Timing and Self Control

Drew Fudenberg* and David K. Levine**

First Version: August 5, 2009
This Version: May 6, 2011

Abstract: The standard dual-self model of self control, with a shorter-run self who cares only about the current period, is excessively sensitive to the timing of decisions and to the interpolation of additional “no-action” time periods in between the dates when decisions are made. We show that when shorter-run self is not completely myopic this excess sensitivity goes away. To accommodate the combination of short time periods and convex costs of self control we introduce a cognitive resource variable that tracks how the control cost depends on the self control that has been used in the recent past. We consider models with both linear and convex control costs, illustrating the theory through a series of examples. We examine when opportunities to consume will be avoided or delayed, and we consider the way in which the marginal interest declines with delay.

* Department of Economics, Harvard University
** Department of Economics, Washington University in St. Louis

This paper was presented as the Fisher-Schultz Lecture at the 2010 World Congress of the Econometric Society in Shanghai. We thank Attila Ambrus, Yuichiro Kamada, Jawaad Noor, Alex Peysakhovich, Tomasz Strzalecki, and Dmitry Taubinsky for helpful comments, and NSF grants SES-0646816, SES-0851315, and SES-0954162 for financial assistance.
1. Introduction

Models of long-run planning and shorter-run\(^1\) impulsive selves suppose that a single patient self makes decisions in each period to maximize the discounted sum of utility net of a cost of self control, where this cost depends on the temptations faced by the impatient impulsive self. These models provide a quantitative explanation of a wide variety of “behavioral” paradoxes, including the Rabin paradox (small stakes risk aversion), the Allais paradox, preferences for commitment in menu choice, violations of the weak axiom of revealed preference, non-exponential discounting, and the effect of cognitive load on decision making and reversals due to probabilistic rewards. However, these models, like the quasi-hyperbolic discounting model,\(^2\) have two implications about the role of timing that are at odds with the data. First, because these models have a fixed horizon for the shorter-run self, they cannot explain overwhelming evidence that the length of delay has a continuous impact on decisions, and they make implausible predictions about the value of a commitment that avoids temptation: There is no value for a commitment that must be made the same period that the temptation is faced, while commitment “one period before” the temptation arises can be highly valued, regardless of the length of a time period. Second, past work has identified the importance of allowing the cost of self control to be a convex as opposed to a linear function of foregone utility, but it has assumed that the cost function in each period is independent of self control used in the recent past. As we point out here, that model has the implausible implication that making decisions simultaneously is different than making them with an arbitrarily small delay. The new version of the dual-self model we propose here addresses both of these issues.

To account for the continuous effect of delay, and to explore the implications of the timing of decisions, we propose that temptation comes from a shorter-run self or selves who are not completely myopic but who value future utility less than the long-run self does, either because there is a succession of shorter-run selves with random lifetimes, or a single shorter-run self whose discount factor is lower than that of the long-run self.

\(^1\) As we are dealing with “short-run selves” who live more than one period we adopt the suggestions of Rajiv Sarin and refer to “shorter-run selves.”

\(^2\) Strotz [1956], Phelps and Polak [1968], Laibson [1997], O'Donoghue and Rabin [1999].
This lets us maintain the strength, simplicity, and applicability of the dual-self model. It also lets us model cases where agents are tempted by future consumption, as in Noor [2007], [2010], and explain why this temptation is most significant with respect to payoffs in the near future.

The key modeling question in extending the dual-self model to non-myopic shorter-run selves is the specification of the cost of self control. In the one-period model this cost depends on the amount of utility the shorter-run self foregoes in the current period. When temptation comes from selves who give non-zero weight to future payoffs expectations about these future payoffs matter; we propose that the control cost of implementing a given action depends on how much that action lowers the highest possible average present value the shorter-run self could obtain from the current period on. This specification is consistent with the interpretation that the shorter-run selves are strategically naïve and evaluate foregone utility assuming that no self control will be used in the future; the web appendix points out some of the complications of an alternative specification.

We begin our analysis with the case where the cost of self control is linear in the foregone value. This is the simplest version of the model, and the one closest to discounted expected utility, as it is consistent with both the independence axiom for choices over lotteries and the weak axiom of revealed preference. Our first application considers the decision of whether to accept or reject a “simple temptation” that gives an initial positive payoff followed by a negative payoff in future periods. We point out that the agent may prefer to resist a simple temptation when it is a once-and-for all choice, and yet prefer to give in when that temptation must be faced in every period unless and until it is accepted; we relate this to the effect of “bundling” of decisions noted by Ainslie [2001] and Kirby and Guatsello [2001]. We then show that the value of a commitment to avoid temptation can be non-monotone in the discount factor of the shorter-run selves. In Example 3 we show that the model explains the fact that the ratio between payment in period \( t \) and payment in period \( t + 1 \) that makes subjects indifferent is typically increasing in \( t \) and not constant; we focus on data from Myerson and Green [1995] but there have been many similar studies.

Although linear costs are a convenient first cut at self-control problems, there is considerable evidence that the costs are often convex, so that it is more than twice as hard
to resist twice the temptation. We therefore extend the model to allow for convex costs. Examples 4 and 5 point out two implications of convex costs for agents faced with simple temptations: First of all, an agent is more likely to resist a temptation that has low probability of being realized than one whose payoff stream is certain. Second, an agent who is faced with two simultaneous simple temptations may choose to accept one of them, even though he would reject both if they were presented in different periods.

This observation suggests that when costs are convex and time periods are short, we expect that the non-linearity of control costs should “spill over” from a decision in one period to a subsequent decision soon afterwards, so that making two decisions in rapid succession is similar to making the two decisions simultaneously. Moreover, since the length of the time period in decision problems is an artificial construct, we want the model to apply to cases where the “time periods” are very short, with decisions made in only a few of them. Adding such “intermediate” no-action periods makes no difference in classic rational models, but can have counterintuitive implications in models of self control.

To capture the effect of changing the period length when there are convex costs, we suppose that self control uses cognitive resources, and that these resources are a stock that can be depleted over short time intervals, as argued by Muraven et al [1998] and modeled by Ozdenoren et al [2009]. From this perspective, the simpler model of the previous sections corresponds to cases where the stock of cognitive resources completely replenishes from one period to the next.

The reason for introducing a stock of cognitive resources is to track variations in the marginal cost of self control and account for the way that using self control in one period can alter the self control cost and decision in a subsequent one. To check that this is all it is doing, we first show in Theorem 3 that when there is a single decision, the stock of cognitive resources is irrelevant, as the agent’s decision will be the same as in a “state-free” model with the appropriately defined cost function. We then show in Theorem 4 that when the marginal cost of self control is constant, the decisions made by an agent with a stock of cognitive resources that partially replenishes over time are again the same as those made in an associated model without cognitive resources. However,

---

3 In multiplayer games the period length models how long it takes one player to respond to the play of another.
when the agent makes multiple decisions and the marginal cost of self control varies, the equivalence with the state-free model fails, precisely because of the link between current decisions and the marginal cost of self control in future periods.

In general, there are three possible sources of non-linearity in the model, any of which can cause variations in the marginal cost of self control: the way the stock of cognitive resources is depleted by using self control; the way the stock is replenished over time; and way the stock provides benefits. Since cognitive resources are not observed directly and have no natural units, there are many equivalent representations of the same preferences, so it is not possible to identify which part of the model is non-linear. Indeed, as Theorem 6 shows, if there is any replenishment at all, it is without loss of generality to specify linear replenishment and lodge all of the non-linearity in the depletion and benefit functions.

After exploring the general properties of the cognitive resources model, we consider a number of examples with linear depletion and linear (or no) replenishment. Example 7 revisits the example of an agent with convex costs facing two temptations in a row, and shows that when resources replenish linearly the agent makes the same choice whether the two decisions are made exactly simultaneously or in rapid succession. Example 8 shows that when the marginal benefits of resources are concave (so the cost of control is convex) it may be optimal to resist a persistent temptation for a while and then take it, a conclusion that is impossible in the stationary model without a stock variable. Example 9 builds on this by adding the option to pay a fee and permanently avoid the temptation. We show that it may be optimal to resist for a while and then pay the fee, which is consistent with the findings of a suggestive recent experiment of Houser et al [2010]. Examples 7 and 8 simplify by assuming no replenishment of resources at all, which is unrealistic but makes it easier to highlight the logic of the argument. Example 10 re-examines persistent temptation with a general depletion function to highlight how the depletion and benefit function interact to determine whether the agent will resist for a while before giving in. Finally, Example 11 shows the issues involved in relaxing our assumption that the “willpower technology” is fixed and cannot be changed by the agent.

Some past work has used the device of random long-run player lifetimes to explain behavioral anomalies: Dasgupta and Maskin [2005] show that uncertain long-run player lifetimes can lead to hyperbolic discounting. Halevy [2008] develops a model
where a single long-run self faces a stopping (or death) risk that is modified by a convex “transformation function” and so is distinct from the agent’s pure time preference. He uses this to explain Keren and Roelofsma’s [1995] data, which shows that “present-biased” preference reversals are much less frequent when both the present and future rewards are uncertain. Epper, Fehr-Duda, and Bruhin [2009] use a similar idea of distorted survival weights to explain present bias as a consequence of prospect theory.

Noor [2007], [2011] develops axioms for infinite-horizon choice problems in which the agent can be tempted by future consumption. His model, like that of Gul and Pesendorfer [2001], is more general than ours in relating the temptation values to the objectives of the long-run player, but less general in imposing the independence axiom; because of this assumption, only the linear-cost model in Section 3 is compatible with his framework. Noor does not investigate how “temptation by the future” depends on the real time between the two periods of his model; his main goal is to show that there can be a self-control problem despite the fact that future temptation results in little demand for commitment, as in our analysis of menu choices in Example 1. In the context of two-period models, Noor and Takeoka [2010a] develop axioms for choices on menus that allow convex costs of self control, and Noor and Takeoka [2010b] develop a more restrictive representation that they extend to multiple periods and temptation by future payoffs. Brocas and Carrillo [2008] explain the covariance of effort and consumption by assuming the long-run self has incomplete information on the shorter-run self’s cost of effort, and Chatterjee and Samuelson [2009] and Dekel, Lipman and Rustichini [2009] axiomatize cases where second period preferences are stochastic and can depend on the first period choice of menu. For the infinite horizon problem Gul and Pesendorfer [2004] develop a recursive extension of their [2001] axioms, including the independence axiom.

2. The Decision Problem

In dual-self models, the agent acts to maximize expected discounted utility subject to a cost of self control that is derived from the preferences of a more impulsive short run self. In this paper we take this control cost as exogenous; the Web Appendix explains how the maximization problem we consider corresponds to the equilibrium of a game.

---

4 Our [2010] paper explains the same data as a consequence of a convex cost of self control.
5 These two papers differ in how the long-run self views the possible second period preferences.
between the agent’s long-run self (a “planner”) and a sequence of shorter-run “temptation selves” (“doers”). To facilitate the exposition and also the comparison of the model with previous work, we will use a discrete-time model, with periods $n = 1, 2, \ldots$. We denote the agent’s discount factor by $\delta$, and suppose that the discount factor of the shorter-run self is $\delta \mu$, where $\mu \in [0, 1)$ can also be interpreted as the survival probability of the current shorter-run selves.

We will frequently be interested in how the solution to the model varies with the period length, which we would like to distinguish from the real time between decisions. To do this we let the period length be $\tau$ units of calendar time, and suppose that $\delta = \exp(-\rho \tau)$ and $\mu = \exp(-\eta \tau)$; for small $\tau$ we will use the approximations $\mu \approx 1 - \eta \tau$, $\delta \approx 1 - \rho \tau$.

The space of states, denoted $Y$, is a measurable subset of a finite-dimensional Euclidean space, as is the space of actions $A$. For each state there is a measurable subset of feasible actions $A(y_n) \subseteq A$ and at least one measurable map $a : Y \rightarrow A$ that satisfies $a(y_n) \in A(y_n)$. Dynamics are Markovian: They are given by probability distributions $\pi(y_{n-1}, a_{n-1})[dy_n]$ over states next period, conditioned by the current state and action, that are measurable functions of $y_{n-1}, a_{n-1}$. In applications, the state can take on many different roles: it can correspond to direct and indirect constraints on feasible actions (such as the presence of tempting consumption opportunities and the agent’s wealth) and also to the agent’s past consumption or other influences on the current period’s utility function.\(^6\)

Each period’s action is taken after that period’s state is known, so the history of play at period $n$ is $h_n = (y_1, a_1, \ldots, y_{n-1}, a_{n-1}, y_n)$; the initial history $h_1 = y_1$ is exogenously given. A strategy or plan for the long-run self is then a measurable map $a$ from histories to actions, so that for each history $h_n$ the strategy specifies an action in $a(h_n) \in A(y_n)$.

The shorter-run self (or selves) have utility $u(y_n, a_n)$ in period $n$ if the action $a_n$ is taken in the state $y_n$. We will work with average present values, so that we hold $u(y_n, a_n)$ fixed as we vary the length $\tau$ of the time period. The objective of the long-run

\(^6\) For example, the model is compatible with the “rational addiction” preference structure of Becker and Murphy [1988] and with short-run consumption commitments, as in the Gabaix and Laibson [2002] explanation of the equity premium puzzle.
player is the average present value of these shorter-run self utilities, minus a cost of self control that is defined with reference to the maximum possible average present value for the shorter-run self. Let $E_{a,h_n}$ be the conditional expectation generated by the plan $a$ conditional on the history $h_n$. The expected average present value of the shorter-run self from period $n$ on under $a$ is given by

$$U(h_n,a) \equiv E_{a,h_n} (1 - \delta \mu) \sum_{\ell=0}^{\infty} (\delta \mu)^\ell u(y_{n+\ell}, a_{n+\ell}),$$

or equivalently

$$U(h_n,a) \equiv E_{a,h_n} (1 - e^{-(\rho+\eta)\tau}) \sum_{\ell=0}^{\infty} e^{-(\rho+\eta)\ell \tau} u(y_{n+\ell}, a_{n+\ell}).$$

In order to focus on the application of the model and not standard technical details we directly impose the following assumption.

