June 1994

TOWARDS A THEORY OF SELF RESTRAINT (°)

N°9421

B. Caillaud*
D. Cohen**
B. Jullien***

* CERAS-ENPC and CEPREMAP, Paris.

** ENS Ulm and CEPREMAP, Paris.

*** ENSAE-CREST (Laboratoire d'économie industrielle), Montrouge and CEPREMAP, Paris.

(°) We thank Philippe Michel for allowing us to draw on some of his previous work with Daniel Cohen. We are also indebted to Olivier Compte for enlightening discussions and comments.

TOWARDS A THEORY OF SELF RESTRAINT

ABSTRACT

The decision problem of a inconsistent agent has usually been formalized as an intrapersonal non-cooperative game between the different selves of the agent, considered as different players. We attempt, in this paper, to offer an alternative route. We will assume the agent is one single self who is able to *self-restrain* his choices, by which we mean that, at every point in time, he obeys the following behavioral constraints: 1) he can willingly commit to a Markovian strategy, but: 2) he also aknowledges the fact that he cannot commit to disregard in the future any alternative commitment to another (Markovian) strategy that will appear to be preferable from this later date (and state) on, this later strategy being itself to be subject to the same constraint. In a simple model with a state variable, we are able to show that there exists one (and only one) way for the agent to comply with our solution concept, i.e. one and only one "self-restrained" way to behave.

JEL: D9
Keywords: Dynamic Inconsistency; self-restrain.

UNE THEORIE DES CHOIX AUTO-RESTREINTS

RESUME

Les choix d'un agent dont les préférences sont temporellement incohérentes ont été représentés, en général, comme un jeu entre différents agents qui incarnent chacun les préférences de l'agent à un moment donné. Nous empruntons, dans cet article, une autre démarche. Nous supposons que l'agent est bel et bien un seul individu capable d'auto-restreindre ses choix. Nous entendons par cela qu'il se soumet aux contraintes suivantes: 1) il s'engage à suivre *une* politique Markovienne; et 2) il soumet cette politique Markovienne à la contrainte de ne jamais être dominée par une autre qui plus tard serait meilleure, *et* qui serait elle-même soumise à ce même critère par la suite. Dans un modèle simple avec variable d'état, nous montrons alors qu'il existe bien une et une seule politique auto-restreinte.

Mots clés: incohérence temporelle; auto-contrôle.

I - INTRODUCTION

In the words of Sacha Guitry (a French humorist) : "It is easy to stop smoking, I've done it a dozen times". This famous schyzophrenia of a smoker who *wants* to quit but cannot make it, is a general feature of intertemporal decisions. Strotz [1955] and Akerlof [1991], for instance, discuss a particularly simple form of preferences that exhibits time inconsistency, namely when discounting is exponential for future periods while a salience effect overemphasizes the weight of current period deicions. More generally, any time an agent's discounting is not exponential, a *dynamic inconsistency* of his choices is bound to arise between the sequence of choices he decides today and the sequence of choices he would want to implement tomorrow. While the agent would like to commit to a sequence of decisions (stop smoking first thing tomorrow morning), he would also like to modify his commitment when times come, and start "afresh" a new program in the future (stop smoking the day after tomorrow).

The decision problem of a inconsistent agent has been formalized as an intrapersonal game between the different selves of the agent, considered as different players who play a non-cooperative game. The Nash equilibrium can be reached either as a Markovian (memoriless) equilibrium (as in the work by Peleg and Yaari [1973]) or as the solution of a repeated game in which the smoker at one period can punish an "earlier" smoker who deviates from a desired policy (see the more recent work by Laibson [1992] and [1993]).

We find the Game Theoretical apparatus unsatisfactory for dealing with these issues. On the one hand, the non-cooperative approach is unsatisfactory because, after all, the smoker is "one" person and it is hard to believe that he would either punish himself or fail to coordinate his actions with himself. On the other hand, a cooperative approach (which is yet to be done), although it would certainly be a welcome alternative, would fail to recognize the fact that the smoker is "free" at any of time to do what he wants, irrespectively of what he may have done or thought in the past, and thus to break any cooperative agreement he may have settled earlier.

We attempt, in this paper, to offer an alternative route. We will

assume the agent is able to *self-restrain* his choices, by which we mean that, at every point in time, he obeys the following behavioral constraints:

1) He can willingly commit to a Markovian strategy, i.e. a plan of actions contingent on the payoff-relevant history (here, a state variable); a complete decision is therefore a behavioral rule that can be applied later on.

2) Yet, he also aknowledges the fact that he cannot commit to disregard in the future any alternative commitment to another (Markovian) strategy that will appear to be preferable from this later date (and state) on. Of course, when contemplating the possibility of switching to an alternative preferred strategy, the agent recognizes that this new strategy has itself to be implementable in the future. A self-restrained decision rule will then correspond to a (Markovian) commitment strategy that has the following fixed-point property: it is the best (Markovian) decision the agent can expect to implement, when this decision is subject to the constraint of never being dominated later on by another decision rule which would itself be subject to similar constraints.

Under these hypotheses (and some others on the utility function) we are able to show that there exists one (and only one) way for the agent to comply with our solution concept, i.e. one and only one "self-restrained" way to smoke (and perhaps to stop smoking).

II - THE MODEL

1 - State variable and dynamics

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 $t = 0, 1, \ldots$ 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 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.

