TOWARDS A THEORY OF SELF-RESTRAINT[1]

Bernard Caillaud[2] Daniel Cohen[3] Bruno Jullien[4]

This version September 2000 [5]

---

[1] This work started while the three authors were at CEPREMAP. We thank Philippe Michel for allowing us to draw on some of his previous work with Daniel Cohen. We are indebted to Olivier Compte for enlightening discussions and for suggesting and proving a key lemma. We also benefited from stimulating discussions with Geir Asheim.

[2] CERAS-ENPC (CNRS URA 2036) and CEPR.

[3] ENS Ulm, CEPREMAP and CEPR.

[4] GREMAQ and IDEI, University of Toulouse, and CEPR.

[5] Correspondence to: D.Cohen, Dept de Sciences Sociales, E.N.S., 48 boulevard Jourdan, 75014 Paris, FRANCE. Email: dcohen@elias.ens.fr.

Abstract

The decision problem of an agent with time-inconsistent preferences has usually been formalized as a game between the different selves of the agent, considered as different players. We propose an alternative route. We assume that the agent is able to coordinate her decisions with herself. In order to model such behavior, we borrow some of the tools of cooperative theory and assume that agents restrict their decisions space to a sub-set of "internally consistent strategies". In a simple model with a state variable, we show that there exists one way for the agent to comply with our solution concept.

JEL: D9

Keywords: Dynamic inconsistency, self-restraint.

1 Introduction

In June 1965, Bill W. goes to Akron, Ohio, for a business trip. The trip is disappointing, and Bill feels crushed. He has only recently stopped drinking and the stress created by the trip makes him feel an oppressing need to drink. In order to resist the temptation, Bill has an idea: he should talk to another alcoholic. He gets the name of a Dr. Bob who lives a daily nightmare because of his drinking habit. They spend the night talking. At the end of the night, Bill has resisted his envy and a few days later Bob stopped drinking for good.

These two alcoholics became the founders of Alcoholic Anonymous, an institution that counts more than 2 million participants all over the world. What is the key to success of such institutions? *Not that it provides "new information"* to its members about either the health consequences of alcoholism or the cost of stopping. It is not even important that you should talk to someone who already stopped drinking. *Not that it provides its members with new commitment devices.* You are free to come, there is no penalty or reward system implemented by the community. It is not even important that you should meet the same people to whom you might be "tied" by friendly relationships or among which your reputation might matter. The recipe is simple. People talk to each other and the key to success is to *acknowledge* that they are sick, that they are alcoholic. The truth to one's identity is the key to success. Quotations abound that depict such feature in detail (extracted from Alcoholic Anonymous' press book generously supplied to us by the Paris branch):

> "(AA) allowed me to accept myself such as I am"
>
> "I recognized myself as an alcoholic similar to the others, my personality was progressively found again."
>
> "So long as one remains in alcohol, one loses one's identity, one is not oneself anymore. Alcohol is a psychotrop, it changes one's behavior,

> not for the better. The fact of being able to say it, in front of people who know what it is, is an extraordinary relief. One can sustain, and find again, one's identity."

In this paper, we propose a new concept of *self-restrained behavior* to analyze how, by being conscious of her own identity, an agent can resolve the fundamental tension between the willingness to stop drinking and the frustration of not being able to do so. We define a self-restrained behavior as the ability of an agent to recognize the recurrent aspect of her choices and to impose on these choices a minimal consistency restriction. In a simple abstract model, we provide a formalization of our concept, we prove general properties (among which existence and payoff-uniqueness), and we relate our solution to other game-theoretical solution concepts.

What is the counterpart, in economic theory, of the tension described above for the alcoholic ? On the one hand, theories of rational addiction (such as in Becker-Murphy (1988) and Boyer (1983)) acknowledge the dependency in which an individual finds herself caught, but they do not capture the frustration that arises when she *tries* to stop. The agent simply does what she views as optimal from her current perspective. On the other hand, theories in which dynamic inconsistency arises in the agent's preferences precisely characterize situations where the course of actions that the agent would like to implement today is not the same as the one she knows she would like to implement tomorrow. These theories then provide an explanation of the frustration that is created when one cannot find the appropriate means to commit oneself (Strotz (1956)). The fundamental phenomenon that they exhibit may simply rely on non-exponential discounting in the standard separable utility model, or more generally on time-varying preferences or recursive utility.

But theories with time-inconsistent preferences have so far postulated a non-cooperative multi-self structure out of which only a non-cooperative equilibrium emerges (see Peleg and Yaari (1973) and the more recent work by Laibson (1992-

1997)): the literature looks for a subgame perfect equilibrium to the multi-agent game form (each self is different from the next to come) so that each self takes as given the strategy of future selves and reacts optimally. In such a framework, an agent may feel some frustration over what she does, but there is nothing she can do about it: other selves strategically constrain her behavior. In contrast, we want to sketch a *theory of the self* which is what seems to form the cornerstone of AA success.

Take indeed the example of the drinker. The Nash equilibrium among the different selves may lead the agent to drink every day indefinitely while *the* agent would actually be better off (in all states of nature) by never drinking. What could then prevent *the* agent from choosing the dominant option of soberness ? Only this: if the agent has convinced herself that she will really stop drinking in the future, then not drinking *today* is not the best choice: she would be better off drinking today and stop drinking tomorrow. In brief, the agent would prefer a *third* option. Yet, such "third options" are self-destructive. Whatever strategy the agent might have in mind, there will indeed always be a "third option" in the future that will do better. In game-theoretical terms, this can be stated as saying that there is no subgame perfect equilibrium of the one-player game (played against nature). Confronted with such incapacity to reach a decision, what should the agent do, how can she restore a "sense of unity" to her decisions? Our answer is that selves (usually) focus on plans which are free from inner contradictions, and reach a "sense of unity" (see below) by expecting that this will be the case in the future as well. In our paper, "internally consistent" strategies are then simply strategies which specify a plan whose implementation can never be dominated by deciding to start that very same strategy all over again rather than pursuing it. We show in a simple model of addiction how restricting the agent's set of decisions to such a sub-set of "internally consistent" decisions delivers an equilibrium path to the agent.

