Timing and Self-Control

Drew Fudenberg* and David K. Levine**

First Version: August 5, 2009

This Version: September 6, 2010

Abstract: The dual self-model of self-control with one-period lived short-run selves is excessively sensitive to the timing of shocks and to the interpolation of additional "no-action" time periods in between the dates when decisions are made. We show that when short-run selves have a random length of time this excess sensitivity goes away. We consider both linear and convex cost of self-control models, illustrating the theory through a series of examples. We examine when opportunities to consume will be avoided or delayed; we consider the way in which the marginal interest declines with delay, and we examine how preference "reversals" depend on the timing of information. To accommodate the combination of short time periods and convex costs of self control we extend the model to treat willpower as a cognitive resource that is limited in the short run.

* Department of Economics, Harvard University

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

We thank Attila Ambrus, Yuichiro Kamada, and Tomasz Strzalecki for helpful comments, and NSF grants SES-0646816 and SES-0851315 for financial assistance.

1. Introduction

Models of long-run planning and short-run impulsive selves 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, have a fixed horizon for the short-run self, and so they cannot explain overwhelming evidence that the length of delay has a continuous impact on decisions. These models also make implausible predictions about the value of a commitment that avoids temptation: If the commitment can be made "the period before" the temptation would be faced, it is essentially costless, while the commitment is valueless if it must be made in the same period that the temptation would occur.

As an example of the continous effect of delay, consider an experiment from Myerson and Green [1995]. Subjects were asked to state how much hypothetical money c_t they would need right now to make the indifferent to receiving a hypothetical $1,000 after a delay of length t. By considering several different delays t_1, t_2 and so forth, we can impute marginal interest rate to the subjects, and with standard geometric discounting these marginal interest rates should be time invariant. In the Myerson and Green data, these marginal interest rates are[1]

months	marginal interest rate
0.23	132
1	82.1
6	40.9
12	42.7

months	marginal interest rate
36	26.0
60	8.0
120	9.4
300	6.6

[1] 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 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.

In the long-run/short-run self model with a fixed period length, or in the quasi-hyperbolic discounting model, the initial marginal interest rate may be high. However, these models predict that all subsequent marginal interest rates are equal and low, as opposed to the gradual decay seen in the data.

To account for the continuous effect of delay, and to explore the implications of the timing of decisions, we suppose that the short-run self or selves are not completely myopic and do value future utility, but less so than the long-run self, either because there is a succession of short-run selves with random lifetimes, or a single short-run self whose discount factor is lower than that of the long-run self. This enables us to maintain the underlying strength and simplicity of the dual-self model and its ability to account for many phenomenon, while accounting for the continuous effect of delay. It also lets us model cases where agents are tempted by future consumption, as in Noor [2007], and explain why this temptation is most significant with respect to payoffs that are relatively soon.

The key modeling question in extending the dual-self model to non-myopic short-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 short run player foregoes in the current period. When short run players have a longer horizon, their expectations about future payoffs can matter as well, and we suppose that the control cost of implementing a given action depends on how much that action lowers the highest possible average present value the short-run self could obtain from the current period on. This specification is consistent with the interpretation that the short-run selves are strategically naïve and evaluate foregone utility assuming that no self control will be used in the future.

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 the standard case, as it is consistent with both the independence axiom for choices over lotteries and the weak axiom of revealed preference. Our first application is to 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 simultaneously prefer to resist a simple temptation when this is a once-and-for all choice, and prefer to give in when the 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]. In Example 2 we show that the model explains the Myerson-Green data mentioned above, and in Example 3 we explain Della Vigna et al's [2010] finding that the people who are willing to incur costs to avoid contributing to charity are the ones who would contribute less when avoidance is not possible.

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. In Example 4 we study the role of the short-run player's effective horizon in determining when the agent will choose a menu that includes tempting choices and when the agent will prefer menus where these temptations are not available. Examples 5 and 6 then 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 raises the following issues: First, 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. Second, since the length of the time period 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 these 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: Here we can assume that the foregone value is simply subtracted from the beginning of period stock, with the cost of self control corresponding to the difference between the benefits obtained from the full stock and the benefits obtained after the reduction, so that a convex cost of self control corresponds to a concave benefit function.

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 5 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 6 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 behavior can be modeled without the use of the cognitive resources .

However when the agent makes multiple decisions and the marginal cost of self control varies then 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. However, since cognitive resources are not observed directly and have no natural units, there are many equivalent representations of the same preferences. Theorem 8 shows that that 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 8 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 9 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 10 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 9 and 10 simplify by assuming no replenishment of resources at all, which is unrealistic but makes it easier to highlight the logic of the argument. Example 11 re-examines the persistent temptation from example 9 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 12 shows the issues involved in relaxing our assumption that the "willpower technology" is fixed and cannot be changed by the agent.

In the first part of the paper we directly assume an objective function with a cost of self-control or benefit from cognitive resources. In Section 7 we show that when we rule out the kind of endogenous changes in willpower explored in Example 12, we can derive the objective function from a game in which a benevolent but patient long-run self faces a sequence of short-run selves who live for a random length of time. In this game decisions are made by the short-run selves, but the long-run self can alter the preferences of the short-run selves by undertaking "self-control" actions that in general lower the utility of the short-run self.

While this paper is the first to study random short-run player lifetimes in a self-control model, 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.[2] 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] develops axioms for a two-period choice problem 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 and not developing the model's recursive extension, and he does not investigate how this "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. This is connected to our example of menu choices with patient short-lived selves. Also in the context of two-period models, Noor and Takeoka [2010] develop axioms for choices on menus that correspond to convex costs of self control, Brocas and Carrillo [2008] explain the covariance of effort and consumption by assuming the long-rub self has incomplete information on the short-run self's cost of effort, and Chaterjee 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.[3] 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 most of the paper we take this control cost as exogenous; Section 7 shows how this maximization problem corresponds to the equilibrium of a game between the agent's "long-run self" and a sequence of short-run selves. 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 δ, and

---

[2] Our [2010] paper explains the same data as a consequence of a convex cost of self control.
[3] These two papers differ in how the long run self views the possible second period preferences.

suppose that the discount factor of the short-run self is $\delta\mu$, where $\mu \in [0,1)$ can also be interpreted as the survival probability of the current short-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 τ units of calendar time, and suppose that $\delta = \exp(-\rho\tau)$ and $\mu = \exp(-\eta\tau)$; for small τ we will make use of the approximations $\mu \approx 1 - \eta\tau$, $\delta \approx 1 - r\tau$.

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 $a : Y \to A$ that satisfies $a(y_n) \in A(y_n)$. Dynamics are Markovian, and described by a probability distribution over states next period conditioned by the current state and action according to a stochastic kernel $\pi(y_{n-1}, a_{n-1})[dy_n]$ which is a measurable function of y_{n-1}, a_{n-1}.

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 $\mathbf{a}$ from histories to actions, so that for each history h_n the strategy specifies an action in $\mathbf{a}(h_n) \in A(y_n)$.

The short-run self (or selves) have get 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 as we consider time periods of various length τ we hold $u(y_n, a_n)$ fixed. The objective of the long run player is the average present value of these short-run self utilities minus a cost of self control that is defined with reference to the maximum possible average present value for the short-run players. To define this maximum we first define the expected average present value of the short-run players. Let $E_{\mathbf{a},h_n}$ be the conditional expectation generated by the plan $\mathbf{a}$ and the stochastic kernel, conditional on the history h_n. The expected average present value of the short-run self from period n on under $\mathbf{a}$ is given by

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

or equivalently

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

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

Assumption SR0*:*
$E_{\mathbf{a},h_n}(1-\delta\mu)\sum_{1=0}^{\infty}(\delta\mu)^{\ell} u(y_{n+\ell},a_{n+\ell}) = (1-\delta\mu)\sum_{\ell=0}^{\infty}(\delta\mu)^{\ell} E_{\mathbf{a},h_n} u(y_{n+\ell},a_{n+\ell})$

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

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

Theorem 1*:* $\max_{\mathbf{a}} U(h_n,\mathbf{a}) = \bar{U}(y_n)$

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

Specifically, we call $\bar{U}(y_n)$ the *temptation value* for the short-run self at n starting at state y_n The *foregone value* is then

$$\Delta(y_n,a_n) = \bar{U}(y_n) - \Big((1-\delta\mu)u(y_n,a_n) + \delta\mu\int_Y \bar{U}(y_{n+1})\pi(y_{n+1} \mid y_n,a_n))[dy_{n+1}]\Big).$$

The foregone value is *recursive* in sense that it depends on the future only through the future temptation utility, and attributes self-control costs to each action as it occurs as opposed to entire contingent plans. The key idea is that the current cost of self control depends on the current foregone utility, and this is only lost because the current choice of action a_n either lowers $u(y_n, a_n)$ or changes the distribution of future states $\pi(y_{n+1} \mid y_n, a_n)$ and thus the best utility the short-run self could hope to get in the future. Future actions that lead to less value than $\bar{U}(y_n)$ incur costs at the time at which they are taken.

One interpretation of the foregone value is that the term

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

is the short-run self's prediction of the expected continuation payoff, and that short-run self predicts that no self control will be used in the future. Under this interpretation the short-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.[4]

Notice that by the principle of optimality any plan that solves $\max_{\mathbf{a}} U(h_n, \mathbf{a})$ must also solve the dynamic programming problems

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

Thus we have

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

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 linear in the foregone value Δ: it is given by $\Gamma\Delta(y_n, a_n))$, where as above Δ is measured in units of average present value, and the scalar constant $\Gamma \geq 0$ is independent of the state.

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

[4] 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.

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 cost of self-control over the next few periods if periods are short.

The agent's objective function is defined by the expected average present value of short-run utility net of the self-control cost

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

Note first that this reduces to the linear-costs version of Fudenberg and Levine [2006], [2010]) when $\mu = 0$. Next, note that the foregone value Δ is not normalized by $1-\delta$. This is in order that the limit as $\tau \to 0$ be well behaved. To understand why, suppose a once and for all action is taken that lowers utility in every period by. This action has foregone value of $\Delta = 1$ independent of period length, and its control cost is Γ regardless of period length. In contrast, 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 is very small when periods are short. If the long-run player undertakes an infinite sequence of such actions the overall cost is $\Gamma(1-\delta\mu)/(1-\delta)$, and for this reason 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. 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 s, except in the case $\mu = 0$, where the short-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; when $\mu = 1$ the long-term commitment is just as costly as the series of short-term ones. Finally, note that sending the time period τ to 0 sends $\mu \to 1$ but changes other parameters as well: Lowering utility by 1 forever starting immediately still costs Γ, 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 long-term commitment and the series of short-term ones stems not from the period length but from the greater impatience of the short-run self.

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

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

Notice that this is a Markov decision problem, so it has a Markov solution. That is, there is an optimal plan in which the choice of action depends only on the current state y_n.

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 short-run self and P for the long-run self of this stream are $(1-(\delta\mu)^N)u_g - (\delta\mu)^N u_b$ and $P = (1-\delta^N)u_g - \delta^N u_b$. Our interest lies in the case $S > 0, P < 0$ so that the short-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 μ to 1 holding all other parameters fixed. However, the reason for interest in the case of μ 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 τ, even though $\mu = e^{-\eta\tau} \to 1$.

Example 1: Resisting Temptation 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 become easier to resist.

Next suppose that the temptation is *persistent:* if it is resisted the same choice is faced again in the next period.[5] This model describes for example the temptation to consume a durable good such as a bottle of wine. Once the agent consumes the substance it is gone, but if he does not consume, the substance is still there the next period,[6] then the problem is stationary; in the state $y_n = 1$ the best possible value for the short-run self is *S* and the best continuation if the temptation is resisted is $\delta\mu S$, so the foregone value is $\Delta = (1-\delta\mu)S$, resisting costs $\Gamma(1-\delta\mu)S$ each period, so 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. Under the same conditions, the agent would choose to face the simple temptation in the next period rather than face the persistent one. This condition gets increasingly difficult to satisfy as $\mu \to 1$ holding all other parameters fixed, which corresponds to sending the birth parameter η to 0.

Since the main reason for large μ is that periods are short, it is more 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

---

[5] 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.

[6] We assume here that the consumption option is all or none, perhaps the wine will spoil once opened.

$(\rho+\eta)\Gamma S > \rho|P| > \rho\Gamma S$. 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 ΓS 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

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

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 whether to take at once or resist forever is monotone in ℓ: There is some $\overline{\ell}$ (possibly 0 or infinity) such that the optimum is to take at once if $\ell < \overline{\ell}$ and resist forever if $\ell > \overline{\ell}$. Intuitively, this is because the short-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 short-run player is increasing in the delay length. If the decision is imminent, there is not much point in trying to avoid temptation by making a commitment, as the temptation already exists, but it may be worthwhile to commit now to resist future temptations. This is related to Noor's [2007] 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.

Finally consider an initial-period choice of whether to accept or reject K simple temptations, the first one in period 1, the second in period $1+K$, the third in period $1+2K$ and 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.[7] 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: Declining Marginal Interest Rates

Now we show how the model generates the sort of declining interest rates seen in the Myerson and Green data. We take utility to be linear in consumption and without further loss of generality we set $u(c) = c$.[8] 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 Appendix we extend this to the interest rate used at time 1 to discount between any two periods n and ℓ, which is closer to the long-run player's rate of time preference ρ because future consumption is less tempting.) Observe that if the long-run self is indifferent between one unit now and c_n units later then since $\mu < 1$ the initial short-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

[7] 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.

[8] 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. Our earlier papers explained this with endogenously-determined "mental accounts;" a similar explanation should be possible here but that remains a project for future work.

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

Note 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, and that as $\mu \to 0$ we have

$$c_n \to \frac{1-\delta+\Gamma}{(1-\delta)\delta^{n-1}}=\frac{1+\Gamma/(1-\delta)}{\delta^{n-1}}.$$ [9]

To relate the model back to Myerson and Green we compute the instantaneous interest rate for consumption decisions at real time $t = n\tau$ rate by letting the period length τ go to 0. In the Appendix we show that

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

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 η. In the case of a short-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 μ we get a more gradual decline, as we see in the data. [10]

Example 3: Door-to-Door Sales

This example has two decisions: First, whether to avoid a tempting opportunity, and second, whether to give in to temptation if it was not avoided. (For concreteness, think of the avoidance activity as avoiding a door-to-door salesman.)

