., $\hat{\beta} = 1$) might.[19] Proposition 1 generalizes this result in two ways: it replicates it for the case where there is a menu of tasks available, and establishes that partial naifs behave "in between" pure naifs and pure sophisticates.

Proposition 1 shows that in a given environment a person who is nearly sophisticated does not procrastinate. But Proposition 2 establishes that for any departure from pure sophistication, there exists an environment where the person procrastinates.

PROPOSITION 2. For all β, δ, and $\hat{\beta} > \beta$, there exists X such that the person procrastinates.

That is, any degree of naivete is sufficient to generate procrastination. A person procrastinates whenever she believes her future tolerance for delay will be at least one period less than her current tolerance. If there is only one task, for instance, and the person barely prefers doing it tomorrow rather than today, then even for $\hat{\beta}$ very close to β the person perceives every day that she will do the task tomorrow, and thus she procrastinates forever.[20]

The terms we have used to describe our results—that people *procrastinate* on a task that is β-*worthwhile*—connote that procrastination harms the person. To see why these terms might be appropriate, we now turn to formal welfare analysis. The meaning of the statement that somebody with time-inconsistent preferences is "hurting herself" has sometimes troubled researchers, since time-inconsistent preferences imply that a person evaluates

main results are driven by such nonmonotonicities, and we therefore downplay their role.

19. This statement is true using the definition of procrastination in this paper.

20. Consider the implications for procrastination of allowing mixed strategies for partial naifs. "Generically" a person cannot be indifferent between doing the β-best task now and doing the $\hat{\beta}$-best task in some future period. Hence, a person can mix only if she has mixed beliefs. But our earlier discussion implies that if $d(\hat{\beta}|\hat{\beta}) > 0$, mixed beliefs indeed exist. However, according to these beliefs the (long-run) continuation payoff beginning next period must be just sufficient to make a person with self-control problem $\hat{\beta}$ indifferent between doing a task now versus waiting. But this means that a person with self-control problem $\beta < \hat{\beta}$ will strictly prefer to wait. We can conclude that whenever $d(\hat{\beta}|\hat{\beta}) > 0$ (under a more general definition) there exists a perception-perfect strategy based on mixed beliefs wherein the person procrastinates. We do not focus on such strategies because they make the analysis somewhat trivial—e.g., procrastination for any $\hat{\beta} > \beta$—and we do not feel that they are particularly realistic. We also remind the reader that such strategies are ruled out by a long, finite horizon.

her well-being differently at different times. Some researchers (e.g., Goldman [1979] and Laibson [1994, 1996, 1997]) have avoided this problem by using a "Pareto criterion," under which one stream of utilities is considered unambiguously better than another only if it is preferred by the person from *all* time perspectives.

DEFINITION 5. A strategy **s** is *Pareto-efficient* if there does not exist an alternative strategy **s**′ such that $U^t(\mathbf{s}',\beta,\delta) \geq U^t(\mathbf{s},\beta,\delta)$ for all t and $U^t(\mathbf{s}',\beta,\delta) > U^t(\mathbf{s},\beta,\delta)$ for some t, where

$$U^t(\mathbf{s},\beta,\delta) \equiv \begin{cases} -c + \beta\delta v/(1-\delta) & \text{if } x(\mathbf{s}) = (c,v) \text{ and } \tau(\mathbf{s}) = t \\ \beta\delta^{\tau(\mathbf{s})-t}(-c + \delta v/(1-\delta)) & \text{if } x(\mathbf{s}) = (c,v) \text{ and } \tau(\mathbf{s}) > t \\ 0 & \text{if } x(\mathbf{s}) = \varnothing \\ v + \beta\delta v/(1-\delta) & \text{if } x(\mathbf{s}) = (c,v) \text{ and } \tau(\mathbf{s}) < t. \end{cases}$$

We shall also judge welfare by a second criterion that allows us to evaluate not just whether a person is hurting herself, but also how severely she is doing so.[21] A person's long-run utility—for which β is irrelevant—is $U^{LR}(\mathbf{s},\delta) \equiv U^1(\mathbf{s},1,\delta)$. We define a person's "welfare loss" as the difference between her actual long-run utility and her best possible long-run utility. We normalize the difference by dividing by c^*, the cost of the long-run-best task, so that the welfare loss does not depend arbitrarily on the unit used to measure costs and benefits.[22]

DEFINITION 6. Let $U^* \equiv \max_{\tilde{\mathbf{s}}} U^{LR}(\tilde{\mathbf{s}},\delta)$, and let c^* be the cost of the task chosen (immediately) to maximize $U^{LR}(\tilde{\mathbf{s}},\delta)$. If a person follows strategy **s**, then her *welfare loss* is $WL(\mathbf{s},\delta) \equiv [U^* - U^{LR}(\mathbf{s},\delta)]/c^*$.

21. More generally, we feel the Pareto criterion is too conservative an approach to intrapersonal welfare analysis. Just as for interpersonal comparisons where the Pareto criterion refuses to call a reallocation that barely hurts one person and enormously helps everyone else an improvement, the Pareto criterion refuses to rank strategies when one perspective barely prefers one strategy and all other perspectives vastly prefer a second strategy. For example, suppose that there are two tasks, x_1, with $c_1 = 0$ and $v_1 = 1$, and x_2, with $c_2 = 1{,}000{,}000{,}000{,}000$ and $v_2 = 1.01$. Unless δ is very close to 1, doing task x_1 immediately is clearly better than doing task x_2 immediately; and yet for any δ, doing task x_2 immediately is not Pareto-dominated by x_1. Furthermore, the Pareto criterion's unwillingness to designate x_2 as inefficient *holds even for time-consistent agents.*

22. If all rewards and costs are multiplied by some factor $k > 0$, the set of perception-perfect strategies does not change.

In what follows, we say that a welfare loss of $WL(\mathbf{s},\delta) < (1-\beta)/\beta$ is "small." A welfare loss of $(1-\beta)/\beta$ corresponds to the maximum possible welfare loss from a single episode of pursuing immediate gratification. For instance, it is the maximum welfare loss a person can suffer when she does the β-best task rather than the long-run-best task in period 1, and it is the maximum welfare loss a person could suffer if she were hypothetically forced to commit in period 1 to her most preferred lifetime behavior path. Our focus is on the more dramatic examples of harmful procrastination where a person repeatedly chooses to pursue immediate gratification rather than long-run welfare, in which case she can suffer welfare losses significantly larger than $(1-\beta)/\beta$.

The following proposition characterizes how a person can hurt herself according to our two welfare criteria.

PROPOSITION 3. For all β, $\hat{\beta}$, δ, and X,

(1) If $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$, then
  (a) $\mathbf{s}^{\varnothing}$ is Pareto-inefficient if and only if it is procrastination; and
  (b) $WL(\mathbf{s}^{\varnothing}) > (1-\beta)/\beta$ only if $\mathbf{s}^{\varnothing}$ is procrastination.

(2) If $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$, then
  (a) there exists $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$ that is Pareto-efficient and has $WL(\mathbf{s}) < (1-\beta)/\beta$;
  (b) any $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$ is Pareto-inefficient if and only if $\tau(\mathbf{s}) > d(\beta|\beta) + 1$; and
  (c) for any $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$, $WL(\mathbf{s}) > (1-\beta)/\beta$ only if $\tau(\mathbf{s}) > d(\beta|\beta) + 1$.

Part 1 describes whether a person hurts herself when she never completes any task. If a person never completes a task merely because no task is β-worthwhile, then she is following her period-1 self's most preferred path of behavior, and therefore does not hurt herself by either criterion. In contrast, if she never completes a task when some task is β-worthwhile, then *every period-self* prefers to complete the β-best task in period 1 as opposed to doing nothing. In this case, never completing a task is clearly Pareto inefficient; it may or may not cause a large welfare loss.

Part 2 describes whether a person hurts herself in cases where she does eventually complete some task. In this case there exists at least one perception-perfect strategy under which the person does not harm herself. In particular, doing the β-best task in period 1 is a perception-perfect strategy, and by definition doing the β-best task right away does not cause severe harm. But there may exist other perception-perfect strategies under which the person does harm herself because the period-1 self would have preferred to do the β-best task right away rather than delay doing the β-best task until period $\tau(\mathbf{s})$—which holds whenever $\tau(\mathbf{s}) > d(\beta|\beta) + 1$. The source of harm is that, although the person correctly predicts when she will do a task, she incorrectly predicts that she will complete the $\hat{\beta}$-best task, which can lead her to tolerate too long a delay.

Combining Proposition 3 with our earlier results yields conclusions about the role of naivete in causing welfare harm. Since a completely sophisticated person never procrastinates, and also correctly predicts which task she would do in the future, Proposition 3 implies that a completely sophisticated person never severely hurts herself. But since Proposition 2 implies that anyone not completely sophisticated can procrastinate, Proposition 3 also implies that anyone not completely sophisticated can behave Pareto inefficiently. Although Proposition 3 does not imply that naive procrastination always causes severe harm, the following Proposition shows that there is no upper bound on the harm caused by naive procrastination.

PROPOSITION 4. For any β and any $\hat{\beta} > \beta$,
(1) for any δ, there exists X such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ and $WL(\mathbf{s}^{\varnothing}) > (1 - \beta)/\beta$; and
(2) for any $Z > 0$, there exist X and δ such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ and $WL(\mathbf{s}^{\varnothing}) > Z$.

Hence, just as the behavioral results above extend earlier results about one-task, fully naive procrastination, so too do these results extend earlier welfare results: a person can severely harm herself if and only if she is to some degree naive.

IV. CHOICE AND PROCRASTINATION

Section III shows that the principles developed in Akerlof [1991] and O'Donoghue and Rabin [1999a] for the case of only one task and extreme naivete extend to multiple tasks and partial naivete. In this section we turn to the core new results of our paper, which illustrate aspects of procrastination that pertain specifically to the presence of more than one option.