**Assumption SR0:**

$$E_{a,h_n} (1 - \delta \mu) \sum_{\ell=0}^{\infty} (\delta \mu)^\ell u(y_{n+\ell}, a_{n+\ell}) = (1 - \delta \mu) \sum_{\ell=0}^{\infty} (\delta \mu)^\ell E_{a,h_n} u(y_{n+\ell}, a_{n+\ell})$$

(the expectation and sum operators can be interchanged) and $U(h_n,a)$ has a maximum for each $n$ and $h_n$.

Because the problem of the shorter-run self is Markov, this maximized value only depends on the state:

**Theorem 1:** $\max_a U(h_n,a) = \overline{U}(y_n)$

Our earlier work [2006, 2010] assumed that the cost of self control depends on the amount of utility foregone by the shorter-run self, which is the difference between the maximum possible utility in the current period and the current-period utility the shorter-run self actually receives. When shorter-run selves live more than one period, we must specify how the cost of self control takes into account the effect of current actions on future payoffs. To do this we suppose that the level used as a benchmark is the highest value that the shorter-run player could hope to receive.

Specifically, we call $\overline{U}(y_n)$ the temptation value in period $n$ starting at state $y_n$. The foregone value is then

$$\Delta(y_n,a_n) = \overline{U}(y_n) - \left(1 - \delta \mu u(y_n,a_n) + \delta \mu \int_y \overline{U}(y_{n+1}) \pi(y_{n+1} \mid y_n,a_n) \, dy_{n+1}\right)$$
The foregone value is *recursive* in the sense that it depends on the future only through the future temptation value, and attributes control costs to each action as it occurs as opposed to entire contingent plans. The idea is that the current cost of self control depends on the foregone value in the current period, which can be non-zero either because the current choice of action $a_n$ lowers $u(y_n, a_n)$ or because it changes the distribution of future states $\pi(y_{n+1} | y_n, a_n)$ and thus the temptation value from the next period on. Future actions that may lower the shorter-run self’s value do not incur a current cost, instead they incur costs at the time at which they are taken. Note in particular that there is no foregone value to postponing a decision, provided that the set of utility possibilities in the next period is the same as in this one, though as we will see the cost of implementing it may change.

One interpretation of the foregone value is that the term

$$\delta \mu \int_Y U(y_{n+1}) \pi(y_{n+1} | y_n, a_n) [dy_{n+1}]$$

is the shorter-run self’s prediction of the expected continuation payoff, and that shorter-run self predicts that no self control will be used in the future. Under this interpretation the shorter-run self is strategically naïve and does not anticipate that today’s actions can change the amount of self control that will be used in the future.

Notice that by the principle of optimality any plan that solves $\max_a U(h_n, a)$ must also solve the dynamic programming problem

$$\max_{a_n} (1 - \delta \mu) u(y_n, a_n) + \delta \mu \int_Y U(y_{n+1}) \pi(y_{n+1} | y_n, a_n) [dy_{n+1}].$$

Thus we have

**Theorem 2:** $\Delta(y_n, a_n) \geq 0$, and if $\hat{a} \in \arg\max_a U(h_n, a)$ then $\Delta(y_n, \hat{a}_n) = 0$.

---

7 This is the same specification of temptation value and foregone value as in Noor [2007]. The web appendix discusses an alternative formulation in which the shorter-run selves have rational expectations about future self control, so each period’s control cost depends on the difference between the best expected present value available to a shorter-run self born in that period and the value that will actually be received. We show that when the cost of self control is linear in the foregone value, as in (1), the two formulations are equivalent, but with non-linear costs the recursive formulation has more plausible implications.

8 This is the same specification of temptation value as in Noor [2007]. Depending on the cost function, other interpretations are sometimes possible as well, for example in the case of linear costs (defined below), the definition of temptation utility is consistent with perfect foresight. But the “naïve” interpretation is valid regardless of the cost function.
The key element of the theory of self-control is the specification of the cost of exerting self-control. We study the linear case first.

**3. Linear Cost of Self Control**

We start our analysis with a particularly simple specification of the cost of self control. We suppose that the cost of self control is $\Gamma \Delta(y_n,a_n)$, so it is linear in the foregone value $\Delta$, and that the scalar constant $\Gamma \geq 0$ is independent of the state and the period length.

The case of linear self-control costs has been the most widely studied. This type of self-control model satisfies the Gul-Pesendorfer axioms, including the independence axiom. Moreover, while nonlinear costs are important in many applications, many insights still arise in the linear case. We examine increasing marginal cost of self control in subsequent sections, along with the idea that willpower is a stock so that exercising self control can increase the control cost over the next few periods if periods are short.

The agent’s objective function is

$$V(h_n,a) \equiv E_{a,h_n} \sum_{\ell=0}^{\infty} \delta^\ell \left( (1-\delta)u(y_{n+\ell},a_{n+\ell}) - \Gamma \Delta(y_{n+\ell},a_{n+\ell}) \right);$$

this is the expected average present value of the per-period utility $u$ minus the discounted sum of control costs.

As in the case of the shorter-run decision problem, we assume existence of a maximum:

**Assumption 0:** $V(h_n,a)$ has a maximum for each $n,h_n$.

Moreover, because this is a Markov decision problem, there is an optimal plan in which the choice of action depends only on the current state $y_n$.

The objective function (1) reduces to the linear-costs version of Fudenberg and Levine [2006], [2010]) when $\mu = 0$. It is also a special case of the functional forms considered in Noor [2007], [2010]; unlike Noor, we assume that the long run and short run selves have the same underlying per-period utility function, and specify how preferences change with the period length.
It is important to note that the term $\Gamma \Delta$ in (1) is not normalized by $1 - \delta$; this is because here we model control cost as an impulse, while we treat utility as a flow. To understand the implications of this, consider the decision to take an action that lowers utility in every period by 1. This action lowers the shorter-run self’s value by 1 and so has foregone value of $\Delta = 1$ independent of period length. Thus its control cost is $\Gamma$ regardless of period length, which seems like the right conclusion. In contrast if the $\Gamma \Delta$ terms were multiplied by $(1 - \delta)$ the cost of implementing this action would go to zero as the periods become short. However, an action that lowers utility by 1 for a single period but has no impact on future utilities has foregone value $\Delta = 1 - \delta \mu$ and a cost of $\Gamma(1 - \delta \mu)$, which does become very small when periods are short.

One implication of this model is that long-term commitments will be more attractive than a series of short-term ones: It is cheaper to resist future temptations now than to resist them as they arise. To see the intuition for this, suppose the long-run player undertakes an infinite sequence of actions each of which lowers utility by 1 in the current period only. The overall cost of this series of actions is $\Gamma(1 - \delta \mu)/(1 - \delta)$, which is greater than the cost $\Gamma$ of a single action that lowers utility by 1 in every period. Moreover, the cost of committing now to forego 1 util in every period from $N$ on is $\Gamma(\delta \mu)^{N-1}$ and in particular is strictly decreasing in $N$, except in the case $\mu = 0$, where the shorter-run self views all future periods $N > 1$ as equally far away. We illustrate these implications in Example 1 below.

Note that the difference between the long-term and short-term commitments is most extreme in the case $\mu = 0$, where the long-term commitment is no more costly than any of the one-period delays. The difference diminishes as $\mu \to 1$, holding other parameters constant; though when $\mu$ is large the preferences of the two selves are closely aligned so the long-run self is has little reason to exert self control.

Finally, note that sending the time period $\tau$ to 0 sends $\mu \to 1$ but changes $\delta$ as well: Lowering utility by 1 forever starting immediately still costs $\Gamma$, lowering utility by 1 forever starting at real time $s = N / \tau$ costs $\Gamma \exp(-(\rho + \eta)s)$, and lowering utility by 1 period-by-period immediately costs $\Gamma(\rho + \eta) / \rho$. Thus the difference between the

---

9 When we introduce cognitive resources, the control cost will the reduction in the value of the associated flow.
long-term commitment and the series of short-term ones stems not from the period length but from the greater impatience of the shorter-run self.

Simple Temptations

Several of our examples will use as a building block what we call a *simple temptation*, which is a choice between either utility 0 in every period or a flow of $u_g > 0$ that is received for a number of periods $N$, with $-u_b < 0$ forever after. This choice represents a stereotypical conflict between short run and long run preferences that is easy to adapt to varying period lengths and to embed in more complicated decision problems.

The average present values $S$ for the shorter-run self and $P$ for the long-run self of this stream are $S = (1 - (\delta \mu)^N) u_g - (\delta \mu)^N u_b$ and $P = (1 - \delta^N) u_g - \delta^N u_b$. Our interest lies in the case $S > 0, P < 0$ so that the shorter-run self would like to take the temptation while the long-run self would prefer to reject it. This conflict arises because the short run self discounts future periods using discount factor $\delta \mu < \delta$, and will not be present if we send $\mu$ to 1 holding all other parameters fixed. However, the reason for interest in the case of $\mu$ near 1 is that it corresponds to very short periods. To analyze this case we fix the calendar length of time $T$ for which the favorable flow lasts, so that $N = T / \tau$. Then when $N$ is an integer we have $S = (1 - e^{-(\rho + \eta)T}) u_g - e^{-(\rho + \eta)T} u_b$, and $P = (1 - e^{-\rho T}) u_g - e^{-\rho T} u_b$, independent of $\tau$, even though $\mu = e^{-\eta \tau} \to 1$.

Example 1: Simple and Persistent Temptations with Linear Cost

To begin, consider a choice between accepting and rejecting a simple temptation in the first period, with no other choices to be made. Then the temptation utility is $S$, so the cost of resisting temptation is $-\Gamma S$, thus temptation will be resisted if $P < -\Gamma S$. Next note that if the decision can be made at date 1 about whether to accept or reject the temptation in period $n$, the cost of resisting is $-\Gamma (\delta \mu)^n S$, so the temptation will be resisted if $\delta^n P < -\Gamma (\delta \mu)^n S$ or $P < -\Gamma \mu^n S$; thus as the decision concerns events further in the future it becomes easier to resist.

Next suppose that the temptation is persistent: Once the agent consumes the substance it is gone, but if he does not consume, the substance is still there the next
period so that the same choice is faced again in the next period.\textsuperscript{10} This model describes for example the temptation to consume a durable good such as a bottle of wine.\textsuperscript{11} If the temptation is available, the best possible value for the shorter-run self is $S$, and the best continuation if the temptation is resisted is $\delta \mu S$. Thus so the foregone value from resisting is $\Delta = (1 - \delta \mu)S$, so resisting costs $\Gamma(1 - \delta \mu)S$ each period, and resisting is optimal if $P(1 - \delta) < -\Gamma(1 - \delta \mu)S$. Consequently the persistent temptation is “harder to resist” than the simple one, and when $(1 - \delta \mu)/(1 - \delta) > |P|/\Gamma S > 1$, the agent would choose to give in to a persistent temptation but resist a simple one. This condition gets increasingly difficult to satisfy as $\mu \to 1$ holding all other parameters fixed, which corresponds to sending the birth parameter $\eta$ to 0.

Since the main reason for large $\mu$ is that periods are short, it is interesting to study the agent’s choice in the limit of short time periods. Here the agent gives in to the persistent temptation but resists the simple one when $(\rho + \eta)/\rho > |P|/\Gamma S > 1$ or $(\rho + \eta)\Gamma S > \rho |P| > \rho \Gamma S$. This last equation has a simple interpretation: $\rho |P|$ is the value of postponing the negative payoff $P$ for an interval $dt$, $(\rho + \eta)\Gamma S$ is the flow cost of resisting the persistent temptation, and $r \Gamma S$ is the flow or average utility resulting from paying the one-time cost of $\Gamma S$ to permanently avoid the temptation.

If declining the temptation in period $n$ means that it will not arise again until period $n + \ell$, the situation is intermediate between a persistent temptation ($\ell = 1$) and a simple one ($\ell = \infty$). Then the best continuation value if the option is resisted is $(\delta \mu)^\ell S$, so $\Delta = (1 - (\delta \mu)^\ell)S$, and resisting forever costs

$$\Gamma \sum_{n=0}^{\infty} \delta^{\ell n} (1 - (\delta \mu)^\ell)S = \frac{\Gamma(1 - (\delta \mu)^\ell)S}{1 - \delta^\ell},$$

so resisting is optimal if $|P|/\Gamma S > (1 - (\delta \mu)^\ell)/(1 - \delta^\ell)$. Consequently resisting is more attractive when the temptation can be avoided for longer, and the decision of

\textsuperscript{10} We can formally model this by assuming that there are two states $Y = \{0, 1\}$, where $y_n = 0$ means that the temptation is not available, and $y_n = 1$ means that it is. In the state $y_n = 0$ no action is possible, $A(y_n) = \{0\}$; in the state $y_n = 1$ the space of actions is $A(y_n) = \{0, 1\}$ where 0 means to resist the temptation and 1 means to give in to the temptation. The transition probabilities in state $y_n = 0$ place probability 1 of remaining in that state, $\pi(0 \mid y_n, a_n) = 1$, while in state $y_n = 1$ the transition probability depends on the action taken: $\pi(0 \mid y_n, 1) = 1$ so that if the action is taken, the temptation is off the table, and $\pi(0 \mid y_n, 0) = 0$ so that if the temptation is resisted it remains for next period.