---

[9] This is the same answer as in Fudenberg-Levine [2006] once we correct for the difference between average and total present value: In Fudenberg-Levine temptations were measured with respect to total and not average utility so the linear coefficient γ in that paper corresponds to $\Gamma/(1-\delta)$ here.

[10] The Myerson-Green data reported above is for money payments and consumption; The Web Appendix explains how to extend the model to allow for consumption financed by income and savings, and shows that for small income increments marginal interest rates are the same as in this linear model. Note that in the data the marginal interest rate declines, but not exponentially. However, if the aggregate is averaged over individuals with different exponential rates of decline, then it will not in fact decline exponentially. The individual data provided us by Myerson and Green is noisy, but more closely exponential than the aggregate data. Note also that the exponential rate of decline follows from the assumption of a constant hazard rate for the death of the short-run self, which, while convenient, is not essential.

The first point, which does not require that short run players can live more than a single period, is that costly self-control leads to a non-monotonicity: If the temptation is very high or very low then the opportunity will not be avoided but it may be avoided for intermediate levels of temptation. The intuition is that when the opportunity is very good there is little conflict between the long-run and short-run self, so the opportunity should be taken advantage of and not avoided. When the opportunity is very bad, the short-run self will not indulge much in it, and so it is not worth paying a fixed cost for avoidance. However, in the intermediate case there may be a more severe conflict between long-run and short-run self, so the long-run self may choose to commit in order to avoid the temptation.

The second substantive point does rely on the short run players sometimes living more than one period: The appeal of the costly avoidance option depends on the time interval between the decision to avoid and when the temptation would arrive, because the current short run self is less willing to go along with avoidance when the possible temptation would arrive soon.

The example is very simple and stylized. In period 1 a cost $F \geq 0$ may be paid or not; think of it as not being at home when the salesman calls. If the cost is paid, the utility in all subsequent periods is 0. If the cost is not paid, then in period 2 a decision must be made on whether to purchase from the salesman. If the purchase is made the utility in period 2 is B, otherwise it is zero. In period 3 if the purchase was made it must be paid for, resulting in a disutility of -1.

To solve the model recursively, we first compute temptation values in each period and state, and then compute the agent's objective function. At that point all that is left is to solve the various inequalities to see when each action is best, which we do in the appendix.

We begin by computing the temptation value in the last period in which action is possible, namely in period 2 when the avoidance cost has not been paid. Here the short-run self's average present value from doing nothing is zero, and that of purchasing is $(1-\delta\mu)(B-\delta\mu)$, so $\bar{U}_2 = (1-\delta\mu)\max\{0, B-\delta\mu\}$. In the initial state if F is chosen the short-run self's value is $-F(1-\delta\mu)$, while if it is not it short-run player value is $\delta\mu\bar{U}_2$.

If $B-\delta\mu<0$, then also $B-\delta<0$ so in period 2 the optimum is not to purchase, which incurs no cost of self-control.

Now suppose $B-\delta\mu>0$. Then if the purchase is made long-run utility is $B-\delta$; if the long run player chooses not to purchase he incurs temptation cost of $\Gamma(1-\delta\mu)(B-\delta\mu)$, so the purchase will be made in period 2 when $(1-\delta)(B-\delta)\geq-\Gamma(1-\delta\mu)(B-\delta\mu)$. If the avoidance cost is paid in the first period, the short-run self's average value is $-(1-\delta\mu)F$, while the temptation value is $\delta\mu\bar{U}_2$, so the average present value of avoidance is

$$-(1-\delta)F-\Gamma\left(F(1-\delta\mu)+\delta\mu\bar{U}_2\right)$$
$$-(1-\delta)F-\Gamma\left(F(1-\delta\mu)+\delta\mu(1-\delta\mu)\max\{0,B-\delta\mu\}\right)$$

and avoidance is optimal if this is higher than the discounted average value of long-run player utility in period 2, which is $\max\{(1-\delta)(B-\delta),-\Gamma(1-\delta\mu)(B-\delta\mu)\}$

Denote the strategy of not paying the avoidance cost and not purchasing as $\mathbf{a}^0$, of not paying the avoidance cost and purchasing as $\mathbf{a}^1$, and of paying the avoidance cost as $\mathbf{a}^F$. In the Appendix we prove the following characterization of the optimal decision rule:

Proposition 3: *Set*

$$F^*=\frac{\Gamma\delta^2(1-\delta\mu)(1-\delta)(1-\mu)^2}{(1-\delta+\Gamma(1-\delta\mu))^2}$$

If $F\geq F^*$ *then* $\mathbf{a}^0$ *is optimal for*

$$B\leq\delta\frac{1-\delta+\Gamma(1-\delta\mu)\mu}{1-\delta+\Gamma(1-\delta\mu)}\equiv B^*$$

and $\mathbf{a}^1$ *is optimal if* $B\geq B^*$. *If* $F\leq F^*$ *then*

$$\bar{B}\equiv\delta\frac{1-\delta+\Gamma(1-\delta\mu)\mu^2}{1-\delta+\Gamma(1-\delta\mu)\mu}-\frac{1-\delta+\Gamma(1-\delta\mu)}{\delta(1-\delta+\Gamma(1-\delta\mu)\mu)}F\geq$$
$$B^*\geq\delta\mu+\frac{1-\delta+\Gamma(1-\delta\mu)}{\Gamma(1-\delta\mu)\delta(1-\mu)}F\equiv\underline{B}$$

and $\mathbf{a}^0$ *is optimal for* $B \leq \underline{B}$, $\mathbf{a}^F$ *is optimal for* $\underline{B} \leq B \leq \overline{B}$, *and* $\mathbf{a}^1$ *is optimal for* $B \geq \overline{B}$.

Note that the RHS inequality in $\underline{B} \leq B \leq \overline{B}$ gets harder to satisfy as $\mu \to 1$ or as $\Gamma \to 0$. In the former case the interests of the short-run self are nearly aligned with those of the long-run self, while in the second the short-run self defers to the wishes of the long-run self. In either case, paying F is just an expensive way to not buy. Paying a small F is attractive as $\mu \to 0$ as here the first SR self is not very tempted by the second period outcome so it is cheap to get him to agree to a commitment that will probably bind on the next self.

We can also use this example to illustrate the effect of changing the amount of time between the two decisions, while keeping other factors constant. Taking the limit as $\tau \to 0$ we find

$$
\begin{aligned}
F^* &= \frac{\Gamma\delta^2(1-\delta\mu)(1-\delta)(1-\mu)^2}{(1-\delta+\Gamma(1-\delta\mu))^2} \approx \frac{\Gamma(1-e^{-(\rho+\eta)s})(1-e^{-\rho s})(1-e^{-\eta s})^2}{(e^{-\rho s}+\Gamma e^{-(\rho+\eta)s})^2} \\
&\approx \frac{\Gamma(\rho+\eta)\rho\eta\tau^4}{\Gamma(\rho+\eta)^2\tau^2} \to 0
\end{aligned},
$$

where the first line holds fixed calendar time s and the second assumes that the decisions are made in consecutive periods. From this we see that it is not worth paying to commit when the two decision are very close together; what matters is not the number of "periods" between the decisions but the real time between them.

DellaVigna *et al* [2010] conduct an experiment on door-to-door charitable fund-raising. They find if an option to avoid the salesperson is available about a quarter of people make use of it, and that if the option is made cheaper by providing a "Do Not Disturb" check box nearly a third of people choose to avoid the salesperson. This is as our model predicts: the lower the cost of avoidance, the more people will choose it. DellaVigna *et al* also find that those who choose avoidance are concentrated among people who donate less when avoidance is not possible. Whether this is the case in our model depends on the distribution of B. If the lowest value of B in the population is greater than or equal to $\underline{B}$, and the highest value exceeds $\overline{B}$, then all those who would not contribute when avoidance is not possible ($F = \infty$) will choose avoidance, while only some of those who would contribute choose avoidance; this is what DellaVigna *et al* find. On the other hand, if the highest value of B in the population is less than or equal to $\overline{B}$ while the lowest is below $\underline{B}$ our model predicts the opposite result. A more elaborate experiment could vary the value of B more systematically – for example in the flier describing the visit, indicating that a level of matching funds are available (three dollars to the charity for every dollar you donate, for example). This would make it possible to test for the non-monotonicity in B that the model predicts.

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 short-run utility net of the self-control cost

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

where we now assume that 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 are known to explain preference "reversals" that arise from failure of the independence axiom, as in the Allais paradox while linear self control costs cannot, as they are consistent with the independence axiom.[11]

Example 4: Menu Dependent Choice

In Fudenberg-Levine [2006] we showed how convex control costs can lead to menu-dependent choice in violation of the weak axiom of revealed preference. We now extend that example to study the role of the short-run player's effective horizon, as measured by $1/(1-\mu)$, in determining when the agent prefers menus that include tempting choices and when the agent prefers more restrictive menus that exclude them. Consider the following three-period model of menu choice. In the original problem there

[11] Noor and Takeoka [2010] 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.

were three possible actions, broccoli b, frozen yogurt y and ice cream i. In the first period a menu consisting of a subset of actions is chosen from a list of menus M. In the second period an action is chosen from the menu; the utility received in periods 2 and 3 depends on the choice x, we denote it by $(u_2,u_3)(x)$. We are interested in the induced preferences over menus. In particular, we would like to model a situation where $\{b,y\} \succ \{y\}$ and $\{b,i,y\} \succ \{b,i\}$. Here the frozen yogurt is a "compromise" option that is appealing in the face of strong temptations but not when faced weaker ones.

This is known from the work of Dekel, Lipman, and Rustichini [2001] to imply that the independence axiom is violated, and since our model with linear g is a form of expected utility it follows that g must be non-linear.

In Fudenberg and Levine [2006] we showed that with a single period lived short-run self ($\mu = 0$) and utilities $(u_2,u_3)(b) = (0,100)$, $(u_2,u_3)(y) = (8,30)$, $(u_2,u_3)(i) = (14,0)$, $\delta = .9$, and $g(u) = .5Gu^2$, and $G = 1$that frozen yogurt is indeed an optimal compromise. We will now hold the vectors (u_2,u_3) fixed, and investigate how the preferences depend on the other parameters, with focus on the role of μ and on the menus $\{b\}$ and $\{b,i\}$, ignoring the compromise choice $\{y\}$.

When $\mu = 0$ there is no self control cost involved in choosing the first-period menu, so the long-run self can do no better than choose the menu $\{b\}$ that consists of the best long-run outcome; indeed this menu is the unique optimum. When $\mu = 1$ the long-run and short-run selves agree on the rankings of both second period choices and first period menus; in particular $\{b\} \sim \{b,i\}$.

To analyze the decision for intermediate values of μ we start at the end, examining the optimal choice from each menu. In the menu $\{b,i\}$ the short-run value of b is $(1-\delta\mu)100\delta\mu$; the short-run value of i is $14(1-\delta\mu)$. Suppose $14 > 100\delta\mu$ so that i is the temptation. The period-2 value from choosing b from $\{b,i\}$ is then $(1-\delta)\left(100 - .5G(1-\delta\mu)(14-100\delta\mu)^2\right)$. We assume that G is large enough that this is negative, so that the optimum is to give in to temptation and choose i.

Working back to period 1, we now consider the choice between the menu $\{b,i\}$ and the menu $\{b\}$. The temptation is $\{b,i\}$, which has value $(1-\delta\mu)14\delta\mu$, while $\{b\}$ gives the short-run player a value of $(1-\delta\mu)100(\delta\mu)^2$. The agent's value from $\{b,i\}$ is then $(1-\delta)14\delta$, while the value from $\{b\}$ is

$$(1-\delta)\delta^2 100 - .5G\Big((1-\delta\mu)14\delta\mu - (1-\delta\mu)100(\delta\mu)^2\Big)^2 =$$
$$(1-\delta)\delta^2 100 - .5G(1-\delta\mu)^2(\delta\mu)^2\left(14-(1-\delta\mu)100(\delta\mu)\right)^2$$

Now set $\mu = .1$; note that $14 > 90\mu$ so the short-run self prefers i to b. Substituting $\delta = .9$, $\delta\mu = .09$, the value of $\{b\}$ is $8.1 - .5G(.91)^2(.09)^2(14-8.19)^2$ and the value of $\{b,i\}$ is 1.26 so if G is large enough then $\{b,i\}$ is preferred. Also for large G we have $90 - .5G(1-\delta\mu)(14-90\mu)^2 < 0$, so i will be chosen from the menu $\{b,i\}$.

Hence we have a non-monotonicity in μ: for $\mu = 0$, the menu $\{b\}$ is strictly preferred, and this remains true for all small enough neighborhood of $\mu = 0$. Similarly, for μ near .1, the menu $\{b,i\}$ is strictly preferred. And finally, for μ equal or close to 1, the two menus are indifferent.

Note that between subjects we can in principle infer both μ and δ. For example, we might conduct an experiment in two stages. In the first stage, subjects are presented with intertemporal choices of the Green-Myerson type discussed in Example 1, and from this we can infer μ, δ for each subject.[12] Then in the second stage we give choices between menus with items that we think will create different degrees of temptation, such as the foods in this example. The theory predicts when we correlate the μ's inferred from the first stage of the experiment we should observe a non-monotonocity in menu choices in the second stage.

Example 5: 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 a simpler 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 instead 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

[12] To avoid the additional complications required to explain temptation by money payoffs we could use pizza, as in Kirby and Guatsello [2001].

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. Note moreover that 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 $-q\delta^{n-1}P < g\left(q(\delta\mu)^{n-1}S\right)$, 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. 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. 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.

Example 6: Two Tempting Choices

We now consider a variation on Example 5, 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 self-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. .

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. [13] 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 Δ 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

[13] 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.

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 $f(w_n, \Delta)$ is non-decreasing in w_n and non-increasing in Δ, continuously differentiable in both arguments and satisfies $f(w_n, 0) = w_n$ and $f(w_n, \Delta) \le w_n$. Note that the stock of cognitive resources depends on the action taken only through the foregone value, so actions that maximize the short-run self's value also maximize the end-of-period stock $\tilde{w}_n$.