If one compares our approach to the multi-self approach, the closest solution concept is that of internal stability of the cooperative game played by the different selves. When the core is empty, agents select strategies within an "internally stable" sub-set, which is one where strategies do not contradict each other (see more on this in the text). Asheim (1997) uses a stability concept based on cooperative theory to select among subgame perfect equilibria of the multi-self game. Another concept which bears some similarity to ours (within multi-self games) is that of renegotiation proofness, which provides a selection device within the set of subgame perfect strategies (see Kocherlakota (1996)). We depart from these works in that we do not impose subgame perfection in the multi-self game.[1]

Before proceeding to the analysis, we wish to relate our work to other more descriptive contributions in various disciplines. In economic literature first, Loewenstein and Thaler (1989) presents some robust behavioral anomalies that cast doubts on the simple discounted utility model (see also Wathieu (1995)). Psychologists are well aware of the problem. In their survey of "self-control" or "impulse control" theories, Ainslie and Haslam (1992) characterize four methods of precommitment (as they themselves call it) that pyschologists have analyzed: i) extrapsychic mechanisms such as those used by Ulysses against the Sirens;[2] ii) control of attention which corresponds to the repression of impulses analyzed by Freud; iii) preparation of emotion which is the Buddhist way of achieving indifference to impulse; iv) and personal rules which are at the core of our paper. Personal rules are characterized by Ainslie and Haslam as follows:

> "an effective personal rule is one that specifies a common interest in such a way that the person never prefers to abandon it".

---

[1] It is however easy to implement our solution as the outcome of a subgame perfect equilibrium of the multi-self game (see Caillaud and Jullien (forthcoming)).

[2] Extrapsychic mechanisms correspond to the idea of irreversible commitment through an outside party, as appears in Strotz (1956), but also in sociology as in Becker (1960).

These personal rules have first been analyzed by Victorian psychologists and Ainslie and Haslam account for these rules by referring to a 19th century author Sully (1884) who wrote:[3]

> "A child consciously or unconsciously begins to refer to a general precept or maxim of action as "maintain health", "seek knowledge", "be good" and so forth. Particular actions are thus united under a common rule, they are viewed as members of a class of actions subserving one comprehensive end. In this way the child will attain a **measure of unity**" (our emphasis).

This "measure of unity" which lies at the root of personal rules is exactly what AA teach their members. Our work fits in this large body of inter-disciplinary contributions and offers a possible formalization to these ideas. It is thus formulated in a model that reflects common concerns among various behavioral sciences, and that is necessarily abstract. Applications to alternative economic situations are left for future work (see Caillaud-Cohen (forthcoming)).

2 The basic model

We consider a Robinson Crusoe economy with only one consumption good. The unique consumer is infinitely lived with an additively time-separable intertemporal utility; periods run from $t = 0, 1, ...$ to infinity. The per-period, or instantaneous, utility depends upon the consumption c_t of the unique good available as well as upon a state variable x_t. The state variable evolves in time according to the consumption path of the consumer. This general setting can be viewed as describing

---

[3] See also Ainslie (1975) who presents such rules as some kind of "private side bets", i.e. as "a self-enforcing contract that the person makes with his own future motivational states" (page 189), where the penalty for a person failing to control an impulsive reaction, i.e. for breaking once the contract, is the loss of faith in her possibility of controlling impulse later on.

the problem of a consumer when the consumption good is addictive: the state variable stands for the level of addiction or the stock of addictive capital, or more generally a measure of the consumer's health (such as analyzed in Becker-Murphy (1988) or Becker-Grossman-Murphy (1991)). But our formulation is general enough to apply to more general problems of consumption/savings decision. In the sequel we shall use the language of the theories of addiction.

In order to simplify the analysis, we focus on the case where the consumer's levels of addiction can be summarized by two different states $X = \{0, 1\}$ in which $x = 0$ corresponds to *not* being addicted, and $x = 1$ corresponds to being addicted.[4] Depending upon $c \in \mathsf{R}_+$, the consumer may shift from one state of addiction to another according to the following Markov transition process. Let

$$p_0(c) = \Pr[x_{t+1} = 1 \mid x_t = 0, c]$$

denote the probability that the consumer becomes addicted when consuming c while she is not (yet) addicted, and

$$p_1(c) = \Pr[x_{t+1} = 1 \mid x_t = 1, c]$$

the probability that she remains addicted when consuming c while already being addicted. We suppose that these probabilities are continuous non-decreasing in consumption and, for each i, $\operatorname{Im} p_i(.) = [0, 1]$. So, for each level of addiction, there exist consumption levels that generate any given probability of becoming / staying addicted.

In the sequel we shall simply take (with no loss of generality) these probabilities themselves as decision variables. In order to do this, let $v(x, c)$ denote the consumer's instantaneous utility in state x when she consumes c. The function v is decreasing in x and increasing concave in c. Following Becker-Murphy (1988), we

---

[4] See our 1996 discussion paper for more general frameworks.

consider that $v_{12} > 0$ characterizes addictive behavior: the more addicted one is, the higher the marginal utility of consumption.

Let us now call

$$u(0, a) = \max_{c \geq 0} v(0, c) \text{ subject to } a = p_0(c)$$

and

$$u(1, a) = \max_{c \geq 0} v(1, c) \text{ subject to } a = p_1(c).$$

$u(x, a)$ represents the utility of an individual in state x who is willing to take the risk a of becoming or staying addicted. u is clearly decreasing in x and increasing in a, and we assume that $u(0, 0) > u(1, 1)$ so that staying addicted forever yields a smaller intertemporal utility than staying sober forever. We also make the natural assumption that, for any state $x \in X$, $u(x, .)$ is strictly concave, although the reader should bear in mind that this implies in fact a comparison between the concavity of $v(x, .)$ in consumption and the shape of the probabilities $p_x(.)$. Finally, we will assume that:

$$u_2(1, 1) \geq u_2(0, 0). \tag{1}$$

This assumption greatly simplifies the exposition of the model as it rules out any cycle in all solution concepts we discussed, as will be clear later on.[5]. A natural extension on $u(x, .)$ of the definition of addiction would be: $u_2(1, a) \geq u_1(0, a)$ for all a. (1) is slightly more restrictive since it implies that the marginal utility of (probability-normalized) consumption is *uniformly* higher when addicted than when non-addicted.

Consider now an agent who expects to move from one state x_t to a trajectory $\{x_{t+1}, x_{t+2}, ...\}$ from some time t on by implementing a sequence of actions $\{a_s\}_{s \geq t}$. We take the intertemporal utility of the consumer to be measured at any point in time by the following intertemporal index :

---

[5] There is no difficulty defining the different solution concepts we investigate if this assumption is not satisfied; the analysis however involves more algebra.

$$J[x_t, \{a_s\}_{s \geq t}] = u(x_{t,} a_t) + \beta \mathsf{E}_t \left[\sum_{s=1} \delta^s u(x_{t+s}, a_{t+s}) \right] \tag{2}$$

We take $\beta \leq 1$. The case $\beta = 1$ is the standard case of a "dynamically consistent" consumer who exponentially discounts the future. The case $\beta < 1$ is the case of a "myopic" consumer (in the words of Strotz (1956)) who gives an additional weight to the present, "a salience" effect in the words of Akerlof (1991).