\textsuperscript{11} We assume here that the consumption option is all or none, perhaps the wine will spoil once opened.
whether to take at once or resist forever is monotone in $\ell$: There is some $\bar{\ell}$ (possibly 0 or infinity) such that the optimum is to take at once if $\ell < \bar{\ell}$ and resist forever if $\ell > \bar{\ell}$.

Intuitively, this is because the shorter-run self is much less concerned about far-off events than the long-run player is, so the gap between the benefit of delay and the cost of buying off the shorter-run player is increasing in the delay length. If the temptation will arrive fairly soon, there is not much point in trying to avoid it by making a commitment now, as the temptation is already important for the short run so committing to reject it is about as costly as waiting and resisting it when it arrives. On the other hand, it is less costly to commit now to resist temptations that will arrive in the distant future. This is related to Noor’s [2007], [2011] point that agents may be tempted by future consumption; the additional structure of our model lets us explain how this effect depends on the real time between the various decisions as opposed to the period length per se.

Next, consider an initial-period choice of whether to accept or reject $K$ simple temptations, the first one in at real time 0, the second at real time period $T$, the third at $2T$ so forth. If the agent is close to indifferent about whether to take the first temptation he will strictly prefer to reject the second if that decision could be made in period 1. For this reason an agent who would accept the simple temptations may choose to reject the “bundle” of them. This is common behavior, as seen for example in the experiments of Kirby and Guatsello [2001] on the “bundling” of decisions. Note that the once-and-for all decision to decline a simple temptation can be seen as a “bundle” of all of the “decline today” decisions, and in each case the agent prefers the bundle for the same reason, namely that he is less tempted by future rewards.

**Example 2: Non-monotonicity of the Value of Commitment**

Here we show how the value of a commitment to avoid temptation can be non-monotone in $\eta$. To do this we consider a menu choice when some of the choices

---

12 See also chapter 5 of Ainslie [2001]. Both Ainslie and Kirby and Guatsello report that merely telling subjects they will face the same decision in the future changes choices as well, which the stationary model of this section cannot explain. Note though that Kirby and Guatsello report a much smaller impact of this “suggested linking” than of actual linking, and that the instructions they used for suggested linking told subjects “the choice you make now is the best indication of how you will choose every time,” which may have induced a spurious effect.

13 Section W1 of the web appendix examines a different sort of non-monotonicity: those who are willing to pay to avoid a temptation can have intermediate utilities from choosing it, as those with very low values
correspond to future commitments. Specifically, suppose that at time 0 the agent’s action is to pick a menu from the list \(\{0\}, \{0,1\}\), where \(0\) means that no future decision are possible, and \(0, 1\) means that at time 1 the agent will face a simple temptation of the form described above. Because the menu \(0,1\) simply postpones the decision, choosing it will have no foregone value and no control cost, while the foregone value associated with \(0\) depends on the preferences of the shorter-run self. We specialize to the case \(u_q = u_b = 1\), \(T = 1\), and assume that \(\Gamma = 3\) and \(\rho = \log(3/2)\). Thus \(P = -1/3\) and \(S = 1 - (4/3)\exp(-\eta)\); note that \(S\) is positive or negative depending on whether \(\eta\) is more or less than \(\log(4/3)\). We define \(\overline{\eta}\) as the unique solution of \(9\exp(-\overline{\eta}) + 6\exp(-2\overline{\eta}) = 1\), and note \(\overline{\eta} > \log(4/3)\).\(^{14}\)

We will show that for \(\eta < \log(4/3)\), the agent is indifferent between \(0\) and \(0,1\), while for \(\log(3/2) < \eta < \overline{\eta}\) the agent strictly prefers \(0,1\) and for \(\eta > \overline{\eta}\) the agent strictly prefers \(0\). The reason for this non-monotonicity is this: for very small \(\eta\) the long run and short run selves agree on whether to take this temptation so no self control is needed to resist it. For larger values of \(\eta\), resisting requires costly self control, and when \(\eta > \log(3/2)\) the self control is so costly that when the decision is made at time 1 (that is, “without precommitment”) it is optimal to give in. Moreover, when \(\eta\) is just a little above \(\log(3/2)\), it is not optimal to avoid the temptation by choosing the menu \(0\) at time 0, because the time interval between 0 and 1 is short enough that choosing the commitment menu requires substantial self control. Finally when \(\eta\) is sufficiently large, the short run self is not tempted by time-1 consumption, so choosing \(0\) is essentially costless.\(^{15}\)

The agent’s choice at time 1 given \(0,1\) was chosen at time 0:

- If \(\eta < \log(4/3)\) then as \(S < 0\) there is no self control problem at time 1; here the optimum is to resist, with payoff 0.
- If \(\eta > \log 4/3\), so that \(S > 0\), then \(S\) is the temptation value in period 1. Resisting gives payoff \(-\Gamma S = 4\exp(-\eta) - 3\), so it is optimal to take the

---

\(^{14}\) This follows from \(9\exp(-\log(4/3)) + 6\exp(-2\log(4/3)) = (27/4) + 6\exp(-2\log(3/2)) > 1\).

\(^{15}\) May find it easy to resist without commitment, while those with high values will have a correspondingly high control cost for choosing the commitment.
temptation if \( P > -\Gamma S \), that is, \( \eta > \log(3/2) \). For \( \log(3/2) > \eta \geq \log(4/3) \) it is optimal to resist.

**The agent’s choice at time 0:**

- If \( \eta < \log(4/3) \) then the shorter-run self’s temptation payoff is 0, so the long-run self is indifferent between choosing \( \{0,1\} \) (and subsequently resisting) or choosing \( \{0\} \).
- If \( \eta > \log(4/3) \) then the shorter-run self’s value is maximized by choosing the menu \( \{0,1\} \) then taking in period 1. The corresponding temptation value in period 0 is \( V = \exp(-\rho)S = (2/3)(\exp(-\eta) - 4/3 \exp(-2\eta)) \). The long-run self’s payoff to \( \{0\} \) is thus \( -\Gamma V = -2(\exp(-\eta) - 4/3 \exp(-2\eta)) \). If the menu \( \{0,1\} \) is chosen it is optimal to take in period 1, resulting in the payoff is \( \exp(-\rho)P = -2/9 \). For \( -2(\exp(-\eta) - 4/3 \exp(-2\eta)) > -2/9 \) or \( \eta > \bar{\eta} \), the long-run self strictly prefers \( \{0\} \) and for \( \bar{\eta} > \eta > \log(4/3) \) the long-run self strictly prefers \( \{0,1\} \).

### Example 3: Declining Marginal Interest Rates

Many studies have agents about their trade-offs between money today and money at various points in the future; see for example the survey in Frederick, Lowenstein and Shane [2002], and the early work of Thaler [1981], who asked subjects to state the dollar amount of they would require at three different future dates to make them indifferent to receiving $15 now. Myerson and Green [1995] used a larger set of delays, allowing a better picture of the delay/interest rate trade-off. Subjects were asked to state how much hypothetical money \( c_t \) they would need right now to make them indifferent to receiving a hypothetical $1,000 after a delay of length \( t \). Using their response, we can define an incremental interest rate between successive queried times \( t_i, t_{i-1} \) as

\[
\rho_i = \log(c_t / c_{t_{i-1}}) / (t_i - t_{i-1}).
\]

---

15 If there is a small fee \( \varepsilon \) for choosing the flexible menu \( \{0,1\} \) then there is a strict failure of non-monotonicity: for \( \eta < \log(4 / 3) \) the menu \( \{0\} \) is strictly preferred, while for \( \log(3 / 2) + \varepsilon < \eta < \bar{\eta} - \varepsilon \) \( \{0,1\} \) is strictly preferred, and for \( \eta > \bar{\eta} \) \( \{0\} \) is again strictly preferred.
If agents treat cash payments like immediate consumption and the utility of consumption is linear, then in both the quasi-hyperbolic discounting model and the long-run/shorter-run self model with a fixed period length these incremental interest rates should be time invariant after the first period. From the Myerson and Green data, the rates are

<table>
<thead>
<tr>
<th>months ($t_i$)</th>
<th>incremental interest rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.23</td>
<td>132%</td>
</tr>
<tr>
<td>1</td>
<td>82.1%</td>
</tr>
<tr>
<td>6</td>
<td>40.9%</td>
</tr>
<tr>
<td>12</td>
<td>42.7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Months ($t_i$)</th>
<th>Incremental interest rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>26.0%</td>
</tr>
<tr>
<td>60</td>
<td>8.0%</td>
</tr>
<tr>
<td>120</td>
<td>9.4%</td>
</tr>
<tr>
<td>300</td>
<td>6.6%</td>
</tr>
</tbody>
</table>

As can be seen, the decline is gradual rather than a one-time decline after the first period.

We will now show how non-myopic short run selves generate this sort of declining interest rate. We take utility to be linear in consumption and without further loss of generality we set $u(c) = c$. We will compute the amount of consumption $c_n$ that makes the long-run self indifferent between a unit of consumption at time 1 and $c_n$ units in period $n$; we will then use this to compute effective marginal interest rates on consumption. (In the Web Appendix we extend this to the interest rate used at time 1 to discount between any two periods $n$ and $\ell$, which is closer to the long-run player’s rate of time preference $\rho$ because future consumption is less tempting.) Observe that if the...

---

16 Andersen et al [2008] find evidence of a smaller but still monotone gradual decline of interest rate with delay when real financial incentives are provided in the form of lotteries and adjustments are made for curvature; Benhabib et al [2010] also find evidence of a gradual monotone decline using (small) cash rewards. Since Keren and Roelsofsma [1995] have already found (in hypothetical experiments) that agents’ intertemporal choices are closer to geometric discounting when rewards are stochastic, it is not clear how much of the difference between Andersen et al’s findings and those of Myerson and Green are due to the fact that subjects were paid.

17 Myerson and Green asked subjects about cash payoffs as opposed to consumption. Both dual-self and quasi-hyperbolic models (need additional structure to explain why subjects (who presumably save) also view cash payoffs as tempting, and why the small-stakes risk aversion seen in lab experiments is so much higher than risk aversion for non-trivial fractions of lifetime wealth. Our earlier papers explained this with endogenously-determined “mental accounts, and showed how this leads to the high level of small stakes risk aversion seen in the lab. We believe that a similar explanation should be possible with non-myopic short run selves, and that the mental accounts model can also explain what is known as the “magnitude effect: implicit interest rate seems to shrink with the stakes (Thaler [1981], Green et al [1997], Benahib et al [2010].) This remains a project for future work. Note also that if utility is a non-linear function of consumption then the marginal rate of substitution between two dates depends on the consumption levels at both dates in addition to the time-preference parameters.
long-run self is indifferent between one unit now and $c_n$ units later then since $\mu < 1$ the initial shorter-run self strictly prefers one unit now. Hence the temptation is to consume now, which incurs no control cost, and provides utility 1 for an average present value of $1 - \delta$. The initial short run self gets average present value of $1 - \delta \mu$ from consuming at time 1, and $(1 - \delta \mu)(\delta \mu)^{n-1} c_n$ from the delayed option, so the control cost of the delayed option is $\Gamma(1 - \delta \mu)(1 - (\delta \mu)^{n-1} c_n)$. Thus the utility of the delayed option is $(1 - \delta)\delta^{n-1} c_n - \Gamma(1 - \delta \mu)(1 - (\delta \mu)^{n-1} c_n)$. Equating the values of the two options determines the consumption level leading to indifference.

$$1 - \delta = (1 - \delta)\delta^{n-1} c_n - \Gamma(1 - \delta \mu)(1 - (\delta \mu)^{n-1} c_n)$$

We can then solve for $c_n$. This is done in the Appendix; from that computation it follows that as $\mu \to 1$, we have $c_n \to 1/\delta^{n-1}$, which is the solution for a single agent without self control costs. To relate this back to Myerson and Green, the Appendix also computes the instantaneous interest rate for consumption decisions at real time $t = n\tau$ rate by letting the period length $\tau$ go to 0; we show that

$$\lim_{\tau \to 0} \log \left( \frac{c_t/\tau^{\eta+1}}{c_t/\tau} \right) / \tau = \rho + \Gamma \frac{\rho + \eta}{\rho} \exp(-\eta t)\eta$$

In this case the marginal interest rate, to a good approximation, is equal to the subjective interest rate of the long-run self, plus a term that declines exponentially at rate $\eta$. In the case of a shorter-run self who lives exactly one period, that is, $\mu = 0$ or $\eta = \infty$, the marginal interest rate declines after a single period to a constant equal to the subjective interest rate of the long-run self. However, for larger values of $\mu$ we get a more gradual decline, as we see in the data.\footnote{This gradual decline is consistent with hyperbolic (as opposed to quasi-hyperbolic) discounting. The hyperbolic model is not widely used in economics, perhaps because it is hard to apply in a dynamic setting.}

4. Convex Costs of Self Control

We now consider a simple extension of the model of linear cost of self control by allowing the cost of self control to be convex. Specifically, we assume that the objective function is defined by the expected average present value of shorter-run utility net of the control cost
where \( g \) is a convex function. Allowing \( g \) to be convex is important both in light of evidence from the psychology literature, and because in the standard dual-self model convex costs can explain preference “reversals” that arise from failure of the independence axiom, such as the Allais paradox while linear control costs cannot.\(^{19}\) In addition, convex costs can generate “compromise effect” violations of the axiom of revealed preference, as shown in Fudenberg and Levine [2006].

**Example 4: Stochastic Temptations**

Another implication of convex control costs is that the agent is more likely to resist “stochastic temptations” than certain ones. This is the basis of the explanation of the Allais paradox in Fudenberg and Levine [2010]; we give an illustration of the idea here using simple temptations. When faced with a single, and certain, simple temptation, with \( S > 0 > P \), it is optimal to choose the temptation if \( P > -g(S) \). Now suppose that the agent is faced with the choice between an action which gives probability \( q \) of the same simple temptation and complimentary probability \( 1 - q \) of 0, or resisting, with utility flow 0. Then resisting the temptation has foregone value \( qS \), so resisting is optimal if \( -qP < g(qS) \), so when \( g \) is convex it may be optimal to give in to the certain temptation but resist the smaller one.

