Econometrica Supplementary Material

SUPPLEMENT TO "TIMING AND SELF-CONTROL"

(*Econometrica*, Vol. 80, No. 1, Januray 2012, 1–42)

BY DREW FUDENBERG AND DAVID K. LEVINE

W1. AN EXAMPLE MOTIVATED BY DELLAVIGNA ET AL.

FIRST WE CONSIDER an example motivated by DellaVigna et al.'s (2010) experiment on door-to-door charitable fund-raising that illustrates a form of nonmonotonicity. DellaVigna et al. (2010) provided homeowners with the option to avoid a fundraiser by using a "Do Not Disturb" check box and found that those who chose avoidance were concentrated among people who donate less when avoidance is not possible. As we explain here, this example is consistent with our model: whether the agent is willing to pay to avoid a temptation can be a nonmonotone function of the temptation's short-run utility. Intuitively, those who are willing to pay to avoid a temptation can have intermediate utilities from choosing it, as those with very low values may find it easy to resist without commitment, while those with high values will have a correspondingly high control cost for choosing the commitment.

EXAMPLE W1—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.) As we will see, costly self-control leads to a nonmonotonicity: 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 self and shorter-run self, so the opportunity should be taken advantage of and not avoided. When the opportunity is very bad, the shorter-run self will not indulge much, 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 self and shorter-run self, so the long-run self may choose to commit so as to avoid the temptation.

The example is very simple and stylized. In period 1, a cost $F \geq 0$ may be paid or not; think of this 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, we first compute temptation values in each period and state, and then compute the agent's objective function. We then solve the various inequalities to see when each action is best.

 DOI: 10.3982/ECTA9655

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 shorter-run self's average present value from doing nothing is 0 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 shorter-run self's value is $-F(1-\delta\mu)$, while if it is not chosen, the shorter-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 we suppose $B-\delta\mu>0$ and compute the agent's decision. Resisting temptation in period 2 will cost $\Gamma(1-\delta\mu)(B-\delta\mu)$, so the purchase will be made when $(1-\delta)(B-\delta)\geq -\Gamma(1-\delta\mu)(B-\delta\mu)$. If avoidance is chosen, the shorter-run self pays the avoidance cost of $-(1-\delta\mu)F$ in the first period. Since the temptation value is $\delta\mu\bar{U}_2$, the average present value of avoidance is

$$\begin{aligned}&-(1-\delta)F-\Gamma(F(1-\delta\mu)+\delta\mu\bar{U}_2)\\&\quad=-(1-\delta)F-\Gamma(F(1-\delta\mu)+\delta\mu(1-\delta\mu)\max\{0,B-\delta\mu\})\end{aligned}$$

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

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 $\mathbf{a}^F$. We have the following characterization of the optimal decision rule:

PROPOSITION W1: *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

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

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

PROOF: (i) As worked out above, the payoff to F is

$$\begin{aligned}&-(1-\delta)F-\gamma(F(1-\delta\mu)+\delta\mu\bar{U}_2)\\&\quad=-(1-\delta)F-\gamma(F(1-\delta\mu)+\delta\mu(1-\delta\mu)\max\{0,B-\delta\mu\}).\end{aligned}$$

(ii) If $\mathbf{a}^0$ (do not avoid, do not purchase) is chosen, the direct utility is 0 and the reduced form utility is the temptation cost incurred in the second period:

$$V(\mathbf{a}^0)=-\Gamma\delta(1-\delta\mu)\max\{0,B-\delta\mu\}.$$

(iii) 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 $-\Gamma(1-\delta\mu)\min\{0,B-\delta\mu\}$, as self-control is needed only when the shorter-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 $B\le\delta\mu$, then the optimum is not to purchase and there is no temptation cost; here 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. 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 made, 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

$$-(1-\delta)F-\Gamma(F(1-\delta\mu)+\delta\mu(1-\delta\mu)(B-\delta\mu)).$$

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

$$\begin{aligned}&-(1-\delta)F-\Gamma(F(1-\delta\mu)+\delta\mu(1-\delta\mu)(B-\delta\mu))\\&\quad\ge\delta(1-\delta)(B-\delta)\end{aligned}$$

or

$$B\le\delta\frac{1-\delta+\Gamma(1-\delta\mu)\mu^2}{1-\delta+\Gamma(1-\delta\mu)\mu}-\frac{1-\delta+\Gamma(1-\delta\mu)}{\delta(1-\delta+\Gamma(1-\delta\mu)\mu)}F\equiv\bar{B}.$$