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 $w_{n+1} = r(\tilde{w}_n) \ge \tilde{w}_n$, 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 $r(\tilde{w}_n) \le \overline{w}$ so that there is an upper bound on the stock of cognitive resources. If $r(\tilde{w}_n) = \overline{w}$ then resources are replenished immediately, which is the usual assumption when short-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.[14] 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

$$V(h_n, \mathbf{a}) \equiv E_{\mathbf{a}, h_n}(1-\delta)\sum\nolimits_{\ell=0}^{\infty} \delta^{\ell} \left(u(y_{n+\ell}, a_{n+\ell}) + m(y_{n+\ell}, f(w_{n+\ell}, a_{n+\ell})) \right)$$

Note that the contribution to utility of the stock of cognitive resources $w_{n+n'}$ is measured in the same units as utility. Thus if there is a fixed stock $\overline{w}$ of cognitive resources, the

[14] 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 $\tilde{w}_n \ge 0$ as do Ozdenoren et al; however we may set $m(y_n, \tilde{w}_n) = -\infty$ if $\tilde{w}_n < 0$ to incorporate that constraint.

stock produces an amount $m(y,\overline{w})$ of utility each period. This $m(y,w)$ 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*: $V(h_n,\mathbf{a})$ has a maximum for each n,h_n.*

Note also that, as in the linear case, this is a Markov decision problem, so it has a Markov solution that depends only the state y_n.

We will say that the value of cognitive resources is *state-independent* if it depends on the state only through the stock of willpower; in a slight abuse of notation we use the same letter write this as: $m(y_n,\tilde{w}_n)=m(\tilde{w}_n)$. State-independent resource valuation has an important implication: it implies that the action most favored by the short-run self maximizes the utility of cognitive resources. To see this, define

$$M(h_n,\mathbf{a})\equiv E^k_{\mathbf{a},h_n}\sum\nolimits_{\ell=0}^{\infty}(\delta\mu)^{\ell}(1-\delta\mu)m(\tilde{w}_{n+\ell}),$$

and note that the value on the RHS is independent of k. If each period's action is chosen to maximize the value $U^k(h_n,\mathbf{a})$ of the current short-run self, the foregone value each period is 0. This implies that the level of resources at each period is as high as possible given the initial value; with state-independent resource valuation, any action plan $\mathbf{a}$ that leads to this highest possible path for $\tilde{w}$ also maximizes the flow of benefits $m(\tilde{w}_n)$ in the strong sense that no other action plan leads to a higher value of m in any period along any history. As a consequence, any action plan that maximizes short-run utility in each period on each history also maximizes the short-run self's expected discounted value M of cognitive resources. We will use this latter implication in relating the model to a game between the long run self and the short run self, so we state it as a theorem:

Theorem 4*: With state-independent resource valuation,*
$\arg\max_{\mathbf{a}} U(h_n,\mathbf{a})=\arg\max_{\mathbf{a}} M(h_n,\mathbf{a})$.

Relation to the Literature

To relate the model with willpower to past work, consider the "cake-eating" problem of Ozdenoren et al [2009].[15] Here there is a cake of fixed size, and the only choice is a consumption level a_n that reduces the size of the cake. We suppose that $\mu = 0$ and that $r(\tilde{w}_n, a_n) = \tilde{w}_n$, so that there is no replenishment. Ozdenoren et al specify that the temptation is a fixed upper bound on consumption $\overline{a}$ if the cake is not exhausted, and 0 if it is, and that the rate of willpower depletion is $\tilde{f}(a_n, w_n)$ for $a_n < \overline{a}$, with f decreasing and strictly convex in a_n. In addition $m(y_n, \tilde{w}_n) = 0$ until the stock of cake runs out or the time horizon is reached, at which time $m(y_n, \tilde{w}_n) = \overline{m}(\tilde{w}_n)$. In our formulation $\Delta(y_n, a_n) = u(\overline{a}) - u(a_n)$, so if we set

$$f(w_n, \Delta(y_n, a_n)) = \tilde{f}(u(\overline{a}) - \Delta(y_n, a_n), w_n),$$

we see that their model a special case of ours.[16] 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 assumption in section 6, along with the possibility that actions have a direct impact on the evolution of cognitive resources; with either of these changes Theorem 4 can fail. For the time being we will assume state-independent cognitive resource utility.

Next, suppose that (i) there is state-independent resource valuation, that (ii) $r(\tilde{w}_n) = \overline{w}$, so that replenishment is immediate, and (iii) that $\mu = 0$ so that short-run selves live only one period. 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 $m(f(\overline{w}, \Delta(y, a_n)))$. We can then define $c(\Delta) = m(\overline{w}) - m(f(\overline{w}, \Delta)) = -m(f(\overline{w}, \Delta)) \geq 0$, and when $\Delta = 0$ we have $c(\Delta) = 0$. Then the objective function is equal to

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

[15] That model is in continuous time, here we give the discrete-time version.

[16] The function f here is not constrained to simply be the difference between $w_n - \Delta$, which allows for lower depletion near $w_n = 0$, as in the multiplicative functional form $f(w_n, \Delta(y_n, a_n)) = \Delta(y_n, a_n) w_n$.

which is equivalent to the one-period of life formulation used in Fudenberg and Levine [2006, 2010]. 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.[17]

Single Decision Problems

Cognitive resources serve to link the decisions in one period to 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 amount of cognitive resources available for decision making may be different. Suppose that $w_1 = \overline{w}$, so that initially cognitive resources are "topped up". In this case we say that resources *start full.*[18]

For any period n define the end of period resource stocks corresponding to an initial shock of Δ and no subsequent shock by $\tilde{w}_n^n(\Delta) = f(\overline{w}, \Delta)$, $\tilde{w}_{n+1}^n(\Delta) = r(f(\overline{w}, \Delta))$, and for $\ell \geq 1$ $\tilde{w}_{n+\ell+1}^n(\Delta) = f(r(\tilde{w}_{n+\ell}^n(\Delta)), 0)$. Then the cost of self-control corresponding to a single shock is

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

The following result is immediate

Theorem 5: *If there is a single decision and cognitive resources start full, the maximization problems*

[17] 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 $m(f(\overline{w}, \Delta(y, a_n) + d_n))$, where d_n is the cognitive load in period n.

[18] 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 $\overline{w}$ with the fixed amount of cognitive resources w^* available when a decision is possible, as this is invariant to $y \in Y^*$.

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

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

have the same set of solutions..

The Linear Case

Next suppose that in addition to conditions (i)-(iii) above, the benefit of cognitive resources is linear in Δ, so that $m(y_n,\tilde{w}_n)=\gamma\tilde{w}_n$, and that resource depletion is linear as well, so that $f(w,\Delta)=w-\Delta$.[19] Then the cost function defined above is $c(\Delta)=(1-\delta)\gamma\overline{w}-(1-\delta)\gamma(\overline{w}-\Delta)=\gamma(1-\delta)\Delta$, so the linear model of the previous section, where the cost of self-control is independent of δ, 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 with the period length. Conversely, if the cost of self-control 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 $\overline{w}$ are independent of τ, while for smaller stocks we have $m_\tau(\overline{w}-\Delta)=m(\overline{w})-g(\Delta)/(1-\delta)$. As with the one-period of life model, in the case of general depletion and full replenishment we can define the cost by $g(\Delta)=(1-\delta)(m(\overline{w})-m_\tau(f(\overline{w},\Delta)))$.

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 $r(\tilde{w}_n)=\tilde{w}_n+\lambda(\overline{w}-\tilde{w}_n)$ where $0\le\lambda\le 1$.

[19] 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.

Theorem 6: *For any model with linear benefit* γ*, depletion and replenishment* λ *define*

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

Then if $\mathbf{a}$ *is a solution to the linear model with parameter* Γ *then it is a solution to the* (λ,γ) *model in which actions are independent of* w_n *and all such solutions to the* (λ,γ) *model are solutions to the* Γ *model.*

Proof: See appendix.

Theorem 6 shows that if depletion, replenishment, and self-control cost are all linear, and state independent, the stock of self control is irrelevant. 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 full replenishment linear model breaks down. Note also that the equivalent linear parameter Γ 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.

Because Theorem 6 maps many (λ,γ) models to the same linear model Γ, it also implies that these models are equivalent in the sense of generating the same decisions. That is, if we change λ to λ' while holding the time period (and thus δ and μ) fixed, the resulting system (λ',γ') will have the same cost for every self-control decision if we set

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

even though the time-path of the willpower stock in the two models will be different. 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.

What happens with linear replenishment when we vary τ, 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 Δ_n is spent in some period n, the state immediately jumps from w_n to $\tilde{w}_n$, and is then replenished according to the continuous-time differential equation $\dot{w}_t = \kappa(\overline{w} - w_t)$. Thus when the period length is τ we have $w_{n+1} = \overline{w} - \exp(-\kappa\tau)(\overline{w} - \tilde{w}_{\boldsymbol{n}})$, so $\lambda(\tau) = 1 - \exp(-\kappa\tau)$. Note that when the period is long, the state almost completely replenishes. Note also that as $\tau \to 0$ we have $\lambda(\tau) \to 0$, $\delta(\tau) \to 1$.

Example 7: Resisting Temptation with Linear Benefits, Depletion and Replenishment

Now we reconsider the persistent, and delayed temptations of Example 1 in the model with linear benefit, linear depletion and linear replenishment. As we will see, when the marginal value of cognitive resources is constant, the optimum is either to take at once or to resist forever. This will help illustrate Theorem 5 and the continuous-time limit. It will also set the stage for our subsequent analysis of these temptations when the benefit function is concave, where it may be optimal to resist for a while, and then take once the marginal value of resources is sufficiently high.

We begin with the case of a persistent temptation, where temptation is present each period unless and until it is accepted. Note that if the agent always resists, the stock of resources evolves according to $r(\tilde{w}_n) = \tilde{w}_n + \lambda(\overline{w} - \tilde{w}_n)$ so

$$\begin{aligned} \overline{w} - w_{n+1} &= \overline{w} - r(\tilde{w}_n) = \overline{w} - (\tilde{w}_n + \lambda(\overline{w} - \tilde{w}_n)) \\ &= (1-\lambda)(\overline{w} - \tilde{w}_n) = (1-\lambda)(\overline{w} - w_n + (1-\delta\mu)S) \end{aligned}$$

Since $0 < \lambda < 1$, the solution to this difference equation $w_t = \overline{w} - (1-\lambda)(1-(1-\lambda)^{n-1})(1-\delta\mu)S/\lambda$, which converges monotonically to the steady state solution $w^* = \overline{w} - ((1-\lambda)/\lambda)(1-\delta\mu)S$, so that

$$\tilde{w}_t \to w^* - (1-\delta\mu)S = \overline{w} - (1/\lambda)(1-\delta\mu)S \;.$$

Note that as the time period goes to 0, the steady state converges to $\overline{w} - ((\rho+\eta)/\kappa)S$ In particular, as $\kappa \to 0$ a steady drain of resources sends the stock to minus infinity, while as $\kappa \to \infty$ resources replenish so quickly that any bounded outflow has negligible impact.

If the benefits are linear, so that $m(w) = \gamma w$, then by Theorem 6 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

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

Substituting we see that the agent resists if

$$P(1-\delta) < -\frac{(1-\delta)\gamma}{1-\delta(1-\lambda)}(1-\delta\mu)S$$
$$\big(1-\delta(1-\lambda)\big)P < -\gamma(1-\delta\mu)S$$

Note that if λ is small, this is about $P(1-\delta) < -\gamma(1-\delta\mu)S$, while if $\lambda = 1$ it is $P < -\gamma(1-\delta\mu)S$. Recall that P is negative so as λ gets bigger this is easier to satisfy as we would expect.

To study what happens when the time period is short, we rewrite $\big(1-\delta(1-\lambda)\big)P < -\gamma(1-\delta\mu)S$ as

$$1-\exp(-\rho\tau)(\exp(-\kappa\tau))) < -\gamma(1-\exp(-(\rho+\eta)\tau)S,$$

and then approximate, yielding

$$\rho P < -\gamma\frac{\rho}{\rho+\kappa}(\rho+\eta)S.$$

Here the LHS is approximately the gain of postponing P by τ, and the RHS is the cost of postponement; this is the reduction of resources of $(\rho+\eta)S$ multiplied by the continous-time limit of the discrete-time cost parameter Γ.

As a check and explanation of Theorem 6, note that we get the same answer working directly with the partial replenishment system in continuous time: If the agent takes at once, the value is $P + \gamma\overline{w}$. If the agents resists forever, then in the associated continuous-time limit the state follows the path

$$w_t = \overline{w} - (\rho+\eta)S/\kappa + \exp(-\kappa t)((\rho+\eta)S/\kappa).$$

The value on the this path in the continuous time model is

$$\int_0^\infty \rho\exp(-\rho t)\gamma\left(\overline{w}-(\rho+\eta)S/\kappa+\exp(-\kappa t)((\rho+\eta)S/\kappa\right)=$$
$$\gamma\left(\overline{w}-(\rho+\eta)S/\kappa\right)+\int_0^\infty \rho\exp(-(\rho+\kappa)t)\left((\rho+\eta)S/\kappa\right)=$$
$$\gamma\overline{w}-\gamma(S(\rho+\eta)/\kappa)(\kappa/(\rho+\kappa))=\gamma\overline{w}-\gamma S(\rho+\eta)/(\rho+\kappa)$$

So again we see that it is optimal to resist if $\rho P < -\gamma\rho S(\rho+\eta)/(\rho+\kappa))$.

Note that when κ is very large compared to the other parameters, the right hand side is near 0 so it is always better to resist: if resources get quickly replenished whenever they dip slightly below the steady state.

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 ℓ periods, it is optimal to resist if $P < -\Gamma(1-(\delta\mu)^\ell)/(1-\delta^\ell)S$, and substituting for Γ yields $P(1-\delta(1-\lambda))/(1-\delta) < -\gamma(1-(\delta\mu)^\ell)/(1-\delta^\ell)S$. If we suppose that the delay is s units of real time, $\ell = s/\tau$, and send τ to zero while holding s fixed, then it is optimal to resist if $P(\rho+\kappa)/\rho < -\gamma\left[(1-\exp(-(\rho+\eta)s)/(1-\exp(-\rho s)\right]S$.