The same qualitative conclusion extends to the case where the agent learns in period 1 that she will need to make the choice in some future period \( n \): Now the temptation value is \( (\delta \mu)^{n-1}S \), so the agents resists a lottery that gives probability \( q \) of the temptation if \( -q\delta^{n-1}P < g(q(\delta \mu)^{n-1}S) \), and it is possible for this inequality to hold for small \( q \) but not larger ones. At the same time, though, since there is less of a self-control problem about future decisions, increasing \( n \) makes it more likely that the agent resists for all values of \( q \). That is, the model predicts an interaction of the effects of risk and

---

\(^{19}\) Linear control costs are consistent with the independence axiom of expected utility theory and so rule out the Allais-type preference reversals. Noor and Takeoka [2010a] weaken independence axiom in Gul-Pesendorfer axioms to allow non-linear control costs, and then develop axioms that correspond to control costs being convex. Since they work in a two period model with a single choice of a menu, they do not address the modeling issues we discuss here. Noor and Takeoka [2010b] develop a more restrictive set of axioms that correspond to the cost having the form \( \phi(\overline{U})\Delta \), so that marginal cost is increasing in the size of the largest temptation as opposed to the amount of temptation resisted or value foregone.
delay: costly self control in the presence of risk can lead to Allais-type behavior, but since future payoffs are less tempting, individual agents will switch from “paradoxical” to expected-value-maximizing choices when the payoffs are sufficiently far in the future. This is consistent with the data of Baucells and Heukamp [2010]: They found that 36% of subjects exhibited preference reversal in a common-ratio Allais paradox, changing from the safer to the expected-value-maximizing choice when the decision is less likely to matter, while only 22% of subjects exhibited this preference reversal when all payoffs were delayed by three months.\textsuperscript{20} Note finally that the dependence of the decision on $q$ holds in the case where the agent is initially uncertain whether she will face the certain temptation or the stochastic one: All that matters is that she knows which temptation she is facing at the moment she decides.\textsuperscript{21}

**Example 5: Two Tempting Choices**

We now consider a variation on Example 4, where instead of a probability of a more or less tempting choice, there is a certainty that two simple temptations will be faced: at both $n_1 = 1$ and $n_2 \geq 1$ the agent has to decide whether to accept or reject a simple temptation with $S > 0, P < 0$.

Our goal is to investigate the sensitivity of the decisions to the timing. Suppose first that $n_2 > 1$ so there is at least some brief delay between the two decisions. Because of the recursive nature of the formulation and the additivity of the utilities, the two decisions are identical. If the option is not taken, utility is 0 and the control cost is $g(S)$. If the option is taken, utility is $P$ and there is no self-control cost, so it is optimal to take at both $n_1$ and $n_2$ if $-P < g(S)$, and not to take if $-P > g(S)$. Notice that the solution is the same for any value of $n_2 > n_1$, and for any period length, so it holds in particular if the periods are arbitrarily short.

\textsuperscript{20} Explaining this data with a one-period horizon for the short-run player requires that the period length of the subjects who switched be at least three months; this is consistent with the theory but not with data from other experiments. Note also that Weber and Chapman [2005] do not find this effect, although their sample was very small.

\textsuperscript{21} Convex costs also imply that it is easier to resist smaller deterministic temptations when (as in example 2) we suppose that utility is a linear function of consumption. As noted in footnote 12, there is evidence that the cost of resisting money temptations does not follow this prediction; we attribute this to the fact that money is not consumed immediately and that its use is governed in part by mental accounts. One advantage of varying the probability of winning and not the prize itself is that the mental accounting for the scale of prize used in lab experiments may well be independent of the probability with which that prize is won.
However, the solution changes if \( n_2 = n_1 \). In this case the possible actions are not to take, \( a = 0 \), to take exactly one of the options, \( a = 1 \), or to take both options \( a = 2 \). The temptation is to take both options, so utility is \( V(a) = aP - g(2S - aS) \).

Then
\[
V(2) - V(1) = 2P - P + g(S) = P + g(S)
\]

and
\[
V(1) - V(0) = P - g(S) + g(2S).
\]

When \( g \) is strictly convex \( g(2S) - g(S) > g(S) \). If \( g(2S) - g(S) > -P > g(S) \) it is optimal to resist each temptation when the options are sequential but it is not optimal to resist both when they are presented simultaneously. This shows that this model of non-linear costs is not suited for analyzing decisions that occur in rapid succession. Intuitively, the problem is that the non-linearity of control costs should “spill over” from one period to the next when time periods are short. The next section extends the model to allow this.

5. Willpower as a Stock and Increasing Marginal Cost of Self Control

The reason that control costs are often convex is that self control can require the use of costly cognitive resources, as argued by Baumeister and various collaborators (for example Baumeister et al [1998] and Muraven et al [1998]). This implies that soon after one tempting choice the marginal cost of another tempting choice will be high; for example two consecutive decisions a microsecond apart should be about the same as two simultaneous decisions. Thus, to develop a model that is consistent with convex control costs and also robust to the timing of decisions and the granularity of the periods, we need to incorporate the way the willpower stock induces a spillover from one period’s self control to self control in the near future. \(^{22}\) To do this, we develop a generalization of the willpower model of Ozdenoren et al [2009].

Specifically, we assume that at the beginning of period \( n \) there is a stock \( w_n \) of cognitive resources or willpower available. Note that this is part of the vector \( y \). Foregone value \( \Delta \) has the same definition as before, and in particular is not affected by \( w_n \); the change in the model is that the cost of self control comes from the fact that it depletes the stock of cognitive resources. Specifically, when \( \Delta(y_n, a_n) \) is the foregone value, the end of period stock is \( \tilde{w}_n = f(w_n, \Delta(y_n, a_n)) \), where \( f(w, \Delta) \) is non-

---

\(^{22}\) In the longer term, it is possible that willpower can be built up, that is, that the “willpower technology” can be improved. This introduces a range of issues that our model does not handle well, and we abstract away from it for most of the paper, section 6 explains some of the complications that arise when willpower can be increased through training.
decreasing in $w_n$ and non-increasing in $\Delta$, continuously differentiable in both arguments
and satisfies $f(w,0) = w$ and $f(w,\Delta) \leq w$. Note that the stock of cognitive resources
depends on the action taken only through the foregone value, so actions that maximize
the shorter-run self’s value also maximize the end-of-period stock $\tilde{w}_n$. We do not require
that cognitive resources are bounded below, and indeed it will be more convenient to
allow them to become arbitrarily negative.

In Ozdenoren et al the stock is depleted, but never replenished. This is a
reasonable approximation for the short-duration problem they analyze, but to adapt the
model to longer horizons we add the possibility that willpower can be replenished.
Specifically we set $w_{n+1} = r(\tilde{w}_n) \geq \tilde{w}_n$, where $r$ is non-decreasing in $\tilde{w}_n$; thus for a
given $w_n$ the highest that $w_{n+1}$ can be is $r(w_n)$, and this maximum is attained by actions
that set $\Delta = 0$. We assume also that $r(\tilde{w}_n) \leq \bar{w}$ so that there is an upper bound on the
stock of cognitive resources. If $r(\tilde{w}_n) = \bar{w}$ then resources are replenished immediately,
which is the usual assumption when shorter-run selves live a single period. If
$r(\tilde{w}_n) = \tilde{w}_n$ resources are never replenished, as in Ozdenoren et al. Self-control costs
arise because cognitive resources have alternative uses. Following Ozdenoren et al, we
assume that an (end of period) stock of cognitive resources $\tilde{w}_n$ yields a utility in other
uses of $m(y_n, \tilde{w}_n)$, non-decreasing in $\tilde{w}_n$, and that this is added to the utility from
consumption. Ozdenoren et al view $\tilde{w}_n$ as representing only the stock of willpower, and
motivate its assumed value as arising from self-control problems that are not directly
modeled. In our earlier work we provide evidence that cognitive resources matter, and
that these resources have alternative uses, so we take a broader view of what the uses of
these resources might be.

The objective function of the long-run self is then to maximize

$$V(h_n, a) \equiv E_{a,b_n} (1-\delta) \sum_{\ell=0}^{\infty} \delta^\ell \left( u(y_{n+\ell}, a_{n+\ell}) + m(y_{n+\ell}, f(w_{n+\ell}, \Delta(y_n, a_n))) \right)$$

Note that the contribution $m$ of the stock of cognitive resources is measured in the same
units as utility. Thus if there is a fixed stock $\bar{w}$ of cognitive resources, the stock produces

---

23 It does not, however, enter into the computation of the temptation utility or the foregone value, as these
are a cause of self-control cost, not a consequence. Note also that we do not impose the restriction $\tilde{w}_n \geq 0$
as do Ozdenoren et al; we can set $m(y_n, \tilde{w}_n) = -\infty$ if $\tilde{w}_n < 0$ to incorporate that constraint, in which
case we modify the differentiability assumptions that we impose below.
an amount \( m(y, \bar{w}) \) of utility each period. The function \( m \) is assumed to be concave, differentiable and strictly increasing in \( w \). Recall that with full replenishment the cost was not normalized by \( 1 - \delta \) as the benefit is here. We will discuss the reason for this difference below.

As in the linear case we assume LR0 which we repeat here for completeness:

**Assumption LR0:** \( V(h_n, a) \) has a maximum for each \( n, h_n \).

Note also that once again the long-run self’s maximization problem has a solution that depends only the current state \( y_n \).

In the “cake-eating” problem of Ozdenoren et al [2009] there is a cake of fixed size, the only choice is a consumption level \( a_n \) that reduces the size of the cake, and the objective of the long-run self is to maximize the discounted value of consumption until the exogenous terminal date, with any remaining cake being useless. They assume a purely myopic short-run self \( (\mu = 0) \) and that there is no replenishment \( r(\tilde{\omega}_n, a_n) = \tilde{\omega}_n \). Ozdenoren et al use a continuous-time model specify that the temptation is a fixed upper bound \( \omega \) on the consumption rate as long as some cake remains, and that the rate of willpower depletion is \( \tilde{f}(a_n, w_n) \) for \( a_n < \omega \), with \( \tilde{f} \) decreasing and strictly convex in \( a_n \). In addition \( m(y_n, \tilde{\omega}_n) = 0 \) until the stock of cake runs out or the time horizon is reached, at which time \( m(y_n, \tilde{\omega}_n) = \tilde{m}(\tilde{\omega}_n) \). In our formulation \( \Delta(y_n, a_n) = u(\omega) - u(a_n) \), so that their model a special case of ours with \( f(w_n, \Delta(y_n, a_n)) = \tilde{f}(u(\omega) - \Delta(y_n, a_n), w_n) \).\(^{24}\) However, their formulation requires that cognitive resource utility is state-dependent. This possibility leads to complications, because it implies that the plan most preferred by the short-run self, which is the plan that has the least temptation, need not minimize the resource cost of self control. We examine this issue in section 6, along with the possibility that actions have a direct impact on the evolution of cognitive resources.

**Assumption [State-Independence]:** \( m \) depends on the state only through the stock of willpower; that is for all \( y_n, y_n', \tilde{\omega}_n, m(y_n, \tilde{\omega}_n) = m(y_n', \tilde{\omega}_n) \)

---

\(^{24}\) The function \( f \) is not constrained to be \( w_n - \Delta \). This allows for lower depletion near \( w_n = 0 \), as in the multiplicative functional form \( f(w_n, \Delta(y_n, a_n)) = \Delta(y_n, a_n)w_n \).
We will maintain this assumption until section 6, and to lighten notation we will write \( m(\bar{w}_n) \) and in place of \( m(y_n, \bar{w}_n) \).