The implications of choice for procrastination derive from the fact that the two aspects of a person's decision—which task to do and when to do it—are determined by two different criteria. The person decides which task to do according to long-term net benefits, choosing the task that maximizes $\beta\delta v/(1 - \delta) - c$. But whether the person delays doing this task has little to do with the long-term net benefit. Rather, it is (primarily) determined by comparing the cost of the chosen task with its *short-run,* per-period benefit.[23] Lemma 3 emphasizes the disjunction between these two criteria.

LEMMA 3. For all β, $\hat{\beta} > \beta$, δ and X such that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$;

(1) Suppose that X' and δ' satisfy $X' = \{g(c,v)|(c,v) \in X\}$ for some function g with $g(c,v) = (c',v')$ only if $\beta\delta' v'/(1 - \delta') - c' = \beta\delta v/(1 - \delta) - c$. Then $x^*(\beta,\delta',X') = g(x^*(\beta,\delta,X))$, but $S^{pp}(\beta,\hat{\beta},\delta',X') = \{\mathbf{s}^{\varnothing}\}$ if $g(x^*(\beta,\delta,X)) \equiv (c^*,v^*)$ is such that $v^*/c^* < (1 - \beta\delta'/\hat{\beta})/(\beta\delta')$.

(2) There exists X' and δ' that satisfy (i) $X' = \{g(c,v)|(c,v) \in X\}$ for some function g with $g(c,v) = (c',v')$ only if $\beta\delta' v'/(1 - \delta') - c' = \beta\delta v/(1 - \delta) - c$ and (ii) $g(x^*(\beta,\delta,X)) \equiv (c^*,v^*)$ is such that $v^*/c^* < (1 - \beta\delta'/\hat{\beta})/(\beta\delta')$.

Lemma 3 examines transformations of the person's choice set that hold constant the long-term net benefits of all tasks. Part 1 states that such a transformation does not change what task the person plans to do, and yet induces procrastination if it makes the β-best task sufficiently more costly relative to its per-period benefit (regardless of how it affects v/c for any other task). Part 2 establishes that there always exists such a transformation that induces procrastination. Lemma 3 therefore implies that no matter how large the long-run net benefit of the β-best task, a person will procrastinate if its cost is sufficiently large relative to its per-period benefit.

Our first main finding regarding the role of choice for procrastination is that providing additional options to a person who is not procrastinating can in fact induce procrastination. Con-

23. The disclaimer "primarily" comes from the fact that the person is not necessarily deciding when to do the β-best task, but rather choosing between doing the β-best task today versus the $\hat{\beta}$-best task later. But since from today's perspective completing the $\hat{\beta}$-best task in the future can only look better than completing the β-best task in the future, this effect only makes procrastination more likely than connoted by the intuition we emphasize.

sider, for example, a person who would complete task $x_1 \equiv (c_1,v_1)$ immediately if it were the only task available. Suppose that we offer this person an additional option $x_2 \equiv (c_2,v_2)$ that becomes both the β-best task (because $\beta\delta v_2/(1-\delta) - c_2 > \beta\delta v_1/(1-\delta) - c_1$) and the $\hat{\beta}$-best task (because $\hat{\beta}\delta v_2/(1-\delta) - c_2 > \hat{\beta}\delta v_1/(1-\delta) - c_1$). In her own mind, the person's decision now boils down to when to do task x_2, and the availability of task x_1 is completely irrelevant. But if c_2 is sufficiently large relative to v_2, the person will never complete task x_2. Hence, the person might procrastinate when tasks x_1 and x_2 are both available, even though she would complete x_1 if it were the only task available.

Of course, this behavior violates one of the core axioms of revealed-preference theory—that additional options should not change choice among existing options. The source of this violation is the person's naive belief that she will soon do the new option when in fact she will not. The person *intends* to adhere to the weak axiom of revealed preference, but fails to follow through. Proposition 5 formalizes the role of naivete in this phenomenon.

PROPOSITION 5. For all β, $\hat{\beta}$, δ, and X such that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$,

(1) if $\hat{\beta} = \beta$, $S^{pp}(\beta,\hat{\beta},\delta,X') \neq \{\mathbf{s}^{\varnothing}\}$ for all $X' \supset X$; and

(2) if $\hat{\beta} > \beta$, there exists task x' such that $S^{pp}(\beta,\hat{\beta},\delta,X \cup \{x'\}) = \{\mathbf{s}^{\varnothing}\}$.

Part 1 establishes that a fully sophisticated person cannot be induced to procrastinate by providing more options. A sophisticated person completes a task whenever there is a task worth completing. If the initial menu contains an option worth completing, so does any superset. Part 2 establishes, however, that for *any* degree of naivete and any menu of tasks there exists a task that when added to this menu induces procrastination. The logic behind this result is exactly as above: to induce procrastination, we merely add an option that yields higher long-term net benefits than existing options but has a high cost relative to its per-period benefit.[24]

Our second main finding regarding the role of choice for procrastination is that people may procrastinate more in pursuit

24. There are of course welfare analogues to the behavioral results of Proposition 5: while additional choices cannot severely harm a sophisticate, there is no limit to how much additional choices can harm a partial naif, because no matter how well off some initial set of choices makes her, adding an option can induce procrastination.

of important goals than unimportant ones, or equivalently that increasing importance can exacerbate procrastination. This result is best demonstrated with a simple numerical example.

EXAMPLE: Suppose that $\beta = 0.6$ and $\hat{\beta} = 1$.

(1) If $\delta = .8$ and $X \equiv \{x_1 = (0,11),\ x_2 = (40,20)\}$, the person completes task x_1 immediately.

(2a) If $\delta = .9$ and $X \equiv \{x_1 = (0,11),\ x_2 = (40,20)\}$, the person procrastinates.

(2b) If $\delta = .8$ and $X \equiv \{x_1' = (0,22),\ x_2' = (40,40)\}$, the person procrastinates.

(3) If $\delta = .8$ and $X \equiv \{x_1'' = (0,44),\ x_2'' = (40,80)\}$, the person completes task x_2'' immediately.

In this example, Cases (2a) and (2b) represent two ways in which it might become more important relative to Case (1) that the person do something. Case (2a) is identical to Case (1) except for increasing δ from .8 to .9; and Case (2b) is identical to Case (1) except for doubling the per-period benefit from each task. Both transformations increase the present discounted value of rewards for each possible cost. How do these transformations affect behavior? In Case (1) the person plans to complete task x_1—it is both the β-best task and the $\hat{\beta}$-best task—and does so immediately. Each of the transformations makes the person plan to complete task x_2 instead. Unfortunately, for each transformation the immediate cost of task x_2 is sufficiently large relative to its per-period benefit (even after the transformation) that the person procrastinates.

This example illustrates the basic intuition behind our importance-exacerbates-procrastination results: the more important a person's goals, the higher the cost she wishes to incur in pursuit of those goals, but she tends to procrastinate more on higher-cost tasks. While the example illustrates that increasing importance can exacerbate procrastination, this phenomenon is of course not universal. For instance, while Case (2b) shows that doubling the per-period benefit of each task induces procrastination by changing the person's preferred task from x_1 to x_2, Case (3) shows that doubling the per-period benefit of each task once more eliminates procrastination—this time by motivating the person to complete x_2 right away. The remainder of this section explores under what conditions increased importance induces procrastination.

The above example illustrates two ways in which a person's goals might become more important: the person might become more patient, or the per-period benefit from each task cost might become larger. The example also clearly shows how the importance-exacerbates-procrastination phenomenon relies on there being a menu of options. Indeed, in the one-task context, increasing importance reduces the likelihood of procrastination (with a minor caveat).

PROPOSITION 6. Consider a person who faces singleton menu $X \equiv \{(c,v)\}$.

(1) When $\hat{\beta} = 1$, $S^{pp}(\beta,\hat{\beta},\delta,(c,v)) \neq \{\mathbf{s}^{\emptyset}\}$ if and only if $v/c \geq (1 - \beta\delta)/(\beta\delta)$, or equivalently if and only if $\delta \geq c/(\beta v + \beta c)$; and

(2) When $\hat{\beta} < 1$, $S^{pp}(\beta,\hat{\beta},\delta,(c,v)) \neq \{\mathbf{s}^{\emptyset}\}$ if $v/c \geq (1 - \beta\delta)/(\beta\delta)$, or equivalently if $\delta \geq c/(\beta v + \beta c)$; and $S^{pp}(\beta,\hat{\beta},\delta,(c,v)) = \{\mathbf{s}^{\emptyset}\}$ if $v/c < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$, or equivalently if $\delta < c/(\beta v + \beta c/\hat{\beta})$.

Part (1) says that when there is just one task, as the person becomes more patient (δ increases) or as the per-period benefit of the task increases relative to the cost (v/c increases), a completely naive person always becomes less likely to procrastinate.[25] A completely naive person always thinks she will do the task next period if she waits now, and as either a person becomes more patient or the magnitude of benefits relative to costs increases, the person becomes more and more likely to prefer doing the task now to doing it next period. Part (2) establishes that a similar result holds when $\hat{\beta} < 1$ in the sense that for δ or v/c large enough the person completes the task and for δ or v/c small enough the person does not complete the task. There is a minor caveat to the general result, however, because for a partially naive person there is a range where whether the person completes the task can be nonmonotonic in either v/c or δ.[26]

Proposition 6 says that when there is a single task available, increasing importance makes a person less likely to procrastinate. This effect is also present when there are multiple options

25. Although Proposition 6 establishes that a completely naive person delays forever if and only if $v/c < (1 - \beta\delta)/(\beta\delta)$, this does not always correspond to procrastination since the task may not be β-worthwhile. Since not completing the task represents procrastination whenever $v/c \geq (1 - \delta)/(\beta\delta)$, a more precise statement is that increasing δ or increasing v/c always decreases the likelihood of procrastination over the range where the task is β-worthwhile.