In the case $s=\tau$, in which the delay is only a single period, this inequality reduces to $P\rho < -\gamma\rho\left[(\rho+\eta)/(\rho+\kappa)\right]S$ 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.

For ease of reference we summarize these findings with a proposition:

Proposition 7: *With linear benefit, depletion and replenishment*

a) When the agent is faced with a persistent temptation, and $\left(1-\delta(1-\lambda)\right)P > -\gamma(1-\delta\mu)S$**,** *he takes it at once; he resists the temptation forever when* $\left(1-\delta(1-\lambda)\right)P < -\gamma(1-\delta\mu)S$. *Thus the agent will resist for a while and then take only in the knife-edge case when he is indifferent between taking at once and resisting forever.*

(b) In the limit of time periods going to zero, the agent takes at once if $P > -\gamma S(\rho+\eta)/(\rho+\kappa))$ *and resists forever if* $P < -\gamma S(\rho+\eta)/(\rho+\kappa))$. *Thus as* $\kappa \to \infty$ *the agent resists forever, and as* $\kappa \to 0$ *the agent takes at once if* $\rho P > -\gamma S(\rho+\eta)$.

(c) If declining the temptation puts it off for s units of calendar time, the agent takes at once or resists forever as P is greater or less than $-[(1-\exp(-(\rho+\eta)s)/(1-\exp(-\rho s)]\gamma\rho S/(\rho+\kappa)$.

The Replenishment Rate in Models with Linear Benefit and Cost

If the system can be represented as linear replenishment for some κ, along with linear depletion and linear benefits, then as we saw in Theorem 5 the system is equivalent to the full replenishment model. Taking the limit as $\tau \to 0$ we find that the equivalent marginal cost is

$$\Gamma = \frac{(1-\exp(-\rho\tau))\gamma}{1-\exp(-(\rho+\kappa)\tau))} \to \frac{\rho\gamma}{\rho+\kappa}$$

In this limit, if κ is very small so that any reduction in the stock is almost permanent, then $\Gamma \to \gamma$. At the other extreme, when κ is very large, the equivalent cost Γ 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 τ.

To construct a limit that where the replenishment rate per period does not go to zero, take $\kappa = 1/\tau$, so that $\lambda(\tau) = 1-\exp(-1)$ for all τ; here the equivalent full-replenishment cost of self control is

$$\Gamma_\tau = \frac{(1-\exp(-\rho\tau))\gamma}{1-\exp(-(\rho\tau+1))} \to 0.$$

This is a different answer than the one we obtained earlier in the full replenishment case when we held Γ fixed, as there we did not scale the cost of cognitive resources by $1-\delta$. Hence we get different limiting results depending on whether we consider a limit in which cost falls relative to utility as we vary κ. The full-replenishment limit we considered earlier, where Γ is constant, corresponds to scaling γ along with κ; to hold Γ constant when $\kappa = 1/\tau$, we need to set

$$\gamma = \frac{\Gamma(1-\exp(-(\rho\tau+1)))}{(1-\exp(-\rho\tau))},$$

so that self control remains costly in the limit, even though the stock of willpower replenishes very quickly. Note that in the limiting model of full replenishment the stock of willpower is irrelevant, so the model is essentially "state-free." However, the limiting argument shows that rapid replenishment with a very high flow value of benefits behaves in a way very similar to that state-free model.

In the linear model the stock of willpower does not play a significant role: as we observed all replenishment rates are equivalent for appropriately chosen values of γ and are equivalent to full replenishment for an appropriate value of Γ. 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, so the rate of replenishment plays a more essential role. This is the case we study next.

6. Cognitive Resources, Non-linearities, and Replenishment

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 Δ_n is governed by the depletion equation $\tilde{w}_n = f(w_n, \Delta_n)$, the replenishment equation $w_{n+1} = r(\tilde{w}_n)$ and of course the benefit function $m(y_n, \tilde{w}_n)$. Putting together the depletion and replenishment equations gives the dynamics of cognitive resources $w_{n+1} = r(f(w_n, \Delta_n))$, where $\overline{w} \geq r(\tilde{w}_n) \geq \tilde{w}_n$ 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 Δ_n.

Linear Replenishment

The units in which w_n are measured are essentially 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 would like to construct a change of units $w' = h^{-1}(w)$ so that the property $h(\overline{w} - \lambda(\overline{w} - w')) = r(h(w'))$ is satisfied, so that the replenishment function r' corresponding to w' has the linear form $w'_{n+1} = r'(\tilde{w}'_n)$. Given such a function we may define $f'(w', \Delta) = h^{-1}(f(h(w'), \Delta))$, $m'(w') = h^{-1}(m(h(w'))$, and in the new units with the new depletion and benefit of self control functions, we have restated the problem so that replenishment is linear.

There are a variety of ways of constructing an h function. One simple method is to consider intervals $I_\lambda = (0, \lambda\overline{w}]$ under the mapping $T(w') \equiv \overline{w} - \lambda(\overline{w} - w')$. Notice that the intervals of the iterated map $T^n(I_\lambda)$ (since T is invertible, we allow negative values of n) form a partition of $(-\infty, \overline{w})$. Hence for any $w' < \overline{w}$ there exists a unique integer $n(w')$ (possibly negative) such that $w' \in T^{n(w')}(I_\lambda)$. The interval I_λ for the units w' corresponds to $(0, r(0)]$ in the original units w'. Define

$$h(w') = r^{n(w')}\left(\frac{T^{-n(w')}(w')}{\lambda\overline{w}} r(0)\right).$$

If $r(0) > 0$ and r is strictly increasing for $\tilde{w} < \overline{w}$, $(0, r(0)]$ is a non-empty interval, so t h maps onto $(-\infty, \overline{w})$. In this case h is strictly increasing, so invertible, and by construction $h(\overline{w} - \lambda(\overline{w} - w')) = r(h(w'))$. If r is continuous, h is continuous and $h(0) = 0$; h extends uniquely to a continuous function on $(-\infty, \overline{w}]$ by defining $h(\overline{w}) = \overline{w}$.

We summarize this as a theorem.

Theorem 8: *Suppose that* r *is continuous and strictly increasing and that* $r(0) > 0$. *The system with replenishment* $r'(w') = \overline{w} - (1-\lambda)(\overline{w} - w')$, *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 function* $r(w)$, *depletion function* $f(w, \Delta)$ *and benefit function* $m(w)$.

Notice that the rescaling of units to linearize r is possible only when there is some replenishment $r(0) > 0$ 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: only in the linear case does Theorem 5 hold – 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 λ', then the equivalent system in units of w' for a different value of λ is not linear. This may seem puzzling in light of our earlier observation that Theorem 5 implies an equivalence between linear models with different replenishment rates. However, the equivalence in Theorem 5 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.

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, $r(\tilde{w}_n) = \tilde{w}_n + \lambda(\overline{w} - \tilde{w}_n)$, 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, $w_{n+\ell} = (1-\lambda^{\ell})\overline{w} + \lambda^{\ell} w_n$. Along this path the average present value of cognitive resources is

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

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

Example 8: Two Tempting Choices with Linear Depletion and Replenishment

Now we re-analyze the two temptations of example 6 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, because the time path of cognitive resources is basically the same, whether the agent resists temptation of value two S in a single period or resists S in two consecutive periods that are close together .

Suppose first that the decisions are made in consecutive periods, so that $n_1 = 1$, $n_2 = 2$. The agent has four possible plans:

Take both options: The first option is provides direct utility of P, the second option provides δP, and no self control is used, so overall utility is $(1+\delta)P + M(w_1)$

.

Take only second option: Self control of S is used in the first period, and none thereafter, so $\tilde{w}_1 = w_1 - S$, $w_2 = \lambda w_1 - \lambda S + (1-\lambda)\overline{w}$, and the overall value is
$0 + (1-\delta)m(w_1 - S) + \delta P + \delta M(w_2)$.

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 $\lambda w_1 + (1-\lambda)\overline{w} - S$, so the stock at the start of the third period is $w_3 \equiv (1-\lambda)\overline{w} + \lambda\left(\lambda w_1 + (1-\lambda)\overline{w} - S\right) = \lambda^2 w_1 + (1-\lambda^2)\overline{w} - \lambda S$. The value of this plan is thus $P + (1-\delta)m(w_1) + \delta(1-\delta)m(\lambda w_1 + (1-\lambda)\overline{w} - S) + \delta^2 M(w_3)$.

Reject both options: Self-control is used in both periods, $w_3 = \lambda^2 w_1 + (1-\lambda^2)\overline{w} - \lambda(1+\lambda)S$ so the value is $m(w_1 - S) + (1-\delta)\delta m(\lambda w_1 - \lambda S + (1-\lambda)\overline{w} - S) + \delta^2 M(w_3)$.

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 $2P + M(w_1)$.
Take one option: Here $w_2 = \lambda w_1 - \lambda S + \lambda \overline{w}$, and overall value is $P + (1-\delta)m(w_1 - S) + \delta M(w_2)$.
Reject both options: Now $w_2 = \lambda w_1 - \lambda 2S + \lambda \overline{w}$ and the overall value is $(1-\delta)m(w_1 - 2S) + \delta M(w_2)$.

We now send τ to 0, and examine the case where two decisions are made in consecutive periods. In the full-replenishment model of the previous section, resisting temptation of S in two consecutive periods reduces the flow of cognitive benefits each period from $m(\overline{w})$ to $m(\overline{w} - S)$, which reduces average value (compared to no self control) by $(1-\delta)(m(\overline{w}) - m(\overline{w} - S))$, while resisting $2S$ in a single period reduces average value by $(1-\delta)(m(\overline{w}) - m(\overline{w} - 2S))$, which is strictly larger, even as $\tau \to 0$, because m is concave.

In the present problem with linear replenishment, we see that the time path of cognitive resources is basically the same, whether the agent resists temptation of two S in a single period or resists S in two consecutive periods that are close together: In the first case resources jump down to $\overline{w} - 2S$ at the end of the first period, in the second case we have $\tilde{w}_1 = \overline{w} - S$; $w_2 = \lambda \tilde{w}_1 + (1-\lambda)\overline{w} \approx (1-\kappa\tau)(w_1 - S) + \kappa\tau\overline{w}$, and

$$\tilde{w}_2 = (1-\kappa\tau)(w_1 - S) + \kappa\tau\overline{w} - S = w_1 - 2S + \kappa\tau(\overline{w} - w_1 + S).$$

Hence as $\tau \to 0$, $\tilde{w}_2 \to w_1 - 2S$, which is the level of $\tilde{w}_1$ in the case of a simultaneous decision. It follows that when the agent has a strict preference for accepting 0 or both options with simultaneous decisions, she will have the same preference when the decision times are close together. If the agent strictly prefers to take one option with simultaneous decisions, she will prefer to take the first one when decisions are sequential

Example 9: A 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 examples 1 and 4 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.

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 ℓ times before giving in, his value is

$$(1-\delta)\Big[\sum_{n=1}^{\ell-1}\delta^{n-1}m(\overline{w}-n(1-\delta\mu)S)+\sum_{n=\ell}^{\infty}\delta^{n-1}m(\overline{w}-\ell(1-\delta\mu)S\Big]+\delta^{\ell}P$$

The bigger is ℓ the smaller is the first term and the larger is the second. This implies that a necessary condition for an optimal ℓ 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. Let us look at the value for ℓ minus the value for $\ell+1$, noting that prior to ℓ the arguments of m are the same in both cases

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

Observe that since m is convex, $D(\ell)$ 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 τ define $d(s,\tau)=D(s/\tau)/\delta^{s/\tau}(1-\delta)$. Note that this has the same sign as D so can equally well be used to characterize the optimum. Observe that

$$d(s)=\lim_{\tau\to 0}d(s,\tau)=m'(\overline{w}-s(\rho+\eta)S)\frac{\rho+\eta}{\rho}S+P$$

We see that if $m'(-\infty) = \infty$ it is always optimal to give in; if

$$m'(\overline{w})\frac{\rho+\eta}{\rho}S + P < 0$$

then it is not optimal to give in right away, and that the optimal time to give in is characterized by $d(s) = 0$.

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 10: Waiting to Commit

Now we add to Example 9 the possibility of taking the temptation off the table for a cost $F < |P|$. For simplicity we analyze only the no-replenishment case with linear depletion.

Using the argument from Example 9, we see that for small enough τ it is not optimal to commit immediately if

$$m'(\overline{w} - F)\frac{\rho+\eta}{\rho}S - F < 0.$$

A sufficient condition is

$$m'(\overline{w} + P)\frac{\rho+\eta}{\rho}S - F < 0, \text{ or}$$

(*)
$$m'(\overline{w} + P) < \frac{\rho}{\rho+\eta}\frac{F}{S}$$

Let us suppose that

$$m'(-\infty)\frac{\rho+\eta}{\rho}S + P < 0$$

so that in the absence of the possibility of commitment, it is optimal never to give in, resulting in at least the value $m(\overline{w}) + P$ .Taking the temptation off the table in the first period gives value

$$\lim_{\tau \to 0}(1-\delta)\Big[\sum_{n=1}^{\infty}\delta^{n-1}m(\overline{w}-(1-\delta\mu)S-F\Big]-F$$
$$= m(\overline{w}-F)-F$$

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 (*) from above 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 (*) becomes

$$m'(\overline{w}+P) < \frac{\rho}{\rho+\eta}\frac{|P|}{2S}.$$

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, 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. This experiment was designed to test previous models with myopic short run selves, and the experimental instructions did not specify when and whether opportunities for giving in might occur in the future. Thus it is not clear what subjects believed about this, and whether the perfect foresight analysis applies. That is, while delay was observed, we can not be certain from these experiments whether it is the type of delay predicted by this model. We are hopeful that future experiments may shed more light on the type of delay that can occur with cognitive resource depletion, but not without it.

Example 11: A Persistent Temptation with Non-Linear Benefits, Partial Linear Replenishment and General Depletion.