The dual-self model of Fudenberg and Levine [2006, 2010] corresponds to the assumptions (i) \( r(\bar{w}_n) = \bar{w}, \) so that replenishment is immediate, and (ii) \( \mu = 0 \) so the short-run self is completely myopic. Here the temptation value is \( \max_a u(y_n, a), \) so 

\[
\Delta(y_n, a_n) = \max_a u(y_n, a) - u(y_n, a_n),
\]

and the benefit derived from cognitive resources in period \( n \) is \( m(f(\bar{w}, \Delta(y_n))) \). We can then define \( c(\Delta) = (m(\bar{w}) - m(f(\bar{w}, \Delta)))/(1 - \delta) \geq 0; \) since \( f(w, 0) = w \) this implies \( c(0) = 0. \)

Substituting this definition of cost into the objective function, we have

\[
V(h_n, a) = E_{a, h_n} \left( 1 - \delta \right) \sum_{\ell=0}^{\infty} \delta^\ell \left( u(y_{n+\ell}, a_{n+\ell}) + m(\bar{w}) - c(\Delta(y_{n+\ell}, a_{n+\ell})) \right).
\]

which is equivalent to the objective function in our earlier papers. Note here that neither the function \( f \) nor the function \( m \) matters on its own: what matters is the composition \( m \circ f, \) for this is what determines the cost function \( c. \)\(^{25} \) There is a similar but more complicated interaction between \( m \) and \( f \) under partial replenishment, as we see in the next subsection and again in the analysis of Example 9.\(^{26} \)

**Single Decision Problems**

Cognitive resources serve to link the decisions in one period to future control costs and thus subsequent decisions. Several of the examples we have considered so far involve a single decision; in those cases the resource variable is superfluous. To make this precise, we define what we mean by a *single decision*. Let \( Y^* \) be the set of states in which a decision is possible, that is \( y \notin Y^* \) implies \( \#(A(y)) \leq 1. \) Then the probability of hitting \( Y^* \) from a state \( y \in Y^* \) must be zero: once a decision is offered, no further decisions are possible. Notice, though, that if \( y \) can occur in different periods, the

---

\(^{25}\) To model the effect of cognitive load (e.g. using short-term memory) on self control, Fudenberg and Levine [2006] assume that the control cost depends on the sum of the foregone value and cognitive load; this corresponds to assuming that the benefit derived from cognitive resources in period \( n \) is \( m(f(\bar{w}, \Delta(y_n) + d_n)), \) where \( d_n \) is the cognitive load in period \( n. \)

\(^{26}\) Note also that cost is normalized differently than \( m. \) When there are cognitive resources, the loss due to a single shock is spread out over many periods, so the per period amount should be in the same units as period utility \( u. \) In contrast when cognitive resources replenish immediately, as implicit in the “cost” formulation, the loss due to a single shock is concentrated in a single period and should be in the same units as the present value of utility.
amount of cognitive resources available for decision making may be different. Suppose that \( w_1 = \bar{w} \), so that initially cognitive resources are “topped up”. In this case we say that resources start full.\(^{27}\)

For any period \( n \) define the end of period resource stocks corresponding to an initial shock of \( \Delta \) and no subsequent shock by \( \tilde{w}_n^n(\Delta) = f(\bar{w}, \Delta) \), and \( \tilde{w}_{n+\ell}^n(\Delta) = r(\tilde{w}_{n+\ell-1}^n(\Delta)) \) for all \( \ell \geq 1 \). Then the cost of self control corresponding to a single shock is

\[
g(\Delta_n) = \sum_{\ell=0}^{\infty} \delta^\ell \left[ m(\bar{w}) - m(\tilde{w}_{n+\ell}^n) \right].
\]

The following result is immediate

**Theorem 3:** If there is a single decision and cognitive resources start full, the maximization problems

\[
E_{a,h_n}(1 - \delta)\sum_{\ell=0}^{\infty} \delta^\ell \left( u(y_{n+\ell}, a_{n+\ell}) + m(y_{n+\ell}, f(w_{n}, \Delta(y_n, a_n))) \right)
\]

and \( E_{a,h_n} \sum_{\ell=0}^{\infty} \delta^\ell \left( (1 - \delta)u(y_{n+\ell}, a_{n+\ell}) - g(\Delta(y_{n+\ell}, a_{n+\ell})) \right) \) have the same set of solutions.

**The Linear Case**

Next suppose that in addition to conditions (i) and (ii) above, the benefit of cognitive resources is linear in \( \Delta \), so that \( m(y_n, \tilde{w}_n) = \gamma \tilde{w}_n \), and that resource depletion is linear as well, so that \( f(w, \Delta) = w - \Delta \).\(^{28}\) Then the cost function \( c(\Delta) = (m(\bar{w}) - m(f(\bar{w}, \Delta))) / (1 - \delta) \) defined above is \( c(\Delta) = (1 - \delta)\gamma \bar{w} - (1 - \delta)\gamma(\bar{w} - \Delta) = \gamma(1 - \delta)\Delta \), so the linear model of the previous section, where the cost of self control is independent of \( \delta \), corresponds to scaling the cost by \( 1/(1 - \delta) \). Intuitively, full replenishment means that all of the cost of self control is borne in the current period, so if foregone utility reduces the flow benefits of cognitive resources by a proportionate amount, the cost of self control goes to zero.

\(^{27}\) The theorem also holds if there is a fixed time at which the decision is possible. That is if \( y_n \in Y^* \) implies \( n = n^* \) then we may replace \( \bar{w} \) with the fixed amount of cognitive resources \( w^* \) available when a decision is possible, as this is invariant to \( y \in Y^* \).

\(^{28}\) Note that we define linear depletion to mean that foregone utility is subtracted one-for-one from the stock of resources. In principle depletion might be linear with a coefficient other than 1, but we can normalize the coefficient to one by choosing appropriate units for \( w \).
with the period length. Conversely, if the control cost is invariant to the period length and there is full replenishment, the flow cost in a period must become large as the periods get small. This is also true when there are convex costs: The convex cost model of Section 4 can be viewed as a model with full replenishment and linear depletion, where the benefits at $\bar{w}$ are independent of $\tau$, while for smaller stocks we have $m_\tau(\bar{w} - \Delta) = m(\bar{w}) - g(\Delta)/(1 - \delta)$.

When benefits are linear as well we have a stronger result: the linear model with partial replenishment is equivalent to the linear model with full replenishment, so that partial replenishment has observable consequences only if at least one of $f, r$ and $m$ is non-linear. Specifically, we say the model has linear replenishment of resources if $r(\tilde{w}_n) = \tilde{w}_n + \lambda(\bar{w} - \tilde{w}_n)$ where $0 \leq \lambda \leq 1$.

**Theorem 4:** For any model with linear benefit $\gamma$, depletion and replenishment $\lambda$ define

$$\Gamma = \frac{(1 - \delta)\gamma}{1 - \delta(1 - \lambda)}.$$ 

Then if $a$ is a solution to the linear model with parameter $\Gamma$ then it is a solution to the $(\lambda, \gamma)$ model in which actions are independent of $w_n$ and all such solutions to the $(\lambda, \gamma)$ model are solutions to the $\Gamma$ model.

Theorem 4, proven in the Appendix, shows that if depletion, replenishment, and self-control cost are all linear, and state independent, the stock of self control is irrelevant. Intuitively, with linear costs all that matters is the average present value of the costs, and not their timing, which is why the stock of willpower resources does not matter. Because the theorem maps many $(\lambda, \gamma)$ models to the same linear model $\Gamma$, it also implies that these models are equivalent in the sense of generating the same decisions. That is, if we change $\lambda$ to $\lambda'$ while holding the time period (and thus $\delta$ and $\mu$) fixed, and set

$$\gamma' = \gamma \left( \frac{1 - \delta(1 - \lambda')} {1 - \delta(1 - \lambda)} \right),$$

the resulting system $(\lambda', \gamma')$ will have the same cost for every self-control decision, even though the time-path of the willpower stock in the two models will be different: the agent has the same preferences with resources that replenish quickly and matter a lot or with resources that replenish very slowly but matter little.
Note that it is important for this result that resources are unbounded below; if there is a lower boundary the model is not linear there, and the equivalence with the full replenishment linear model breaks down. Note also that the equivalent linear parameter $\Gamma$ depends on the replenishment rate: when resources are replenished very quickly ($\lambda = 1$), the cost of self control is of the order $(1 - \delta)$ of a single period’s utility, while when replenishment is slow, self control has a long-term cost of order 1.

When we vary $\tau$, the length of the period, we will want to hold fixed the amount of calendar time required for a given amount of replenishment, so we set $\lambda(\tau) = 1 - \exp(-\kappa \tau)$. This corresponds to assuming that self control in a given period reduces the stock of willpower at the start of that period. That is, we suppose that when $\Delta_n$ is spent in some period $n$, the stock immediately jumps from $w_n$ to $\hat{w}_n$, and is then replenished according to the continuous-time differential equation $\dot{w} = \kappa (\bar{w} - w)$. Thus when the period length is $\tau$ we have $w_{n+1} = \bar{w} - \exp(-\kappa \tau)(\bar{w} - \hat{w}_n)$ so $\lambda(\tau) = 1 - \exp(-\kappa \tau)$. Note that when the period is long, the stock almost completely replenishes. Note also that as $\tau \to 0$ we have $\lambda(\tau) \to 0$, $\delta(\tau) \to 1$.

To characterize the behavior of the linear model as $\tau \to 0$, we compute the limit of the equivalent marginal cost, which is

\begin{equation}
\Gamma^* \equiv \lim_{\tau \to 0} \frac{(1 - \exp(-\rho \tau))\gamma}{(1 - \exp(-(\rho + \kappa)\tau))} = \frac{\rho \gamma}{\rho + \kappa}.
\end{equation}

If $\kappa$ is very small, so that any reduction in the stock is almost permanent, then $\Gamma^* \approx \gamma$. At the other extreme, when $\kappa$ is very large, the equivalent cost is near 0. Intuitively, in this case reductions in resources are replaced so quickly that they are virtually costless, even though the amount of replenishment in a given period goes to 0 with $\tau$.

Example 6: Resisting Temptation with in the Linear Case

To illustrate Theorem 4 we reconsider the persistent and delayed temptations of Example 1 in the model with linear benefit, linear depletion and linear replenishment.

---

29 In the Web Appendix we consider alternate specifications of the timing of the reduction in the resources stock caused by self control, for example the reduction might happen at the end of the period instead of the start. We show that with linear replenishment the timing does not matter when periods are small, though of course it does matter when periods are longer.
This will let us illustrate the continuous-time limit, and set the stage for our subsequent analysis of these temptations when the benefit function is concave.

We begin with the case of a persistent temptation, where temptation is present each period unless and until it is accepted. If the benefits are linear, so that \( m(w) = \gamma w \), then by Theorem 4 the solution is the same as in the linear case: The agent will resist the temptation if \( P(1 - \delta) < -\Gamma(1-\delta\mu)S \) and accept it when the reverse inequality is satisfied, where \( \Gamma = (1 - \delta)\gamma / (1 - \delta(1 - \lambda)) \). Using equation (4) we see that the when the time period is short the agent resists the temptation when

\[
\rho P < -\gamma \rho S (\rho + \eta) / (\rho + \eta). \tag{5}
\]

Here the LHS is approximately the gain of postponing \( P \) by \( \tau \), and the RHS is the cost of postponement; this is the foregone value of \((\rho + \eta)S\) multiplied by the continuous-time cost parameter \( \Gamma^* \). Note that when \( \kappa \) is very large compared to the other parameters, the right hand side is near 0 so it is always better to resist.

To extend this analysis to the case where declining the temptation delays it for a real time \( s \), recall that when declining puts off the temptation for \( \ell \) periods, it is optimal to resist if \( P < -\Gamma(1-(\delta\mu)^\ell)/(1-\delta^\ell)S \); substituting for \( \Gamma \) yields \( P(1 - \delta(1 - \lambda)) / (1 - \delta) < -\gamma S(1-(\delta\mu)^\ell) / (1 - \delta^\ell) \). If we suppose that the delay is \( s \) units of real time, \( \ell = s / \tau \), and send \( \tau \) to zero while holding \( s \) fixed, then it is optimal to resist if \( P(\rho + \kappa) / \rho < -\gamma S \left[ \left( 1-\exp(-(\rho + \eta)s) \right) / \left( 1 - \exp(-\rho s) \right) \right] \).

In the case \( s = \tau \), in which the delay is only a single period, this inequality reduces to \( \rho P < -\gamma \rho S \left[ (\rho + \eta) / (\rho + \kappa) \right] \) which is what we had before. As \( s \to \infty \), the condition reduces to \( P < -\gamma \rho S / (\rho + \kappa) \) which is easier to satisfy. More generally, as in the discrete-time model the decision of whether to take at once or resist forever is monotone in \( s \): There is some \( \overline{s} \) (possibly 0 or infinity) such that the optimum is to take at once if \( s < \overline{s} \) and resist forever if \( s > \overline{s} \). When we re-examine this problem with a concave benefit function, we will see that it can be optimal to resist for a finite length of time and then take, and that the optimal time to give in is monotone in the length \( s \) of the delay.

6. Cognitive Resources, Non-linearities, and Replenishment

In the linear model the stock of willpower does not play a significant role. When there are non-linearities in \( f, r \) and/or \( m \), the stock of willpower determines the way the
cost of foregone utility is allocated between different periods. The main reason for introducing the cognitive resources variable is to allow for the possibility that (1) the cost of self control depends on the stock, (2) the stock does not completely replenish from one period to the next, so that exerting willpower in one period can have a carry-over effect on decisions made soon afterwards, and (3) the agent faces more than one decision so using self control in an earlier decision can alter the control cost in a subsequent one. The simplest way to capture this is to suppose that there is no replenishment at all, so that the stock evolves according to \( w_{n+1} = f(w_n, \Delta(y_n, a_n)) \). This stark assumption is sufficient for demonstrating many of the implications of willpower as a resource that is limited in the short run, but it is not necessary, and many of the same results obtain provided that replenishment is incomplete.

In general, the dynamic value of cognitive resources given the foregone utility process \( \Delta_n \) is governed by the depletion equation \( \bar{w}_n = f(w_n, \Delta_n) \), the replenishment equation \( w_{n+1} = r(\bar{w}_n) \) and of course the benefit function \( m(y_n, \bar{w}_n) \). Putting together the depletion and replenishment equations gives the dynamics of cognitive resources \( w_{n+1} = r(f(w_n, \Delta_n)) \), where \( \underline{w} \geq r(\bar{w}_n) \geq \bar{w}_n \) is non-decreasing and \( f(w_n, \Delta_n) \leq w_n, f(w_n, 0) = w_n \) is non-decreasing in \( w_n \) and non-increasing in \( \Delta_n \).

### Linear Replenishment

The units in which \( w_n \) are measured are arbitrary; by changing them we change \( f, r \) and \( m \). As we shall see there is redundancy in these three functions, meaning that we can choose one of them to normalize.

Specifically, let \( 0 < \lambda < 1 \) be a fixed number. We will construct a change of units \( w' = h^{-1}(w) \) and \( \bar{w}' = h^{-1}(\bar{w}) \) so that the replenishment function \( r' \) corresponding to \( w' \) is linear, with \( r'(\bar{w}') = \bar{w} - (1 - \lambda)(\bar{w} - \bar{w}') \). Given such a function we may define \( f'(w', \Delta) = h^{-1}(f(h(w'), \Delta)) \), \( m'(w') = h^{-1}(m(h(w'))) \), and the model with the new units and new depletion and benefit functions is the same as the original one.