26. See our earlier discussion of nonmonotonicities in footnote 18. Once more, the results that follow are *not* driven by these nonmonotonicities.

available in the sense that increasing importance makes a person less likely to procrastinate on any specific option. But with multiple options, increasing importance also makes costly tasks more attractive, which can make procrastination more likely. Whether increased importance increases procrastination depends on the relative importance of these two forces. As illustrated by Cases (1) to (3) of our example, either force can dominate.

To further explore the effects of increasing importance, we consider the limit cases where a person's goals become really important; i.e., as either $\delta \rightarrow 1$ or $v/c \rightarrow \infty$ for all tasks. A crucial question for the limit cases is whether there exists a maximal productive task. Formally, the *maximal productive task in menu* X is

$$(c^{\max},v^{\max}) \equiv \{(c,v) \in X | \ \forall \ (c',v') \in X \text{ either } v' < v \text{ or } v' = v \text{ and } c' \geq c\}.$$

The maximal productive task is the task that yields the maximum possible benefit at the lowest possible cost. There does not exist a maximal productive task if X is such that the person is always able to incur a larger cost to receive a larger benefit.

If the menu of tasks X is finite, then a maximal productive task $(c^{\max},v^{\max})$ exists, and moreover, as the person's goals become sufficiently important, $(c^{\max},v^{\max})$ becomes both the β-best task and the $\hat{\beta}$-best task. Hence, for sufficiently high importance, the person makes her decision about whether to delay as if she were facing the single task $(c^{\max},v^{\max})$. Even when the menu of tasks X is infinite, a similar logic holds if a maximal productive task exists. Proposition 7 summarizes this logic.

PROPOSITION 7. Consider a menu X with a maximal productive task $(c^{\max},v^{\max})$.

(1) If $v^{\max}/c^{\max} > (1-\beta)/\beta$, then there exists $\delta^* < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ for all $\delta > \delta^*$; and if $v^{\max}/c^{\max} < (1-\beta/\hat{\beta})/\beta$, then there exists $\delta^{**} < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ for all $\delta > \delta^{**}$; and

(2) For any strictly increasing function f: $\mathbb{R}_+ \rightarrow \mathbb{R}_+$ that satisfies $f(v^{\max})/c^{\max} > (1-\beta\delta)/\beta\delta$, $S^{pp}(\beta,\hat{\beta},\delta,X'(X)) \neq \{\mathbf{s}^{\varnothing}\}$ where $X'(X) \equiv \{(c,f(v)) | (c,v) \in X\}$.

Proposition 7 extends the single-task results in Proposition 6. Implicit in Proposition 6 is that a sufficiently patient person would complete task (c,v) if $v/c > (1-\beta)/\beta$, and that a sufficiently patient person would procrastinate task (c,v) if $v/c < (1 - \beta/\hat{\beta})/\beta$. Part (1) of Proposition 7 therefore establishes that for any menu with a maximal productive task $(c^{\max},v^{\max})$, whether a sufficiently patient person completes a task depends on whether a sufficiently patient person would complete task $(c^{\max},v^{\max})$ if it were the only task available. Also implicit in Proposition 6 is that for any singleton menu a person will for sure complete the task if the per-period benefit is made large enough. Part (2) of Proposition 7 establishes that for any menu with a maximal productive task $(c^{\max},v^{\max})$, a person will for sure complete some task if the per-period benefits of all tasks are made large enough.[27]

Proposition 7 reflects that when there is a maximal productive task, there is a limit to how much increasing importance can exacerbate procrastination. Intuitively, importance can exacerbate procrastination only because increasing importance leads a person to choose a costlier task. When there is a maximal productive task, however, once performance is important enough the person plans to complete the maximal productive task, and further increasing importance only makes procrastination less likely.

When there is no maximal productive task, in contrast, increasing importance always leads a person to choose a costlier task, and hence the person is more likely to procrastinate for sufficiently important goals. Indeed, increasing patience is quite likely to induce procrastination in this case. To formalize this claim, define $L(X) \equiv \lim_{c\to\infty} \sup \{v'/c' | (c',v') \in X \text{ and } c' \geq c\}$. $L(X)$ is the "limit ratio" of per-period benefits to costs as the person expends ever more effort, and can be loosely interpreted as the limit of the marginal return to additional effort.

PROPOSITION 8. Consider a menu X with no maximal productive task. If $L(X) = 0$, then for all β and $\hat{\beta}$ such that $\hat{\beta} > \beta$, there exists $\delta^* < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ for all $\delta > \delta^*$.

Proposition 8 is the clearest and most striking example of how high importance can exacerbate procrastination: under the reasonable assumption that there is no upper bound on how much effort a person can productively put into a task but the marginal return to additional effort eventually becomes arbitrarily small, a sufficiently patient person with any degree of naivete surely

27. We remind the reader that these results imply nothing about how the person would behave when her goals are only mildly important.

procrastinates. As δ approaches 1, the value of even a small increase in per-period benefits becomes enormous, and hence the optimal task involves a very large cost. But the per-period benefit of the optimal task becomes very small relative to its cost, and therefore the person procrastinates.[28] Proposition 8 is clearer when the task menu can be represented by a continuous function. The following corollary follows directly from Proposition 8.

COROLLARY 1. Suppose that a function $v\colon \mathbb{R}_+ \to \mathbb{R}_+$ is continuously differentiable with $v'(c) > 0$ for all c and $\lim_{c\to\infty} v'(c) = 0$. If $X = \{(c,v(c)) | c \in \mathbb{R}_+\}$, then for all β and $\hat{\beta}$ such that $\hat{\beta} > \beta$, there exists $\delta^* < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{s^{\varnothing}\}$ for all $\delta > \delta^*$.

Corollary 1 indicates that there are some natural classes of task menus such that a person always procrastinates when sufficiently patient. Examples include $v(c) = (a + bc)^d$, where $a \geq 0$, $b > 0$ and $d \in (0,1)$; $v(c) = \ln(c + 1)$; and $v(c) = c/(ac + 1)$, where $a > 0$.

While sufficiently high patience unambiguously leads to procrastination in the case where there is no maximal productive task but $L(X) = 0$, the implication of increasing the per-period benefits of all tasks is more ambiguous. To illustrate, we consider multiplicative transformations of the benefits of all tasks, defining $X(k) \equiv \{(c,kv) | (c,v) \in X\}$. Multiplicative transformations are of particular interest. For instance, if we interpret the task cost as the effort expended to find a good investment opportunity and the task benefits as the per-dollar return, then the factor k corresponds to the quantity of funds that a person plans to invest. The examples above which satisfy the conditions in Corollary 1 illustrate that increasing k has an indeterminate effect on procrastination. If $v(c) = (a + bc)^d$, where $a, b > 0$ and $d \in (0,1)$, then there exists $k^* > 0$ such that $S^{pp}(\beta,\hat{\beta},\delta,X(k)) = \{\mathbf{s}^{\varnothing}\}$ for all $k > k^*$.[29] If $v(c) = c/(ac + 1)$, where $a > 0$, then there exists $k^* > 0$ such that $S^{pp}(\beta,\hat{\beta},\delta,X(k)) \neq \{\mathbf{s}^{\varnothing}\}$ for all $k > k^*$.

We have unfortunately found no useful general characterization of when multiplicative transformations of the benefits induce

28. Proposition 8 restricts attention to $L(X) = 0$ to avoid existence issues. When there is no maximal productive task, existence requires that $L(X) < (1 - \delta)/(\hat{\beta}\delta)$; $L(X) > 0$ therefore implies that no perception-perfect strategy exists for δ close enough to 1.

29. When $a = 0$, increasing per-period benefits has no effect on procrastination; that is, $S^{pp}(\beta,\hat{\beta},\delta,X(k))$ is independent of k.

procrastination. But it is worth exploring why in these examples increasing δ has unambiguous effects on procrastination whereas increasing k has ambiguous effects. Both increasing δ and increasing k cause the person to plan on a more costly task. But whereas for δ close to 1 increasing δ has little impact on the short-term benefits of completing a task immediately, for k large increasing k obviously has a significant impact on the short-term benefits of completing a task immediately. To see this formally, let $(c^*(\delta,k), v^*(\delta,k))$ denote the β-best task as a function of δ and k. Proposition 8 follows because $\lim_{\delta\to 1} [v^*(\delta,k)/c^*(\delta,k)]$ is small. The corresponding condition for increasing k is $\lim_{k\to\infty} [kv^*(\delta,k)/c^*(\delta,k)]$ being small, and clearly this limit can remain large even as $v^*(\delta,k)/c^*(\delta,k)$ becomes small.

The indeterminate effects of increasing per-period benefits is quite general. Indeed, for any X such that a person completes a task, there exists a monotonic transformation of the benefits that induces procrastination—by making the β-best task significantly more costly without drastically increasing the per-period benefits of tasks. And for any X such that a person procrastinates, there exists a monotonic transformation of the benefits that induces completing a task—by significantly increasing the per-period benefits of all tasks without significantly increasing the cost of the β-best task. But one class of transformations always eliminates procrastination no matter X.

LEMMA 4. Let $X(\eta) = \{(c,v+\eta)|(c,v) \in X\}$. Then for all β, $\hat{\beta}$,δ, and X, there exists η^* such that $S^{pp}(\beta,\hat{\beta},\delta,X(\eta)) \neq \{\mathbf{s}^{\varnothing}\}$ for all $\eta > \eta^*$.

A sufficiently large additive transformation in which the benefits of all tasks are increased by the same absolute amount eliminates procrastination. This result drives home one final time the underlying logic behind the importance-exacerbates-procrastination results. Increased importance may induce procrastination when it induces a person to plan to exert more effort. But an additive transformation of the benefits does not affect the optimal amount of effort to exert: the cost associated with the β-best and $\hat{\beta}$-best tasks are unchanged. Hence, when the benefits become large enough, the person for sure completes a task.