In the sequel we shall denote $\hat{J}[x_t, \{a_s\}_{s \geq t}]$ the value function which is obtained when $\beta = 1$, so that:

$$J[x_t, \{a_s\}_{s \geq t}] = (1 - \beta)\ u(x_t, a_t) + \beta \hat{J}[x_t, \{a_s\}_{s \geq t}] \tag{3}$$

where the first term is the source of dynamic inconsistency and only bears on current consumption. The salience effect of consumption a_τ at date τ disappears when viewed from earlier on, i.e. when viewed at date $t < \tau$.

3 Game-theoretical approaches

3.1 Strategies of the self

Within the framework just described, consider the individual decision problem of choosing a sequence of actions as a dynamic game. Let h denote an history: it consists of an initial state x_0 at date 0, subsequent states at each date t between 0 and the actual date denoted $t(h)$ and actions taken at each date before $t(h)$; we let $x(h)$ denote the current state. Finally, let (h, a) denote an history consisting of h and action a chosen at date $t(h)$.

A general strategy, denoted θ, specifies an action $a(h)$ for any feasible history h. This includes time-varying strategies, where the action taken, given a particular value of the state variable, depends upon the actual time at which the decision must be taken. This also includes bootstrap strategies where actions can be based upon the complete sequence of past actions and not only the state variable, and can therefore serve as deterring punishment, itself rationalized by fear of future punishment, etc...in the usual logic of folk theorems. Let $\theta \mid_h$ denote the continuation strategy following an history h. Finally, for any history h, replacing the sequences of actions $\{a_s\}_{s \geq t(h)}$ in criteria $J[.,.]$ and $\hat{J}[.,.]$ defined in (2) by the continuation strategy $\theta \mid_h$, let $J(x(h), \theta \mid_h)$ (respectively $\hat{J}(x(h), \theta \mid_h)$) denote the consumer's intertemporal utility (respectively utility with $\beta = 1$), evaluated at date $t(h)$, resulting from the indefinite implementation of strategy θ from date $t(h)$ on.

We shall also denote $\mathcal{H}(x, \theta)$ the set of histories that occur with positive probability when θ is applied with initial state x.

3.2 Multi-selves equilibria

The traditional approach to the problem of the choice of an individual with time-inconsistent preferences has been to consider each temporal incarnation of the individual as a distinct player (a self) in a game with an infinite horizon. The game is then solved by using standard non-cooperative game-theoretical concepts.

Among various types of multi-self, subgame-perfect equilibrium, a central solution corresponds to the concept of Markov perfect equilibrium (hereafter MPE)[6] of the game played by the different selves of the agent.[7] This is simply a subgame perfect equilibrium of the multi-self game, where strategies only depend upon station-

---

[6] See Maskin - Tirole [1998] for a general definition and characterization of MPE.

[7] As emphasized by Laibson (1993), the set of equilibria may be large, including bootstrap strategies. Choosing a MPE is one possible selection that rules out threats involving strong punishment out of the equilibrium. Such threats, while credible in a game theoric context, may be seen as unreasonable for the multi-selves approach. This highlights the difficulties that arise in using non-cooperative multi-agent game theory in the context of a single-agent decision problem, even with multiple selves. An alternative approach would be to rely on ideas related to renegotiation-proofness (see Kocherlakota (1996) who defines a reconsideration-proof concept).

ary payoff-relevant history.[8] For a given history h, the stationary payoff-relevant history reduces to the current state of addiction $x(h)$, as it is the only element of history that determines the flow payoffs for the coming periods of consumption. It follows that a MPE is characterized by a Markov strategy θ^B that determines an action to be played contingent on the current state variable, i.e. here a pair of actions (a_0^B, a_1^B), that solves the following Bellman equation:

$$\forall x \in \{0,1\}, J_x^B = \max_{a \in A_x} \left\{ u(x,a) + \beta\delta a \hat{J}_1^B + \beta\delta(1-a)\hat{J}_0^B \right\} \tag{4}$$

with: $\beta \hat{J}_x^B = J_x^B - (1-\beta)u(x, a_x^B)$ for all x.

In words, the consumer takes her future behavior as given and reacts optimally given her current state of addiction. A MPE is a strategy that is purely state-contingent and that is a best-response to the conjecture that it will always be implemented later on, in each state.

To simplify the analysis and concentrate on the main issues, we focus on the specific situation where the MPE behavior *eventually* induces addiction and the consumer thereafter stays addicted forever.

Proposition 1 : *Suppose that the salience effect is such that:*

$$\beta^B = \frac{(1-\delta)u_2(0,0)}{\delta\left[u(0,0) - u(1,1)\right]} > \beta. \tag{5}$$

then, there exist Markov perfect equilibria and all MPE strategies are characterized by $a_0^B > 0$ *and* $a_1^B = 1$ *with*

$$u_2(0, a_0^B) \geq \beta \frac{\delta\left[u(0, a_0^B) - u(1,1)\right]}{1 - \delta(1 - a_0^B)}$$

with an equality if $a_0^B < 1$.

---

[8] Two histories are stationary equivalent if they generate strategically equivalent continuation game forms. A stationary Markov strategy is a strategy that is measurable with respect to the partition of the set of histories generated by the stationary equivalence classes of the above-defined equivalence relation.

As β decreases, the consumer values more and more current consumption compared to its consequences in terms of probability of becoming / staying addicted. Moreover the marginal utility of consumption is larger once addicted than when still non-addicted. So, as β decreases, it first becomes more and more costly to refrain from consuming the addictive good once addicted and the addictive state becomes stable; the consumer does not try to reduce addiction and she is trapped in an addictive state. For moderate β, it can still be the case that staying non-addicted is also a stable state. But for small enough β, the consumer eventually becomes addicted and stays in such a state.[9]

A MPE that eventually leads to addiction yields the following intertemporal utility index to the consumer:

$$J_0^B = \left(1+\beta\frac{\delta(1-a_0^B)}{1-\delta(1-a_0^B)}\right)u(0,a_0^B)+\beta\frac{\delta a_0^B}{(1-\delta(1-a_0^B))}\frac{u(1,1)}{(1-\delta)}, \tag{6}$$

$$J_1^B = \left(1+\beta\frac{\delta}{1-\delta}\right)u(1,1). \tag{7}$$

Compared to such a MPE strategy, the Markov state-preserving strategy ($a_0^O = 0, a_1^O = 1$) delivers the following intertemporal utility index : $J_0^O \equiv \left(1+\beta\frac{\delta}{1-\delta}\right)u(0,0)$ and $J_1^O = J_1^B$. Note first that, if $\beta = 1$, that is if preferences are time-consistent, it is easy to differentiate the function $J_0^B(a_0)$ defined by (6) and to find that under (5) it is increasing in a_0 for $a_0 \in [0, a_0^B]$. In other words, for $\beta = 1$, $J_0^O \leq J_0^B$. If, however, $\beta < 1$ there is no such general result and the inequality may be reversed so that, by committing to the state-preserving strategy, the consumer could be

---

[9] As illustrated by the proof, we could relax the assumption that $u_2(0,0) \leq u_2(1,1)$ and replace (5) by:

$$\frac{\delta\,[u(0,0)-u(1,1)]}{(1-\delta)\inf\{u_2(0,0),u_2(1,1)\}} < 1.$$

The proof would be modified only to the extent that there could now be cycles, instead of state-preserving strategies, that would be ruled out with the same method.

better off than under a MPE strategy. When the salience effect is not too large, it is proved in the appendix that the welfare is higher under the state-preserving strategy than under the MPE strategy, as stated below.