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 nonempty interval of B where $\mathbf{a}^F$ 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 $\underline{B}\leq B\leq \bar{B}$, 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; if $B\geq\bar{B}$, 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. *Q.E.D.*

Note that the right-hand side inequality in $\underline{B}\leq B\leq\bar{B}$ gets harder to satisfy as $\mu\to 1$ or as $\Gamma\to 0$. In the former case, the interests of the shorter-run self are nearly aligned with those of the long-run self, while in the second case, the shorter-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.

DellaVigna et al. (2010) found that if an option to avoid the fundraiser 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. If we imagine that without checking the box there is a cost of avoiding, then this is as our model predicts: the lower the cost of avoidance, the more people will choose it. As noted above, DellaVigna et al. found that those who chose avoidance were 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 $B/\underline{B}$ in the population is greater than or equal to 1 and the highest value of $B/\bar{B}$ also exceeds 1, 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. found. On the other hand, if the highest value of $B/\bar{B}$ in the population is less than or equal to 1 while the lowest value of $B/\underline{B}$ is below 1, our model predicts the opposite result. A more elaborate experiment could vary the value of B and the cutoffs more systematically—for example, in the flier describing the visit, indicating that a level of matching funds are available ($3 to the charity for every $1 you donate, for example). This would make it possible to test for the nonmonotonicity in B that the model predicts.

W2. STATE DEPENDENCE

In the text, we defined the value of cognitive resources to be *state-independent* if it depends on the state only through the stock of willpower; in a slight abuse of notation, we write this as $m(y_n, \tilde{w}_n) = m(\tilde{w}_n)$. State-independent resource valuation implies that the action most favored by the shorter-run self maximizes the utility of cognitive resources. To see this, define

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

and note that the value on the right-hand side is independent of k. If each period's action is chosen to maximize the value $U^k(h_n, \mathbf{a})$ of the current shorter-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 shorter-run utility in each period on each history also maximizes M.

THEOREM W2: *With state-independent resource valuation*, $\arg\max_{\mathbf{a}} U(h_n, \mathbf{a}) = \arg\max_{\mathbf{a}} M(h_n, \mathbf{a})$.

W3. THE GAME BETWEEN LONG-RUN AND SHORTER-RUN SELVES

Here we show that the optimization problem in the text can be identified with the outcome of a game between the long-run self and a sequence of shorter-run selves. To do this, we introduce an augmented state variable Y_k that is defined in any period n in which a new shorter-run self is born and includes along with y_n the value of n as well as available cognitive resources; that is, $Y_k = (y_n, w_n, n)$. Notice that any strategy $\mathbf{a}$ that maps histories to actions induces a well defined stochastic kernel $\Pi(\mathbf{a}, Y_k)[dY_{k+1}]$ based on the original stochastic kernel and the laws of motion for cognitive resources.

In the game formulation, the actions are taken by the shorter-run selves, and the long-run self chooses self-control actions that influence the preferences of the shorter-run self; the control cost we specified in the text will now correspond to a reduction in the utility function as opposed to an additional term.

Each shorter-run self can be thought of as choosing an $\mathbf{a}$: Although this contains irrelevant information such as how the shorter-run self will behave after he "dies," we will ignore this in computing the shorter-run self's payoff. Following Fudenberg and Levine (2006), we assume that before the shorter-run self moves, the long-run self chooses a *self-control action* $e \in E$. It is convenient to take $E = 0 \cup \mathbf{A}$; we then interpret $e \in \mathbf{A}$ as the "suggested action" and $e = 0$ as "no recommendation" or "no self-control." In the game formulation, cognitive resources follow the exact same equations of motion as in the reduced form model and depend only on the action actually taken by the short-run self and not on the self-control action e. We consider a sequence of stage games between the long-run self and the kth shorter-run self. The kth stage game consists of a choice of self-control action e by the long-run self and a response $\mathbf{a}$ by the shorter-run self. The utility of the kth shorter-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, $\mathbf{a}_k$. A strategy from the long-run self is a map $\mathbf{e}$ from the previous history to a self-control action, and a strategy for the kth shorter-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 shorter-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.