V. Discussion and Conclusion

This paper identifies a number of lessons about naive procrastination. We believe that these lessons apply beyond our abstract model, and may be quite relevant in important economic contexts. In O'Donoghue and Rabin [1999c], for instance, we calibrate a model of whether and how a person invests her savings for retirement. We argue that people may significantly delay transferring savings from their checking accounts into higher-interest accounts, even when the long-term benefits of doing so are enormous. For example, suppose that a person is saving $10,000 for retirement 30 years from now. If the person currently earns 1-percent interest in her checking account, and knows of an easy opportunity to earn 6-percent interest instead, it is well worth making the transfer. While the person may or may not procrastinate when the 6-percent account is her sole alternative, choice can greatly exacerbate her procrastination for the same reasons developed in this paper. Because investing for retirement is so important, she may decide that she should put in the effort to do it right—to find (say) a 6.2-percent account. Even if it is very cheap to transfer her money to the 6-percent account, she may not do so because she persistently plans to transfer directly to the 6.2-percent account in the near future. But she may procrastinate searching for the 6.2-percent account for years, making herself much poorer in retirement than she would have been had she settled for the very good option of investing in the 6-percent account.[30]

Moreover, reflecting our importance-exacerbates-procrastination arguments, such procrastination can be exacerbated when the person has more money to invest. For example, the person may severely procrastinate when her principal is $10,000, but not

30. Madrian and Shea [2000] analyze 401(k) savings decisions for two groups of employees at a single firm, those who must elect participation and those who are automatically enrolled unless they opt out. The groups differ in whether the employees were hired before or after a change in the company 401(k) plan. For employees with similar tenure at the firm, the 401(k)-participation rate is 86 percent for the latter group compared with 37 percent for the former group. Moreover, in the automatic-enrollment group, 61 percent choose the default option of a 3-percent contribution rate allocated entirely into a money market fund, whereas very few people who elect participation choose this option. Such "default" behavior is consistent with people procrastinating on their retirement preparation. On the other hand, Madrian and Shea also find inertia effects—people in the automatic-enrollment group who do not choose the default option tend to choose something close to the default option—that are hard to reconcile with procrastination. See also Choi, Laibson, and Metrick [2000] for a similar study.

when her principal is only $1,000. The logic is as in this paper: she plans and executes a quick-and-easy investment strategy for the $1,000, while she plans—but does *not* execute—a more ambitious investment strategy for the $10,000. The calibration exercises in O'Donoghue and Rabin [1999c] also support our claim in Section III that it need not take much naivete to generate procrastination, and reinforce our view that it would be a mistake for economists studying self-control problems to focus solely on models of complete sophistication.

We conclude with some conjectures about some realism-enhancing extensions of our model. For instance, if instead of assuming that a person can complete only one task, we assumed that she might be working on a number of unrelated projects at the same time, then an interesting and ironic result can arise. When there are other projects worth doing tomorrow, not doing a task today means the person must either delay the task for more than one day or delay the other projects. The logic of procrastination says that a person procrastinates because she perceives the cost of delay to be small; if the person is busy, she sees the cost of delaying as higher, and is therefore less likely to procrastinate.[31]

Notice that being busy is not the same as having a larger immediate cost of completing a task. Indeed, if the immediate cost of doing a task is exogenously increased in all periods, a person is *more* likely to procrastinate. If every day a person chooses between doing her taxes versus playing tennis, the more she likes tennis, the higher her immediate cost, and hence the less likely she is to do her taxes. If in addition to paying her taxes she must also paint the workbench, adjust the carburetor, or do other household chores, she might pay her taxes soon. Having only to pay her taxes means that by playing tennis today she is only delaying completion of her taxes. Having to do these other chores too means that by playing tennis today she is delaying completion of all chores.[32]

A second realistic extension is to suppose that a person need not or cannot complete a task all at once. On many projects, a person can do a quick, cheap fix, initiating some benefits in the

31. As Voltaire might have said, "Se vuole aver' fatto una cosa immediatamente, la dia in mano a una persona molto ocupata."

32. While assuming that the person has many tasks to complete might suggest a decrease in procrastination, the importance-exacerbates-procrastination results are likely to generalize. For instance, if for every project there is no maximal productive task but $L(X) = 0$, then no matter how many projects a person faces, she will surely procrastinate on all projects if she is patient enough.

short run, and later come back and do a proper job to yield the rest of the benefits. If a person is writing a research paper, she need not wait until she has the final version, with all the desired results, before distributing it. She can distribute a preliminary draft, labeled as such, telling readers that she intends to produce a more complete paper in the near future. If a person is deciding how best to invest her money, she can put her money in an easy-to-initiate, relatively good investment in the short run, and then continue to search for the ideal investment.

Some preliminary analysis of such situations suggests two implications. First, if a person can improve on what she has done in the past, it becomes more likely that she does at least a quick fix. In our earlier investment example, for instance, if it is costless to transfer first to the 6-percent account and then later to the 6.2-percent account (once it is found), then the person may immediately make the first transfer, which she perceives as a short-term fix. Since the model in this paper precludes the possibility of a short-run fix, it may overstate the likelihood that a person does absolutely nothing.

While the presence of quick fixes makes it less likely that a person does nothing, however, it also makes it less likely that a person completes the task in the ideal way. Once a person has done a quick fix, the short-term damage caused by delay in completing the task is relatively small, and therefore procrastination is more likely. Once the person has put her savings in the 6-percent account, the short-term damage caused by delay in finding the 6.2-percent account is smaller than if her savings were still in the 1-percent checking account. Similarly, once a person has taken half an hour to cover the roof with a tarp to effectively stop the leaks, the cost of delay in fixing the roof is smaller than if the roof were uncovered. Hence, while our model overstates the likelihood that a person does nothing, it also overstates the likelihood that a person actually completes the task.

Conventional economic theory says that a person does something if she believes the benefits outweigh the costs. Models with present-biased preferences assume that people engage in conventional cost/benefit analysis in formulating their *plans.* But they posit that people use a sort of immediate-cost/immediate-benefit analysis in deciding whether to do something *now.* This alternative conception of when people take action challenges the traditional economic notion that behavior reflects one's preferences. As an alternative to the conventional Weak Axiom of Revealed Preference, previous papers generate what might be called the *Weak Axiom of Revealed Procrastination*: if we observe somebody never doing a task when it is the only one she is considering, we learn little about whether she prefers to do that task. This paper generates what can be called the *Strong Axiom of Revealed Procrastination*: if we observe somebody never doing a task when she has a menu of tasks to choose from, we learn even less.

Appendix: Proofs

Proof of Lemma 1. (1) If $a_t = (c,v) \in X$, then $V^t(a_t,\hat{\mathbf{s}},\hat{\beta},\delta) = \hat{\beta}\delta v/(1-\delta) - c$. Because $x^*(\hat{\beta},\delta,X) \equiv \arg\max_{(c,v)\in X} [\hat{\beta}\delta v/(1-\delta) - c]$, $x^*(\hat{\beta},\delta,X) = \arg\max_{a\in A\setminus\{\varnothing\}} V^t(a,\hat{\mathbf{s}},\hat{\beta},\delta)$. Since any dynamically consistent beliefs $\hat{\mathbf{s}}(\hat{\beta},\delta)$ must satisfy $\hat{a}_t(\hat{\beta},\delta) = \arg\max_{a\in A} V^t(a,\hat{\mathbf{s}},\hat{\beta},\delta)$ for all t, it follows that for all t either $\hat{a}_t(\hat{\beta},\delta) = \varnothing$ or $\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)$.

(2) If $x^*(\hat{\beta},\delta,X) \equiv (c,v)$ is not $\hat{\beta}$-worthwhile, then $\hat{\beta}\delta v/(1-\delta) - c < 0$, which implies that $\hat{\beta}\delta v/(1-\delta) - c < \hat{\beta}\delta^\tau[\delta v/(1-\delta) - c]$ for all $\tau \in \{1,2, \ldots\}$. Given $\hat{a}_t(\hat{\beta},\delta) \in \{\varnothing, x^*(\hat{\beta},\delta,X)\}$ for all t, the latter inequality implies that $\arg\max_{a\in A} V^t(a,\hat{\mathbf{s}},\hat{\beta},\delta) = \varnothing$ for all t, and so $\hat{a}_t(\hat{\beta},\delta) = \varnothing$ for all t.

Suppose that $x^*(\hat{\beta},\delta,X) \equiv (c,v)$ is $\hat{\beta}$-worthwhile. Given the definition of $d(\hat{\beta}|\hat{\beta})$, for any $d' \in \{1, \ldots, d(\hat{\beta}|\hat{\beta})\}$, if $\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)$ and $\hat{a}_{t-d}(\hat{\beta},\delta) = \varnothing$ for all $d \in \{1, \ldots, d'-1\}$, then $\arg\max_{a\in A} V^{t-d'}(a,\hat{\mathbf{s}},\hat{\beta},\delta) = \varnothing$. For $d' = d(\hat{\beta}|\hat{\beta}) + 1$, if $\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)$ and $\hat{a}_{t-d}(\hat{\beta},\delta) = \varnothing$ for all $d \in \{1,2, \ldots, d'-1\}$, then $\arg\max_{a\in A} V^{t-d'}(a,\hat{\mathbf{s}},\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)$. It follows that $\hat{\mathbf{s}}(\hat{\beta},\delta)$ must have $\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)$ every $d(\hat{\beta}|\hat{\beta}) + 1$ periods, and $\hat{a}_t(\hat{\beta},\delta) = \varnothing$ otherwise. This condition can be satisfied only if $\min\{t \in \{2,3, \ldots\}|\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)\} \in \{2, \ldots, d(\hat{\beta}|\hat{\beta}) + 2\}$. The result follows.