There are a variety of ways of constructing an \( h \) function. One simple method is to start with an interval \( I_\lambda = (0, \lambda \bar{w}] \) and a mapping \( T(w') \equiv \bar{w} - \lambda(\bar{w} - w') \) and then consider the images of the iterated map \( T^n(I_\lambda) \) (since \( T \) is invertible, we allow negative values of \( n \)). Notice that these intervals form a partition of \(( -\infty, \bar{w}) \).
Hence for any \( w' < \bar{w} \) there exists a unique integer \( n(w') \) (possibly negative) such that \( w' \in T^{n(w')}(I_\lambda) \). The interval \( I_\lambda \) for the units \( w' \) corresponds to \( (0, r(0)] \) in the original units \( w \). Define \( h(w') = r_{n(w')}(T^{-n(w')}(w')r(0) / \lambda \bar{w}) \). If \( r(0) > 0 \) and \( r \) is strictly increasing for \( \tilde{w} < \bar{w} \), \( (0, r(0)] \) is a non-empty interval, so \( h \) maps onto \((-\infty, \bar{w})\). In this case \( h \) is directly increasing, so invertible, and by construction \( h(\bar{w} - \lambda(\bar{w} - w')) = r(h(w')) \). If \( r \) is continuous, \( h \) is continuous and \( h(0) = 0 \); \( h \) extends uniquely to a continuous function\(^{31}\) on \((-\infty, \bar{w}]\) by defining \( h(\bar{w}) = \bar{w} \).

We summarize this as a theorem.

**Theorem 5:** Suppose that \( r \) is continuous and strictly increasing and that \( r(w) > w \) for at least one \( w \). Then the system with replenishment \( r'(\tilde{w}') = \bar{w} - (1 - \lambda)(\bar{w} - \tilde{w}') \), depletion \( f'(w', \Delta) = h^{-1}(f(h(w'), \Delta)) \), and benefit function \( m'(w') = h^{-1}(m(h(w'))) \) maps strategies to values of the agent’s objective function exactly as does the system with replenishment \( r(\tilde{w}) \), depletion \( f(w, \Delta) \) and benefits \( m(w) \).

Notice that the rescaling of units to linearize \( r \) is possible only when there is some replenishment (\( r(w) > w \)) and less than full replenishment (\( r \) strictly increasing). When there is full replenishment, we cannot change the units to spread the foregone utility shock over time: as soon as there is non-linearity partial replenishment spreads the marginal cost of self control over time.

Note also that if we start with a system where benefits and depletion are linear, and replenishment is linear with some \( \lambda' \), then the equivalent system in units of \( w' \) for a different value of \( \lambda \) is not linear. This may seem puzzling in light of our observation that Theorem 3 implies an equivalence between linear models with different replenishment rates. However, this is only for average present values, not the stronger sort of equivalence established here, which tracks the moment-by-moment movement of the flow benefit of cognitive resources. The weaker form of equivalence is sufficient when benefits and depletion are linear, but once these functions are allowed to be non-linear the stronger sort of equivalence is needed, and this equivalence requires a non-linear change of units.

---

\(^{30}\) The same construction would work with a different origin provided there is some \( w \) with \( r(w) > w \).

\(^{31}\) Notice that \( h \) need not be differentiable everywhere, it may have kinks at the boundaries of the intervals.
Non Linear Costs and Linear Replenishment

Now we investigate the implications of non-linear costs when the agent faces multiple decisions, so that self control in one period can increase the marginal cost of self control in the next one. To make the computations easier we pick units so there is linear replenishment, \( r(\hat{w}_n) = \hat{w}_n + \lambda(\bar{w} - \hat{w}_n) \). In the examples that follow we will frequently need to compute the average present value of cognitive resources when the stock at the start of period \( n \) is some arbitrary \( w_n \) and no self control is used from period \( n \). With linear replenishment and no foregone utility, \( w_{n+\ell} = (1 - \lambda^\ell)\bar{w} + \lambda^\ell w_n \). Along this path the average present value of cognitive resources is

\[
M(w_n) = (1 - \delta)\sum_{\ell=0}^{\infty} \delta^\ell m(w_{n+\ell}).
\]

Recall that to study the effect of varying the period length recall we take \( \lambda(\tau) = 1 - \exp(-\kappa\tau) \), so for small \( \tau \) we have \( \lambda(\tau) \approx \kappa \tau \).

Example 7: Two Tempting Choices with Linear Depletion and Replenishment

Now we re-analyze the two temptations of example 5 assuming partial linear replenishment and linear depletion. We show that the agent makes the same decision whether the decisions on the two temptations are made simultaneously or in very rapid succession.

Suppose first that the decisions are made in consecutive periods, so that \( n_1 = 1 \), \( n_2 = 2 \), and to simplify notation suppose that the initial stock of resources is \( \bar{w} \). The agent has four possible plans:

*Take both options:* The first option is provides direct utility of \( P \), the second option provides \( \delta P \), and no self control is used, so overall utility is \( (1 + \delta)P + M(\bar{w}) \). As \( \tau \to 0 \) this converges to \( 2P + M(\bar{w}) \).

*Take only second option:* Self control of \( S \) is used in the first period, and none thereafter, so \( \bar{w}_1 = \bar{w} - S \), and the overall value is

\[
(1 - \delta)m(\bar{w} - S) + \delta P + \delta M(\bar{w} - (1 - \lambda)S) = \delta P + M(\bar{w} - S).
\]

As \( \tau \to 0 \) this converges to \( P + M(\bar{w} - S) \).

*Take only the first option:* No self control is used in the first period; in the second period the foregone SR utility (and thus the expenditure of cognitive resources) is \( S \) and no self
control is used thereafter, so cognitive resources at the end of the second period are \( \bar{w} - S \), and the overall value is

\[
P + (1 - \delta)m(\bar{w}) + \delta(1 - \delta)m(\bar{w} - S) + \delta^2 M(\bar{w} - (1 - \lambda)S) = P + (1 - \delta)m(\bar{w}) + \delta M(\bar{w} - S)
\]

As \( \tau \to 0 \) this converges to \( P + M(\bar{w} - S) \).

**Reject both option:** Self control is used in both periods, so the value is

\[
(1 - \delta)m(\bar{w} - S) + (1 - \delta)\delta m(\bar{w} - (2 - \lambda)S) + \delta^2 M(\bar{w} - (1 - \lambda)(2 - \lambda)S) =
\]

\[
(1 - \delta)m(\bar{w} - S) + \delta M(\bar{w} - (2 - \lambda)S)
\]

As \( \tau \to 0 \) this converges to \( M(\bar{w} - (2 - \lambda)S) \).

Because the two projects are identical, when the decisions are made simultaneously there are only three plans to consider:

**Take both options:** Here the overall value is \( 2P + M(\bar{w}) \). As \( \tau \to 0 \) this does not change.

**Take one option:** Here overall value is \( P + (1 - \delta)m(\bar{w} - S) + \delta M(\bar{w} - (1 - \lambda)S) \).

As \( \tau \to 0 \) this converges to \( P + M(\bar{w} - S) \)

**Reject both options:** Now the overall value is

\[
(1 - \delta)m(\bar{w} - 2S) + \delta M(\bar{w} - (1 - \lambda)2S) = M(\bar{w} - 2S)
\]

As \( \tau \to 0 \) this converges to \( M(\bar{w} - 2S) \).

Thus as \( \tau \to 0 \) the value for taking both converges to \( 2P + M(\bar{w}) \), the value for taking either one of the two options converges to \( P + M(\bar{w} - S) \), and the value for rejecting both options to \( M(\bar{w} - 2S) \), regardless of whether the options are sequential or simultaneous, and regardless of which options is taken when only one is taken. Hence simultaneity does not matter in the limit, and the optimum is governed by whichever of these three numbers is largest. In the sequential case, if it is optimal to take just one option, the advantage of taking the second option over the first is given by

\[
(1 - \delta)[-P + M(\bar{w} - S) - m(\bar{w})]
\]

Notice that \( m(\bar{w}) = M(\bar{w}) \), and that the first option is strictly better if and only if \( -P < M(\bar{w}) - M(\bar{w} - S) \). On the other hand, for the first option to be a strict optimum, it must also be strictly better than taking both options, which implies \( M(\bar{w}) - M(\bar{w} - S) < -P \). Thus if it is strictly best to take one option, it must be the second option. Intuitively, when the second option occurs cognitive resources are depleted, so it makes more sense to given in then rather than to first give in and then resist.
Example 8: Persistent Temptation with Non-Linear Benefits and No Replenishment

To further explore the implications of willpower being a stock that can be depleted over time, we now revisit the persistent temptation of example 1 in a setting with no replenishment of cognitive resources, linear depletion, and non-linear benefits. One of the main differences is that with it may now be optimal to resist a while and then take the temptation once the marginal benefit of resources becomes sufficiently high. We emphasize that this sort of non-stationary behavior is consistent with perfect foresight and thus need not be interpreted as a sign that agents misperceive their own future intentions.

Because there is no replenishment at all, the stock decreases by \( \Delta = (1 - \delta\mu)S \) each time the agent resists, and if the agent resists \( \ell \) times before giving in, his value is

\[
(1 - \delta) \left[ \sum_{n=1}^{\ell-1} \delta^{n-1} \mu (\bar{w} - n(1 - \delta\mu)S) + \sum_{n=\ell}^{\infty} \delta^{n-1} m(\bar{w} - (\ell + 1)(1 - \delta\mu)S) + \delta\mu \right]
\]

The bigger is \( \ell \) the smaller is the first term and the larger is the second. This implies that a necessary condition for an optimal \( \ell \) is that the value for resisting \( \ell + 1 \) time is not bigger, and that a sufficient condition for \( \ell = 0 \) optimal is that the value for \( \ell = 1 \) is lower.

The value for \( \ell \) minus the value for \( \ell + 1 \) is

\[
D(\ell) = \delta^\ell \left[ m(\bar{w} - \ell(1 - \delta\mu)S) - m(\bar{w} - (\ell + 1)(1 - \delta\mu)S) + (1 - \delta)P \right] \]32

Observe that since \( m \) is concave, \( D(\ell) \) is strictly increasing, so it is optimal to take at the first time this expression is positive, and never to take if it is always negative.

To characterize the solution for small \( \tau \) define

\[
d(s, \tau) = D(s / \tau) / \left( \delta^{(s/\tau)}(1 - \delta) \right) = \left[ m(\bar{w} - (s / \tau) \exp(-((\rho + \eta)S\tau) - m(\bar{w} - (s / \tau + 1)(\exp(-((\rho + \eta)\tau)S + \exp(-\rho\tau)P) / \exp(-\rho\tau)) \right)
\]

where \( s / \tau \) is the number of periods until real time \( s \). Note that \( d \) has the same sign as \( D \), so it can be used to characterize the optimum. Observe also that

\[
d(s) = \lim_{\tau \to 0} d(s, \tau) = m'(\bar{w} - s(\rho + \eta)S)/(\rho + \eta)S / \rho + P.
\]

Thus if \( \lim_{w \to -\infty} m'(w) < -\rho P / S(\rho + \eta) \) it is optimal to never give in. If \( m'(\bar{w}) > -\rho P / (\rho + \eta)S \) it is optimal to give in right away, and if neither corner

---

32 Here we make use of the fact that the two decisions lead to the same time path of resources up to time \( \ell \).
solution applies the optimal time to give in is characterized by
\[ m'(w - s(\rho + \eta)S) = -\rho P / (\rho + \eta)S. \]

This case of no replenishment is extreme, and we will soon revisit this example to allow not only partial linear replenishment but also a general depletion function. First, though, we want to make a simpler point about the possibility it is optimal to “wait to commit.”

The case where declining the temptation postpones the decision for a number of periods \( T \) is also of interest. Using an analogous argument to the one above, it can be shown that the stopping time is increasing in \( T \); we omit the details.

### Example 9: Waiting to Commit

Without cognitive resources (or with full replenishment) it is always cheaper to commit now to avoid a future temptation that to do so later when the temptation is more imminent. However, because there is no foregone value associated with delaying the commitment, it is easy to construct examples where it is optimal to postpone a decision if the current stock of resources is quite low. The next example expands on this theme of optimal delay: we now add to Example 8 the possibility of taking the temptation off the table for a cost \( F < |P| \), while maintaining the simplifying assumptions of no replenishment and linear depletion. We will show that it can be optimal to wait for a while and then pay to commit; once again a non-stationarity in behavior is consistent with perfect foresight.

Using the argument from Example 8, we see that for small enough \( \tau \) it is not optimal to commit immediately if
\[ m'(\bar{w} - F)(\rho + \eta)S / \rho < F. \]

A sufficient condition is
\[ m'(\bar{w} + P) < \rho F / (\rho + \eta)S \]

Suppose that \( m'(-\infty)(\rho + \eta)S / \rho + P < 0 \), so that in the absence of the possibility of commitment it is optimal never to give in, resulting in at least the value \( m(\bar{w}) + P \). Taking the temptation off the table in the first period gives value
\[ \lim_{\tau \to 0} (1 - \delta) \left[ \sum_{n=1}^{\infty} \delta^{n-1} m(\bar{w} - (1 - \delta \mu)S - F) \right] = F = m(\bar{w} - F) - F \]

---

33 A naïve interpretation of these inequalities is that if \( \rho \) is large compared to the other terms the agent resists forever while if \( \rho \) is small the agent accepts at once, but as both \( P \) and \( S \) depend on \( \rho \) the comparative statics require more care.
Hence a sufficient condition for committing is \( m(\overline{w}) - m(\overline{w} - F) + P + F < 0 \).