Fudenberg and Levine (2006) specified the procedure for deriving a utility function from an underlying objective function and a "cost of self-control" function. We mimic that procedure here to show that the equilibria of the game are equivalent to those of a particular optimization problem that we define below; we then show that the solution to this optimization problem is the same as the solution to the optimization problem in the text .

To do so, we first define $E^k_{\mathbf{e},\mathbf{a},Y_n}$ to be the conditional expectation when k is alive. Write[1]

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

Parallel to the definition of M in Section W2, define

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

[1]Since the right-hand side of this equality does not depend on $\mathbf{e}$, we write the expectation conditional on $\mathbf{e} = 0$ to facilitate later steps.

where again the right-hand side does not depend on $\mathbf{e}$. Following Fudenberg and Levine (2006), we now define the SR objective function in the game:

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

Notice that with this objective function the shorter-run self cares about resources, but except when the utility $\tilde{U}(Y_k, \mathbf{a}) + \tilde{M}(Y_k, \mathbf{a})$ from the chosen action is smaller than the utility $\tilde{U}(Y_k, \mathbf{e}) + \tilde{M}(Y_k, \mathbf{e})$ from the suggested action, the shorter-run self views those resources as being outside of his control.

The long-run self is completely benevolent: her payoff in the game is the discounted sum of shorter-run self utilities

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

We assume that $\tilde{U}(Y_k, \mathbf{a}), \tilde{M}(Y_k, \mathbf{a})$ are continuous in $\mathbf{a}$ and define the cost of self-control to be

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

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

As in Fudenberg and Levine (2006), we now consider equilibria in which the shorter-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 shorter-run self: with one-period lifetimes for the shorter-run players, this Stackelberg equilibrium is equivalent to subgame-perfect equilibrium in which the long-run player moves first against each shorter-run player and is equivalent 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 shorter-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 shorter-run self k is stronger when the shorter-run self lives multiple periods. First, the self-control action changes the preferences of the shorter-run self over many periods. Second, the self-control action cannot be "changed" as long as the particular shorter-run self is alive. As we note below, if the long-run self were unable to commit for the life of the shorter-run self, then there would be a nontrivial strategic interaction between the two.

As is the case in which the shorter-run self lives only for a single period, the expectations of the shorter-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) defined a SR-perfect Nash equilibrium profile to be *equivalent* to a solution to the reduced form optimization problem of maximizing

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

if the reduced strategy induced from the shorter-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. Provided that $\tilde{U}(Y_k,\mathbf{a})$ and $\tilde{M}(Y_k,\mathbf{a})$ are continuous in $\mathbf{a}$, the conditions of Fudenberg and Levine's (2006) Theorem 1 are satisfied, so this equivalence does indeed hold.[2]

THEOREM W3: *If $\tilde{U}(Y_k,\mathbf{a})$ and $\tilde{M}(Y_k,\mathbf{a})$ are continuous in a, then SR-perfect Nash equilibria are equivalent to solutions to the reduced form optimization problem.*

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

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

[2]These conditions are costly and unlimited self-control, limited indifference, and continuity.

to those of

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

the agent's objective function that we used as the starting point in this paper.

THEOREM W4: $(*)$ *and* $(**)$ *have the same solutions*.