We assume that the state variable can take values in a finite grid of potential state values, $X \equiv \{0,\ldots,I\}$. Consumption of the good can however take a continuum of values in a compact interval $[\underline{c},\bar{c}]$. The law of motion of the state variable is given by a stochastic process with the following properties:

- for any state variable i, there exists a unique level of consumption γ_i that maintains the state variable unchanged with probability 1;
- when consumption exceeds (resp. is lower than) the state-preserving level γ_i, the state variable can either stay unchanged or jump to the neighboring state $i+1$ (resp. $i-1$);
- the probability of jumping from state i to the relevant neighboring state is given by $a_i(c)$, a function of consumption, from $[\underline{c},\bar{c}]$ into $[0,1]$, smooth strictly increasing on $]\gamma_i,\bar{c}[$, smooth strictly decreasing on $]\underline{c},\gamma_i[$, and equal 0 at γ_i.

Formally, if at time t the state is $x_t=i$, the probability distribution of x_{t+1} at time t+1 is given by:

if $c = \gamma_i$, Prob $[x_{t+1} = i] = 1$;

if $c > \gamma_i$, Prob $[x_{t+1} = i+1] = a_i(c) = 1$ - Prob $[x_{t+1} = i]$;

If $c < \gamma_i$, Prob $[x_{t+1} = i-1] = a_i(c) = 1$ - Prob $[x_{t+1} = i]$;

The consumer can maintain his state variable unchanged, with certainty, if he chooses the state preserving consumption γ_i. If, however, he increases his consumption beyond this level, he also increases the risk of switching to the immediately higher state (e.g. getting more addicted in the case of consumption of an addictive good). Symmetrically, if he reduces his consumption below this level, he increases the chances of lowering his state variable (e.g. getting partly disintoxicated).

Given the assumptions made on the jump probability, we will make a convenient change of variable. We will consider the consumer as directly choosing the probability of switching to a higher or lower state variable. The consumer in state i will then be viewed as choosing $a \in A_i \subseteq [-1,1]$, where A_i is compact and contains 0, where a negative value corresponds to the unique level of consumption $c<\gamma_i$ such that $a_i(c) = a$, a positive value to the unique level of consumption $c>\gamma_i$ such that $a_i(c) = a$, and $a = 0$ corresponds to $c = \gamma_i$.[1]

2 - Instantaneous utility

The per-period, or instantaneous, utility reached by a consumer in state x who decides to consume a quantity c of the good is denoted by $U(x,c)$. We assume that U is smooth, defined over $\mathbb{R} \times [\underline{c},\bar{c}]$, although the state variable only takes values on the finite grid X. With the change of variable suggested earlier, we will denote $u(x,a)$ the utility reached by a consumer in state x who decides to consume a quantity of the good that gives rise to a probability of switching a, above or below the state-preserving consumption depending upon the sign of a. In the following, we will often call a itself, the consumption decision.

We make an assumption of decreasing marginal utility of "normalized" consumption, i.e. consumption rescaled in terms of probability of jump of the state variable: for any state $x \in X$, $u(x,.)$ is increasing, concave in a on $\{a \in A_x, a<0\}$ and on $\{a \in A_x, a>0\}$. The assumption holds under decreasing marginal utility of (true) consumption provided that the probability function $a_i(.)$ is not too much curved.

3 - Intertemporal utility

Consider an agent who expects to move from one state x_t to a trajectory

---

[1] Note that we could easily allow for a whole range of consumption for which the probability of jump is null, i.e. γ_i could be an interval. In what follows, $a=0$ should then be understood as the utility-maximising consumption within the interval γ_i.

$(x_{t+1}, x_{t+2}, \ldots)$ from some time t onto the future by implementing a sequence of actions $(a_s)_{s \geq t}$. We take the utility of the consumer to be measured at any point in time by the following intertemporal index:

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

We take $\delta \geq 1$. $\delta=1$ is the standard case of a "dynamically consistent" consumer who exponentially discounts the future. $\delta > 1$ is the case of a "myopic" consumer (in the words of Strotz) 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 $\delta=1$. With this notation, one can write the link between J and $\hat{J}$ at some date t:

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

where the first term is the source of dynamic inconsistency and only bears on current consumption. From a dynamic perspective:

$$J[x_t, (a_s)_{s \geq t}] = \delta \, u(x_t, a_t) + \beta \, \mathbb{E}\left[\hat{J}[x_{t+1}, (a_s)_{s \geq t+1}] \right]$$

the salience effect of consumption a_{t+1} at date $t+1$ disappears when viewed from earlier on, i.e. when viewed at date t.

III - STRATEGIES AND SOLUTION CONCEPT

In this section, we start by a discussion of the time-consistency issue in game-theoretical terms. We then develop the behavioral assumptions that we impose on the consumer and we present the definition of our solution concept.

1 - Time-inconsistency and sub-game perfection

Consider the individual decision problem of choosing a sequence of actions as a dynamic game, with initial node at time $t=0$ and infinite

horizon. Define an history, denoted h, as the specification of a date $t(h)$, a state $x(h)$ at date t, and a sequence of states and actions from date 0 to $t-1$. A time-dependent 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.... We denote $\theta|_h$, the continuation strategy following an history h, in other words the restriction of the strategy θ to the set of histories that coincide with h between dates 0 and $t(h)$. Finally, let $J(x,\theta)$ denote the value of criterion (1) when the strategy θ is used and when the initial state is x.

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 in the multi-agent game form, or equivalently in the game with an infinite number of players, one for each period; we rather insist on equilibrium concepts with respect to the <u>one player, multi-stage game</u>.

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). When preferences are time-consistent ($\delta = 1$), 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 ($\delta > 1$), 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 $\theta|_h$, which may not coincide with the ex-ante evaluation (obtained with $\hat{J}[.]$ if $t(h) > 1$).[2]

---

[2] 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.

A strategy θ is then a subgame perfect equilibrium if:

For any history h and any strategy θ', $J(\theta|_h, x(h)) \geq J(\theta', x(h))$.