Observe that
\[
m(\overline{w}) - m(\overline{w} - F) + F + P \leq m'(\overline{w} - F)F + F + P \leq m'(\overline{w} + P)F + F + P.
\]

Thus a sufficient condition for committing is \( m'(\overline{w} + P) < \frac{|P|}{F - 1} \). This together with (7) are sufficient for it to be optimal to wait a while then commit. To see that both conditions can be satisfied simultaneously, take \( F = |P|/2 \). Then the sufficient condition for committing is \( m'(\overline{w} + P) < 1 \) while (7) becomes \( m'(\overline{w} + P) < \rho |P| / 2(\rho + \eta)S \). In other words, if the marginal benefit of cognitive resources is low when resources are higher than \( \overline{w} + P \), and if \( F \) is small but not too small, then it pays to use cognitive resources for self control until the marginal benefit of cognitive resources is sufficiently high, then commit to taking the temptation off the table.

Houser et al [2010] have a very suggestive experiment indicating that delay in commitment may occur in practice: Subjects were paid for completing certain tasks while they had access to a web browser, and they could pay a cost to remove web access. The experimental instructions did not specify when and whether opportunities for commitment might occur in the future, so it is not clear what subjects believed about this, and whether the perfect foresight analysis applies. Thus although delay was observed, we cannot be certain from these experiments whether it is the type of delay predicted by this model. We hope that future experiments will shed more light on the type of delay that can occur with cognitive resource depletion.

Example 10: A Persistent Temptation with Non-Linear Benefits, Partial Linear Replenishment and General Depletion.

Our final variation on the persistent temptation problem drops the option to commit and examines the interplay between the benefit function \( m \) and the depletion function \( f \). To begin, note that regardless of the form of the benefit and depletion functions, if there is full replenishment the problem is stationary, so it is never optimal to wait for a while and then take. Defining the cost of self control to be \( g(\Delta) = (1 - \delta)(m(\overline{w}) - m_*(f(\overline{w}, \Delta))) \) (as in Section 4) we see that the policy of taking at once gives payoff \( P + m(\overline{w}) \), and resisting forever gives payoff \( m(f(\overline{w}, (1 - \delta \mu)S)) \), so resisting forever is optimal if
\[ P < m(f(\bar{w}, (1 - \delta \mu)S)) - m(\bar{w}) = -g((1 - \delta \mu)S) / (1 - \delta). \]

Next suppose that \( m(w) = \gamma w \) for \( w \geq 0 \) and \( m(w) = -\infty \) for \( w < 0 \). If \( P < -\gamma(1 - \delta \mu)S \) then as in our earlier analysis the optimum is to take the temptation immediately. If \( (1 - \delta(1 - \lambda))P > -\gamma(1 - \delta \mu)S \) the solution is to resist forever if this is consistent with resources remaining non-negative; this is the case if \( 1 - (1 - \delta \mu)S / \bar{w} > \lambda \). Otherwise, if \( (1 - \delta(1 - \lambda))P > -\gamma(1 - \delta \mu)S \) and \( (1 - (1 - \delta \mu)S) / \bar{w} < \lambda \) the solution is to resist until one more period of resistance would “exhaust the stock” (that is, make \( \tilde{w} < 0 \)) and then give in to the temptation.

Now consider the general case, normalizing to have linear replenishment. We assume that \( m \) and \( f \) are twice continuously differentiable, and that

\[
-\frac{dm}{dw}((1 - A)\bar{w} + Aw) \frac{\partial f}{\partial \Delta}(w, 0)
\]

(which is positive) is decreasing in \( w \); we call this “increasing marginal cost of self control.” Note that the marginal cost of control is increasing if \( m \) is strictly concave and \( f \) is linear. Equivalently we can write this condition in terms of second derivatives as

\[
-\frac{d^2m}{dw^2}((1 - A)\bar{w} + Aw) \frac{\partial f}{\partial \Delta}(w, 0) - \frac{dm}{dw}((1 - A)\bar{w} + Aw) \frac{\partial^2 f}{\partial \Delta \partial \bar{w}}(w, 0) < 0,
\]

If \( m \) is concave, this says that the cross-partial derivative of \( f \) should not be “too negative.” If the cross partial is strongly positive, then \( m \) need not be concave.

**Proposition 6:** Suppose there is increasing marginal cost of self control and there is strictly partial linear replenishment, \( 0 < \kappa \).

a) There is \( \tau > 0 \) such that if \( \tau < \tau \), there are \( |P_\tau| > |P| > 0 \) such that it is optimal to resist forever if \( |P| > P_\tau \); it is optimal to resist until period \( \infty > \ell > 1 \) then take if \( |P_\tau| > |P| > |P_\ell| \), and it is optimal to take immediately if \( |P_\ell| > P \).

b) Let \( P_0, P_\ell \) denote the limits of \( P_\tau, P_\ell \) as \( \tau \to 0 \). Let \( W_s \) be the solution to the differential equation

\[ \dot{W}_s = \kappa(\bar{w} - W_s) + \frac{\partial f}{\partial \Delta}(W_s, 0)(\rho + \eta)S, \]

and let \( W_\infty \) be the solution to \( 0 = \kappa(\bar{w} - W_\infty) + \frac{\partial f}{\partial \Delta}(W_\infty, 0)(\rho + \eta)S \). Then
\[ |\overline{P}_0| = -(\rho + \eta)S \int_0^\infty e^{-(\rho + \kappa)t} m'(W_\infty) \frac{\partial f}{\partial \Delta} (W_\infty, 0) dt, \]
\[ |\overline{P}_0| = -(\rho + \eta)S \int_0^\infty e^{-(\rho + \kappa)t} m'(\overline{w}) \frac{\partial f}{\partial \Delta} (\overline{w}, 0) dt \]

and if \( |\overline{P}_0| > |P| > |\overline{P}_0| \) then \( \hat{s} = \lim_{\tau \to 0} \tau \hat{\ell} \) is finite and strictly positive, and is determined by
\[ |P| = (\rho + \eta)S \int_0^\infty e^{-(\rho + \kappa)t} m'(\overline{w} - e^{-\kappa t}W_\kappa) \frac{\partial f}{\partial \Delta} (W_\kappa, 0) dt. \]

Remark 1: One way of reading this result is that the agent’s choice depends on the magnitude of \( P \), but recall that \( P = (1 - \delta^T)u_g - \delta^T u_b \) and \( S = (1 - (\delta \mu)^T)u_g - (\delta \mu)^T u_b \), so changing \( P \) implies changes in \( S \) and/or in \( \delta \) and \( \mu \) (or \( \rho \) and \( \eta \) in the continuous-time formulation) and any of these other changes will also matter for the decision.

Remark 2: To better understand the formulas given above, note that when depletion and benefits are both linear,
\[ |\overline{P}_0| = |\overline{P}_0| = (\rho + \eta)\gamma S \int_0^\infty e^{-(\rho + \kappa)t} dt = \frac{(\rho + \eta)}{(\rho + \kappa)} \gamma S \]
which is the same as the condition for the critical value of \( P \) given in equation (4).

Remark 3: To illustrate the fact that “concavity of the optimization” can come from any of the 3 functions \( f, m \) and \( r \), consider the case where \( f \) and \( m \) are linear, and \( r \) is piecewise linear:
\[ r(\overline{w}_n) = \overline{w}_n + \lambda_1(\overline{w} - \overline{w}_n) \text{ for } \overline{w}_n \in \overline{w}^*, \overline{w} \]
\[ r(\overline{w}_n) = \overline{w}_n + \lambda_2(\overline{w} - \overline{w}_n) \text{ for } \overline{w}_n < \overline{w}^*, \overline{w} \]
where \( \lambda_1 = 1 - \exp(-\kappa_1 \tau) \), \( \lambda_2 = 1 - \exp(-\kappa_2 \tau) \), and
\[ -\gamma S(\rho + \eta)/(\rho + \kappa_2) < P < -\gamma S(\rho + \eta)/(\rho + \kappa_1) \]

Then if the replenishment rate was fixed at \( \kappa_1 \) the agent would always resist, while if it was fixed at \( \kappa_2 \) the agent would take the temptation at once. We claim that the short-time-period solution with the piecewise linear replenishment function is to resist until resources fall to \( \overline{w}^* \). To see why, first consider the agent’s problem when the resource level is \( \overline{w}^* \). Resisting forever gives exactly the same payoff as when the replenishment rate is fixed at \( \kappa_2 \), and taking gives a higher payoff than with replenishment fixed at \( \kappa_2 \).
so since taking gives a higher payoff here than with $\kappa = \kappa_2$ the agent takes. When $w > w^*$, the gain from resisting for a short interval and then taking the temptation, instead of taking it now, is exactly as in the case $\kappa = \kappa_1$, so the agent resists; since this causes resources to decrease the agent will resist until resources fall to $w^*$.

The proof of Proposition 6 is in the appendix, but the intuition is simple. We first show that because of the increasing marginal cost of self control, and because resisting temptation lowers the stock next period, the gain to waiting one more period is monotone in the number of periods $\ell$ that the temptation has been resisted. Thus, if $P$ is small enough (sufficiently bad) relative to all the other parameters it is optimal to wait forever, if $P$ is close enough to 0 it is optimal to take at once, and for intermediate $P$ it is optimal to wait a while and then take. For an arbitrary length $\tau$ of the time period this intermediate region may be empty, but when $\tau$ is very small the concavity assumption ensures that it is non-empty.

Example 11: State-Dependent Marginal Cost

We assume full replenishment of willpower each period, so the stock of willpower is constant and thus irrelevant. As in the case of constant marginal cost, we assume both linear resource depletion and linear value of cognitive resources. However, we drop the assumption that the marginal benefit of cognitive resources are constant, and instead let them depend on the state. Specifically, we assume $m(y_n, \tilde{w}_n) = \Gamma(y_n)\tilde{w}_n$ and more specifically that the marginal benefit $\Gamma(y_n)$ in period 1 is $\Gamma > 0$ while from period 2 it is either $\Gamma$ or 0 depending on the first period choice.

In period 1 there is a choice of whether or not to pay a cost $F$; think of it as spending time learning self control, perhaps with the aid of a counselor or religious or spiritual advisor. If the cost is paid then there is no problem of self control at all in future periods, that is $\Gamma(y_n) = 0$; if the cost is not paid, the marginal benefit remains equal to $\Gamma$.

In period 2 the agent can decide whether to take or resist a simple temptation, with shorter-run player value $S$ and direct value $P$ for the long-run player, with $S > 0 > P$ and $P < -\Gamma S$, so that if the agent does not pay in period 1, it will be optimal to take in period 2.

Now we examine the decision in period 1. The future best value for the short run self is $\delta \mu S$, regardless of whether $F$ is paid today or not. Thus the temptation utility is
\( \delta \mu S \), the utility the SR associates with “pay” is \( \delta \mu S - F(1 - \delta \mu) \), so the foregone value of “pay” is \( F(1 - \delta \mu) \) and the self-control cost for this action is \( \bar{\Gamma}F(1 - \delta \mu) \). Hence it is optimal in the reduced form problem to “pay” whenever \( (\bar{\Gamma}(1 - \delta \mu) + (1 - \delta))F < -\delta P \).

In contrast, if paying \( F \) today makes taking tomorrow impossible, the foregone value of “pay” is \( \delta \mu S + F(1 - \delta \mu) \), so for some parameters (such as \( \mu \) close to 1) the commitment will not be optimal even though the arguably equivalent “training” action would be. The difference between commitment and lowering control costs is a consequence of our assumption that the shorter-run selves are strategically naïve, so that the shorter-run player is unconcerned by any action that leaves the feasible set unchanged. Models with non-naïve shorter-run players may also be of interest, but they are much more complicated. \(^{34}\)

To make this example simple, we kept the stock of willpower constant and assumed that the first-period action had a direct effect on the cost of self control in the second period. Similar effects could be obtained if we allowed the replenishment function \( r \) to depend on the action as well as on the end of period willpower, and let the benefits of cognitive resources be slightly concave (so that the cost is slightly convex). Specifically, suppose that acting in the first period increases the willpower stock from 1 to \( 1 + w^* \), and that the benefit of cognitive resources \( w \) is \( w^\alpha \) for some \( \alpha \in (0,1) \). Then if the agent does not act in the first period, the cost of resisting second period temptation is \( 1 - (1 - \delta \mu S)^\alpha \geq \alpha \delta \mu S \), while the cost if the agent acts goes to 0 with \( w^* \).

In the online appendix we discuss some of the complications that arise when the evolution of resources can be state-dependent, and can depend directly on actions and not only on the foregone value. That appendix also shows that when we rule out the kind of endogenous changes in willpower explored in Example 11, we can show that the actions that maximize the long-run self’s objective function correspond to the equilibrium of a game in which a benevolent but patient long-run self faces a sequence of shorter-run selves who live for a random length of time. In this game decisions are made by the shorter-run selves, but the long-run self can alter the preferences of the shorter-run selves.

\(^{34}\) This example suggests that non-naivety is necessary to capture St. Augustine’s request “give me chastity and continence, but not yet.”
by undertaking “self-control” actions that in general lower the utility of the shorter-run selves.

7. Conclusion

Allowing shorter-run selves to live more than a single period provides a natural way to capture the way preferences change as the “period” becomes shorter. This lets us explain why commitments to avoid far-off temptations are less costly, and more attractive, than commitments to avoid more imminent ones, and lets us explain the subjective interest rates decline with delay; it also provides a natural parameterization of the effect of the length of the time interval between potential decision nodes. This is important because the concept of a discrete time period in these decision problems is simply a convenient construction; the real-time length of delay is what should matter for decision making.