PROOF: Observe that since we have assumed state independent resource valuation, by Theorem W2, $\tilde{U}(Y_k,\mathbf{a})$ and $\tilde{M}(Y_k,\mathbf{a})$ have the same arg max, while $\tilde{u}(Y_k,0,\mathbf{a})=\tilde{U}(Y_k,\mathbf{a})+\max_{\mathbf{a}'}\tilde{M}(Y_k,\mathbf{a}')$ trivially has the same arg max. It follows that if $\mathbf{a}\in\arg\max_{\mathbf{a}'}\tilde{u}(Y_k,0,\mathbf{a})$, then $\mathbf{a}\in\arg\max_{\mathbf{a}'}\tilde{M}(Y_k,\mathbf{a})$ and, by definition, $C(Y_k,\mathbf{a})=0$. The former implies $\max_{\mathbf{a}'}\tilde{M}(Y_k,\mathbf{a}')-\tilde{M}(Y_k,\mathbf{a})=0$, so $C(Y_k,\mathbf{a})=\max_{\mathbf{a}'}\tilde{M}(Y_k,\mathbf{a}')-\tilde{M}(Y_k,\mathbf{a})$. Since this also holds by definition for $\mathbf{a}\notin\arg\max_{\mathbf{a}'}\tilde{u}(Y_k,0,\mathbf{a})$, it holds for all $\mathbf{a}$. Hence

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

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

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

Q.E.D.

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 shorter-run self. This is intended to capture the strategic naivete of the shorter-run self as a passive actor. 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 shorter-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 the nature of the shorter-run self.

W4. RECURSIVE VERSUS OPPORTUNITY COST

The text supposes that the control cost depends on the foregone value, which is defined with respect to the maximum possible SR utility from tomorrow onward given tomorrow's state. This specification 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

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

where, as in Section W3, $\tilde{U}(Y_k, \mathbf{a}) \equiv (1-\delta)E^k_{0,\mathbf{a},Y_k}\sum_{n=0}^{\infty}(\delta\mu)^{n-k}u(a^S_{k+n}, y_{k+n})$. With this specification, the control cost is computed each period by the difference between the best expected present value available to a shorter-run self born in that period and the present value actually received, taking into account what will actually happen in future periods.

With this alternative specification of the cost, the alternative objective function for the long-run self in period 1 is

$$\begin{aligned}\tilde{V}(h_1, \mathbf{a}) \equiv E_{\mathbf{a},h_1}\sum_{n=0}^{\infty}\delta^n\bigl(&u(y_{1+n}, a_{1+n})\\ &-(1-\mu)\Gamma[\bar{U}(y_{1+n}) - U^{1+n}(h_{1+n}, \mathbf{a})]\bigr)\\ &-\mu\Gamma[\bar{U}(y_1) - U^1(h_1, \mathbf{a})].\end{aligned}$$

Here the 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. It has the form of an *opportunity cost* that depends only on the best present value available to the short-run selves and the actual utility they receive. Note that the name of the short-run self born in period $1+n$ does not matter, so we may compute the self-control cost without loss of generality for $k = 1 + n$. This expected present value cost is weighted by $1 - \mu$, which is the probability of a new short-run self being born in a given period. In period 1, however, the weight must be taken to be 1, since the optimization problem always begins with the birth of a new short-run self.

In contrast, the formulation in the paper computes the foregone value and thus the cost in each period "as if" no self-control will be used in future periods. However, in the linear case, these two formulations are equivalent.

THEOREM W5: *If* $\tilde{C}(Y_k, \mathbf{a}) = \Gamma(\bar{U}(Y_k) - \tilde{U}(Y_k, \mathbf{a}))$, *then* $\tilde{C}(Y_k, \mathbf{a}) = C(Y_k, \mathbf{a})$.