Corollary 1 *: If the salience effect is such that:*

$$1 > \beta^B > \beta > \frac{1-\delta}{1-\delta\beta^B}\beta^B, \tag{8}$$

then $J_0^O > J_0^B$*.*

This condition is perfectly compatible with condition (5). Therefore, when (8) holds, the previous proposition characterizes MPE that yield strictly lower intertemporal utility in state $x = 0$ than the state-preserving strategy: committing to remain non-addicted would yield a higher utility level for the consumer than following the MPE behavior that leads ultimately to addiction. This raises some concern about the relevance of the concept of MPE in our setting.

Furthermore, for a range of values of β, it may even be the case that a Markov strategy that recommends to reduce consumption so as to become eventually non-addicted and then stay sober forever ($a_0 = 0, a_1 < 1$) yields larger intertemporal utility (state-by-state) than the MPE strategy. This makes our concern for the MPE solution more acute: MPE recommends to get addicted while becoming sober in some specific way would unambiguously constitute a preferable alternative for the consumer. We omit the precise characterization of when this happens as it is lengthy and useless at this point.

3.3 Parting with the multi-self approach

Why is it that the consumer does not decide, whenever her addiction is $x = 0$, that $a_0 = 0$ is her best choice in program (4) ? That is, why is it that the state-preserving strategy is not a MPE. The consumer maximizes her intertemporal utility taking as given the continuation strategy $(a_0^B, 1)$ and associated continuation values $\hat{J}_x$; a

MPE is attained when the choice indeed coincides with the Markov strategy that is conjectured for the future. If the consumer conjectures that she will implement the state-preserving strategy later on, she expects continuation values $\hat{J}_0^O$ and $\hat{J}_1^O$ associated to $a_0 = 0$, and then she has to decide upon her current consumption. The salience effect being large enough by (5), she prefers to consume a positive amount of the good today, as it brings her much satisfaction compared to the small risk of becoming addicted next period:

$$u_2(0,0) - \beta\delta \left[\hat{J}_0^O - \hat{J}_1^O\right] = u_2(0,0) - \beta\frac{\delta}{1-\delta}\left[u(0,0) - u(1,1)\right] > 0$$

and so the choice is $a_0' > 0 = a_0$.

The critical point behind this reasoning is that the consumer does not realize that her present decision to postpone abstinence until tomorrow (technically, to choose $a_0' > 0$ today faced with $a_0 = 0$ and $a_1 = 1$ from tomorrow on) will be replicated at every date in the future as she will be the very same person facing the very same problem and so making the very same choice. Put in another way, if the consumer seriously plans to implement a state-preserving continuation strategy ($a_0 = 0, a_1 = 1$), she anticipates that, from tomorrow on, her utility evaluated at one of these future dates will be J_0^O if she does not become addicted from today's consumption. And, as a positive consumption maximizes her currently evaluated intertemporal utility, she can expect $J_0' > J_0^O$ as of today. So, the theory supposes that the consumer plans to obtain a smaller intertemporal utility tomorrow than today although both decision problems are identical and *she is the very same person who has to solve them both.*

As explained in the introduction, we wish to emphasize the fact that the same individual is choosing the action at all dates, as opposed to contributions based on the premise that there is a different player at each date. Consequently, we do not consider equilibrium concepts such as MPE, that is equilibrium concepts in the multi-agent game form or equivalently in the game with an infinite number of players, one for each period. We rather wish to explore alternative solution concepts that reintroduce some unity or coordination between the different selves of the consumer.

One natural possibility is then to consider the problem as a *one-player, multi-stage game*.[10] A strategy θ of the single player game is a subgame perfect equilibrium if it is a Nash equilibrium of any subgame (a subgame is here the dynamic choice problem faced by the individual after some history h). Formally, a strategy θ is a subgame perfect equilibrium if, for any history h and any strategy θ',

$$J(x(h), \theta \mid_h) \geq J(x(h), \theta').$$

When preferences are time-consistent ($\beta = 1$), this reduces to the optimality of the strategy θ evaluated at date 0, because this choice remains optimal at any future date. But when preferences are time-inconsistent ($\beta < 1$), one has to take account of the fact that in the subgame starting at history h, the individual uses criterion $J[.,.]$ to evaluate the strategy $\theta \mid_h$, which may not coincide with the ex-ante evaluation (obtained with $\hat{J}[.,.]$ if $t(h) > 1$).[11] It is then a well-known result that the quest for a one-agent subgame-perfect equilibrium with dynamic inconsistency is vain:

Proposition 2 *: When $\beta < 1$, there exists no subgame perfect equilibrium to the one-agent game.*

The existence of a subgame-perfect equilibrium is in fact formally equivalent to the existence of an optimal time-consistent history-contingent plan. Therefore, going all the way towards a complete unification of the different selves of the consumer

---

[10] Piccione-Rubinstein [1994] also addresses issues related to the comparison between multi-selves and one-self game forms, but in the context of imperfect recall.

[11] Technically, with time-consistent preferences, the one-stage deviation principle (see e.g. Fudenberg-Tirole [1991]) applies and an equilibrium in the multi-agent game form is a perfect equilibrium of the one-player game. With time-inconsistent preferences, this is not true anymore.

does not constitute either a satisfactory theory of the decision under dynamically inconsistent preferences. We content that, faced with this situation, the individual will adopt a decision procedure that she can expect to build upon not only today but in the future as well. This is the line that we now explore.

4 Self-restrained strategies

What we intend to capture is that *any reasoning that is used at the start of the decision process must incorporate the very mere possibility to use the same reasoning later on.* This leads first to a notion of internal consistency.

Definition 1 : *A strategy* θ *is* **internally consistent (IC)** *if and only if:*

for any history $h \in \cup_{x \in X} \mathcal{H}(x, \theta)$, $J(x(h), \theta\, |_h) \geq J(x(h), \theta)$.