Using this definition, we show in appendix that *when $\delta > 1$, there exists no subgame perfect equilibrium.* Indeed the existence of a subgame perfect equilibrium is formally equivalent to the existence of an optimal time-consistent history contingent plan. We contend that faced with this situation, the individual will design an optimal strategy by coordinating the decision processes at successive dates. This coordination must however be based on logical grounds.

2 - Self-restrained strategies

Subgame perfection tests a strategy against deviations from which the agent could switch to *any* other action plan in the future. This is too demanding for the problem that we analyze. What we want to do here is then to restrict test deviations to be themselves subject to a constraint of "consistency". Suppose that the individual has decided to implement some strategy θ_1 considered as a "consistent" choice. The agent may find on its way along the induced state path that at some state x, a better alternative θ_2 occurs to him. However in order for θ_2 to be a "consistent" alternative, it must be indefinetely implemented. Now if θ_1 is itself a "consistent" choice, the agent is free to choose θ_1 at any time. Therefore θ_2 cannot be expected to be indefinitely implemented if along the induced orbit, it is optimal to revert to the choice of θ_1. The strategy θ_1 should thus only be tested against strategies with this property. To capture this notion, we shall first define a binary relation on the set of Markovian strategies. We shall return below on the reason why we restrict our analysis to such strategies.

Formally, call Θ the set of all Markovian strategies, with generic element θ. By definition, a strategy $\theta \in \Theta$ maps the set of state variables X into the set of available (probability normalized) consumption: $\theta(i) \in A_i$ indicates the consumption recommended by strategy θ when the consumer's state variable has value $x=i$. For any $\theta \in \Theta$, let us then define $O[\theta, x]$ to be the trajectory (the orbit) that is generated by θ, with positive

probability, when the initial state variable is x: formally,

if $\theta(x) > 0$, $O[\theta,x] \equiv \{y \in X,\ y \geq x \text{ and } \forall z \in X,\ x \leq z < y,\ \theta(z) > 0\}$
if $\theta(x) < 0$, $O[\theta,x] \equiv \{y \in X,\ y \leq x \text{ and } \forall z \in X,\ y < z \leq x,\ \theta(z) < 0\}$
if $\theta(x) = 0$, $O[\theta,x] \equiv \{x\}$

Call $J_\theta(.)$ the value of criterion (1) when the strategy θ is *indefinitely* implemented, where the argument stands for the starting value of the state variable.

This is no surprise that if $\delta > 1$, trying to apply subgame perfection to the set of Markovian strategies will not solve the basic problem underlying the inexistence of a subgame perfect equilibrium: choosing action a at state y may be seen as optimal when the initial state is y, but dominated when considered from an initial state is $x \neq y$ (See the appendix).

Definition 1: A strategy θ_1 is said to be dominated by a strategy θ_2 if for some x: $J_{\theta_1}(x) < J_{\theta_2}(x)$ and $\forall y \in O[\theta_2,x]$, $J_{\theta_1}(y) \leq J_{\theta_2}(y)$.

Notice that this binary relation may not be transitive or antisymetric. We then define a solution as a maximal point for this binary relation. We should point out that such a solution concept was first investigated in a somewhat different setting by Cohen-Michel [1990].

Definition 2: A "self-restrained" strategy is any Markovian decision rule θ^* which is not dominated by any other Markovian strategy.

Our definition suggests another related interpretation of a self-restrained strategy in terms of the best (state-contingent) *reservation* welfare level. A reservation welfare characterizes a level of (state-dependent) intertemporal utility, that a rational agent can expect from his sequence of choices: formally, it corresponds to a value function $J(.)$ $(= J_{\theta^*}(.))$ such that the consumer will disregard any strategy θ that leads (with positive probability), if implemented indefinitly, to a state z where the consumer would enjoy a strictly lower intertemporal utility, $J_\theta(z) < J(z)$. A reservation utility reflects maximin considerations and could

correspond to a simple rule of behavior that the consumer might want to follow to start his logical reasoning. For a given reservation welfare level $J(.)$, the consumer will then try to find the best strategy that does not violate, at some point in the future, the minimal welfare level that the consumer expects from the situation.

Call $\mathcal{T}$ the operator that associates to any option welfare function $J(.)$ the best option which is consistent with (i.e. not contradicted by) $J(.)$:

$$\forall x \in X, \quad \mathcal{T}J(x) \equiv \underset{\theta \in \Theta}{\text{Max}} \; J_\theta(x) \quad \text{s.t. } \forall y \in O[\theta,x], \; J_\theta(y) \geq J(y) \tag{2}$$

The agent will clearly solve the program (2) and choose a strategy that attains the maximum $\mathcal{T}J(.)$. The equivalence between this concept and definition 1 is immediate.

Proposition 1: A "self-restrained" strategy is any Markovian decision rule θ^* whose value function $J_{\theta^*}(.)$ is a fixed point of the mapping $\mathcal{T}$ defined in (2).

Proof: Suppose that θ^* is a "self-restrained" strategy. Consider a state x and a strategy θ such that for all $y \in O[\theta,x]$, $J_{\theta^*}(y) \leq J_\theta(y)$. Since θ cannot dominate θ^*, it must be the case that $J_{\theta^*}(x) \geq J_\theta(x)$. Therefore $J_{\theta^*}(.)$ is solution of program (2). Now suppose that $J_{\theta^*}(.)$ is a fixed point of $\mathcal{T}$. Any strategy θ such that for all $y \in O[\theta,x]$, $J_{\theta^*}(y) \leq J_\theta(y)$ must give a payoff $J_\theta(x) \leq J_{\theta^*}(x)$. Therefore θ^* is not dominated.

QED

Our solution concept thus allows the consumer to choose the best possible reservation welfare, i.e. a welfare level that cannot be improved at some initial state variable without later violating this reservation level in some other state.