PROOF: Since the identity of the SR self does not matter, is suffices to prove this for $k = 1$. So we must show

$$E_{\mathbf{a},h_1} \sum_{n=0}^{\infty} \delta^n \Delta(y_{1+n}, a_{1+n})$$

$$= E_{\mathbf{a},h_1} \sum_{n=0}^{\infty} \delta^n \big((1-\mu)[\bar{U}(y_{1+n}) - U^{1+n}(h_{1+n}, \mathbf{a})]\big)$$

$$+ \mu[\bar{U}(y_1) - U^1(h_1, \mathbf{a})].$$

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

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

$$= E^n_{\mathbf{a},h_n} \left(\sum_{\ell=0}^{\infty} (\delta\mu)^\ell (\Delta(y_{n+\ell}, a_{n+\ell})) \right).$$

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

$$E_{\mathbf{a},h_1} \sum_{\ell=0}^{\infty} \delta^\ell \big((1-\mu)[\bar{U}(y_{1+\ell}) - U^{1+\ell}(h_{1+\ell}, \mathbf{a})]\big) + \mu[\bar{U}(y_1) - U^1(h_1, \mathbf{a})]$$

$$= E_{\mathbf{a},h_1} \sum_{\ell=0}^{\infty} \delta^\ell \left((1-\mu) \sum_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} (\Delta(y_{1+\ell+\ell'}, a_{1+\ell+\ell'})) \right)$$

$$+ \mu \left[\sum_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} (\Delta(y_{1+\ell'}, a_{1+\ell'})) \right].$$

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

$$
\begin{aligned}
&E_{\mathbf{a},h_1} \sum_{\ell''=0}^{\infty} \sum_{\ell' \le \ell:}^{\infty} \big((1-\mu)\big(\delta^{\ell''} \mu^{\ell'} \varDelta(y_{1+\ell''}, a_{1+\ell''})\big)\big) \\
&\qquad + \mu \left[\sum_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} (\varDelta(y_{1+\ell'}, a_{1+\ell'})) \right] \\
&\quad = E_{\mathbf{a},h_1} \sum_{\ell''=0}^{\infty} (1-\mu) \delta^{\ell''} \varDelta(y_{1+\ell''}, a_{1+\ell''}) \sum_{\ell' \le \ell''}^{\infty} \mu^{\ell'} \\
&\qquad + \mu \left[\sum_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} (\varDelta(y_{1+\ell'}, a_{1+\ell'})) \right] \\
&\quad = E_{\mathbf{a},h_1} \sum_{\ell''=0}^{\infty} \delta^{\ell''} \varDelta(y_{1+\ell''}, a_{1+\ell''}) \big(1 - \mu^{\ell''+1}\big) \\
&\qquad + \mu \left[\sum_{\ell'=0}^{\infty} (\delta\mu)^{\ell'} (\varDelta(y_{1+\ell'}, a_{1+\ell'})) \right] \\
&\quad = E_{\mathbf{a},h_1} \sum_{\ell''=0}^{\infty} \delta^{\ell''} \varDelta(y_{1+\ell''}, a_{1+\ell''}),
\end{aligned}
$$

which is the desired result.

The key idea is that the principle of optimality for the shorter-run self enables us to write the overall loss to the shorter-run self as a sum of recursively computed losses,

$$
\begin{aligned}
&\bar{U}(y_n) - U^n(h_n, \mathbf{a}) \\
&\quad = \bar{U}(y_n) - (1-\delta\mu) E^n_{\mathbf{a},h_n} \sum_{\ell=0}^{\infty} (\delta\mu)^{\ell} u(y_{n+\ell}, a_{n+\ell}) \\
&\quad = E^n_{\mathbf{a},h_n} (1-\delta\mu) \left(\sum_{\ell=0}^{\infty} (\delta\mu)^{s} (\varDelta(y_{n+\ell}, a_{n+\ell})) \right).
\end{aligned}
$$

Hence the opportunity cost is just a weighted sum of the increments $\varDelta(y_{n+\ell}, a_{n+\ell})$, and the proof simply consists of bookkeeping to verify that the weights are the same as in the recursive case. *Q.E.D.*

In the linear case, in other words, it does not matter whether the cost of imposing self-control on the shorter-run self arises from recursive considerations or from an opportunity cost. The reason we adopted the recursive formulation in the text is that in the nonlinear case, the model of opportunity cost leads to implausible predictions about timing, such as changes in behavior when a shorter-run self "dies."