This definition means that the consumer should not regard as viable an option that she will want to renege upon in the future by invoking this very same option. If the consumer thought that an internally *inconsistent* strategy would be implemented, she would have to assume that her "future selves" could not follow the same decision process as herself, that she would be denied the very same option as she is now choosing. This would contradict the basic premise of our approach which is, based on the previous discussion, that the consumer understands that she will face the same problem repeatedly and could implement the same action each time. We then contend that an internally inconsistent plan of consumption cannot be expected to be implemented indefinitely unless the consumer fools herself on her future ability to evaluate later on the options she has access to.[12] When evaluating her intertemporal utility, the consumer can restrict attention to consumption

---

[12] In a similar vein, Kocherlakota (1996) imposes symmetry, but focusing on sub-game perfect equilibria of the multi-self game.

plans that can be the actual outcome of her repeated decisions, that is to internally consistent plans of consumption.

Definition 2 *: For any* θ_1, θ_2 *in* Θ, θ_1 *is dominated by* θ_2 ($\theta_1 \prec \theta_2$) *if there exists some history* $h_0 \in \cup_X \mathcal{H}(x, \theta_1)$ *such that,*

$$J(x(h_0), \theta_2) > J(x(h_0), \theta_1 \mid_{h_0})$$

and that for all history $h \in \mathcal{H}(x(h_0), \theta_2)$,

$$J(x(h), \theta_2 \mid_h) \geq J(x(h), \theta_1)$$

This definition means that the consumer should be willing to switch from one policy to another one, if the latter option takes her into a path that the former option can never contradict in the future. This captures the idea that the initial choice of the individual act as a "standard of behavior" that can be called upon at any future date. It embeds some asymmetry between the actual plan made at the start of the decision and subsequent potential deviations, reflecting the prevalence of the current self. It is conceivable that the individual decides to shift to another strategy (as in a multi-self context), but when doing so the date t self will be bound by the possibility of any future self to apply the initial reasoning from the start, disregarding the past choice of deviation made at date t.

Definition 3 *: A strategy is a self-restrained strategy if and only if it is an extremal point in the set of internally consistent policies with respect to the dominance criterion* $\prec$.

Our approach bears some resemblance to cooperative concepts of stability (see VonNeumann-Morgenstern (1947) and Greenberg (1990)). Our definition of internal consistency is inspired by the notion of internal stability, applied to the set generated through re-initialization by a given strategy. However, we do not consider a criterion of external stability with respect to all strategies, including non internally consistent ones, as it would lead to no equilibria in our setting. We rather look for an extremal strategy within the set of internally consistent policies.

5 Characterization

In this section we present the main results concerning self-restraint strategies in our model. These results are formally derived in the appendix. Let us first state our central result:

Theorem 1 *: There exists a self-restrained strategy and, for each self-restrained strategy, there exists a payoff-equivalent, self-restrained Markov strategy.*

The theorem is proved through a sequence of lemmas that are interesting in their own rights. The first lemma is quite general and links self-restrained strategies to Markov strategies.

Lemma 1 *: Any internally consistent strategy that is dominated by another internally consistent strategy is also dominated by a Markov strategy.*

This critical Lemma shows that it is sufficient to restrict the set of deviations that challenge any given potential equilibrium to the set of Markov strategies. Along with the theorem itself, it shows that for all practical matters, it is sufficient to look for and to test strategies within or through Markov strategies.[13] As the proof makes it clear, this is far from a trivial outcome of our assumptions.

The characterization of the Markov self-restrained strategy then follows several steps. The main point is that under condition (1) a Markov self-restrained strategy

---

[13] As it appears from the proof, this result is quite general. This does not mean that a self-restrained strategy has to be a Markov strategy. Any strategy that yields the same value function is also a self-restrained strategy. Among those, some may seem more appealing than Markov strategies, depending on the context.

must converge surely towards *one* level of addiction and to stay there once it has been reached. We call this state the *stable state of addiction.* Moreover, given one stable state of addiction, a self-restrained Markov strategy must necessarily propose the best way to reach this stable state (and stay there), starting from the other state of addiction. Consequently, only two types of Markov strategies can possibly be self-restrained, θ_0^* and θ_1^*, which we now characterize.

Let θ_0^* denote the following Markov strategy associated with stable state $x = 0$:

- The action at $x = 0$ is $\theta_0^*(0) = 0$.
- At the state $x = 1$, the action maximizes the payoff among all the Markov strategies that preserve state 0 with probability 1 $(\theta(0) = 0)$,that is: $\theta_0^*(1) \in \arg\max_a J_0^*(a)$ with

$$J_0^*(a) = (1 + \beta\frac{\delta a}{1 - \delta a})u(1, a) + \beta\frac{\delta(1 - a)}{(1 - \delta a)}\frac{u(0, 0)}{(1 - \delta)}.$$

Similarly, let θ_1^* denote the "best" Markov strategy that leads to state $x = 1$ and preserves it:

- $\theta_1^*(1) = 1$
- $\theta_1^*(0) \in \arg\max_a J_1^*(a)$ with

$$J_1^*(a) = (1 + \beta\frac{\delta(1 - a)}{(1 - \delta + \delta a)})u(0, a) + \beta\frac{\delta a}{(1 - \delta + \delta a)}.\frac{u(1, 1)}{(1 - \delta)}.$$

At least one of these two Markov strategies is a self-restrained strategy and no other Markov strategy can be self-restrained. The theorem above provides an existence result; uniqueness is not guaranteed and this should not come as a surprise since even for MPE strategies, uniqueness is not guaranteed. It is straightforward to see that if J_0^* and J_1^* are such that $1 \notin \arg\max_a J_0^*(a)$ and $0 \notin \arg\max_a J_1^*(a)$, then both strategies are self-restrained.[14]

---

[14] Given Corollary 2 in the Appendix, it is sufficient to test one against the other according to the dominance criterion of Definition 2.

As can be expected, uniqueness of the self-restrained Markov strategy can be established with quasi-concavity assumptions on J_0^* and J_1^*. Given the case under scrutiny in Proposition 1, the following result is particularly relevant.

Proposition 3 *: Suppose $J_0^*(.)$ and $J_1^*(.)$ are strictly quasi-concave, and that the salience effect is such that $1 > \beta > \frac{1-\delta}{1-\delta\beta^B}\beta^B$, then θ_0^* is the unique Markov self-restrained strategy.*

Typically the self-restrained strategy allows the individual to escape from the logical trap that prevented her to credibly maintain state 0 in a MPE. Starting from state 0 and considering the possibility of a one-shot deviation, the individual must acknowledge that any deviation in the current period will be replicated in the future. One can see the current action as providing clues for future behavior. Contemplating the prospect of constant deviations from the strategy that maintains addiction to its stable state, the individual then prefers to stick to $a_0 = 0$.

>From the previous analyses of MPE and of self-restrained strategies, it first appears that for $\beta^B > \beta > \frac{1-\delta}{1-\delta\beta^B}\beta^B$, state 0 is a stable-state of addiction under self-restraint while the MPE behavior eventually leads to addiction ($x = 1$). As witnessed by the AA quotation in the introduction, realizing one's own unity helps the consumer escape the inconsistency trap leading to addiction.