QED

Proof of Lemma 2. A logic analogous to that in the proof of Lemma 1 implies that for any $\hat{\mathbf{s}}$, $\arg\max_{a\in A} V^t(a,\hat{\mathbf{s}},\beta,\delta) \in \{\varnothing, x^*(\beta,\delta,X)\}$ for all t, which implies that any perception-perfect strategy must satisfy for all t either $a_t(\beta,\hat{\beta},\delta) = \varnothing$ or $a_t(\beta,\hat{\beta},\delta) = x^*(\beta,\delta,X)$. Moreover, $a_t(\beta,\hat{\beta},\delta) = x^*(\beta,\delta,X)$ if and only if $V^t(x^*(\beta,\delta,X), \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) \geq V^t(\varnothing,\hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta)$.

We next prove $V^t(\varnothing, \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) \geq 0$ for all t. If $x^*(\hat{\beta},\delta,X)$ is not $\hat{\beta}$-worthwhile, then Lemma 1 implies the unique $\hat{\mathbf{s}}(\hat{\beta},\delta)$ is $\mathbf{s}^{\varnothing}$, in which case $V^t(\varnothing,\hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) = 0$. If $x^*(\hat{\beta},\delta,X) \equiv (c',v')$ is $\hat{\beta}$-worthwhile, then Lemma 1 implies that any $\hat{\mathbf{s}}(\hat{\beta},\delta)$ must yield for all t, $V^t(\varnothing, \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) = \beta\delta^d[\delta v'/(1-\delta) - c']$ for some $d \in \{1,2, \ldots, d(\hat{\beta}|\hat{\beta}) + 1\}$, which is nonnegative since $\hat{\beta}\delta v'/(1-\delta) - c' \geq 0$ implies that $\delta v'/(1-\delta) - c' \geq 0$.

Suppose that $x^*(\beta,\delta,X)$ is not β-worthwhile. Then for any $\hat{\mathbf{s}}(\hat{\beta},\delta)$ and for all t, $V^t(x^*(\beta,\delta,X), \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) < 0 \leq V^t(\varnothing, \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta)$, and therefore $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$.

Suppose that $x^*(\beta,\delta,X) \equiv (c^*,v^*)$ is β-worthwhile but $d(\hat{\beta}|\hat{\beta}) + 1 \leq d(\beta|\hat{\beta})$. If $x^*(\beta,\delta,X)$ is β-worthwhile then $x^*(\hat{\beta},\delta,X)$ must be $\hat{\beta}$-worthwhile, in which case for any $\hat{\mathbf{s}}(\hat{\beta},\delta)$ and for all t, $V^t(\varnothing,\hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) = \beta\delta^d[\delta v'/(1-\delta) - c']$ for some $d \in \{1, 2, \ldots, d(\hat{\beta}|\hat{\beta}) + 1\}$. This implies that $V^t(\varnothing, \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) \geq \beta\delta^{d(\hat{\beta}|\hat{\beta})+1}[\delta v'/(1-\delta) - c']$ for all t. Since $V^t(x^*(\beta,\delta,X), \hat{\mathbf{s}}(\hat{\beta},\delta),\beta,\delta) = \beta\delta v^*/(1-\delta) - c^*$, $a_t(\beta,\hat{\beta},\delta) = x^*(\beta,\delta,X)$ only if $\beta\delta v^*/(1-\delta) - c^* \geq \beta\delta^{d(\hat{\beta}|\hat{\beta})+1}[\delta v'/(1-\delta) - c']$. But since the definition of $d(\beta|\hat{\beta})$ implies that $\beta\delta v^*/(1-\delta) - c^* < \beta\delta^d[\delta v'/(1-\delta) - c']$ for all $d \leq d(\beta|\hat{\beta})$, $d(\hat{\beta}|\hat{\beta}) + 1 \leq d(\beta|\hat{\beta})$ implies that $a_t(\beta,\hat{\beta},\delta) = \varnothing$ for all t. Hence, if $x^*(\beta,\delta,X)$ is β-worthwhile but $d(\hat{\beta}|\hat{\beta}) + 1 \leq d(\beta|\hat{\beta})$, for any $\hat{\mathbf{s}}(\hat{\beta},\delta)$ the associated perception-perfect strategy is $\mathbf{s}^{\varnothing}$, and thus $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$.

Finally, suppose that $x^*(\beta,\delta,X)$ is β-worthwhile and $d(\hat{\beta}|\hat{\beta}) = d(\beta|\hat{\beta})$ (it is straightforward to show that $d(\hat{\beta}|\hat{\beta}) > d(\beta|\hat{\beta})$ is not possible). The definition of $d(\beta|\hat{\beta})$ implies that for any $\hat{\mathbf{s}}(\hat{\beta},\delta)$ the associated perception-perfect strategy must satisfy $a_t(\beta,\hat{\beta},\delta) = x^*(\beta,\delta,X)$ if and only if $\min\{d \in \{1,2, \ldots\}|\hat{a}_{t+d}(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)\} = d(\beta|\hat{\beta}) + 1 = d(\hat{\beta}|\hat{\beta}) + 1$, and otherwise $a_t(\beta,\hat{\beta},\delta) = \varnothing$. Hence, if $x^*(\beta,\delta,X)$ is β-worthwhile and $d(\hat{\beta}|\hat{\beta}) = d(\beta|\hat{\beta})$, then $\mathbf{s}^{\varnothing} \notin S^{pp}(\beta,\hat{\beta},\delta,X)$, and any $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$ must satisfy $x(\mathbf{s}) = x^*(\beta,\delta,X)$ and $\tau(\mathbf{s}) = \min\{t \in \{1,2, \ldots\}|\ \min\{d \in \{1,2, \ldots\}|\hat{a}_{t+d}(\hat{\beta},\delta) = x^*(\hat{\beta},\delta,X)\} = d(\hat{\beta}|\hat{\beta}) + 1\}$. Clearly, either $\tau(\mathbf{s}) = 1$ (when $\tau(\hat{\mathbf{s}}) = d(\hat{\beta}|\hat{\beta}) + 2$), or $\tau(\mathbf{s}) = \tau(\hat{\mathbf{s}})$.

QED

Proof of Proposition 1. This proof develops a series of properties that will be used in this and other proofs. Throughout this proof we use notation $x^*(\beta,\delta,X) \equiv (c^*(\beta),v^*(\beta))$.

(AA): $c^*(\beta)$ and $v^*(\beta)$ are nondecreasing in β.

Proof. Consider any β and $\beta' > \beta$. The definition of $x^*(\beta,\delta,X)$ implies that $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) \geq \beta\delta v^*(\beta')/(1-\delta) - c^*(\beta')$; the definition of $x^*(\beta',\delta,X)$ implies that $\beta'\delta v^*(\beta)/(1-\delta) - c^*(\beta) \leq \beta'\delta v^*(\beta')/(1-\delta) - c^*(\beta')$; and combining these inequalities yields $[\beta\delta/(1-\delta)][v^*(\beta') - v^*(\beta)] \leq c^*(\beta') -$

$c^*(\beta) \leq [\beta'\delta/(1-\delta)][v^*(\beta') - v^*(\beta)]$. This condition can hold only if $[v^*(\beta') - v^*(\beta)][c^*(\beta') - c^*(\beta)] \geq 0$. But $v^*(\beta') - v^*(\beta) < 0$ and $\beta' > \beta$ imply that $[\beta\delta/(1-\delta)][v^*(\beta') - v^*(\beta)] > [\beta'\delta/(1-\delta)][v^*(\beta') - v^*(\beta)]$ and the condition cannot be satisfied. Hence, $v^*(\beta') - v^*(\beta) \geq 0$ and $c^*(\beta') - c^*(\beta) \geq 0$.

(BB): $[\delta v^*(\beta)/(1-\delta) - c^*(\beta)]$ is nondecreasing in β.

Proof. Consider any β and $\beta' > \beta$. The proof of Property (AA) establishes that $v^*(\beta') - v^*(\beta) \geq 0$ and $c^*(\beta') - c^*(\beta) \leq [\beta'\delta/(1-\delta)][v^*(\beta') - v^*(\beta)]$, which together imply that $c^*(\beta') - c^*(\beta) \leq [\delta/(1-\delta)][v^*(\beta') - v^*(\beta)]$, which in turn yields $\delta v^*(\beta)/(1-\delta) - c^*(\beta) \leq \delta v^*(\beta')/(1-\delta) - c^*(\beta')$.

(CC): If $\hat{\beta} = 1$, then $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if and only if $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) < \beta\delta[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})]$.

Proof. When $\hat{\beta} = 1$, the unique $\hat{\mathbf{s}}(\hat{\beta},\delta)$ has $\hat{a}_t(\hat{\beta},\delta) = x^*(\hat{\beta},\delta, X)$ for all t. Given this $\hat{\mathbf{s}}(\hat{\beta},\delta)$, $a_t(\beta,\hat{\beta},\delta) = \varnothing$ for all t if and only if $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) < \beta\delta[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})]$.

(DD): For any $\hat{\beta}$, $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if $\delta v^*(\beta)/(1-\delta) - c^*(\beta)/\beta < \delta(\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})/\hat{\beta})$.

Proof. The proof of Lemma 2 establishes that $S^{pp}(\beta,\hat{\beta},\delta, X) = \{\mathbf{s}^{\varnothing}\}$ if $d(\hat{\beta}|\hat{\beta}) + 1 \leq d(\beta|\hat{\beta})$. Since $d(\beta|\hat{\beta})$ must satisfy $\delta^{d(\beta|\hat{\beta})}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})] > \delta v^*(\beta)/(1-\delta) - c^*(\beta)/\beta \geq \delta^{d(\beta|\hat{\beta})+1}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})]$ and $d(\hat{\beta}|\hat{\beta})$ must satisfy $\delta^{d(\hat{\beta}|\hat{\beta})}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})] > \delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})/\hat{\beta} \geq \delta^{d(\hat{\beta}|\hat{\beta})+1}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})]$, $d(\hat{\beta}|\hat{\beta}) + 1 \leq d(\beta|\hat{\beta})$ if $\delta v^*(\beta)/(1-\delta) - c^*(\beta)/\beta < \delta(\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})/\hat{\beta})$.