When the marginal cost of self control is constant, the agent’s decision problem is not affected by the timing of when self-control costs are incurred, and there is no need for the model to track the stock of the agent’s cognitive resources: As we saw, the model with linear replenishment, benefits, and depletion is equivalent to the “state-free” model. However, when non-linearities matter, so does the timing of self-control decisions and costs, and the willpower stock provides a way to model the “spillover” from one period’s self control to future control costs. Tracking the stock of willpower allows simultaneous decisions to be about the same as almost simultaneous ones, and lets us explain why agents may choose to “resist and then give in” and “wait to commit.”

We explored some but far from all of the many possible ways to model these non-linearities, and there is ample scope for future work on this. In particular we have looked for plausible properties, such as insensitivity to minor changes in timing; it would be useful to compile these properties in a set of clear and readily interpretable axioms to better understand the universe of models that satisfy them. Also, it would be good to extend the qualitative analysis here by exploring the extent to which we can find, for each individual agent, a stable constellation of preference parameters that fits that agent’s quantitative behavior across a range of problems. This was done to a limited extent in Fudenberg and Levine [2010] for the model where shorter-run selves live a single period, although the calibration was for the median subject across a number of different
experiments as opposed to individual subjects observed in many different treatments. However, several of the experiments studied there are better fit by allowing shorter-run selves to have random lifetimes. Finally, while our analysis here has presumed that the cost function $c$ (in the first part of the paper) or the functions governing the evolution of cognitive resources (in the second part) are fixed and constant, it is straightforward to allow it to have random variation, so long as this variation is exogenous: For example, each period $t$ the cost of foregoing $\Delta$ units of value could be $\Gamma_t$ where the $\Gamma_t$ are i.i.d. according to some known distribution.

References


Appendix

Example 3: Derivation of marginal interest rates

From the indifference condition in the text, we may compute

\[
1 - \delta + \Gamma(1 - \delta \mu) \over (1 - \delta)\delta^{n-1} + \Gamma(1 - \delta \mu)(\delta \mu)^{n-1} = c_n, \quad \text{so setting } n = t / \tau \text{ and taking the limit } \tau \to 0 \text{ shows that}
\]

\[
MR_t = \lim_{\tau \to 0} \log \left( \frac{c(t/\tau) + 1}{c(t/\tau)} \right) / \tau
\]

\[
= \lim_{\tau \to 0} \log \left( \frac{1 - \delta + \Gamma(1 - \delta \mu)}{(1 - \delta)\delta^t + \Gamma(1 - \delta \mu)(\delta \mu)^t} \right) / \tau
\]

\[
= \lim_{\tau \to 0} \log \left( \frac{1 (1 - \delta) + \Gamma(1 - \delta \mu)(\delta \mu)^t}{\delta (1 - \delta) + \Gamma(1 - \delta \mu)(\delta \mu)^t} \right) / \tau
\]

\[
= \rho + \Gamma \lim_{\tau \to 0} (1 - \delta \mu)(\exp(-\eta t)(\mu^{-1} - 1))/(1 - \delta) \tau = \rho + \Gamma (\rho + \eta) \exp(-\eta t) \eta / \rho
\]

Proof of Theorem 4

\textbf{Proof:} Recall that } \bar{w}_n = w_n - \Delta_n. \text{ With linear replenishment } w_{n+1} = \bar{w}_n + \lambda(\bar{w} - \bar{w}_n) = w_n - \Delta_n + \lambda(\bar{w} - w_n + \Delta_n). \text{ Define the willpower deficit } D_n = \bar{w} - w_n, \text{ then}

\[
D_{n+1} = (1 - \lambda)D_n + (1 - \lambda)\Delta_n = (1 - \lambda)^{n+1}D_0 + \sum_{n'=0}^{n-1} (1 - \lambda)^{n'+1}\Delta_n.
\]

Recall that the average value of cognitive resources in the linear case is

\[
M = (1 - \delta)\sum_{n=0}^{\infty} \delta^n \gamma w_n
\]

It follows that the total value of cognitive resources is
\[ M / (1 - \delta) = \gamma \left[ \bar{w} - \sum_{n=0}^{\infty} \delta^n D_n \right] \]
\[ = \gamma \left[ \bar{w} - \sum_{n=0}^{\infty} \delta^n \left( (1 - \lambda)^n D_0 + \sum_{n'=1}^{n} (1 - \lambda)^{n-n'} \Delta_{n'} + \Delta_{n+1} \right) \right] \]
\[ = \gamma \left[ \bar{w} - \sum_{n=0}^{\infty} \delta^n \left( (1 - \lambda)^n D_0 - \sum_{n'=0}^{\infty} \delta^{n'} \sum_{n'=0}^{n'} (1 - \lambda)^{n-n'} \Delta_{n'} \right) \right] \]
\[ = \gamma \left[ \bar{w} - \sum_{n=0}^{\infty} \delta^n (1 - \lambda)^n D_0 - \sum_{n'=0}^{\infty} \delta^{n'} \Delta_n \sum_{n=n'}^{\infty} (\delta(1 - \lambda))^{n-n'} \right] \]
\[ = \gamma \left[ \bar{w} - \sum_{n=0}^{\infty} \delta^n (1 - \lambda)^n D_0 - \frac{1}{1 - \delta(1 - \lambda)} \sum_{n'=0}^{\infty} \delta^{n'} \Delta_n \right] \]

Hence if we define \( \Gamma = (1 - \delta) \gamma / (1 - \delta(1 - \lambda)) \) we see the equivalence to the simple linear model without replenishment. \( \square \)

**Proof of Proposition 6**

**Proof:** Suppose the agent resists for \( \ell \) periods then gives in. Let \( w^\ell_n \) be the corresponding time path of cognitive resources. Note that this is a weakly decreasing function of \( \ell \), strictly decreasing for \( n > \ell \). The resulting average value is

\[ (1 - \delta) \left( \sum_{n=0}^{\ell-1} \delta^n m(f(w^\ell_{n+1}, (1 - \delta \mu))S) + \left( \sum_{n=\ell}^{\infty} \delta^n m(w^\ell_{n+1}) \right) + \delta^\ell P, \]

and the bigger is \( \ell \) the smaller is the first term and the larger is the second (recall that \( P \) is negative). This implies that a necessary condition for an optimal \( \ell \) is that the value for \( \ell + 1 \) is no bigger, and that a sufficient condition for \( \ell = 1 \) optimal is that the value for \( \ell + 1 \) is lower. Let us look at the value at \( \ell \) minus the value at \( \ell + 1 \)

\[ D(\ell) = (1 - \delta) \left( \sum_{n=0}^{\ell-1} \delta^n m(f(w^\ell_{n+1}, (1 - \delta \mu))S) + \left( \sum_{n=\ell}^{\infty} \delta^n m(w^\ell_{n+1}) \right) + \delta^\ell P \right) - (1 - \delta) \left( \sum_{n=0}^{\ell} \delta^n m(f(w^{\ell+1}_{n+1}, (1 - \delta \mu))S) + \left( \sum_{n=\ell+1}^{\infty} \delta^n m(w^{\ell+1}_{n+1}) \right) \right) - \delta^{\ell+1} P \]

Note that for \( n \leq \ell \) we have \( w^\ell_n = w^{\ell+1}_n \), so we can write this difference as

\[ D(\ell) = (1 - \delta) \delta^\ell \left[ (m(w^\ell_{\ell+1}) - m(f(w^\ell_{\ell+1}, (1 - \delta \mu))S)) + \sum_{n=\ell+1}^{\infty} \delta^{n-\ell} (m(w^\ell_{n+1}) - m(w^{\ell+1}_{n+1}) + P \right] \]

Observe that because there is partial replenishment \( w^\ell_{\ell+n} \) strictly decreases in \( \ell \).

We now use the assumption of increasing cost of self control to conclude there is a \( \tau \) such that for \( \tau < \tau \) each individual term in \( D(\ell) \) is strictly increasing in \( \ell \), and hence that \( D(\ell) \) is strictly increasing. The first term

\[ m(w^\ell_{\ell+1}) - m(f(w^\ell_{\ell+1}, (1 - \delta \mu))S)) = m(f(w^\ell_{\ell+1}, 0)) - m(f(w^\ell_{\ell+1}, (1 - \delta \mu))S) \]

strictly
decreases in \( w_{\ell+1} \) from increasing marginal cost of self control at \( A = 1 \) and the fact that \( (1 - \delta \mu)S \to 0 \) as \( \tau \to 0 \). Since \( w_{\ell+1} \) decreases in \( \ell \), these differences increase.

For the terms in the sum, since \( n \) runs from \( \ell + 1 \) to \( \infty \), the arguments \( w_{\ell+1} + \ell \) have the form \( 1 - \lambda \ell^{\ell-1}(\bar{w} - w_{\ell+1}) \). The individual terms have the form:

\[
(1 - \lambda \ell^{\ell-1}(\bar{w} - w_{\ell+1}))
\]

Putting this together, we have

\[
m(\ell) = m(\ell) - m(\ell + 1) = m(\ell - (1 - \lambda)\ell^{\ell-1}(\bar{w} - w_{\ell+1})) - m(\ell - (1 - \lambda)\ell^{\ell-1}(\bar{w} - f(w_{\ell+1},(1 - \delta \mu)S)))
\]

\[
m(A\bar{w} + (1 - A)w_{\ell+1}) - m(A\bar{w} + (1 - A)f(w_{\ell+1},(1 - \delta \mu)S)))
\]

where \( A = (1 - \lambda)\ell^{\ell-1} \). When \( \tau \) is small enough, increasing marginal cost of self control implies that this expression is decreasing in \( w_{\ell+1} \) and so increasing in \( \ell \) when \( \tau \) is small enough.

Because \( w_n \) is bounded below by the steady state, \( D \) is bounded above as a function of \( \ell \). If \( P \) is large enough in absolute value (it is negative) given all the other parameters then this expression is negative for all \( \ell \), and it is optimal to wait forever; let \( P_\tau \) be the smallest such \( P \) in absolute value. If \( P \) is small enough in absolute value, this expression is positive for all \( s \) and it is optimal to take immediately, let \( P_\tau \) be the largest such \( P \) in absolute value.

Next we assume that \( \tau \) is small, and show that \( |P_\tau| > |P_\tau| \). Observe that

\[
w_{\ell+1} = \bar{w} - (1 - \lambda)\ell^{\ell-1}(\bar{w} - w_{\ell+1})
\]

\[
w_{\ell+1} = \bar{w} - (1 - \lambda)\ell^{\ell-1}(\bar{w} - f(w_{\ell+1},(1 - \delta \mu)S))
\]

Let \( d(s, \tau) = D(s / \tau) / \delta^{s / \tau}(1 - \delta) \) with \( W_s^t \equiv w_{s / \tau}^t \). Then

\[
d(s, \tau) = (m(W_{s+\tau}^s) - m(f(W_{s+\tau}^s,(1 - \delta \mu)S))) + \sum_{n=2}^{\infty} e^{-\rho \mu}(m(W_{s+n\tau}^s) - m(W_{s+n\tau}^{s+\tau}))) + P
\]

where \( W_{s+\tau}^s - W_{s+n\tau}^{s+\tau} = (1 - \lambda)^{n-1}[W_{s+\tau}^s - f(W_{s+\tau}^s,(1 - \delta \mu)S)] \) and

\[
W_{s+\tau}^s = \bar{w} - (1 - \lambda)^{n-1}(\bar{w} - W_{s+\tau}^s)).
\]

The first term of \( d \) converges to zero as \( \tau \to 0 \), and since \( m, f \) are differentiable the sum converges to
\[ d(s) \equiv (\rho + \eta)S \int_0^\infty e^{-(\rho + \kappa)t} m'(\bar{w} - e^{-\kappa t}(\bar{w} - W_s)) \frac{\partial f}{\partial \Delta}(W_s,0)dt \]

where \( W_s \) is the solution to the differential equation

\[ \dot{W}_t = \kappa(\bar{w} - W_t) + \frac{\partial f}{\partial \Delta}(W_t,0)(\rho + \eta)S \]

with initial condition \( W_0 = \bar{w} \). Thus we have \( d(s) \equiv \lim_{\tau \to 0} d(s,\tau) \). Recall that \( D \) is strictly increasing, and that \( \hat{\ell} = 1 \) is optimal if and only if \( D(1) \geq 0 \). As \( \tau \to 0 \) this is equivalent to

\[ d(0) = -(\rho + \eta)S \int_0^\infty e^{-(\rho + \kappa)t} m'(\bar{w}) \frac{\partial f}{\partial \Delta}(\bar{w},0)dt + P = -\frac{(\rho + \eta)m'(\bar{w})}{(\rho + \kappa)} \frac{\partial f}{\partial \Delta}(\bar{w},0) \geq 0 \]

Similarly \( \hat{\ell} = \infty \) is optimal if and only if \( \lim_{\ell \to \infty} D(\ell) \leq 0 \), and so when \( d(\infty) \leq 0 \).

Finally, resisting for a while and the taking, that is, \( 1 < \hat{\ell} < \infty \), is optimal if and only if \( D(\hat{\ell} - 1) \leq 0, D(\hat{\ell}) \geq 0 \), hence \( d(\hat{s}) = 0 \). This gives the characterization of the optimum in the Proposition. Finally, the assumption that the marginal cost of self control is increasing implies \( d(s) \) is strictly increasing, so

\[ |P_{\hat{\ell}}| = -(\rho + \eta)S \int_0^\infty e^{-(\rho + \kappa)t} m'(\bar{w}) \frac{\partial f}{\partial \Delta}(\bar{w},0)dt < \]

\[ -(\rho + \eta)S \int_0^\infty e^{-(\rho + \kappa)t} m'(W_\infty) \frac{\partial f}{\partial \Delta}(W_\infty,0)dt = |P_{\hat{\ell}}|, \]

and hence \( |P_{\hat{\ell}}| < |P_{\hat{\ell}}| \) must hold for all sufficiently small \( \tau \).