6 Summary and conclusion

We have characterized what may be called a "state-consistent" decision rule. In standard Markov approaches, an agent takes as given the decision that she will find optimal to implement in the *future*, and correspondingly solves the Bellman equation. We have characterized instead the decision rule which optimizes the agent's welfare when she takes as given what she will do in other states of nature, but yet internalizes the fact that whatever is done today, will also be done tomorrow if the agent finds herself in the same state of nature. In brief, we have attempted to describe which equilibrium could be reached when the agent internalizes the unity of her self.

References

[1] **Akerlof, G.** (1991) : ”Procrastination and Obedience”, Richard T. Ely Lecture, *American Economic Review*, Vol. 81, n○ 2,1-19.

[2] **Ainslie, G.** (1975): ”Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control”, *Psychological Bulletin*, Vol 82(4), 463-496.

[3] **Ainslie, G.** and **N.Haslam** (1992): ”Self-Control”, Chapter 8 in *Choice over Time*, Loewenstein G. and J.Elster (Eds), Russell Sage Foundation, New-York.

[4] **Asheim, G.** (1997): ”Individual and Collective Time-Consistency”, *Review of Economics Studies,* 64, 427-443.

[5] **Becker,G., M. Grossman** and **K.Murphy** (1991) : ”Rational Addiction and the Effects of Price on Consumption”, *American Economic Review*, Papers and Proceedings, Vol.81, n○2, 237-241.

[6] **Becker, G.** and **K.Murphy** (1988) : ”A Theory of Rational Addiction”, *Journal of Political Economy*, Vol. 96, 675-700.

[7] **Becker, H**, (1960): ”A note on the concept of commitment”, *American Journal of Sociology*, Vol 66, 32-40.

[8] **Boyer M** (1983): ”Rational Demand and Expenditures Patterns Under Habit Formation”, *Journal of Economic Theory, XXXVI*, 27-53.

[9] **Caillaud, B., Cohen, D.** and **B. Jullien** (1996): ”Towards a Theory of Self-Restraint”, CEPREMAP Working Paper.

[10] **Caillaud, B.** and **D.Cohen** (forthcoming): ”Intergenerational transfers and common values in a society”, *European Economic Review, Papers and Proceedings.*

[11] **Caillaud B.** and **B. Jullien, (**forthcoming): "Modelling Time Inconsitent Preferences", *European Economic Review, Papers and Proceedings.*

[12] **Cohen, D.** and **P. Michel** (1990) : "Which Rules Rather than Discretion in a Democracy? I -An Axiomatic Approach with Open-Loop Commitments", CEPREMAP WP 9008.

[13] **Fudenberg, D.** and **J. Tirole** (1991) : *Game Theory*, MIT Press, Cambridge.

[14] **Greenberg, J.** (1990), *The Theory of Social Situations: An Alternative Game-Theoretic Approach*, Cambridge University Press, Cambridge.

[15] **Kocherlakota, N.R.** (1996): "Reconsideration-Proofness: A Refinement for Infinite Horizon Time Inconsistency, *Games and Economic Behavior,* 15, 33-54.

[16] **Laibson, D, A..** (1992): "Self-Control and Saving", Mimeo MIT.

[17] **Laibson, D.** (1997): "Golden Eggs and Hyperbolic Discounting", *Quarterly Journal of Economics*, Vol 112(2), 443-77.

[18] **Loewenstein, G.** and **R.H.Thaler** (1989): "Anomalies: Intertemporal Choice", *Journal of Economic Perspectives,* Vol 3, 181-193.

[19] **Maskin, E.** and **J.Tirole** (1998): "Markov Perfect Equilibrium I: Observable Actions", mimeo IDEI Toulouse.

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

[21] **Piccione, M.** and **A. Rubinstein** (1994) : "On the Interpretation of Decision Problems with Imperfect Recall", Working Paper 24-94, Tel Aviv Univ.

[22] **Sully, J.** (1884), *Outlines of Psychology,* Appleton, New-York.

[23] **Strotz, R.** (1956) : "Myopia and Inconsistency in Dynamic Utility Maximization", *Review of Economic Studies*, 23, 165-180.

[24] **Von Neumann, J.** and **O. Morgenstern** (1947) : *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.

[25] **Wathieu, L.** (1995): "Habits and the Anomalies in Intertemporal Choice", mimeo Hong-Kong University.

A Proof of Proposition 1 and Corollary

Consider the mapping that associates $\mathcal{F}Z = (J_0', J_1', a_0', a_1')$ to $Z = (J_0, J_1, a_0, a_1)$, as follows:

$$J_x' = \max_{\alpha \in [0,1]} \left\{ u(x,\alpha) + \delta\alpha \left[J_1 - (1-\beta)u(1,a_1)\right] + \delta(1-\alpha)\left[J_0 - (1-\beta)u(0,a_0)\right] \right\}$$

with a_x' the argument maximum in this program. $u(x,.)$ being defined, continuous and concave on a compact set, $\mathcal{F}$ maps a compact set of R^4 into itself, and it is a continuous function: hence, it has a fixed point, which coincides with a MPE.[15]

The concavity of $u(x,.)$ allows to characterize the necessary and sufficient conditions for an interior solution (a_0, a_1), in particular with $0 < a_0$ and $a_1 < 1$:

$$u_2(0,a_0) = u_2(1,a_1) = \beta\delta \left[\hat{J}_0^B - \hat{J}_1^B \right]$$

which is impossible since $u_2(1,a_1) > u_2(1,1) \geq u_2(0,0) > u_2(0,a_0)$. Therefore, either $a_0 = 0$ or $a_1 = 1$.[16]

For a MPE with $a_0 = 0$ and $0 < a_1 < 1$, the following should hold:

$$\hat{J}_0 - \hat{J}_1 = \frac{u(0,0) - u(1,a_1)}{1 - \delta a_1}$$

$$u_2(0,0) \leq u_2(1,1) \leq u_2(1,a_1) = \beta \frac{\delta\left[u(0,0) - u(1,a_1)\right]}{1 - \delta a_1}.$$

The implicit equation determining a_1^B can be written as:

$$(1 - \delta a_1)u_2(1,a_1) + \beta\delta\left[u(1,a_1) - u(0,0)\right] = 0.$$

The LHS of this equation is strictly decreasing in a_1. But condition (5) trivially implies that the LHS of this equation evaluated at $a_1 = 1$ is strictly positive. So, there cannot exist such a MPE.

---

[15] Note, however, that the need to keep track of the argmax actions forbids the conclusion that $\mathcal{F}$ is a contraction.

[16] Corner solutions at $a_1 = 1$ and $a_0 > 0$, or $a_0 = 0$ and $a_1 < 1$, or $a_0 = 1$ and $a_1 = 0$ are trivially also impossible.