The recursive formulation of the text defines a cost of self-control for each state and action to be a convex function g of the difference between the most utility obtainable for any shorter-run self starting in that state and the utility that would be obtained if the shorter-run self next period gets the most possible utility,

$$C(y,a) = g\left(\bar{U}(y) - \left(u(y,a) + \delta \int_Y \bar{U}(y')\pi(y'|y,a)[dy']\right)\right).$$

The long-run self then solves the recursive problem directly without reference to which shorter-run self is born at a particular time, resulting in reduced form utility

$$\tilde{V}^n_{\mathbf{a},h_n} \equiv E^n_{\mathbf{a},h_n} \sum_{\ell=0}^{\infty} \delta^\ell [u(y_{n+\ell}, a_{n+\ell}) - C(y_{n+\ell}, a_{n+\ell})].$$

In the opportunity cost formulation, the cost of self-control for shorter-run self n born in period n is a convex function g of the difference between the most utility obtainable for the shorter-run self and the utility actually obtained,

$$C^n_{\mathbf{a},h_n} = g(\bar{U}(y_n) - U^n_{\mathbf{a},h_n}).$$

Note that there will typically be many shorter-run selves and that the long-run player pays the control cost for the nth shorter-run when that self is "born."

EXAMPLE W2: We now explore the difference between the two formulations through a simple example. In the example, the recursive formulation captures the simple intuition that when the probability of a temptation is reduced, it becomes less tempting. In the opportunity cost formulation, reducing the probability of a temptation has a complicated effect that depends on the exact timing of when the temptation occurs relative to the "birth" of a new short-run self. We argue that this is both unintuitive and inconsistent with experimental results.

In the example it is known that at some point in the future, a simple temptation will arrive, and it will have value P for the long-run self and value S for the shorter-run self. The exact nature of this opportunity is initially unknown; it is equally likely to be highly tempting (H) or less so (L). The agent learns which opportunity he will face at a time $t_1 + 1$, where t_1 follows an exponential waiting time with parameter p_1, which is the conditional probability of arrival each period. Once t_1 arrives, the agent is informed which state prevails: there is a second exponential waiting time with parameter p_2 until the time t_2 when the agent can choose whether to take the tempting action or decline it. Notice that t_1 and t_2 take on the value zero with positive probability, meaning, for example, that if $t_1 = t_2 = 0$, the agent learns immediately that he faces an opportunity and he takes an action during the same period. Note also that when p_1 and p_2 are large compared to the birth rate $1 - \mu$ of the short-run players, the initial short-run self will be the one who makes the eventual decision, while in the reverse case, a short-run self will probably be born after t_1 and soon before t_2.

In the highly tempting state H, the simple temptation if chosen is received for sure. In the less tempting state L, there is only a probability q if the temptation is chosen that it will be received.

EXAMPLE W2a—Recursive Cost of Self-Control: In the recursive case, the only relevant cost of self-control is in the period in which action is taken. The maximization problem can be written conditional on whether state H or L has occurred, so the condition for taking is exactly as in the text: in H the optimum is to take if

$$P < g(S),$$

while in L the condition is

$$qP < g(qS).$$

Note moreover that this solution holds regardless of the values of p_1 and p_2.

EXAMPLE W2b—Opportunity Cost of Self-Control: Here the long-run self faces one decision problem before t_1, a different decision problem once the information is revealed at t_1 but before action at t_2, and yet a third at the decision time t_2. Each of these problems is stationary and independent of the past history. Each one corresponds to a shorter-run self who faces a different temptation. To compute that temptation, we simply compute the probability distribution over arrivals of the event that the action becomes available, conditional on which type the shorter-run self is.