We now examine a final variation on the persistent temptation problem. We consider a persistent temptation without the option to commit, and now assume partial linear replenishment, and allow general depletion. This lets us highlight the interplay of 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 as in Section 4, by $g(\Delta) = (1-\delta)(m(\overline{w}) - m_{\tau}(f(\overline{w},\Delta)))$ 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 $P < m(f(\overline{w},(1-\delta\mu)S)) - m(\overline{w}) = -g((1-\delta\mu)S))/(1-\delta)$.

Now let us generalize the linear case with partial replenishment to allow for a lower bound on the stock of cognitive resources. That is, in place of linear m we have $m(w) = \gamma w$ for $w \geq 0$ and $m(w) = -\infty$ for $w < 0$ (which is our way of modeling a lower bound of 0 on resources). If $P < -\gamma(1-\delta\mu)S$ then as in our earlier analysis the optimum is to take the temptation immediately. If $\left(1-\delta(1-\lambda)\right)P > -\gamma(1-\delta\mu)S$ the solution is to resist forever if it is feasible to do so: $w^{*} - (1-\delta\mu)S = \overline{w} - \left(1/(1-\lambda)\right)(1-\delta\mu)S \geq 0 \; 1-(1-\delta\mu)S/\overline{w} > \lambda$. Otherwise, if $\left(1-\delta(1-\lambda)\right)P > -\gamma(1-\delta\mu)S$ and $1-(1-\delta\mu)S/\overline{w} < \lambda$ 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 let us consider the general case, normalizing to have linear replenishment. We assume that m and f are twice continuously differentiable, and that

$$-\frac{dm}{dw}((1-A)\overline{w} + Aw)\frac{\partial f}{\partial \Delta}(w,0)$$

(which is positive) is decreasing in w; we call this "increasing marginal cost of self control." Note that there is increasing marginal cost of self-control if m is strictly concave and f is linear. We may also write this condition in terms of second derivatives as

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

If m is concave, this says that the cross-partial of f should not be "too negative." If the cross partial is strongly positive, then m need not be concave.

Proposition 7: *Suppose there is increasing marginal cost of self control and there is strictly partial linear replenishment,* $0<\kappa$. *Then there is* $\underline{\tau}>0$ *such that if* $\tau<\underline{\tau}$, *there are* $\left|\overline{P}_\tau\right|>|\underline{P}_\tau|>0$ *such that it is optimal to resist forever if* $|P|>\overline{P}_\tau$, *it is optimal to resist until period* $\infty>\hat{\ell}>1$ *then take if* $\left|\overline{P}_\tau\right|>|P|>|\underline{P}_\tau|$, *and it is optimal to take immediately if* $|\underline{P}_\tau|>P$. *Let* $\overline{P}_0,\underline{P}_0$ *denote the limits as* $\tau\to 0$. *Let* W_s *be the solution to the differential equation*

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

and let W_∞ *be the solution to* $0=\kappa(\overline{w}-W_\infty)+\frac{\partial f}{\partial\Delta}(W_\infty,0)(\rho+\eta)S$. *Then*

$$\left|\overline{P}_0\right|=-(\rho+\eta)S\int_0^\infty e^{-(\rho+\kappa)t}m'(W_\infty)\frac{\partial f}{\partial\Delta}(W_\infty,0)dt,$$

$$|\underline{P}_0|=-(\rho+\eta)S\int_0^\infty e^{-(\rho+\kappa)t}m'(\overline{w})\frac{\partial f}{\partial\Delta}(\overline{w},0)dt$$

and if $\left|\overline{P}_0\right|>|P|>|\underline{P}_0|$ *then* $\hat{s}=\lim_{\tau\to 0}\tau\hat{\ell}$ *is finite and strictly positive, and is determined by*

$$|P|=(\rho+\eta)S\int_0^\infty e^{-(\rho+\kappa)t}m'(\overline{w}-e^{-\kappa t}W_{\hat{s}})\frac{\partial f}{\partial\Delta}(W_{\hat{s}},0)dt.$$

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 δ and μ (or ρ and η 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,

$$\left|\overline{P}_0\right| = |\underline{P}_0| = (\rho+\eta)\gamma S\int_0^\infty e^{-(\rho+\kappa)t}dt = \frac{(\rho+\eta)}{(\rho+\kappa)}\gamma S$$

which is the same as the condition for the critical value of *P* given in part (b) of Theorem 6 for the linear case.

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:

$$r(\tilde{w}_n) = \tilde{w}_n + \lambda_1(\overline{w} - \tilde{w}_n) \text{ for } \tilde{w}_n \in [w^*, \overline{w}]$$
$$r(\tilde{w}_n) = \tilde{w}_n + \lambda_2(\overline{w} - \tilde{w}_n), \text{ for } \tilde{w}_n < w^*,$$

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 κ_1 the agent would always resist, while if it was fixed at κ_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 w^*. To see why, first consider the agent's problem when the state is at w^*. Resisting forever gives exactly the same payoff as when the replenishment rate is fixed at κ_2, and taking gives a higher payoff than with replenishment fixed at κ_2, so since taking gives a higher payoff here than with $\kappa = \kappa_2$ the agent takes. Next consider the agent's problem 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.

The proof of Proposition 7 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 ℓ 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 τ of the time period, this intermediate region may be empty, but when τ is very small the concavity assumption ensures that it is non-empty.

7. The Game Between Long-run and Short-run Selves

We now want to show that the optimization problem we have been considering can be identified with the outcome of a game between the long-run self and a sequence of short-run selves. To do this we introduce an augmented state variable Y_k that is defined in any period n in which a new short-run self is born, and includes along with y_n the value of n; that is $Y_k = (y_n, n)$. Notice that any strategy $\mathbf{a}$ mapping histories to actions induces a well-defined stochastic kernel $\Pi(\mathbf{a}, Y_k)[dY_{k+1}]$.

In the game formulation the "actions" are taken by the short-run selves, and the long-run self chooses "self-control" actions that influence the preferences of the short-run self. Each short-run self can be thought of as choosing an $\mathbf{a}$: Although this contains irrelevant information such as how the short-run self will behave after he "dies" we will ignore this in computing the short-run self's payoff. Following Fudenberg and Levine [2006] we assume that before the short-run self moves, the long-run self chooses a *self-control action* $e \in \Xi$. It is convenient to take $E = 0 \bigcup \mathbf{A}$.We wish to study a sequence of stage game between the long-run self and the k th short-run self. The k th stage game consists of a choice of self control action e by the long-run self and a response $\mathbf{a}$ by the short-run self. The utility of the k th short-run self has the form $u(Y_k, e, \mathbf{a})$, which we specify below.

Histories in this game are sequences of augmented states Y_k along with the chosen actions $e_k, \mathbf{a}_k$, while a strategy from the long-run self is a map $\mathbf{e}$ from the previous history to a self-control action, a strategy for the kth short-run self is a map $\mathbf{a}_k$ from the previous history and choice of the long-run self to an action. The vector of strategies for all short-run players is denoted $\vec{\mathbf{a}}$. We define the conditional expectation operator $E_{\mathbf{e},\mathbf{a},Y_k}$ given the strategies $\mathbf{e},\mathbf{a}$ and state Y_k. The long-run self is completely benevolent and maximizes the discounted sum of short-run self utilities:

$$\tilde{V}(Y_1,\mathbf{e},\vec{\mathbf{a}}) \equiv (1-\delta)E_{\mathbf{e},\vec{\mathbf{a}},Y_1}\sum\nolimits_{k=0}^{\infty}\delta^k u(Y_k,e_k,\mathbf{a}_k)\,.$$

We now wish to specify the utility function of the short-run self to satisfy the assumptions of costly and unlimited self-control and limited indifference in Fudenberg and Levine [2006]. In that paper we give a procedure for deriving a utility function from an underlying objective function and a "cost of self control" function. We mimic that procedure here – the goal being to define the objective function for the short-run self so that the reduced form optimization problem is that of maximizing

$$E_{\mathbf{a},Y_1}\sum\nolimits_{n=0}^{\infty}\delta^n(1-\delta)\big(u(a_n,y_n)+m(y_n,\tilde{w}_n)\big)\,.$$

To this end we first define $E^k_{\mathbf{e},\mathbf{a},Y_n}$ to be the conditional expectation when k is alive. Write

$$\tilde{U}(Y_k,\mathbf{a}) \equiv (1-\delta)E^k_{0,\mathbf{a},Y_k}\sum\nolimits_{n=0}^{\infty}(\delta\mu)^{n-k}u(a^S_{k+n},y_{k+n})$$

and as above

$$\tilde{M}(Y_k,\mathbf{a}) \equiv (1-\delta)E^k_{0,\mathbf{a},Y_k}\sum\nolimits_{n=0}^{\infty}(\delta\mu)^{n-k}m(y_{k+n},\tilde{w}_{k+n}),$$

Following Fudenberg and Levine [2006] we define

$$u(Y_k,e,\mathbf{a}) = \begin{cases} \tilde{U}(Y_k,\mathbf{a}) + \max_{\mathbf{a}'}\tilde{M}(Y_k,\mathbf{a}') & r=0 \\ \tilde{U}(Y_k,e) + \tilde{M}(Y_k,e) - \|e-\mathbf{a}\| & \tilde{U}(Y_k,\mathbf{a}) \ge \tilde{U}(Y_k,e) \\ \tilde{U}(Y_k,\mathbf{a}) + \tilde{M}(Y_k,\mathbf{a}) - \|e-\mathbf{a}\| & \tilde{U}(Y_k,\mathbf{a}) < \tilde{U}(Y_k,e) \end{cases}$$

The cost of self-control is now defined to be

$$C(Y_k, \mathbf{a}) \equiv u(Y_k, 0, \mathbf{a}) - \max_{e|u(Y_k,e,a)\geq u(Y_k,e,\cdot)} u(Y_k, e, \mathbf{a})$$
$$= \begin{cases} 0 & \mathbf{a} \in \arg\max_{\mathbf{a}'} u(Y_k, 0, \mathbf{a}) \\ \max_{\mathbf{a}'} \tilde{M}(Y_k, \mathbf{a}') - \tilde{M}(Y_k, \mathbf{a}) & \mathbf{a} \notin \arg\max_{\mathbf{a}'} u(Y_k, 0, \mathbf{a}) \end{cases}$$

which has the property that $C(Y_k, \mathbf{a}) \geq 0$ and $C(Y_k, \mathbf{a}) = 0$ if and only if $\mathbf{a} \in \arg\max_{\mathbf{a}'} \tilde{U}(Y_k, \mathbf{a}')$.

Notice that the cost of self-control does not necessarily satisfy the property of being an *opportunity cost.* In general an opportunity cost for the short run self would have the form

$$\tilde{C}(Y_k, \mathbf{a}) = G(\bar{U}(Y_k) - \tilde{U}(Y_k, \mathbf{a})).$$

Here the self-control cost is computed each period by the difference between the best expected present value available to a short-run self born in that period and the present value actually received, taking into account what will actually happen in future periods, in contrast to our definition of the recursive cost, which is "as if" no self-control will be used in future periods. However, as this dependence is only through the variation of the marginal cost over time: In the linear case, where marginal cost is constant, the recursive cost is an opportunity cost.

Theorem 10: *If* $\tilde{C}(Y_k, \mathbf{a}) = \Gamma(\bar{U}(Y_k) - \tilde{U}(Y_k, \mathbf{a}))$ *then* $\tilde{C}(Y_k, \mathbf{a}) = C(Y_k, \mathbf{a})$.

The proof is in the Appendix. They key idea is that the principle of optimality for the short-run self enables us to write the overall loss to the short-run self as a sum of recursively computed losses

$$\bar{U}(y_n) - U^n(h_n, \mathbf{a}) =$$
$$\bar{U}(y_n) - (1-\delta\mu) E^n_{\mathbf{a},h_n} \sum_{\ell=0}^{\infty} (\delta\mu)^{\ell} u(y_{n+\ell}, a_{n+\ell}) =$$
$$E^n_{\mathbf{a},h_n} (1-\delta\mu) \Big(\sum_{\ell=0}^{\infty} (\delta\mu)^{s} \left(\Delta(y_{n+\ell}, a_{n+\ell}) \right) \Big)$$

Hence the opportunity cost is just a weighted sum of the increments $\Delta(y_{n+\ell}, a_{n+\ell})$, and the proof simply consists of bookkeeping to verify that the weights are the same as in the recursive case. In the linear case, in other words, it does not matter whether the cost of imposing self-control on the short-run self arises from recursive considerations or from an opportunity cost. However, in the non-linear case, the model of opportunity cost leads to implausible predictions about timing, such as changes in behavior when a short-run self "dies." Hence we focus on the recursive model of self-control cost in the non-linear case.

As in Fudenberg and Levine [2006] we wish to consider equilibria in which the short-run selves optimize following every history and the long-run player anticipates this reaction and plays like a Stackelberg leader. This is designed to capture what we imagine is the strategic naivete of the short-run self: With one-period lifetimes for the short-run players, this Stackelberg equilibrium is equivalent to subgame-perfect equilibrium in which the long-run player moves first against each short-run player, and to the weaker concept of "SR-perfect Nash equilibrium" defined in Fudenberg and Levine [2006]. If we assume that the long-run player can choose a self-control action e_k that is observed by short-run self k before choosing plan $\mathbf{a}_k$ SR-perfect Nash equilibrium has the same implication here. However, the assumption that e_k is chosen once and for all at the beginning of the life of short-run self k is stronger when the short-run self lives multiple periods. First, the self-control action changes the preferences of the short-run self over many periods. Second, the self-control action cannot be "changed" as long as the particular short-run self is alive. Again, this assumption is intended to capture the strategic naivete of the short-run self.

As is the case in which the short-run self lives only for a single period, the expectations of the short-run self about play by the long-run self do not matter, because the long-run self has already moved. For this reason, the situation does not correspond to a repeated game (which it would in the absence of the commitment assumption.) Moreover, the case for subgame perfection may be stronger here than it is in general, as when the long-run self can commit, the predictions of subgame perfections are less sensitive to changes in the information structure.

Fudenberg and Levine [2006] defines a SR-perfect Nash equilibrium profile to be *equivalent* to a solution to the reduced form optimization problem of maximizing

$$E_{\mathbf{a},Y_1} \sum_{k=0}^{\infty} \delta^k \left(u(Y_k, 0, \mathbf{a}) - C(Y_k, \mathbf{a}) \right)$$

if the reduced strategy induced from the short-run players strategy profile is a solution to the optimization problem. Conversely, if there exists a SR-perfect Nash equilibrium profile with this property for a particular solution to the optimization problem, we say that this solution of the reduced form optimization problem is equivalent to the SR-perfect Nash equilibrium profile. As the conditions of Fudenberg and Levine [2006] Theorem 1 are satisfied, this equivalence does indeed hold.