(EE): For any $\hat{\beta}$, $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ if $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) \geq \beta\delta(\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta}))$.

Proof. The proof of Lemma 2 establishes that $S^{pp}(\beta,\hat{\beta},\delta, X) \neq \{\mathbf{s}^{\varnothing}\}$ if $d(\hat{\beta}|\hat{\beta}) = d(\beta|\hat{\beta})$, which must hold if $d(\beta|\hat{\beta}) = 0$. $d(\beta|\hat{\beta}) = 0$ if and only if $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) \geq \beta\delta(\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta}))$.

Continuity Property A: $[\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta)]$ is continuous in β.

Proof. $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) \equiv \max_{(c,v)\in X}[\beta\delta v/(1-\delta) - c]$, which is continuous if X is closed.

Continuity Property B: For every $\epsilon > 0$ there exists $\beta' > \beta$ such that for all $\hat{\beta} \in (\beta,\beta')$, $[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})] - [\delta v^*(\beta)/(1-\delta) - c^*(\beta)] < \epsilon$.

Proof. The definition of $x^*(\beta,\delta,X)$ implies that $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) \geq \beta\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})$, which implies that $[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})] - [\delta v^*(\beta)/(1-\delta) - c^*(\beta)] \leq [(1-\beta)/\beta][c^*(\hat{\beta}) - c^*(\beta)]$. It is therefore sufficient to show that for any $\bar{\epsilon} > 0$ there exists $\beta' > \beta$ such that for all $\hat{\beta} \in (\beta,\beta')$, $c^*(\hat{\beta}) - c^*(\beta) < \bar{\epsilon}$. Define $\beta^* \equiv \inf\{\beta' > \beta | c^*(\beta') > c^*(\beta)\}$. If either β^* does not exist (because $c^*(\beta') = c^*(\beta)$ for all $\beta' > \beta$) or $\beta^* > \beta$, the result follows. Suppose that $\beta^* = \beta$. Let $\hat{c} \equiv \lim_{\beta' \to \beta^+} c^*(\beta')$ and $\hat{v} \equiv \lim_{\beta' \to \beta^+} v^*(\beta')$, both of which exist since c^* and v^* are nondecreasing. We must have $\beta\delta\hat{v}/(1-\delta) - \hat{c} \geq \beta\delta v^*(\beta)/(1-\delta) - c^*(\beta)$, since otherwise there would exist a neighborhood $\hat{X}$ of $(\hat{c},\hat{v})$ such that $(c^*(\beta),v^*(\beta))$ is $\hat{\beta}$-preferred to any $x \in \hat{X}$ for $\hat{\beta}$ close enough to β. Given the definition of $x^*(\beta,\delta,X)$, we can conclude that $(\hat{c}(\beta),\hat{v}(\beta)) = (c^*(\beta),v^*(\beta))$, and the result follows.[33]

Proof of Part (1). Follows directly from the proof of Lemma 2.

Proof of Part (2). Suppose that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ for $\hat{\beta} = 1$, in which case Property (CC) implies that $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) \geq \beta\delta[\delta v^*(1)/(1-\delta) - c^*(1)]$. Since Property (BB) implies that $[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})] \leq [\delta v^*(1)/(1-\delta) - c^*(1)]$ for all $\hat{\beta} < 1$, it follows that $\beta\delta v^*(\beta)/(1-\delta) - c^*(\beta) \geq \beta\delta[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})]$ for all $\hat{\beta}$, in which case Property (EE) implies that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ for all $\hat{\beta}$.

Suppose that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ for $\hat{\beta} = 1$, in which case Property (CC) implies that $\delta v^*(\beta)/(1-\delta) - c^*(\beta)/\beta < \delta[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})/\hat{\beta}]$ for $\hat{\beta} = 1$. Property (DD) and Continuity Property A imply that there exists $\beta^{**} < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ for all $\hat{\beta} > \beta^{**}$.

Lemma 1 implies that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ for $\hat{\beta} = \beta$. "Generically,"[34] $d(\beta|\beta)$ satisfies $\delta^{d(\beta|\beta)}[\delta v^*(\beta)/(1-\delta) - c^*(\beta)] > \delta v^*(\beta)/(1-\delta) - c^*(\beta)/\beta > \delta^{d(\beta|\beta)+1}[\delta v^*(\beta)/(1-\delta) - c^*(\beta)]$. Since $d(\beta|\hat{\beta})$ must satisfy $\delta^{d(\beta|\hat{\beta})}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})] >$

33. This last step relies on our assumption that if the set $\mathbf{B} \equiv \arg\max_{(c,v)\in X} [\beta\delta v/(1-\delta) - c]$ is not a singleton, then $x^*(\beta,\delta,X)$ is the task $(c^*,v^*) \in \mathbf{B}$ such that $v^* = \max\{v|(c,v) \in \mathbf{B}\}$.

34. By "generically," we mean ruling out knife-edge parameters where $\delta v^*(\beta)/(1-\delta) - c^*(\beta)/\beta = \delta^{d(\beta|\beta)+1}[\delta v^*(\beta)/(1-\delta) - c^*(\beta)]$.

$\delta v^*(\beta)/(1-\delta) - c^*(\beta)/\beta \geq \delta^{d(\beta|\hat{\beta})+1}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})]$, Continuity Property B implies that there exists $\beta' > \beta$ such that $d(\beta|\hat{\beta}) = d(\beta|\beta)$ for all $\hat{\beta} \in (\beta,\beta')$. Similarly, since $d(\hat{\beta}|\hat{\beta})$ must satisfy $\delta^{d(\hat{\beta}|\hat{\beta})}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})] > \delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})/\hat{\beta} \geq \delta^{d(\hat{\beta}|\hat{\beta})+1}[\delta v^*(\hat{\beta})/(1-\delta) - c^*(\hat{\beta})]$, Continuity Properties A and B imply that there exists $\beta'' > \beta$ such that $d(\hat{\beta}|\hat{\beta}) = d(\beta|\beta)$ for all $\hat{\beta} \in (\beta, \beta'')$. If $\beta^* \equiv \min\{\beta', \beta''\}$, then $d(\beta|\hat{\beta}) = d(\hat{\beta}|\hat{\beta}) = d(\beta|\beta)$ for all $\hat{\beta} < \beta^*$, and therefore $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ for all $\hat{\beta} < \beta^*$. Since it is clear that $\beta^* \leq \beta^{**}$, the result follows.

QED

Proof of Proposition 2. It is sufficient to show there exists a singleton such $X \equiv \{(c,v)\}$. The task is β-worthwhile if $\beta\delta v/(1-\delta) - c > 0$ or $v/c > (1-\delta)/(\beta\delta)$. Property (DD) from the proof of Proposition 1 implies that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if $\delta v/(1-\delta) - c/\beta < \delta[\delta v/(1-\delta) - c/\hat{\beta}]$, or $v/c < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$. $\hat{\beta} > \beta$ implies that $(1 - \beta\delta/\hat{\beta})/(\beta\delta) > (1-\delta)/(\beta\delta)$, and thus there exists (c,v) such that $(1-\beta\delta/\hat{\beta})/(\beta\delta) > v/c > (1-\delta)/(\beta\delta)$. The result follows.

QED

Proof of Proposition 3. (1a) If $\mathbf{s}^{\varnothing}$ is not procrastination, then no task in X is β-worthwhile. For any $\mathbf{s} \neq \mathbf{s}^{\varnothing}$, $U^{\tau(\mathbf{s})}(\mathbf{s},\beta,\delta) < 0 = U^{\tau(\mathbf{s})}(\mathbf{s}^{\varnothing},\beta,\delta)$, which implies that $\mathbf{s}^{\varnothing}$ is Pareto-efficient. If $\mathbf{s}^{\varnothing}$ is procrastination, then task $x^*(\beta,\delta,X) \equiv (c,v)$ is β-worthwhile. For any $\mathbf{s}$ satisfying $\tau(\mathbf{s}) = 1$ and $x(\mathbf{s}) = x^*(\beta,\delta,X)$, $U^1(\mathbf{s},\beta,\delta) \geq 0 = U^1(\mathbf{s}^{\varnothing},\beta,\delta)$ and $U^t(\mathbf{s},\beta,\delta) > 0 = U^t(\mathbf{s}^{\varnothing},\beta,\delta)$ for all $t \in \{2,3,\ldots\}$, which implies that $\mathbf{s}^{\varnothing}$ is Pareto-inefficient.

(1b) Letting $x^*(1,\delta,X) \equiv (c^*,v^*)$, $WL(\mathbf{s}^{\varnothing},\delta) = \max\{0, [\delta/(1-\delta)]v^*/c^* - 1\}$. If $\mathbf{s}^{\varnothing}$ is not procrastination, then no task in X is β-worthwhile. (c^*,v^*) not β-worthwhile implies that $\beta\delta v^*/(1-\delta) - c^* < 0$ or $[\delta/(1-\delta)]v^*/c^* < 1/\beta$. Hence, $WL(\mathbf{s}^{\varnothing},\delta) < 1/\beta - 1 = (1-\beta)/\beta$.

(2a) Lemma 2 implies that if $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ then there exists $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$ such that $\tau(\mathbf{s}) \leq d(\beta|\beta) + 1$. The result is then a direct implication of parts (2b) and (2c).