To calculate the temptation utility for any shorter-run self born at or after t_1 and strictly before t_2 in state H, we compute the expectation of the discounted value of S over different values of t_2:

$$\begin{aligned} U^{2H} &= E(\mu\delta)^{t_2} S = \sum_{t_2=1}^{\infty} (\mu\delta)^{t_2} S p_2 (1 - p_2)^{t_2 - 1} \\ &= \frac{\mu\delta p_2 S}{(1 - \mu\delta(1 - p_2))}. \end{aligned}$$

In state L, the temptation utility is just $U^{2L} = qU^{2H}$. For the short-run self born at exactly t_2, the temptation utilities are $U^{2H} = S$ and $U^{2L} = qS$. In either case, the temptation costs are given by $g(U^{2H})$ and $g(qU^{2H})$. If p_1 is large relative to p_2, then it is far more likely that the shorter-run self is born after t_1 than before, so most of the weight in the objective function of the long-run self is on these temptation costs. This leads to an analysis that is qualitatively similar to the recursive case. In the laboratory, however, the time between when the subjects learn they will face a particular decision problem and the time they are asked to take an action is usually on the order of minutes, while we expect the mean length of life of a shorter-run self[3] to be on the order of a day, so that it much more likely that the shorter-run self is born before t_1 rather than after.

The temptation utility for a self born before t_1 is the expected value of taking the temptation, averaged over the arrival times t_1 and t_2. Since these times are independent, the temptation value is

$$\begin{aligned} U^1 &= E\big[(\mu\delta)^{(t_1+t_2)} 0.5(1+q)S\big] \\ &= \frac{0.5\mu\delta p_1 p_2 (1+q)S}{(1-\delta\mu(1-p_1))(1-\delta\mu(1-p_2))}. \end{aligned}$$

Let us focus on the case where most of the weight is on shorter-run selves born before t_1. Let a^J be an indicator function of whether the long-run self takes at $J \in \{H, L\}$. Let $Q = 0.5(a^H + qa^L)$ be the overall probability of the long-run self taking. Then the objective of the long-run self is

$$\begin{aligned} V(Q) = {} & \frac{\delta p_1 p_2}{(1-\delta(1-p_1))(1-\delta(1-p_2))} QP \\ & - g\left(\frac{\mu\delta p_1 p_2 S}{(1-\delta\mu(1-p_1))(1-\delta\mu(1-p_2))}(0.5(1+q)-Q)\right). \end{aligned}$$

In particular, the objective function depends only on Q. Notice that there are four relevant values of Q: $0, 0.5q, 0.5, 0.5(1+q)$. Note moreover that since g is convex,

$$\begin{aligned} V_0' \equiv \frac{V(0.5q) - V(0)}{0.5Q} &\geq V_{0.5q}' \equiv \frac{V(0.5) - V(0.5q)}{0.5 - 0.5q} \\ &\geq V_{0.5}' \equiv \frac{V(0.5(1+q)) - V(0.5)}{0.5(1+q) - 0.5}. \end{aligned}$$

It follows that the optimal choice of Q is 0 if $V_0' \leq 0$; it is $0.5q$ if $V_0' \geq 0$, $V_{0.5q}' \leq 0$, it is q if $V_{0.5q}' \geq 0$, $V_{0.5}' \leq 0$, and it is $0.5(1+q)$ if $V_{0.5}' \geq 0$, and since

[3]See Fudenberg and Levine (2010).

increasing P lowers the slope of V, the optimal Q is a decreasing function of P. If P is close to zero, then Q should be equal to 1, meaning that the long-run self takes regardless of the state. If P is highly negative, then Q should be equal to zero, meaning that the long-run self never takes. As P is reduced from zero, eventually the probability Q must be reduced from 1 to 0.5, meaning that the optimum is to take only in the state H. This is the same behavior exhibited in the case in which most shorter-run selves are born after t_2 or that of a recursive shorter-run self. As we reduce P further, the probability Q must drop again from 0.5 to $0.5q$, meaning that the optimum is to take only in the state L. This is the opposite direction from that of a shorter-run self born after t_2 or a recursive shorter-run self. It also is contrary to indicating that reversals have the form that as the probability of a prize is reduced, behavior favored by the long-run self is more likely to be observed.[4]

Dept. of Economics, Harvard University, 310 Littauer Center, Cambridge, MA 02138, U.S.A.; dfudenberg@harvard.edu

and

Dept. of Economics, Washington University in St. Louis, St. Louis, MO 63130-4899, U.S.A.; david@dklevine.com.

Manuscript received November, 2010; final revision received May, 2011.

[4]See Fudenberg and Levine (2010).