3 - Beyond Markovian strategies

We may now come back on the reasons why we decided to restrict the analysis to Markovian strategies. If we were to consider a larger set of

strategies Θ, whose generic term is a strategy θ that maps any past history h into current action a, one would want to look for a fixed point of the following mapping:

$$\forall x \in X, \quad \mathcal{T}J(x) \equiv \underset{\theta \in \Theta \text{ s.t. } \forall\, h \in O[\theta,x],\ J_\theta(h) \geq J[x(h)]}{\text{Max}}\ J_\theta(x) \qquad (2')$$

Assume that the agent picks up a set Θ which has the property that any $\theta|_h \in \Theta$ for any $\theta \in \Theta$ and any history h. Then, if there exists a fixed point $J^*(.)$ of (2'), any θ^* which generates such a fixed point from an initial history $h_0 = \{x\}$ onwards must clearly have the property that $J_{\theta^*}(h) = J_{\theta^*}[x(h)]$ for all histories h[3]. Although they need not be Markovian in the general case, all solutions must then nevertheless deliver a Markovian pay off. Although we were not capable of showing in the general case that any such Θ would always yield the same equilibria as Markovian policies, we conjecture that this is likely to be the case in most instances. In particular, one can readily prove that any set Θ for which the equilibrium trajectories are always monotonic (as indeed happens in the Markovian case) will have this property (see footnote 4). We leave it to future research to explore the general case.

IV - SELF-RESTRAINED STRATEGIES

We now provide the construction of a *stable state*, to which a self-restrained strategy must necessarily converge; then we build a particular Markovian strategy that we call *the recursive strategy*, and analyze is property. We show finally that the recursive policy is the unique self-restrained strategy in our model.

---

[3] The proof is straigthforward. Take θ^* one strategy that delivers the fixed point J^*. From the condition in (2'), one must have on any history h that is reached: $J^*[x(h)] \leq J_{\theta^*}(h)$. But it cannot be that a strict inequality prevails: otherwise, starting from $h_0 = \{x(h)\}$, the strategy $\theta^*|_h$ would violate the constrained optimality of $J^*(.)$.

1 - The "stable" state

We shall first define a "stable" state i^* as one at which the agent prefers to stay forever rather than to move to a slightly higher or a slightly lower state variable, i.e. rather than to move to its immediate right i^*+1 or to its immediate left i^*-1.

For this purpose, let us first analyze a strategy which simply amounts to move from state i to the neighboring state $i+1$ (or $i-1$) and to stay at $i+1$ (or $i-1$) *indefinitely*. This only defines a path in the space of state variables and the agent's decision is a probability distribution upon the trajectories following this path, i.e. the agent indeed controls the speed at which he switches from i to $i+1$ (or $i-1$). The value associated to the optimal strategy that implements such a path is denoted by $J[i,i+\varepsilon]$, where $\varepsilon = 1$ or -1. This value function can be characterized by:

$$\begin{aligned} J[i,i+\varepsilon] &= \underset{\substack{a\in A_i,\ \text{s.t. } \varepsilon a\geq 0}}{\text{Max}}\ J_a(i,i+\varepsilon) \\ &\equiv \underset{\substack{a\in A_i,\ \text{s.t. } \varepsilon a\geq 0}}{\text{Max}}\ \left(\delta u(i,a) + \beta|a|\hat{J}_0(i+\varepsilon) + \beta(1-|a|)\hat{J}_a(i,i+\varepsilon)\right) \end{aligned} \qquad (3)$$

with

$$\hat{J}_0(i) = \frac{1}{1-\beta}\, u(i,0)$$

$$\begin{aligned} \hat{J}_a(i,i+\varepsilon_a) &\equiv u(i,a) + \beta|a|\ \hat{J}_0(i+\varepsilon_a) + \beta(1-|a|)\ \hat{J}_a(i,i+\varepsilon_a) \\ &= J_a(i,i+\varepsilon_a) - (\delta - 1)\, u(i,a) \end{aligned} \qquad (4)$$

or, in closed form:

$$\hat{J}_a(i,i+\varepsilon_a) = \frac{1}{1-\beta(1-|a|)}\left(u(i,a) + \frac{\beta}{1-\beta}|a|\ u(i+\varepsilon_a,0)\right) \qquad (5)$$

$J_a(i,i+\varepsilon)$ is the agent's intertemporal utility of adopting consumption a in state i, which also determines the speed (or probability) of switching to state $i+\varepsilon$, anticipating that once he reaches state $i+\varepsilon$, he will consume his state preserving consumption, namely $a=0$. $J_0(i)$ can be defined as the agent's intertemporal utility of choosing the state-preserving consumption

at state i. The "$\hat{}$"-functions have the same interpretation when δ is taken equal to *1*. It should be noted that in program (3), the consumer anticipates that a not only determines the probability of switching to state $i+\varepsilon$ at the current date, but that it also commits the consumer to choose the very same level of consumption tomorrow, or later, if he has not switched to state $i+\varepsilon$ at this later date; (3) charaterizes a more sophisticated behavior than a simple best response to what will be done in the future.

We shall now simply define a *stable state* as one at which the agent prefers to stay forever rather than to move (in a Markovian way) and stay indefinitely either to the right or to the left. Formally:

Definition 3: State i^* is a *stable state* if and only if, for $\varepsilon = 1$ and -1, the solution of program (3) at i^* is reached for $a = 0$.