We now wish to relate solutions to the optimization problem equivalent to SR-perfect Nash equilibria

$$E_{\mathbf{a},Y_1} \sum_{k=0}^{\infty} \delta^k \left(u(Y_k, 0, \mathbf{a}) - C(Y_k, \mathbf{a}) \right)$$

to those of

$$E_{\mathbf{a},Y_1} \sum_{n=0}^{\infty} \delta^n \left((1-\delta) u(a_n, y_n) + m(y_n, \tilde{w}_n) \right)$$

the agent's objective function that we have used as the starting point in this paper. To this end, observe that since we have assumed state independent resource valuation, by Theorem 4 $C(Y_k, \mathbf{a}) = \max_{\mathbf{a}'} \tilde{M}(Y_k, \mathbf{a}') - \tilde{M}(Y_k, \mathbf{a})$ holds for $\mathbf{a} \in \arg\max_{\mathbf{a}'} u(Y_k, 0, \mathbf{a})$, hence for all $\mathbf{a}$. Hence

$$\begin{aligned} &E_{\mathbf{a},Y_1} \sum_{k=0}^{\infty} \delta^k \left((\tilde{U}(Y_k, \mathbf{a}) + \max_{\mathbf{a}'} \tilde{M}(Y_k, \mathbf{a}')) - (\max_{\mathbf{a}'} \tilde{M}(Y_k, \mathbf{a}') - \tilde{M}(Y_k, \mathbf{a})) \right) = \\ &E_{\mathbf{a},Y_1} \sum_{t=0}^{\infty} \delta^k \left(\tilde{U}(Y_k, \mathbf{a}) + \tilde{M}(Y_k, \mathbf{a}) \right) \end{aligned}$$

Let A_{kt} be the probability k is alive at t then we may write

$$\begin{aligned} &E_{\mathbf{a},Y_1} \sum_{t=0}^{\infty} \delta^k \left(\tilde{U}(Y_k, \mathbf{a}) + \tilde{M}(Y_k, \mathbf{a}) \right) = \\ &E_{\mathbf{a},Y_1} \sum_{k=0}^{\infty} \delta^k (1-\delta) \left(\sum_{n=k}^{\infty} (\delta\mu)^{n-k} A_{k,k+n} u(a_{k+n}^S, y_{k+n}) + \sum_{n=k}^{\infty} (\delta\mu)^{n-k} A_{k,k+n} m(y_{k+n}, \tilde{w}_{k+n}) \right) = \\ &E_{\mathbf{a},Y_1} \sum_{n=0}^{\infty} \delta^n \sum_{k=1}^{n} (\mu)^{n-k} A_{k,k+n} (1-\delta) \left(u(a_{k+n}^S, y_{k+n}) + m(y_{k+n}, \tilde{w}_{k+n}) \right) \\ &E_{\mathbf{a},Y_1} \sum_{n=0}^{\infty} \delta^n (1-\delta) \left(u(a_{k+n}^S, y_{k+n}) + m(y_{k+n}, \tilde{w}_{k+n}) \right) \end{aligned}$$

Thus the reduced form of the game is the same agent's objective function that we used in our analysis, hence our study of the solutions of the agent's objective function can be interpreted as an equilibrium of this game.

Notice that we have assumed that the long-run self can commit for the lifetime of the short-run self. This is intended to capture the strategic naivete of the short-run self as a passive actor. Notice that if the long-run self simply moves first each period but cannot commit to contingent plans for future periods, the equilibrium here is still a SR-perfect equilibrium, since we have shown that the solution to the reduced form optimization problem is Markov, so that the long-run self has no wish to renege on his commitment. However, without commitment there can be other equilibria in which the short-run self chooses a plan different from that suggested by the long-run self as part of a repeated game equilibrium. However, we regard such equilibria as inconsistent with our notion of nature of the short-run self.

6. Actions That Increase Cognitive Resources

So far, except when discussing the Ozdenoren et al model, we have assumed that cognitive resources depend on actions only through the foregone utility Δ, and that cognitive resources are maximized by setting $\Delta = 0$. In particular, we have assumed that the replenishment rate $r(\tilde{w}_n)$ does not depend directly on the action taken, and that the value $m(y_n, \tilde{w}_n)$ of cognitive resources is state independent, so actions do not indirectly change the effective level or value of cognitive resources. Notice that Theorem 4 fails if the value of cognitive resources m depends either directly or indirectly on actions or elements of the state other than w_t. This is a key ingredient in the game between the long-run and short-run self: it means that there is no intrinsic conflict between the two over the use of cognitive resources, which is why the reduced form derived from the game is the same agent's objective function we have examined in detail. In the examples of this section, Theorem 4 fails, which raises the question of which is the "correct" objective function: the agent's objective function used above and in the examples, or the reduced form that is derived from the game between long-run and short-run self. Indeed it raises broader questions about what is the appropriate model.

Observe that the issue of value of cognitive resources depending on the history also arises in the linear case. Here in addition to the failure of Theorem 5, Theorem 4 also fails: it is no longer true that the stock of self-control serves merely to allocate the marginal utility of self-control between periods.

The next example illustrates the complexities that can occur when actions can directly influence future marginal costs of self-control: current choices can have the same implications for future choice as a commitment to avoid temptation, but, unlike such a commitment, lowering the marginal cost does not lower the short-run self's highest attainable payoff, so it has no foregone value and thus does not require a control cost to implement.

Example 12: 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 $\overline{\Gamma} > 0$ while from period 2 it is either $\overline{\Gamma}$ 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 $\overline{\Gamma}$.

In period 2 the agent can decide whether to take or resist a simple temptation, with short-run player value S and direct value P for the long run player, with $S > 0 > P$ and $P < -\overline{\Gamma}S$, 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 $\overline{\Gamma}F(1-\delta\mu)$. Hence it is optimal in the reduced form problem to "pay" whenever $\overline{\Gamma}F(1-\delta\mu) < -\delta P$.

In contrast, if paying F today made taking tomorrow impossible, the foregone value of "pay" is $\delta\mu S + F(1-\delta\mu)$, so for some parameters (such as μ close to 1) the commitment will not be optimal even though the arguably equivalent "training" action would be. The difference between commitment and lowering self control costs is a consequence of our assumption that the short-run selves are strategically naïve, so that the short-run player is unconcerned by any action that leaves the feasible set unchanged. Models with non-naïve short-run players may also be of interest, but they are much more complicated.[20]

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^α 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 $1^\alpha - (1-\delta\mu S)^\alpha \geq \alpha\delta\mu S$, while the cost if the agent acts goes to 0 with w^*.

7. Conclusion

The random-lifetimes extension of the dual-self models allows for short-run selves who live more than a single period, and 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. The random-lifetime version of the model also provides a natural way to examine the effect of the length of the periods between potential decision nodes. This is important because the concept of a discrete time period in these decision problems is simply a convenient construction, and the extended model shows how the delay between consecutive decisions should matter for whether agents exhibit "preference reversals."

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

[20] This example suggests that non-naivety is necessary to capture St. Augustine's request "give me chastity and continence, but not yet."

with linear replenishment, benefits, and depletion is equivalent to the "state-free" model. However, once non-linearities become important, so does the timing of self-control decisions and costs; the willpower stock provides a way to model the "spillover" from one period's self control to future self control costs.

We explored some but far from all of the many possible ways to model these non-linearities, and these 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 axiomatic form 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 short-run selves live a single period, not for individual subjects but for the median subject across a number of different experiments. However, several of the experiments studied there are better fit by allowing short-run selves to have random lifetimes; for example Baucell et el [2007] show that paradoxical choices in Allais-type problems are reduced but not eliminated when the payoffs are delayed.

References

Ainslie, G. [2001] *Breakdown of Will*. Cambridge: Cambridge University Press.

Andersen, S., G. Harrison, M. Lau, E. Rutström [2008] "Eliciting risk and time preferences," *Econometrica,* **76**: 583-618.

Baucells, M. and F. H. Heukamp [2010], "Common Ratio using Delay," *Theory and Decision,* 68:149-158.

Baumeister, R. F., E. Bratslavsky, M. Muraven, and D. M. Tice, [1998]. "Ego Depletion: Is the Active Self a Limited Resource?" Journal of Personality and Social Psychology **74**: 1252–1265

Benhabib, J., A. Bisin, and A. Schotter [2010] "Present-bias, quasi-hyperbolic discounting, and fixed costs," *Games and Economic Behavior* 69: 205-223.

Brocas, J. and I. Carrilo [2008] "The Brain as a Hierarchical Organization," *American Economic Review*, 98: 1312-1346.

Chatterjee, K. and V. Krishna [2009] "'A Dual Self' Representation for Stochastic Temptation," American Economic Journal: Microeconomics, 1: 148-167.

Dasgupta, P. and E. Maskin, [2005] "Uncertainty and Hyperbolic Discounting," American Economic Review, 95: 1290–99.

Dekel, E., B. Lipman, and A. Rustichini, [2001] "Representing Preferences with a Unique Subjective State Space," *Econometrica* **69**: 891-934.

Dekel, E., B. Lipman, and A. Rustichini [2009] "Temptation-Driven Preferences," *Review of Economic Studies*,

DellaVigna, S., J. A. List and U. Malmendier [2010] "Testing for Altruism and Social Pressure in Giving," mimeo.

Epper, T., H Fehr-Duda, and A. Bruhin [2009] "Uncertainty Breeds Decreasing Impatience: The Role of Risk Preferences in Time Discounting," mimeo.

Fudenberg, D. and D.K. Levine [2006], "A Dual Self Model of Impulse Control," *American Economic Review* 96: 1449-1476.

Fudenberg, D. and D. K. Levine [2010] "Risk, Delay, and Self Control," mimeo.

Green, I. And J. Myerson [1996], "Exponential Versus Hyperbolic Discounting of Delayed Outcomes: risk and Waiting Time," *American Zoologist* **36**: 496-505.

Gul, F. and W. Pesendorfer [2001]: "Temptation and Self Control," *Econometrica*, **69**: 1403-1436.

Gul, F. and W. Pesendorfer [2004]: "Self Control and the Theory of Consumption," *Econometrica*, **72**, 110-158.

Halevy, Y. [2008] "Strotz meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review* 98:233-265.

Houser, D., D. Schunk, J. Winter, and E. Xiao [2010] "Temptation and Commitment in the Laboratory," mimeo.

Keren, G. and P. Roelofsma [1995], "Immediacy and Certainty in Intertemporal Choice," *Organizational Behavior and Human Decision Making*, **63** 297-297.

Kirby, K. and B. Guatsello [2001] "Making Choices in Anticipation of Similar Future Choices can Increase Self Control," *Journal of Experimental Psychology: Applied* 7: 154-164.

Muraven, M. M. D. Tice and R. Baumeister [1998], "Self-Control as a Limited Resource." *Journal of Personality and Social Psychology* 74, 774-789.

Myerson, J. and L. Green [1995] "Discounting of Delayed Rewards: Models of Individual Choice", *Journal of the Experimental Analysis of Behavior*, 64: 263-276.

Noor, J. [2007] "Commitment and Self Control," *Journal of Economic Theory* 135: 1-34.

Noor, J. and N. Takeoka [2010] "Uphill Self Control," forthcoming in *Theoretical Economics.*

Ozdenoren, E., S. Salant, and D. Silverman [2009] "Willpower and the Optimal Control of Visceral Urges," forthcoming in the *Journal of the European Economic Association*

Rabin, M. [2000] "Risk Aversion and Expected utility Theory: A Calibration Theorem." *Econometrica*, *68*:1281-1292.

Shiv, B. and A. Fedorikhin, [1999] "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making." *Journal of Consumer Research*, *26*: 278-292.

Appendix

Example 2: Derivation of marginal interest rates

$$
\begin{aligned}
MR_t &= \lim_{\tau \to 0} \log\left(\frac{c_{t/\tau+1}}{c_{t/\tau}}\right)/\tau \\
&= \lim_{\tau \to 0} \log\left(\frac{\dfrac{1-\delta+\Gamma(1-\delta\mu)}{(1-\delta)\delta^{t/\tau}+\Gamma(1-\delta\mu)(\delta\mu)^{t/\tau}}}{\dfrac{1-\delta+\Gamma(1-\delta\mu)}{(1-\delta)\delta^{n-1}+\Gamma(1-\delta\mu)(\delta\mu)^{n-1}}}\right)/\tau \\
&= \lim_{\tau \to 0} \log\left(\frac{(1-\delta)+\Gamma(1-\delta\mu)\mu^{t/\tau-1}}{(1-\delta)\delta+\Gamma(1-\delta\mu)\delta\mu^{t/\tau}}\right)/\tau \\
&= \lim_{\tau \to 0} \log\left(\frac{1}{\delta}\frac{(1-\delta)+\Gamma(1-\delta\mu)\mu^{t/\tau-1}}{(1-\delta)+\Gamma(1-\delta\mu)\mu^{t/\tau}}\right)/\tau \\
&= \lim_{\tau \to 0} \frac{\log \delta^{-1} + \log(1+\Gamma\dfrac{1-\delta\mu}{1-\delta}\mu^{n-1}) - \log(1+\Gamma\dfrac{1-\delta\mu}{1-\delta}\mu^{n}))}{\tau} \\
&= \rho + \lim_{\tau \to 0} \frac{\log(1+\Gamma\dfrac{1-\delta\mu}{1-\delta}\mu^{t/\tau-1}) - \log(1+\Gamma\dfrac{1-\delta\mu}{1-\delta}\mu^{t/\tau}))}{\tau} \\
&= \rho + \lim_{\tau \to 0} \frac{\log(1+\Gamma\dfrac{1-\delta\mu}{1-\delta}\mu^{t/\tau-1}) - \log(1+\Gamma\dfrac{1-\delta\mu}{1-\delta}\mu^{t/\tau}))}{\tau} \\
&= \rho + \Gamma \lim_{\tau \to 0} \frac{1-\delta\mu}{1-\delta}\frac{\exp(-\eta t)(\mu^{-1}-1)}{\tau} = \rho + \Gamma\frac{\rho+\eta}{\rho}\exp(-\eta t)\eta
\end{aligned}
$$

Example 2 Extended

Now extend the analysis of example 2 to consider the choice between consuming an amount at time $n+\ell$ and a single unit of consumption at time $1+\ell$. If the agent is informed of the choice at time 1 but cannot commit to a decision until time ℓ, then because there is a Markov solution that depends only on the current state, it follows that the choice will be the same as if the choice only became available in period ℓ. Moreover because the preferences of both long-run and short-run selves are stationary, the amount of consumption at $n+\ell$ that makes the long-run self indifferent between consuming 1 unit at $1+\ell$ or waiting to consume at $n+\ell$ is the same as the amount that would make the agent indifferent between consuming 1 unit at time 1 between a unit at time 1 or waiting to consume more at time $1+n$.