MPE with $a_0 = 0$ and $a_1 = 1$ are ruled out similarly, using the inequalities that determine a_0 and a_1.

The conditions for an MPE with $a_1 = 1$ and $0 < a_0 < 1$ are:

$$\hat{J}_0 - \hat{J}_1 = \frac{u(0,a_0) - u(1,1)}{1-\delta(1-a_0)}$$

$$u_2(1,1) \geq u_2(0,0) \geq u_2(0,a_0) = \beta \frac{\delta\left[u(0,a_0) - u(1,1)\right]}{1-\delta(1-a_0)}. \tag{9}$$

The implicit equation determining a_0^B can be written as:

$$(1-\delta(1-a_0))u_2(0,a_0) + \beta\delta\left[u(1,1) - u(0,a_0)\right] = 0.$$

Condition (5) implies that the LHS is positive for $a_0 = 0$. Therefore, either there exists an interior solution to this equation or $a_0 = 1$. Note that the LHS is decreasing in a_0 for $\beta = 1$ which implies that for time-consistent preferences there exists a unique solution (possibly a corner solution). For $\beta < 1$, however, the LHS may be non-monotonic, hence the possibility of several solutions.

This completes the proof of Proposition 1.

The state preserving strategy yields higher intertemporal utility starting in state $x = 0$ than the MPE if:

$$\max_{a_0} \left((1+\beta\frac{\delta(1-a_0)}{1-\delta+\delta a_0})u(0,a_0) + \beta\frac{\delta a_0}{1-\delta+\delta a_0}\frac{u(1,1)}{1-\delta} - (1+\beta\frac{\delta}{1-\delta})u(0,0) \right) \leq 0$$

From the envelope theorem and $u_2 > 0$, the product of $\frac{1}{\beta}$ by the LHS is nonincreasing with β. Therefore, if the condition holds for some β, it also holds for all $\beta' > \beta$. Using the concavity assumption on $u(x,.)$, a sufficient condition for the condition to hold is that, for all a_0,

$$\frac{\delta}{1-\delta}u(0,0) > \frac{\delta(1-a_0)}{1-\delta+\delta a_0}(u(0,0) + a_0 u_2(0,0)) + \frac{\delta a_0}{1-\delta+\delta a_0}\frac{u(1,1)}{1-\delta} + \frac{1}{\beta}a_0 u_2(0,0)$$

or, rearranging the terms,

$$\frac{\delta}{1-\delta+\delta a_0}(\frac{u(0,0) - u(1,1) + a_0 u_2(0,0)}{u_2(0,0)}) > \frac{1}{\beta}(1-\delta) + \delta$$

The LHS reaches a minimum at $a_0 = 0$ so that we obtain

$$\frac{\delta}{1-\delta}\left(\frac{u(0,0)-u(1,1)}{u_2(0,0)}\right) = \frac{1}{\beta^B} > \left(\frac{1-\delta}{\beta}\right) + \delta,$$

as a sufficient condition for $J_0^O > J_0^B$. ■

B Characterization: proof the theorem

Proof of lemma 1:

Let θ_2 dominates θ_1 at $x_0 = x(h_0)$. For $x \in \{x(h), h \in \mathcal{H}(x_0, \theta_2)\}$, define:

$$\hat{J}_2(x) = \max_{h \in \mathcal{H}(x_0,\theta_2), x(h)=x} \{\hat{J}(x, \theta_2 \mid_h)\}.$$

For a given x, let $h_\varepsilon(x)$ denote an history in $\mathcal{H}(x_0, \theta_2)$ such that $x = x(h_\varepsilon(x))$ and $\hat{J}(x, \theta_2 \mid_{h_\varepsilon(x)}) \geq \hat{J}_2(x) - \frac{\varepsilon}{\beta}$. Let also $a_\varepsilon(x)$ denote the action specified by θ_2 at $h_\varepsilon(x)$.

Notice that the choice of θ_2 for $x \notin \{x(h), h \in \mathcal{H}(x_0, \theta_2)\}$ is arbitrary, so set θ_2, and $a_\varepsilon(x)$ correspondingly, so that the state is maintained at such an x forever.

Then

$$J(x, \theta_2 \mid_{h_\varepsilon(x)}) \leq u(x, a_\varepsilon(x)) + \delta \mathsf{E}\left[\beta \hat{J}_2(y) \mid x, a_\varepsilon(x)\right]$$

where y denotes the subsequent state.

Notice that

$$\beta \hat{J}_2(y) \leq J(y, \theta_2 \mid_{h_\varepsilon(y)}) - (1-\beta) u(y, a_\varepsilon(y)) + \varepsilon$$

Let T_ε denote the probability transition matrix induced by the action a_ε. Denote J_ε the vector of values, u_ε the vector of instantaneous utility. Denote Id the identity matrix and e the unit vector. Then

$$J_\varepsilon \leq u_\varepsilon + \delta T_\varepsilon \left(J_\varepsilon - (1-\beta) u_\varepsilon + \varepsilon.e\right)$$

or

$$J_\varepsilon \leq (Id - \delta T_\varepsilon)^{-1} \left((Id - (1-\beta)\delta T_\varepsilon) u_\varepsilon + \delta \varepsilon T_\varepsilon.e\right)$$

Let a_ε converges to a (choose a converging sequence), with transition matrix T and utility u. Denote θ the corresponding Markov strategy ($\theta(x) = a(x)$) and J_θ the value function of θ. Taking the limit we have by continuity:

$$\lim_{\varepsilon \to 0} J_\varepsilon \leq (Id - \delta T)^{-1}(Id - (1-\beta)\delta T)u$$

and

$$J_\theta = (Id - \delta T)^{-1}(Id - (1-\beta)\delta T)u$$

Notice that $\mathcal{H}(x_0, \theta) \subset \mathcal{H}(x_0, \theta_2)$.

Thus for $x \in \mathcal{H}(x_0, \theta)$, $J(x, \theta) \geq \lim J(x, \theta_2 \mid_{h_\varepsilon(y)}) \geq J(x, \theta_1)$. Moreover, because θ_2 is internally consistent, $J(x_0, \theta) \geq J(x_0, \theta_2) > J(x_0, \theta_1 \mid_{h_0})$.

It follows that θ dominates θ_1. ■

As an immediate consequence:

Corollary 2 *: A Markov strategy is a self-restrained strategy if and only if it is not dominated by another Markov strategy.*

Proof. Immediate. ■

Lemma 2 *: If θ_1 is internally consistent, then there exists a Markov strategy θ such that for all x, $J(x, \theta) \geq J(x, \theta_1)$.*

Proof. The proof is similar to the proof of Lemma 1. For x define

$$\hat{J}_1(x) = \max_{h \in \cup_{y \in X} \mathcal{H}(y, \theta_2), x(h) = x} \{\hat{J}(x, \theta_1 \mid_h)\}.$$

Define a_ε for θ_1 as in the previous proof and θ as a limit Markov strategy.