The first order condition associated to program (3) can readily be written, for an interior solution, as:

$$\frac{d}{da}\left(J_a(i,i+\varepsilon)\right) = 0$$

From (4), one finds:

$$\frac{d}{da}\left(J_a(i,i+\varepsilon)\right) = \frac{d}{da}\left(\hat{J}_a(i,i+\varepsilon)\right) + (\delta - 1)\, u_2(i,a)$$

which yields as first order condition for (3):

$$u_2(i,a)\left(1 + (\delta-1)(1-\beta(1-|a|)\right) = \beta\,\varepsilon\left(\hat{J}_a(i,i+\varepsilon) - \hat{J}_0(i+\varepsilon)\right) \qquad (6)$$

In the simple case in which δ=1 (which corresponds to the case when the agent's choices are dynamically consistent) this first order condition can simply be read as balancing the marginal benefit (say if $a > 0$) from a small increase in consumption with the marginal cost that corresponds to the marginal increase in the risk of shifting from state i to state $i+1$.

This stable state will be reached when the benefit of consuming any additional small amount is less than the cost of an increased risk of switching from state i to the neighboring state $i+1$; and when the disutility

of reducing one's consumption by any small amount is larger than the benefit of increased probability of getting in a lower state (by one unit). Such a stable state is consequently characterized by the following inequalities, stating that $a=0$ is a local optimum in (3):

$$\beta\left(\hat{J}_0(i-1) - \hat{J}_0(i)\right) \leq u_2(i,0)\left(1 + (\delta-1)(1-\beta)\right) \leq \beta\left(\hat{J}_0(i) - \hat{J}_0(i+1)\right) \qquad (7)$$

It is now easy to motivate the following assumption:

A: i) For $\varepsilon=-1$ and 1 and for any i, program (3) is concave;
ii) $\lambda(i) \equiv u(i,0) - u(i+1,0)$ is increasing in i;
iii) $u_2(i,0)$ is constant in i.

Assumption A simply validates the first order approach that we have followed to define the stable state. Ai) basically requires that utility be concave enough with respect to consumption a so as to damp possible convexities with respect to a resulting from (5). Aii) and iii) are sufficient conditions that ensure that system (7) defines one and only one value for the state variable. They are clearly stronger than necessary. Aii) requires that the intertemporal loss of staying forever in state $i+1$ rather than in state i increases with state i; in some sense, higher states are increasingly domageable in the long run. Aiii) simply states that the marginal utility of consumption at the state-perserving level is independent of the state.

It should be readily noted that if Aii) and iii) were violated but $\left\{\frac{\beta}{1-\beta}\right.$ $\left.\lambda(i) - u_2(i,0)\left(1 + (\delta-1)(1-\beta)\right)\right\}$ is still increasing in i, then, provided the grid is sufficiently fine, there would exist a (range of) stable state(s); the whole analysis would carry through (with only minor changes within the stable range). However, if the previous expression becomes decreasing, the analysis would lead to stable states at the corners of the grid.

Taking A into account, it is immediate to check that the system of inequalities (7) characterizes the stable state.

Lemma 1: Under A, there exists a unique stable state; it is characterized by (7).

The system of inequalities characterizing the stable state under A also delivers additional properties of the stable state i^*. In particular, the local characterization of the stable state implies, under the concavity assumption Ai), a global stability property: for any initial value i of the state variable, the consumer always prefers to stay at this state rather than to move forever one step further in the direction opposite to i^*, i.e. rather than to move to $i+\varepsilon_i$ and stay there, where ε_i is defined by:

$\forall i \neq i^*$, $\varepsilon_i(i-i^*)>0$, i.e. $\varepsilon_i \equiv 1$ if $i>i^*$, -1 if $i<i^*$.

This global property of the stable state is crucial to the proof of our main result and is, indeed, the only property of the stable state that is used in our theorem; we summarize this property in Lemma 2.

Lemma 2: Under A, $\forall i \neq i^*$, $\forall a \in A_i$ such that $a\varepsilon_i>0$, $J_0(i) > J_a(i,i+\varepsilon_i)$.

Proof: Under the concavity assumption, it is sufficient to prove that the derivative of $J_a(i,i+\varepsilon_i)$ has a sign opposite to ε_i at $a=0$. It comes:

$$\frac{d}{da}\left(J_a(i,i+\varepsilon_i)\right)\Big|_{a=0} = u_2(i,0)\left(1 + (\delta-1)(1-\beta)\right) + \beta\varepsilon_i\left(\hat{J}_0(i+\varepsilon_i) - \hat{J}_0(i)\right)$$

If $\varepsilon_i=1$, the second term is proportional to $-\lambda(i)$ and therefore decreasing in i above i^*; since the above derivative is non-positive at i^*, the result follows. The proof is similar for $i<i^*$.

QED

Note that Lemma 2 is also valid at i^* for any a and ε; this is simply the definition of the stable state. We are going to use the ideas behind Lemma 2 to construct a Markovian policy over the whole grid, that will never tend to move the consumer to a state variable further away from the stable state.

2 - The recursive policy

In this subsection, we construct the Markovian policy that will turn out to be the unique self-restrained strategy, i.e. our proposed solution to the dynamic inconsistency of the consumer: we call it the recursive policy.

The general idea of this construction can be simply explained. First, the recursive policy requires that the consumer stays at the stable state, once it has been reached. Then, the recursive policy is defined in recursive terms with respect to the state variable. Given Lemma 2, it is natural to try to construct a policy that never goes to the wrong direction. Then, at some value of the state variable i, the recursive policy is built as the best Markovian way (optimal probability) of moving into the right direction, namely towards i^*, taking as given the recommendation of the recursive policy for subsequent state variables, that will necessarily be closer to the stable state. In some sense, instead of finding a best action today given actions in future times, we replace the abstract chronology by its physical counterpart, i.e. the state variable, and we find a best action for the current state, given actions at subsequent states. This approach clearly corresponds to our notion of a self-restained way to converge to the stable state.