The situation is different when the agent can make an irrevocable choice at time 1: the first short-run self faces a temptation cost but no other short-run self does. This raises the possibility that the choice may depend upon ℓ. We will compute the value of $c_{1+\ell,n+\ell}$ that makes the long run self indifferent between a unit of consumption at time $1+\ell$ and $c_{1+\ell,n+\ell}$ units at time $n+\ell$; we will then use this to compute effective marginal interest rates on consumption.

Observe that as in the case $n=1$ if the long-run self is indifferent between one unit at time $1+\ell$ and $c_{1+\ell,n+\ell}$ at time $n+\ell$, then since $\mu<1$ the initial short-run self strictly prefers one unit at time $1+\ell$, so the temptation is to consume at time $1+\ell$. Thus consuming the one unit at time $1+\ell$, incurs no control cost, so the long-run player's utility of utility of this choice is $(1-\delta)\delta^{\ell}$. (Since the decision is taken at time 1, no other short-run self feels any temptation.) The initial short run self gets average value utility $(1-\delta\mu)(\delta\mu)^{\ell}$ from consumption at time $1+\ell$, and $(1-\delta\mu)(\delta\mu)^{n+\ell-1}c_{1+\ell,n+\ell}$ from the delayed option, so the control cost of the delayed option is $\Gamma(1-\delta\mu)(\delta\mu)^{\ell}(1-(\delta\mu)^{n-1}c_{1+\ell,n+\ell})$. The direct utility of consuming $c_{1+\ell,n+\ell}$ at $n+\ell$ is $(1-\delta)\delta^{n+\ell-1}c_{1+\ell,n+\ell}$, so the reduced form utility is $\left((1-\delta)\delta^{n+\ell-1}c_{1+\ell,n+\ell}-\Gamma(1-\delta\mu)(\delta\mu)^{\ell}(1-(\delta\mu)^{n-1}c_{1+\ell,n+\ell})\right)$ Equating the present value from choosing the earlier consumption of 1 at $1+\ell$ to that from $c_{1+\ell,n+\ell}$ at determines the consumption level leading to indifference.

$$(1-\delta)\delta^{\ell}=(1-\delta)\delta^{n+\ell-1}c_{1+\ell,n+\ell}-\Gamma(1-\delta\mu)(\delta\mu)^{\ell}(1-(\delta\mu)^{n-1}c_{1+\ell,n+\ell})$$

Dividing both sides by δ^{ℓ} yields

$$1-\delta=(1-\delta)\delta^{n-1}c_{1+\ell,n+\ell}-\Gamma(1-\delta\mu)\mu^{\ell}(1-(\delta\mu)^{n-1}c_{1+\ell,n+\ell}),$$

so we see that the effect of moving both consumption dates forward by ℓ is to reduce the marginal cost of self control from Γ to $\Gamma\mu^{\ell}$.

We can then solve for consumption giving

$$c_{1+\ell,n+\ell}=\frac{(1-\delta)+\Gamma(1-\delta\mu)\mu^{\ell}}{\delta^{n-1}((1-\delta)+\Gamma(1-\delta\mu)\mu^{n+\ell-1})}$$

Note that this says that the effect of a delay $\ell > 0$ is to decrease the effective cost of self control by the factor η^{ℓ}: Intuitively, the initial short run self is less tempted as the delay increases so the self-control cost goes down. If we suppose that the delay is some fixed real time, so the choice is between 1 unit at time t or more consumption at time $t+s$, and send the time period to 0, the approximation used in the text shows that the marginal interest rate converges to

$$\rho + \Gamma \frac{\rho+\eta}{\rho} \exp(-\eta(t+s)).$$

Proof of Proposition 3

Proposition 3: *Set*

$$F^* = \gamma\left(\frac{\delta(1-\mu)}{1+\gamma}\right)^2$$

If $F > F^*$ *then* $\mathbf{a}^0$ *is optimal for*

$$B \leq \delta \frac{1+\gamma\mu}{1+\gamma} \equiv B^*$$

and $\mathbf{a}^1$ *if* $B \geq B^*$. *If* $F \leq F^*$ *then*

$$\overline{B} \equiv \delta \frac{1+\gamma\mu^2}{1+\gamma\mu} - \frac{1+\gamma}{\delta(1+\gamma\mu)} F \geq B^* \geq \delta\mu + \frac{1+\gamma}{\gamma\delta(1-\mu)} F \equiv \underline{B}$$

and $\mathbf{a}^0$ *is optimal for* $B \leq \underline{B}$, $\mathbf{a}^F$ *is optimal for* $\underline{B} \leq B \leq \overline{B}$, *and* $\mathbf{a}^1$ *is optimal for* $B \geq \overline{B}$.

Proof:

1) As worked out in the text, the payoff to F is

$$-(1-\delta)F - \gamma\left(F(1-\delta\mu) + \delta\mu\overline{U}(2,0,0)\right) =$$
$$-(1-\delta)F - \gamma\left(F(1-\delta\mu) + \delta\mu(1-\delta\mu)\max\{0, B-\delta\mu\}\right)$$

2) If $\mathbf{a}^0$ (don't avoid, don't purchase) is chosen, the direct utility is 0, and the reduced form utility is the temptation cost incurred in the second period: $V(\mathbf{a}^0) = -\gamma\delta(1-\delta\mu)\max\{0, B-\delta\mu\}$.

3) If $\mathbf{a}^1$ is chosen, the direct utility is $(1-\delta)(\delta B - \delta^2)$, while the cost of self-control is in period 2 and is $-\gamma(1-\delta\mu)\min\{0, B-\delta\mu\}$ – self-control is needed only when the short-run player does not want to purchase. Thus
$V(\mathbf{a}^1) = (1-\delta)\delta(B-\delta) + \gamma\delta(1-\delta\mu)\min\{0, B-\delta\mu\}$

If $B \le \delta\mu$ as noted in the text the optimum is not to purchase and there is not temptation cost. So it is also not optimal to avoid in the first period, and the optimum is $\mathbf{a}^0$. Next suppose that $B > \delta\mu$ and consider the period 2 choice assuming the avoidance cost was not paid. As noted, if the purchase is not made, the average value from period 2 on is $-\gamma(1-\delta\mu)(B-\delta\mu)$, while if it is, the average value is $(1-\delta)(B-\delta)$. So the optimum is not to purchase when

$$B \le \delta \frac{1-\delta+\gamma(1-\delta\mu)\mu}{1-\delta+\gamma(1-\delta\mu)} \equiv B*$$

Next observe that since $B > \delta\mu$ the present value of utility from avoiding is given by

$$-(1-\delta)F - \gamma\left(F(1-\delta\mu) + \delta\mu(1-\delta\mu)(B-\delta\mu)\right).$$

Then $V(\mathbf{a}^F) \ge V(\mathbf{a}^0)$ if and only

$$B \ge \delta\mu + \frac{1-\delta+\gamma(1-\delta\mu)}{\gamma(1-\delta\mu)\delta(1-\mu)} F \equiv \underline{B}.$$

Since $B - \delta\mu > 0$ this implies there is a range of sufficiently small F where $\mathbf{a}^F$ is better and a range of F so large that $\mathbf{a}^0$ is better.

Finally, $V(\mathbf{a}^F) \ge V(\mathbf{a}^1)$ if
$-(1-\delta)F - \gamma\left(F(1-\delta\mu) + \delta\mu(1-\delta\mu)(B-\delta\mu)\right) \ge \delta(1-\delta)(B-\delta)$, or

$$B \le \delta \frac{1-\delta+\gamma(1-\delta\mu)\mu^2}{1-\delta+\gamma(1-\delta\mu)\mu} - \frac{1-\delta+\gamma(1-\delta\mu)}{\delta(1-\delta+\gamma(1-\delta\mu)\mu)} F \equiv \overline{B}$$

We conclude that $\mathbf{a}^F$ is best when

$$\delta\frac{1-\delta+\gamma(1-\delta\mu)\mu^2}{1-\delta+\gamma(1-\delta\mu)\mu}-\frac{1-\delta+\gamma(1-\delta\mu)}{\delta(1-\delta+\gamma(1-\delta\mu)\mu)}F \geq B$$
$$\geq \delta\mu+\frac{1-\delta+\gamma(1-\delta\mu)}{\gamma(1-\delta\mu)\delta(1-\mu)}F$$

Straightforward algebra shows that there is a non-empty interval of B where $\vec{\mathbf{a}}^F$ is best when

$$F \leq \frac{\gamma(1-\delta\mu)(1-\delta)\delta^2(1-\mu)^2}{(1-\delta+\gamma(1-\delta\mu))^2} = F^*$$

If $F > F^*$ it is not optimal to use $\mathbf{a}^F$; in this case the optimum is determined from the condition for $V(\mathbf{a}^0) \geq V(\mathbf{a}^1)$ above. If $F \leq F^*$, and if $B \leq \underline{B}$ then $V(\mathbf{a}^F) \leq V(\mathbf{a}^0)$ and $V(\mathbf{a}^1) \leq V(\mathbf{a}^0)$, so $\mathbf{a}^0$ is optimal; if $\underline{B} \leq B \leq \overline{B}$ then $V(\mathbf{a}^F) \geq V(\mathbf{a}^0)$ and $V(\mathbf{a}^F) \geq V(\mathbf{a}^1)$, so $\mathbf{a}^F$ is optimal; while if $B \geq \overline{B}$ then $V(\mathbf{a}^F) \leq V(\mathbf{a}^1)$ and $V(\mathbf{a}^1) \geq V(\mathbf{a}^0)$, so $\mathbf{a}^1$ is optimal. Finally note that $\delta\mu \leq \underline{B}, B^*$, so that the case $B \leq \delta\mu$ where $\mathbf{a}^0$ is optimal is included in this result.

☑

Timing of Temptation

Here we consider how the timing of the foregone utility shock Δ_n within a period interacts with linear replenishment. In particular we find that the timing does not matter when periods are small, though of course it does matter when periods are longer.

Model 1: $\dot{W}_t = \kappa(\overline{w}-W_t)-\tilde{\Delta}_t$, where $\tilde{\Delta}_t$ is a constant flow during the period. The solution is $W_t = \overline{w}-\tilde{\Delta}/\kappa-\exp(-\kappa t)(\overline{w}-\tilde{\Delta}/k-w_1)$. For one period this gives
$w_2 \approx \overline{w}-\tilde{\Delta}/\kappa-(1-\kappa\tau)(\overline{w}-\tilde{\Delta}/\kappa-w_1) = \lambda\overline{w}+(1-\lambda)\omega_0-\tau\tilde{\Delta}$.

Model 2: If $\tau\tilde{\Delta}$ is incurred at the beginning of the period, then cognitive resources jump down immediately from w_1 to $w_1^+ = w_1-\Delta$, and then follows $\dot{W}_t = \kappa(\overline{w}-W_t)$. Hence

$$
\begin{aligned}
w_2 &= \overline{w} - \exp(-\kappa\tau)(\overline{w} - w_1 - \tau\tilde{\Delta}) \\
&\approx \lambda\overline{w} + (1-\lambda)w_1 - \tau\tilde{\Delta} + \kappa\tau^2\tilde{\Delta}\,. \\
&\approx \lambda\overline{w} + (1-\lambda)w_1 - \tau\tilde{\Delta}
\end{aligned}
$$

Note that this is the same as if the resources were withdrawn at the end of the period.

Turning from jumps to flows, we observe that if the shock Δ is fixed rather than proportional to $\tilde{\Delta}$ then the difference between the terminal stocks is of order τ rather than τ^2. However, this is consistent with the difference between discrete shocks and flows. In the case of shocks on order $\tilde{\Delta}\tau$ there may be order $n = 1/\tau$ shocks per unit of calendar time, so that the overall error per unit of calendar time goes to zero when the error per period is of order τ^2. However, fixed shocks Δ cannot occur too frequently, typically only a finite number K of such shocks per unit of calendar time. Hence if the per period error is of order τ, the error over calendar time is of order $K\tau$, which vanishes as the length of period goes to zero.

Proof of Theorem 5

Theorem 5: *For any model with linear benefit* γ*, depletion and replenishment* λ *define*

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

Then if $\mathbf{a}$ *is a solution to the linear model with parameter* Γ *then it is a solution to the* (λ,γ) *model in which actions are independent of* w_n *and all such solutions to the* (λ,γ) *model are solutions to the* Γ *model.*

Proof: Recall that $\tilde{w}_n = w_n - \Delta_n$. With linear replenishment $w_{n+1} = \tilde{w}_n + \lambda(\overline{w} - \tilde{w}_n) = w_n - \Delta_n + \lambda(\overline{w} - w_n + \Delta_n)$. Define the *willpower deficit* as $D_n = \overline{w} - w_n$, then $D_{n+1} = (1-\lambda)D_n + (1-\lambda)\Delta_n$. Hence

$$D_{n+1} = (1-\lambda)^n D_1 + \sum\nolimits_{n'=1}^{n} (1-\lambda)^{n'} \Delta_{n'}\,.$$

Recall that the average value of cognitive resources in the linear case is $M = (1-\delta)\sum\nolimits_{n=0}^{\infty} \delta^n \gamma w$.