Then

$$\lim_{\varepsilon \to 0} J_\varepsilon \leq (Id - \delta T)^{-1}(Id - (1-\beta)\delta T)u = J_\theta$$

Because θ_1 is internally consistent, for all x :

$$J(x, \theta_1) \leq \lim_{\varepsilon \to 0} J(\varepsilon, \theta_1 \mid_{h_\varepsilon(x)}) \leq J(x, \theta).$$

■

Corollary 3 *: If θ_1 is a self-restrained strategy, there exists a payoff-equivalent Markov self-restrained strategy θ.*

Proof. Choose θ as in the previous lemma. It must be payoff equivalent because otherwise it would dominate θ_1 according to $\prec$. Generating the same payoff as θ_1, it is then self-restrained. ■

The two corollaries imply proposition 2 apart from existence.

Lemma 3 *: A Markov self-restrained strategy cannot be cycling, i.e. either $a_0 = 0$ or $a_1 = 1$.*

Proof. Suppose that $a_0 > 0$ and $a_1 < 1$. Then

$$\begin{aligned}
J(0,\theta) &= u(0,a_0) + \delta a_0 \beta \hat{J}(1,\theta) + \delta(1-a_0)\beta\hat{J}(0,\theta) \\
J(1,\theta) &= u(1,a_1) + \delta a_1 \beta \hat{J}(1,\theta) + \delta(1-a_1)\beta\hat{J}(0,\theta) \\
\beta\hat{J}(0,\theta) &= J(0,\theta) - (1-\beta)u(0,a_0) \\
\beta\hat{J}(1,\theta) &= J(1,\theta) - (1-\beta)u(1,a_1)
\end{aligned}$$

Now

$$(1-a_1)J(0,\theta) + a_0 J(1,\theta) = (1+\beta\frac{\delta}{1-\delta})((1-a_1)u(0,a_0) + a_0 u(1,a_1))$$

By concavity of u :

$$\begin{aligned}
u(0,a_0) &< u(0,0) + a_0 u_2(0,0) \\
u(1,a_1) &< u(1,1) - (1-a_1)u_2(1,1)
\end{aligned}$$

Therefore

$$\begin{aligned}
(1-a_1)u(0,a_0) + a_0 u(1,a_1) &< (1-a_1)u(0,0) + a_0 u(1,1) \\
&\quad + a_0(1-a_1)(u_2(0,0) - u_2(1,1)).
\end{aligned}$$

Using condition (1),

$$(1-a_1)u(0,a_0)+a_0u(1,a_1)<(1-a_1)u(0,0)+a_0u(1,1)$$

Staying at state x with certainty yields payoffs $J_0^O=(1+\beta\frac{\delta}{1-\delta})u(0,0)$ and $J_1^O=(1+\beta\frac{\delta}{1-\delta})u(1,1)$. Therefore

$$(1-a_1)J(0,\theta)+a_0J(1,\theta)<(1-a_1)J_0^O+a_0J_1^O$$

Thus either $J(0,\theta)<J_0^O$ or $J(1,\theta)<J_1^O$, which means that θ is dominated according to $\prec$ either by choosing $a_1=1$ at 1 or by choosing $a_0=0$ at 0. ■

According to this lemma, the level of addiction must converge with probability one. So, depending on the stable state, a self-restrained Markov strategy can only be of the following form. Either it converges to state $x=0$, and then it is characterized by $a_0=0$ and a_1, with the highest possible valuation while at state $x=1$ among the Markov strategies that preserve $x=0$. So, necessarily,

$$\begin{aligned}
a_1 &\in \arg\max_{\alpha} J(1,\theta)\\
\text{s.t. } \theta &= (0,\alpha)\\
J(1,\theta) &= u(1,\alpha)+\delta\beta\alpha\hat{J}(1,\theta)+\delta\beta(1-\alpha)\frac{u(0,0)}{1-\delta}\\
\hat{J}(1,\theta)\beta &= J(1,\theta)-(1-\beta)u(1,\alpha).
\end{aligned}$$

The charactrization of θ_0^* given in the text follows. Or it converges to $x=1$, and one obtains easily a similar expression for θ_1^*.

Lemma 4 *: The Markov strategy θ_0^* is self-restrained if and only if $J_0^*(1)>J_1^O$ or if θ_0^* and θ_1^* coincide with the state-preserving strategy θ^O.*

Proof. It is sufficient to show that θ_0^* is then undominated by another Markov strategy. First, by construction, it cannot be dominated by a Markov strategy that converges to $x=0$. Second, it is not dominated by the Markov strategy θ_1^*, since by assumption, either both strategies are equivalent or, when state $x=1$ is reached on the continuation path of θ_1^*, the valuation following θ_1^* falls short of the valuation of switching back to θ_0^*. Third, a fortiori, it cannot be dominated by any Markov strategy that converges to state $x = 1$. Finally, θ^* yields a payoff that is larger or equal to J_0^O and J_1^O. The proof of lemma (3) then implies that it cannot be dominated by a cycle. ■

A similar lemma can be derived for θ_1^*.

We conclude this appendix by proving the last proposition, namely that if $J_0^*(.)$ and $J_1^*(.)$ are strictly quasi-concave and $\beta > \frac{1-\delta}{1-\delta\beta^B}\beta^B$, then θ_0^* is the unique self-restrained Markov strategy.

Strict quasi-concavity makes it sufficient to look at the following derivatives:

$$\frac{dJ_0^*}{da}(1) = \left(1 + \beta\frac{\delta}{1-\delta}\right) u_2(1,1) - \beta\delta\frac{[u(0,0) - u(1,1)]}{(1-\delta)^2}$$

$$\frac{dJ_1^*}{da}(0) = \left(1 + \beta\frac{\delta}{1-\delta}\right) u_2(0,0) - \beta\delta\frac{[u(0,0) - u(1,1)]}{(1-\delta)^2}.$$

Our assumption on $u(.)$ imply that: $\frac{dJ_1^*}{da}(0) \leq \frac{dJ_0^*}{da}(1)$. Using the previous lemma (and its counterpart for θ_1^*), if $0 < \frac{dJ_1^*}{da}(0)$, then θ_1^* is the only self-restrained Markov strategy; if $\frac{dJ_0^*}{da}(1) < 0$, θ_0^* is the only self-restrained Markov strategy; and if $\frac{dJ_1^*}{da}(0) \leq 0 \leq \frac{dJ_0^*}{da}(1)$, the state-preserving strategy is the only self-restrained Markov strategy, and it coincides with θ_0^* and θ_1^*. Under the additional assumption that $\beta \geq \frac{1-\delta}{1-\delta\beta^B}\beta^B$, $\frac{dJ_1^*}{da}(0) \leq 0$. Hence Proposition 3.