Let us now characterize formally the "self-restrained" way to converge to i^*. Call θ^* the Markovian strategy that we iteratively characterize using steps a, b and c.

a/ <u>At the stable state:</u> the recursive policy requires to stay at the stable state:

-- $\theta^*(i^*) = 0$

■

b/ <u>Above the stable state:</u> for the immediate neighbor i^*+1, $\theta^*(i^*+1)$ recommends the optimal consumption that allows to jump to the stable state, given that the consumer will stay there:

$$-- \quad \theta^*(i^*+1) = \underset{a \leq 0}{\text{Argmax}} \; J_a(i^*+1, i^*)$$

Simply call $J^*(i^*+1) = J_{\theta^*}(i^*+1, i^*)$, and $\hat{J}^*(i^*+1)$ the corresponding value when δ=1. The recursive policy then recommends at i^*+2 the best consumption that can lead to the lower state i^*+1, given that it will still apply if the jump does not take place and that $\hat{J}^*(i^*+1)$ will measure the continuation utility if the jump does occur:

$$-- \quad \theta^*(i^*+2) \equiv \underset{a \leq 0}{\text{ArgMax}} \left\{ \delta u(i^*+2, a) + \beta |a| \hat{J}^*(i^*+1) + \beta(1-|a|) \hat{J}_a(i^*+2) \right\} \qquad (8)$$

with:

$$\hat{J}_a(i^*+2) = u(i^*+2, a) + \beta |a| \hat{J}^*(i^*+1) + \beta(1-|a|) \hat{J}_a(i^*+2)$$

$$= \frac{1}{1-\beta(1-|a|)} \left[u(i^*+2, a) + \beta |a| \hat{J}^*(i^*+1) \right]$$

and $J^*(i^*+2)$ the maximum value in (8). The same construction can be iterated and we consequently define recursively a Markovian policy which never leads the consumer to choose to jump to a higher state variable on the subset $\{i^*, \ldots, I\}$ and which delivers a value function $J^*(i) \geq J_0(i)$ for any $i \in \{i^*, \ldots, I\}$.

■

c/ Below the stable state: on the left hand side of i^*, one can define the recursive policy in a similar way:

$$-- \quad \theta^*(i^*-1) = \underset{a \geq 0}{\text{Argmax}} \; J_a(i^*-1, i^*)$$

Again, define $J^*(i^*-1) = J_{\theta^*}(i^*-1, i^*)$, and $\hat{J}^*(i^*-1)$ the corresponding value when δ=1. The recursive policy then recommends at i^*-2 the best consumption that can lead to jump to i^*-1, given that it will still apply if the jump does not take place and that $\hat{J}^*(i^*-1)$ will measure the continuation utility if the jump does occur:

$$-- \quad \theta^*(i^*-2) \equiv \underset{a \geq 0}{\text{ArgMax}} \left\{ \delta u(i^*-2, a) + \beta a \hat{J}^*(i^*-1) + \beta(1-a) \hat{J}_a(i^*-2) \right\} \qquad (9)$$

with:

$$\hat{J}_a(i^*-2) = u(i^*-2, a) + \beta \, a \, \hat{J}^*(i^*-1) + \beta(1-a) \, \hat{J}_a(i^*-2)$$

$$= \frac{1}{1-\beta(1-a)}\left[u(i^*-2,a) + \beta\, a\, \hat{J}^*(i^*-1)\right]$$

and $J^*(i^*-2)$ the maximum value in (9). The construction then proceeds by iteration and defines $\theta^*(i)$ for $i\in\{0,\ldots,i^*-1\}$, which completes the whole constructive definition of the recursive policy.

■

Since for each value of the state variable, the recursive policy elicits the best action that *weakly* leads towards i^*, given its recommendation for subsequent states, it is clear that it delivers a value function that is weakly preferred to a state-perserving policy:

$$\forall i\in X,\ J^*(i) \geq J_0(i)$$

But we also know from the existence and uniqueness of the stable state that for $\varepsilon=-1$ and $\varepsilon=1$, $\theta^*(i^*+\varepsilon)\neq 0$, which together with the concavity assumption in A2 implies that: $J^*(i^*+\varepsilon) > J_0(i^*+\varepsilon)$. It is interesting to note that this strict inequality generalizes to the right of the stable state: the recursive policy is *strictly* converging to the stable state from above.

Lemma 3: Under A, the recursive policy is such that for any $i>i^*$, $\theta^*(i) < 0$ and $J^*(i) > J_0(i)$.

Proof: The proof is by induction. The result holds for $i=i^*+1$. Suppose it holds for any j within $\{i^*+1,\ldots,i\}$. Then,

$$\hat{J}^*(i) = J^*(i) - (\delta-1)u(i,\theta^*(i)) > J_0(i) - (\delta-1)u(i,\theta^*(i))$$
$$> J_0(i) - (\delta-1)u(i,0) = \hat{J}_0(i)$$

where the second inequality comes from $\theta^*(i)<0$ and $u(i,.)$ increasing. Using this result, it comes:

$$J^*(i+1) = \max_{a\leq 0}\left\{\delta\, u(i+1,a) + \beta|a|\hat{J}^*(i) + \frac{\beta(1-|a|)}{1-\beta(1-|a|)}\left[u(i+1,a) + \beta|a|\hat{J}^*(i)\right]\right\}$$
$$> \max_{a\leq 0}\left\{\delta\, u(i+1,a) + \beta|a|\hat{J}_0(i) + \frac{\beta(1-|a|)}{1-\beta(1-|a|)}\left[u(i+1,a) + \beta|a|\hat{J}_0(i)\right]\right\}$$

$$= \underset{a \le 0}{\text{Max}}\ J_a(i+1,i) \quad \geq \quad J_0(i+1).$$

Hence the result for $i+1$.