(2b) Any $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$ with $\tau(\mathbf{s}) \leq d(\beta|\beta) + 1$ is Pareto-efficient because (i) any $\mathbf{s}' \neq \mathbf{s}$ with $\tau(\mathbf{s}') = \tau(\mathbf{s})$ and $x(\mathbf{s}') = x^*(\beta,\delta,X)$ yields $U^t(\mathbf{s}',\beta,\delta) = U^t(\mathbf{s},\beta,\delta)$ for all t; (ii) any $\mathbf{s}' \neq \mathbf{s}$ with $\tau(\mathbf{s}') = \tau(\mathbf{s})$ and $x(\mathbf{s}') \neq x^*(\beta,\delta,X)$ yields $U^{\tau(\mathbf{s})}(\mathbf{s}',\beta,\delta) < U^{\tau(\mathbf{s})}(\mathbf{s},\beta,\delta)$; (iii) any $\mathbf{s}' \neq \mathbf{s}$ with $\tau(\mathbf{s}') > \tau(\mathbf{s})$ yields $U^{\tau(\mathbf{s}')}(\mathbf{s}',\beta,\delta) < U^{\tau(\mathbf{s}')}(\mathbf{s},\beta,\delta)$ because having completed $x^*(\beta,\delta,X)$ in the past is better than completing any task now; and (iv) any $\mathbf{s}' \neq \mathbf{s}$ with $\tau(\mathbf{s}') < \tau(\mathbf{s})$

yields $U^{\tau(\mathbf{s}')}(\mathbf{s}',\beta,\delta) < U^{\tau(\mathbf{s}')}(\mathbf{s},\beta,\delta)$ because $\tau(\mathbf{s}) \leq d(\beta|\beta) + 1$ implies that completing $x^*(\beta,\delta,X)$ in $\tau(\mathbf{s}) - \tau(\mathbf{s}')$ periods is better than completing any task now.

Any $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$ with $\tau(\mathbf{s}) > d(\beta|\beta) + 1$ is Pareto-dominated by any strategy $\mathbf{s}' \neq \mathbf{s}$ with $\tau(\mathbf{s}') = 1$ and $x(\mathbf{s}') = x^*(\beta,\delta,X)$. The definition of $d(\beta|\beta)$ implies that $U^1(\mathbf{s}',\beta,\delta) > U^1(\mathbf{s},\beta,\delta)$, and since having completed $x^*(\beta,\delta,X)$ in the past is better than completing $x^*(\beta,\delta,X)$ now or in the future, $U^t(\mathbf{s}',\beta,\delta) \geq U^t(\mathbf{s},\beta,\delta)$ for all $t \geq 2$.

(2c) $WL(\mathbf{s},\delta) = (1/c^*)[(\delta v^*/(1-\delta) - c^*) - \delta^{\tau(\mathbf{s})-1}(\delta v/(1-\delta) - c)]$ where $x^*(\beta,\delta,X) \equiv (c,v)$ and $x^*(1,\delta,X) \equiv (c^*,v^*)$. If $\tau(\mathbf{s}) = 1$, $\delta^{\tau(\mathbf{s})-1}(\delta v/(1-\delta) - c) = \delta v/(1-\delta) - c \geq \delta v/(1-\delta) - c/\beta$. If $\tau(\mathbf{s}) \in \{2, \ldots, d(\beta|\beta) + 1\}$, then the definition of $d(\beta|\beta)$ implies that $\delta^{\tau(\mathbf{s})-1}(\delta v/(1-\delta) - c) > \delta v/(1-\delta) - c/\beta$. Hence, for any $\mathbf{s} \in S^{pp}(\beta,\hat{\beta},\delta,X)$ with $\tau(\mathbf{s}) \leq d(\beta|\beta) + 1$, $WL(\mathbf{s},\delta) \leq (1/c^*)[(\delta v^*/(1-\delta) - c^*) - (\delta v/(1-\delta) - c/\beta)]$. $x^*(\beta,\delta,X) \equiv (c,v)$ implies that $\delta v/(1-\delta) - c/\beta \geq \delta v^*/(1-\delta) - c^*/\beta$, which yields $WL(\mathbf{s},\delta) \leq (1-\beta)/\beta$.

QED

Proof of Proposition 4. (1) It is sufficient to prove each result for a singleton $X \equiv \{(c,v)\}$. Property (DD) from the proof of Proposition 1 implies that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if $\delta v/(1-\delta) - c/\beta < \delta(\delta v/(1-\delta) - c/\hat{\beta})$, or $v/c < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$. As long as $v/c > (1-\delta)/\delta$, $WL(\mathbf{s}^{\varnothing},\delta) = [\delta/(1-\delta)]v/c - 1$ (and otherwise $WL(\mathbf{s}^{\varnothing},\delta) = 0$). Hence, for any $\epsilon > 0$ there exists $X \equiv \{(c,v)\}$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ and $WL(\mathbf{s}^{\varnothing},\delta) > ([\delta/(1-\delta)][(1-\beta\delta/\hat{\beta})/(\beta\delta)] - 1) - \epsilon$. Since $\hat{\beta} > \beta$ implies that $([\delta/(1-\delta)][(1-\beta\delta/\hat{\beta})/(\beta\delta)] - 1) > (1-\beta)/\beta$, the result follows.

(2) This result is a straightforward extension of the proof of part (1)—if we can choose X and δ, then we can make $WL(\mathbf{s}^{\varnothing},\delta)$ arbitrarily large by choosing δ sufficiently close to 1.

QED

Proof of Lemma 3. (1) Given that $x^*(\beta,\delta,X) \equiv \arg\max_{(c,v)\in X} [\beta\delta v/(1-\delta) - c]$, it is clear that $\arg\max_{(c,v)\in X'} [\beta\delta' v/(1-\delta') - c] = g(x^*(\beta,\delta,X))$. That $S^{pp}(\beta,\hat{\beta},\delta',X') = \{\mathbf{s}^{\varnothing}\}$ if $g(x^*(\beta,\delta,X)) \equiv (c^*,v^*)$ is such that $v^*/c^* < (1 - \beta\delta'/\hat{\beta})/(\beta\delta')$ follows directly from the following property.

Property (FF): For any menu X, if $x^*(\beta,\delta,X) \equiv (c,v)$ then for any $\hat{\beta}$, $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if $\delta v/(1-\delta) - c/\beta < \delta(\delta v/(1-\delta) - c/\hat{\beta})$.

Proof. Letting $x^*(\hat{\beta},\delta,X) = (c',v')$, Property (DD) from the proof of Proposition 1 says that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if $\delta v/(1 - \delta) - c/\beta < \delta(\delta v'/(1 - \delta) - c'/\hat{\beta})$. Since $x^*(\hat{\beta},\delta,X) = (c',v')$ implies that $\delta v'/(1 - \delta) - c'/\hat{\beta} \geq \delta v/(1 - \delta) - c/\hat{\beta}$, the result follows.

(2) Let $\delta' = \delta$, and let $g(c,v) = (c + n,\ v + n(1 - \delta)/(\beta\delta))$. X' and δ' satisfy condition (i) for any n. Letting $x^*(\beta,\delta,X) \equiv (c_o, v_o)$ and $x^*(\beta,\delta',X') \equiv (c^*,v^*)$, part (1) implies that $c^* = c_o + n$ and $v^* = v_o + n(1 - \delta)/(\beta\delta)$. Because $\lim_{n\to\infty} [v^*/c^*] = (1 - \delta)/(\beta\delta) < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$ given $\hat{\beta} > \beta$, for n large enough X' and δ' satisfy condition (ii) as well.

QED

Proof of Proposition 5. (1) Proposition 1 establishes $S^{pp}(\beta,\beta,\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ if and only if there exists $x \in X$ that is β-worthwhile. $X' \supset X$ implies that if there exists $x \in X$ that is β-worthwhile, then there exists $x \in X'$ that is β-worthwhile, and the result follows.

(2) Define $x^*(\beta,\delta,X) \equiv (c^*,v^*)$, and consider $x' \equiv (c',v')$ with $c' > c^*$. If $\beta\delta v'/(1 - \delta) - c' > \beta\delta v^*/(1 - \delta) - c^*$, then $x^*(\beta,\delta,X \cup x') = x'$. If in addition $\delta v'/(1 - \delta) - c'/\beta < \delta(\delta v'/(1 - \delta) - c'/\hat{\beta})$, then Property (FF) from the proof of Lemma 3 implies that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$. We can rewrite the first inequality as $(v' - v^*)/(c' - c^*) > (1 - \delta)/(\beta\delta)$, and the second inequality as $v'/c' < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$. Since $\hat{\beta} > \beta$ implies that $(1 - \beta\delta/\hat{\beta})/(\beta\delta) > (1 - \delta)/(\beta\delta)$, for any (c^*,v^*) there exists $(c',v') \in \mathbb{R}^2_+$ with $c' > c^*$ that satisfies both properties.

QED

Proof of Proposition 6. (1) When $X \equiv \{(c,v)\}$, Property (CC) from the proof of Proposition 1 becomes if $\hat{\beta} = 1$, then $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if and only if $\beta\delta v/(1 - \delta) - c < \beta\delta[\delta v/(1 - \delta) - c]$, which can be rearranged as $v/c < (1 - \beta\delta)/(\beta\delta)$ or $\delta < c/(\beta v + \beta c)$.

(2) When $X \equiv \{(c,v)\}$, Property (EE) from the proof of Proposition 1 becomes for any $\hat{\beta}$, $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ if $\beta\delta v/(1 - \delta) - c \geq \beta\delta[\delta v/(1 - \delta) - c]$, which can be rearranged as $v/c \geq (1 - \beta\delta)/(\beta\delta)$ or $\delta \geq c/(\beta v + \beta c)$. When $X \equiv \{(c,v)\}$, Property (DD) from the proof of Proposition 1 becomes for any $\hat{\beta}$, $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if $\delta v/(1 - \delta) - c/\beta < \delta[\delta v/(1 - \delta) - c/\hat{\beta}]$, which can be rearranged as $v/c < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$ or $\delta < c/(\beta v + \beta c/\hat{\beta})$.