It follows that the total value of cognitive resources is

$$
\begin{aligned}
M/(1-\delta) &= \gamma\Big[\,\overline{w} - \sum\nolimits_{n=0}^{\infty}\delta^n D_n\,\Big] \\
&\gamma\Big[\,\overline{w} - \sum\nolimits_{n=0}^{\infty}\delta^n\left((1-\lambda)^n D_1 + \sum\nolimits_{n'=1}^{n}(1-\lambda)^{n-n'}\Delta_{n'} + \Delta_{n+1}\right)\Big] \\
&= \gamma\Big[\,\overline{w} - \sum\nolimits_{n=0}^{\infty}\delta^n\left((1-\lambda)^n D_1 + \sum\nolimits_{n'=0}^{n}(1-\lambda)^{n-n'}\Delta_{n'}\right)\Big] \\
&= \gamma\Big[\,\overline{w} - \sum\nolimits_{n=0}^{\infty}\delta^n(1-\lambda)^n D_1 - \sum\nolimits_{n=0}^{\infty}\delta^n\sum\nolimits_{n'=0}^{n}(1-\lambda)^{n-n'}\Delta_{n'}\Big] \\
&= \gamma\Big[\,\overline{w} - \sum\nolimits_{n=0}^{\infty}\delta^n(1-\lambda)^n D_1 - \sum\nolimits_{n'=0}^{\infty}\delta^{n'}\Delta_{n'}\sum\nolimits_{n=n'}^{\infty}(\delta(1-\lambda))^{n-n'}\Big] \\
&= \gamma\left[\,\overline{w} - \sum\nolimits_{n=0}^{\infty}\delta^n(1-\lambda)^n D_1 - \frac{1}{1-\delta(1-\lambda)}\sum\nolimits_{n'=0}^{\infty}\delta^{n'}\Delta_{n'}\right]
\end{aligned}
$$

Hence if we define

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

we see the equivalence to the simple linear model without replenishment.

☑

Proof of Proposition 7

Proposition 7: *Suppose there is increasing marginal cost of self control and there is strictly partial linear replenishment,* $0<\kappa$. *Then there is* $\underline{\tau}>0$ *such that if* $\tau<\underline{\tau}$, *there are* $\left|\overline{P}_\tau\right| > |\underline{P}_\tau| > 0$ *such that it is optimal to resist forever if* $|P| > \overline{P}_\tau$, *it is optimal to resist until period* $\infty > \hat{\ell} > 1$ *then take if* $\left|\overline{P}_\tau\right| > |P| > |\underline{P}_\tau|$, *and it is optimal to take immediately if* $|\underline{P}_\tau| > P$. *Let* $\overline{P}_0, \underline{P}_0$ *denote the limits as* $\tau\to 0$. *Let* W_s *be the solution to the differential equation*

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

and let W_∞ *be the solution to* $0 = \kappa(\overline{w} - W_\infty) + \frac{\partial f}{\partial \Delta}(W_\infty, 0)(\rho+\eta)S$. *Then*

$$\left|\overline{P}_0\right| = -(\rho+\eta)S\int_0^\infty e^{-(\rho+\kappa)t} m'(W_\infty)\frac{\partial f}{\partial \Delta}(W_\infty, 0)dt,$$

$$|\underline{P}_0| = -(\rho+\eta)S\int_0^\infty e^{-(\rho+\kappa)t}m'(\overline{w})\frac{\partial f}{\partial \Delta}(\overline{w},0)dt$$

and if $\left|\overline{P}_0\right| > |P| > |\underline{P}_0|$ *then* $\hat{s} = \lim_{\tau\to 0}\tau\hat{\ell}$ *is finite and strictly positive, and is determined by*

$$|P| = (\rho+\eta)S\int_0^\infty e^{-(\rho+\kappa)t}m'(\overline{w}-e^{-\kappa t}W_{\hat{s}})\frac{\partial f}{\partial \Delta}(W_{\hat{s}},0)dt$$

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

$$(1-\delta)\Big(\sum_{n=0}^{\ell-1}\delta^n m(f(w_{n+1}^\ell,(1-\delta\mu))S) + (\sum_{n=\ell}^{\infty}\delta^n m(w_{n+1}^\ell))\Big) + \delta^\ell P \text{, and}$$

the bigger is ℓ 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 ℓ 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 ℓ minus the value at $\ell+1$

$$D(\ell) = (1-\delta)\Big(\sum_{n=0}^{\ell-1}\delta^n m(f(w_{n+1}^\ell,(1-\delta\mu)S)) + (\sum_{n=\ell}^{\infty}\delta^n m(w_{t+1}^\ell))\Big) + \delta^{n'}P$$
$$-(1-\delta)\Big(\sum_{n=0}^{\ell}\delta^n m(f(w_{n+1}^{\ell+1},(1-\delta\mu)S)) + (\sum_{n=\ell+1}^{\infty}\delta^n m(w_{n+1}^{\ell+1}))\Big) - \delta^{\ell+1}P$$

Note that for $n \le \ell$ we have $w_n^\ell = w_n^{\ell+1}$, so we can write this difference as

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

Observe that because there is partial replenishment $w_{\ell+n}^\ell$ strictly decreases in ℓ.

We now use the assumption of increasing cost of self control to conclude there is a $\underline{\tau}$ such that for $\tau < \underline{\tau}$ each individual term in $D(\ell)$ is strictly increasing in ℓ, and hence that $D(\ell)$ is strictly increasing. The first term $m(w_{\ell+1}^\ell) - m(f(w_{\ell+1}^\ell,(1-\delta\mu)S)) = m(f(w_{\ell+1}^\ell,0)) - m(f(w_{\ell+1}^\ell,(1-\delta\mu)S)$, strictly decreases in $w_{\ell+1}^\ell$ 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 $w_{\ell+1}^\ell$ decreases in ℓ, these differences increase.

For the terms in the sum, since n runs from $\ell+1$ to ∞, the arguments $w_{n+1}^{\ell}, w_{n+1}^{\ell+1}$ have the form $w_{\ell+\ell'}^{\ell}, w_{\ell+\ell'}^{\ell+1}$, and the former decrease with ℓ. The individual terms have the form $m(w_{\ell+\ell'}^{\ell}) - m(w_{\ell+\ell'}^{\ell+1})$ where $w_{\ell+1}^{\ell+1} = f(w_{\ell+1}^{\ell},(1-\delta\mu)S)$ and $w_{\ell+\ell'}^{\ell+i} = \overline{w} - (1-\lambda)^{\ell'-1}(\overline{w} - w_{\ell+1}^{\ell+i})$. Putting this together, we have

$$\begin{aligned}
&m(w_{\ell+\ell'}^{\ell}) - m(w_{\ell+\ell'}^{\ell+1}) = \\
&m(\overline{w} - (1-\lambda)^{\ell'-1}(\overline{w} - w_{\ell+1}^{\ell})) - m(\overline{w} - (1-\lambda)^{\ell'-1}(\overline{w} - f(w_{\ell+1}^{\ell},(1-\delta\mu)S))) = \\
&m(A\overline{w} + (1-A)w_{\ell+1}^{\ell}) - m(A\overline{w} + (1-A)f(w_{\ell+1}^{\ell},(1-\delta\mu)S))) = \\
&m(A\overline{w} + (1-A)f(w_{\ell+1}^{\ell},0)) - m(A\overline{w} + (1-A)f(w_{\ell+1}^{\ell},(1-\delta\mu)S)))
\end{aligned}$$

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

Notice that w_n is bounded below by the steady state. Hence D is bounded above as a function of ℓ. If P is large enough in absolute value (it is negative) given all the other parameters then this expression is negative for all ℓ, and it is optimal to wait forever; let $\overline{P}_{\tau}$ 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 $\underline{P}_{\tau}$ be the largest such P in absolute value.

Next we assume that τ is small, and show that $\left|\overline{P}_{\tau}\right| > |\underline{P}_{\tau}|$. Observe that

$$\begin{aligned}
w_{\ell+\ell'}^{\ell} &= \overline{w} - (1-\lambda)^{\ell'-1}(\overline{w} - w_{\ell+1}^{\ell}) = \overline{w} - (1-\lambda)^{\ell'-1}(\overline{w} - w_{\ell+1}^{\ell}) \\
w_{\ell+\ell'}^{\ell+1} &= \overline{w} - (1-\lambda)^{\ell'-1}(\overline{w} - w_{\ell+1}^{\ell+1}) = \overline{w} - (1-\lambda)^{\ell'-1}(\overline{w} - f(w_{\ell+1}^{\ell},(1-\delta\mu)S))
\end{aligned}$$

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

$$\begin{aligned}
d(s,\tau) = {} & (m(W_{s+\tau}^{s}) - m(f(W_{s+\tau}^{s},(1-\delta\mu)S))) \\
& + \sum\nolimits_{n=2}^{\infty} e^{-\rho n}(m(W_{s+n\tau}^{s}) - m(W_{s+n\tau}^{s+\tau})) + P
\end{aligned}$$

where

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

$$W_{s+n\tau}^{s} = \overline{w} - (1-\lambda)^{n-1}(\overline{w} - W_{s+\tau}^{s}).$$

The first term of d converges to zero as $\tau \to 0$, and since m, f are differentiable the sum converges to

$$d(s) \equiv (\rho + \eta)S\int_0^\infty e^{-(\rho+\kappa)t} m'(\overline{w} - e^{-\kappa t}(\overline{w} - W_s))\frac{\partial f}{\partial \Delta}(W_s, 0)dt$$

where W_s is the solution to the differential equation

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

with initial condition $W_0 = \overline{w}$. 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

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

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

$$|\underline{P}_0| = -(\rho + \eta)S\int_0^\infty e^{-(\rho+\kappa)t} m'(\overline{w})\frac{\partial f}{\partial \Delta}(\overline{w}, 0)dt <$$
$$-(\rho + \eta)S\int_0^\infty e^{-(\rho+\kappa)t} m'(W_\infty))\frac{\partial f}{\partial \Delta}(W_\infty, 0)dt = \left|\overline{P}_0\right|,$$

and hence $|\underline{P}_\tau| < \left|\overline{P}_\tau\right|$ must hold for all sufficiently small τ.

☑

Proof of Theorem 8

Theorem 8: *If* $\tilde{C}(Y_k, \mathbf{a}) = \Gamma(\overline{U}(Y_k) - \tilde{U}(Y_k, \mathbf{a}))$ *then* $\tilde{C}(Y_k, \mathbf{a}) = C(Y_k, \mathbf{a})$.

Proof: We must show

$$E_{\mathbf{a},h_1}\sum_{n=0}^\infty \delta^n \Delta(y_{1+n}, a_{1+n}) =$$
$$E_{\mathbf{a},h_1}\sum_{n=0}^\infty \delta^n \left((1-\mu)[\overline{U}(y_{1+n}) - U^{1+n}(h_{1+n}, \mathbf{a})]\right) + \mu[\overline{U}(y_1) - U^1(h_1, \mathbf{a})]$$

We do so by showing that we can apply the principle of optimality for the short-run self to compute the opportunity cost as a sum of current and future foregone utilities, then rearrange the resulting sum to get the desired result. As noted in the text, the principle of optimality for the short-run self gives the opportunity cost as a sum of weighted increments:

$$\begin{aligned}
&\bar{U}(y_n) - U^n(h_n, \mathbf{a}) = \\
&\bar{U}(y_n) - E^n_{\mathbf{a},h_n} \sum\nolimits_{\ell=0}^{\infty} (\delta\mu)^{\ell} u(a_{n+\ell}, y_{n+\ell}) = \\
&E^n_{\mathbf{a},h_n} \left(\sum\nolimits_{\ell=0}^{\infty} (\delta\mu)^{\ell} \left(\Delta(y_{n+\ell}, a_{n+\ell}) \right) \right)
\end{aligned}$$

Writing out the full average present value of opportunity costs we can in turn express that as a weighted sum of foregone utilities.

$$\begin{aligned}
&E_{\mathbf{a},h_1} \sum\nolimits_{\ell=0}^{\infty} \delta^{\ell} \left((1-\mu)[\bar{U}(y_{1+\ell}) - U^{1+\ell}(h_{1+\ell}, \mathbf{a})] \right) + \mu[\bar{U}(y_1) - U^1(h_1, \mathbf{a})] = \\
&E_{\mathbf{a},h_1} \sum\nolimits_{\ell=0}^{\infty} \delta^{\ell} \left((1-\mu)[\sum\nolimits_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} \left(\Delta(y_{1+\ell+\ell'}, a_{1+\ell+\ell'}) \right) \right) \\
&+\mu[\sum\nolimits_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} \left(\Delta(y_{1+\ell'}, a_{1+\ell'}) \right)]
\end{aligned}$$

Set $\ell'' = \ell + \ell'$. The final step is to rearrange this sum of increments to get the recursive cost

$$\begin{aligned}
&E_{\mathbf{a},h_1} \sum\nolimits_{\ell''=0}^{\infty} \sum\nolimits_{\ell' \leq \ell:}^{\infty} \left((1-\mu) \left(\delta^{\ell''} \mu^{\ell'} \Delta(y_{1+\ell''}, a_{1+\ell''}) \right) \right) \\
&\quad + \mu[\sum\nolimits_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} \left(\Delta(y_{1+\ell'}, a_{1+\ell'}) \right)] = \\
&E_{\mathbf{a},h_1} \sum\nolimits_{\ell''=0}^{\infty} (1-\mu) \delta^{\ell''} \Delta(y_{1+\ell''}, a_{1+\ell''}) \sum\nolimits_{\ell' \leq \ell''}^{\infty} \mu^{\ell'} \\
&\quad + \mu[\sum\nolimits_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} \left(\Delta(y_{1+\ell'}, a_{1+\ell'}) \right)] = \\
&E_{\mathbf{a},h_1} \sum\nolimits_{\ell''=0}^{\infty} \delta^{\ell''} \Delta(y_{1+\ell''}, a_{1+\ell''}) (1 - \mu^{\ell''+1}) \\
&\quad + \mu[\sum\nolimits_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} \left(\Delta(y_{1+\ell'}, a_{1+\ell'}) \right)] = \\
&E_{\mathbf{a},h_1} \sum\nolimits_{\ell''=0}^{\infty} \delta^{\ell''} \Delta(y_{1+\ell''}, a_{1+\ell''})
\end{aligned}$$

which is the desired result.