QED

But, contrarily to the case discussed above relative to the right-hand side segment $\{i^*,\ldots I\}$, we cannot guarantee that $\theta^*(i)>0$ for any $i<i^*$. We will call i a *stopping point* if the recursive policy is such that $\theta^*(i)=0$. Following the recursive policy, the agent may stop on his way of increasing his state variable toward i^*; however, one can show that these stopping points are isolated.

Lemma 4: Under A, if i is a stopping point, then $i \le i^*$ and $i-1$ is not a stopping point.

Proof: The first part is a consequence of Lemma 3. Suppose that i is a stopping point. Then $\theta^*(i-1)$ is determined as the argument maximum of:

$$\delta\ u(i-1,a) + \beta|a|\hat{J}^*(i) + \frac{\beta(1-|a|)}{1-\beta(1-|a|)}\left(u(i-1,a) + \beta|a|\hat{J}^*(i)\right)$$

If i is a stooping point, then $\hat{J}^*(i)=\hat{J}_0(i)$ and the above maximand is nothing else than $J_a(i-1,i)$. Lemma 1 then implies that $\theta^*(i-1)>0$.

QED

3 - Main Theorem

We can now prove our main result, which addresses the issue of the characterization of self-restrained strategies. From the construction of the recursive policy, it is natural to expect that the recursive policy is indeed self-restrained. Moreover, it turns out to be the only self-restrained strategy.

Theorem: The recursive policy is the only self-restrained strategy.

Proof: In order to prove this theorem, we first have to contemplate the possibility of strategies that would converge to a "cycle". Our discrete

model with Markovian strategies implies that the only possible cycles that ought to be considered are two-state cycles. The following lemma is proved in the Appendix and rules out such cycles.

Lemma 5: Any Markovian strategy θ that "oscillates" between two states $\{i,i+1\}$ is necessarily dominated at either state i or state $i+1$ by the state-preserving policy: $\exists j \in \{i,i+1\}$, $J_\theta(j) < J_0(j)$.

The recursive policy is a self-restrained strategy

Let us then prove, first, that the recursive policy is a self-restrained strategy. If not, there would exist another Markovian policy θ and one trajectory $O[\theta,x]$ upon which, for any state y on the trajectory $O[\theta,x]$, $J_\theta(y) \geq J^*(y)$, with at least one strict inequality.

Because of Lemma 5 (and because of the fact that $J^*(y) \geq J_0(y)$ everywhere), one knows that θ cannot converge to a cycle; it must be converging towards some x^*. At that point, we must then have:

$$J_\theta(x^*) = J_0(x^*) \geq J^*(x^*)$$

which shows that x^* must be a stopping point of θ^*. From the constructive definition of $J^*(.)$, we know that necessarily either $x^* = i^*$ or $x^* \leq i^*-2$.

Suppose that $x^* \leq i^*-2$. The last state x through which x^* can be reached on a trajectory of θ is either x^*-1 or x^*+1. However, one knows that:

$$J^*(x^*+1) \geq J_0(x^*+1) > J[x^*+1,x^*] = J_\theta(x^*+1)$$

The second inequality follows from $x^*+1 \leq i^*-1$ and Lemma 2; the last equality is necessary if θ if a self restrained policy that stops at x^*. It follows that θ cannot lead from x^*+1 to x^*.

So θ must go through x^*-1 before x^* is reached and the trajectory is necessarily a segment $\{x,x+1,\ldots,x^*\}$. Since x^* is a stopping of θ^*, θ must by construction be the recursive policy itself. Similarly if $x^*=i^*$. Hence, a contradiction with the requirement that a strict inequality is needed for θ^* not to be a solution of the fixed point problem.

Uniqueness of the self-restrained strategy

Let us now similarly show that θ^* is the only self-restrained strategy. If not, call θ another self-restrained strategy. Again, it cannot be that θ is cycling between two states, otherwise there would be one state x for which $J_0(x) > J_\theta(x)$ which would clearly contradict the fact that θ is a solution to (2).

The trajectory generated by θ must then be monotonic[4]. Clearly, θ must be "pointing" toward i^*. Otherwise there would be some state x^*, such that: $\theta(x^*-1) > 0 = \theta(x^*) > \theta(x^*+1)$, which implies:

$$J_0(x^*-1) < J_\theta(x^*-1)$$
$$J_0(x^*+1) < J_\theta(x^*+1)$$

If $x^* \neq i^*$, one of these inequalities contradicts Lemma 2. Hence $x^*=i^*$.

Now, if θ is "pointing" towards i^*, it must necessarily coincide with θ^* and share its stopping points. This is immediate by induction on each side $\{0,\ldots,i^*-1\}$ and $\{i^*+1,\ldots,I\}$.

QED

To summarize, our concept of self-restraint leads to a unique and interpretable strategy: the recursive policy. In a situation where, at each date, the agent tries to find the best commitment to a state-dependent strategy that will not be replaced in the future by an alternative commitment, that will not itself be replaced in the future by still another commitment, ..., he will choose the recursive policy. The recursive policy can only be dominated at some date, by a strategy that incorporates some inconsistency in the sense that the consumer cannot believe he will stick forever to this later strategy.

---

[4]The forthcoming proof would apply with little changes to the case when the set of strategies is a set Θ that includes the Markovian policies, and for which the equilibrium trajectories are monotonic.

V - COMPARISON TO THE BELLMAN EQUATION

Let us finally compare the "self-restrained" equilibrium to the solution of the Bellman equation. The Bellman solution $J(.)$ solves the following equation :

$$\forall x,\ J(x) = \underset{a\varepsilon \geq 0}{\text{Max}} \left\{ \delta\ u(x,a) + \beta|a|\ \hat{J}(x+\varepsilon) + \beta(1-|a|)\ \hat{J}(x) \right\}$$

The crucial difference with programs characterizing our solution concept is that the consumer at time t and state x does not internalize that the very same choice of action at state x will occur if later he still happens to be in state x: $\hat{J}(x)$ is taken as given.