QED

Proof of Proposition 7. (1) Define $x^*(\beta,\delta,X) \equiv (c^*(\delta), v^*(\delta))$. We first prove that $\lim_{\delta\to 1} [v^*(\delta)/c^*(\delta)] = v^{\max}/c^{\max}$. c^* and v^* must be nondecreasing in δ (the logic is exactly analogous to that used to prove Property (AA) in the proof of Proposition 1), which implies that $\lim_{\delta\to 1} c^*(\delta)$ and $\lim_{\delta\to 1} v^*(\delta)$ both exist. Note that for any (c,v) and (c',v') satisfying $c' > c$ and $v' > v$, there exists $\delta' < 1$ such that $\beta\delta v'/(1 - \delta) - c' > \beta\delta v/(1 - \delta) - c$ for all $\delta > \delta'$. Since for any $(c,v) \in X$ with $c < c^{\max}$ there exists $(c',v') \in X$ with $c' > c$ and $v' > v$, for any $(c,v) \in X$ with $c < c^{\max}$ we must have $\lim_{\delta\to 1} c^*(\delta) > c$. Then $c^*(\delta) \leq c^{\max}$ for all δ implies that $\lim_{\delta\to 1} c^*(\delta) = c^{\max}$. It follows directly that $\lim_{\delta\to 1} v^*(\delta) = v^{\max}$ and therefore $\lim_{\delta\to 1} [v^*(\delta)/c^*(\delta)] = v^{\max}/c^{\max}$.

Suppose that $v^{\max}/c^{\max} < (1 - \beta/\hat{\beta})/\beta$. Property (FF) from the proof of Lemma 3 implies that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ if $v^*(\delta)/c^*(\delta) < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$. Because $(1 - \beta/\hat{\beta})/\beta < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$, $\lim_{\delta\to 1} [v^*(\delta)/c^*(\delta)] = v^{\max}/c^{\max} < (1 - \beta/\hat{\beta})/\beta$ implies that there exists $\delta^* < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$ for all $\delta > \delta^*$.

Suppose that $v^{\max}/c^{\max} > (1 - \beta)/\beta$. To prove that there exists $\delta^* < 1$ such that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$ for all $\delta > \delta^*$, we prove that $v^{\max}/c^{\max} > (1 - \beta\delta)/(\beta\delta)$ implies that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$, from which the result follows because $v^{\max}/c^{\max} > (1 - \beta)/\beta$ implies that there exists $\delta^* < 1$ such that $v^{\max}/c^{\max} > (1 - \beta\delta)/(\beta\delta)$ for all $\delta > \delta^*$. Let $x^*(\beta,\delta,X) \equiv (c^*,v^*)$ and $x^*(\hat{\beta},\delta,X) \equiv (c',v')$. $x^*(\hat{\beta},\delta,X) = (c',v')$ implies that $\hat{\beta}\delta v'/(1 - \delta) - c' \geq \hat{\beta}\delta v^{\max}/(1 - \delta) - c^{\max}$, or $([\hat{\beta}\delta/(1 - \delta)]v'/c' - 1)c' \geq ([\hat{\beta}\delta/(1 - \delta)]v^{\max}/c^{\max} - 1)c^{\max}$. $c' \leq c^{\max}$ implies that $v'/c' > v^{\max}/c^{\max}$ and so $v'/c' > (1 - \beta\delta)/(\beta\delta)$, which implies that $\beta\delta v'/(1 - \delta) - c' \geq \beta\delta[\delta v'/(1 - \delta) - c']$. Then $x^*(\beta,\delta,X) = (c^*,v^*)$ implies that $\beta\delta v^*/(1 - \delta) - c^* \geq \beta\delta v'/(1 - \delta) - c' \geq \beta\delta[\delta v'/(1 - \delta) - c']$, in which case Property (EE) from the proof of Proposition 1 implies that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$.

(2) Let $(c^{\max\prime},v^{\max\prime})$ denote the maximal productive task in menu $X'(X)$. f increasing implies that $c^{\max\prime} = c^{\max}$ and $v^{\max\prime} = f(v^{\max})$. Hence, $v^{\max\prime}/c^{\max\prime} = f(v^{\max})/c^{\max} > (1 - \beta\delta)/(\beta\delta)$, in which case it follows from the proof of part (1) that $S^{pp}(\beta,\hat{\beta},\delta,X) \neq \{\mathbf{s}^{\varnothing}\}$.

QED

Proof of Proposition 8. $L(X) = 0$ implies that there exists $\underline{c} < \infty$ such that any $(c,v) \in X$ with $c > \underline{c}$ has $v/c < (1 - \beta/\hat{\beta})/\beta$. Given $(1 - \beta/\hat{\beta})/\beta < (1 - \beta\delta/\hat{\beta})/(\beta\delta)$ for any $\delta < 1$, and given Property (FF) from the proof of Lemma 3, if $x^*(\beta,\delta,X) = (c,v)$ for some $c > \underline{c}$ then $S^{pp}(\beta,\hat{\beta},\delta,X) = \{\mathbf{s}^{\varnothing}\}$. Defining $x^*(\beta,\delta,X) \equiv (c^*(\delta),v^*(\delta))$ as in the proof of Proposition 7, it remains to show that there exists $\delta^* < 1$ such that $c^*(\delta) > \underline{c}$ for all $\delta > \delta^*$. Extending the logic from the proof of Proposition 7 to the case where for every $(c,v) \in X$ there exists $(c',v') \in X$ with $c' > c$ and $v' > v$, we conclude that $\lim_{\delta \to 1} c^*(\delta) = \infty$. The result follows.

QED

Proof of Lemma 4. First note that $x^*(\beta,\delta,X) = (c^*,v^*)$ implies that $x^*(\beta,\delta,X(\eta)) = (c^*,v^* + \eta)$ and $x^*(\hat{\beta},\delta,X) = (c',v')$ implies that $x^*(\hat{\beta},\delta,X(\eta)) = (c',v' + \eta)$. Property (EE) from the proof of Proposition 1 then implies that for any η, $S^{pp}(\beta,\hat{\beta},\delta, X(\eta)) \neq \{\mathbf{s}^{\varnothing}\}$ if $\beta\delta(v^* + \eta)/(1 - \delta) - c^* \geq \beta\delta(\delta(v' + \eta)/(1 - \delta) - c')$, or $\eta \geq \delta v'/(1 - \delta) - v^*/(1 - \delta) + c^*/(\beta\delta) - c'$. The result follows.

QED

CORNELL UNIVERSITY
UNIVERSITY OF CALIFORNIA AT BERKELEY

REFERENCES

Ainslie, George, "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control," *Psychological Bulletin,* LXXXII (1975), 463–496.

——, "Derivation of 'Rational' Economic Behavior from Hyperbolic Discount Curves," *American Economic Review,* LXXXI (1991), 334–340.

——, *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person* (New York, NY: Cambridge University Press, 1992).

Ainslie, George, and Nick Haslam, "Self-Control," in George Loewenstein and Jon Elster, eds., *Choice over Time* (New York, NY: Russell Sage Foundation, 1992a).

Ainslie, George, and Nick Haslam, "Hyperbolic Discounting," in George Loewenstein and Jon Elster, eds., *Choice over Time* (New York, NY: Russell Sage Foundation, 1992b).

Akerlof, George, "Procrastination and Obedience," *American Economic Review,* LXXXI (1991), 1–19.

Choi, James, David Laibson, and Andrew Metrick, "The Effect of Negative Election in 401(k) Plans," Harvard University and University of Pennsylvania, 2000.

Fischer, Carolyn, "Read This Paper Even Later: Procrastination with Time-Inconsistent Preferences," University of Michigan, 1997.

Goldman, Steven, "Intertemporally Inconsistent Preferences and the Rate of Consumption," *Econometrica,* XLVII (1979), 621–626.

Laibson, David, "Essays in Hyperbolic Discounting," Ph.D. Thesis, Massachusetts Institute of Technology, 1994.

——, "Hyperbolic Discount Functions, Undersaving, and Savings Policy," NBER Working Paper No. 5635, 1996.

——, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics,* CXII (1997), 443–477.

Laibson, David, Andrea Repetto, and Jeremy Tobacman, "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity,* 1 (1998), 91–196.

Loewenstein, George, and Drazen Prelec, "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics,* CVII (1992), 573–597.

Madrian, Brigitte, and Dennis Shea, "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," University of Chicago, 2000.

O'Donoghue, Ted, and Matthew Rabin, "Doing It Now or Later," *American Economic Review,* LXXXIX (1999a), 103–124.
O'Donoghue, Ted, and Matthew Rabin, "Incentives for Procrastinators," *Quarterly Journal of Economics,* CXIV (1999b), 769–816.
O'Donoghue, Ted, and Matthew Rabin, "Procrastination in Preparing for Retirement," in Henry Aaron, ed., *Behavioral Dimensions of Retirement Economics* (Washington, DC, and New York: Brookings Institution Press and Russell Sage Foundation, 1999c).
Phelps, E. S., and Robert Pollak, "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies,* XXXV (1968), 185–199.
Pollak, Robert, "Consistent Planning," *Review of Economic Studies,* XXXV (1968), 201–208.
Prelec, Drazen, "Decreasing Impatience: Definition and Consequences," Harvard Business School, 1989.
Strotz, Robert, "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies,* XXIII (1956), 165–180.
Thaler, Richard, "Some Empirical Evidence on Dynamic Inconsistency," in Richard Thaler, ed., *Quasi Rational Economics* (New York, NY: Russell Sage Foundation, 1991).
Thaler, Richard, and George Loewenstein, "Intertemporal Choice," in Richard Thaler, ed., *The Winner's Curse: Paradoxes and Anomalies of Economic Life* (New York, NY: Free Press, 1992).
Voltaire, Jean Francois Marie Arouet de, "L'art dramatique," in *Œuvres completes de Voltaire,* Volume 17 (Paris: Garnier Frères, 1878).