The first-order condition yields :

$$\delta\ u_2(x,a) = \beta\ \varepsilon\ \left(\hat{J}(x) - \hat{J}(x+\varepsilon) \right)$$

Again, the difference with the "self-restrained" strategy is apparent. The Bellman equation does not take into account the fact that the Markovian strategy θ that is implemented may be committed from one period to the next on a state-by-state basis. Any point x at which it is optimal to stay, satisfies:

$$\beta \left(\hat{J}_0(x-1) - \hat{J}_0(x) \right) \leq \delta\ u_2(x,0) \leq \beta \left(\hat{J}_0(x) - \hat{J}_0(x+1) \right) \qquad (11)$$

It then follows immediately:

Proposition 2: Under A, there exists a unique stable point i^B of the solution to the Bellman equation and: $i^B \geq i^*$.

Proof: Existence and uniqueness of i^B are direct conclusions of the concavity of Bellman program and A. A simple comparison to (7) provides the final result since:

$$1 + (\delta-1)(1-\beta) = \delta - (\delta-1)\beta < \delta$$

QED

This shows that the stable-state of the Bellman equation is necessarily higher than the stable-state of the self-restrained strategy. The logic of the Bellman solution does not lead the consumer to internalize the future consequences of his present choice, when he will be in the same state; therefore, his behavior is more myopic than if he self-restains.

If one starts from the steady-state of the Bellman equation, it should be clear that the self-restrained agent improves upon this level of welfare by switching to a self-restrained way of smoking. Indeed, when at i^B, it is better to implement the self-restrained policy rather than staying indefintely to the right of i^* (lemma 3); then at $i^B - 1$, it is better to remain there rather than switching to i^B (lemma 2), and it is better to implement the self-restrained policy rather than to stand still. Following the proof of lemma 3, one can then get by induction that the self-restrained polciy dominates the Bellman choice all along the trajectory that goes from i^B to i^*.

To come back to our opening example, if the smoker does not internalize the fact that by smoking less today, given his current level of addiction, he also commits to smoke less tomorrow if he still is at the same level of addiction, then he has much less incentives to move to a lower level of addiction; he ends up more intoxicated and unhappier than if he had internalized such commitment aspect of his state-dependent behavior.

VI - SUMMARY AND CONCLUSION

We have characterized what may be called a "state-consistent" decision rule. In standard Markovian approaches, an agent takes as given the decision that he 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 he takes as given what he 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 himself in the same state of nature. Starting from the steady-state of his time-consistent choices, a "self-restrained" agent could raise his welfare, not only

initially, but uniformly across the new trajectory that he would set in motion (by say, smoking less all the time). Conversely, he would move away from this self-restrained steady-state only by indulging himself into thinking that he can deviate "only once" from this diet.

REFERENCES

Akerlof,G., [1991]: "Procastination and Obedience", Richard T.Ely Lecture, American Economic Review, Vol 81, N°2, 1-19.

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.

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

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.

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

Laibson,D., [1992]: "Self-Control and Saving", mimeo MIT.

Laibson,D., [1993]: "Golden Eggs and Hyperbolic Discounting", mimeo MIT.

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.

Strotz,R., [1955]: "Myopia and Inconsistency in Dynamic Utility Maximization", Review of Economic Studies, 23, 165-180.

APPENDIX

INEXISTENCE OF A SUBGAME PERFECT EQUILIBRIUM

Suppose that θ_1 is a subgame perfect equilibrium. Consider an history h such that $t(h) = 1$, the path of states is x_0 followed by x_1. Denote a_1 the action called for by $\sigma_1|_h$ at date *1*. To avoid trivial cases, assume that it is interior. Denote $W(a) = J(x_1,\theta)$ the value of a strategy θ that calls for a as the first action when the initial state is x_1 and coincides with $\theta_1|_h$ otherwise. From subgame perfection, $\theta_1|_h$ reaches the maximum of $J(x_1,\theta)$ over strategies θ, therefore a_1 must reach the maximum of $W(a)$. Moreover θ_1 reaches the maximum of $J(x_0,\theta)$: this implies that $\theta_1|_h$ reaches the maximum of $J(x_1,\theta) + (1-\delta)u(x_1,a)$, and in particular that a_1 reaches the maximum of $W(a) + (1-\delta)u(x_1,a)$. But then $u_2(x_1,a_1) = 0$ which is impossible. Therefore a subgame perfect equilibrium does not exist.

The same proof holds if one rectricts strategies to be Markovian. The only difference is that one should take $x_1 \neq x_0$, hold the action prescribed at states $x \neq x_1$ constant, and consider the choice of the action a_1 prescribed at state x_1.

QED

PROOF OF THE LEMMA 5

Consider a Markovian policy that is cycling between i and $i+1$. There exists $a_1 > 0$ and $a_2 < 0$, such that:

$$J_\theta(i) = \delta\, u(i,a_1) + \beta\, a_1\, \hat{J}_\theta(i+1) + \beta\, (1-a_1)\, \hat{J}_\theta(i)$$

$$J_\theta(i+1) = \delta\, u(i+1,a_2) + \beta|a_2|\, \hat{J}_\theta(i) + \beta\, (1-|a_2|)\, \hat{J}_\theta(i+1)$$

and

$$\hat{J}_\theta(i) = J_\theta(i) - (\delta-1)\, u(i,a_1)$$

$$\hat{J}_\theta(i+1) = J_\theta(i+1) - (\delta-1)\, u(i+1,a_2)$$

Now

$$|a_2|\ J_\theta(i) + a_1\ J_\theta(i+1) = \delta\ \left(|a_2|\ u(i,a_1) + a_1\ u(i+1,a_2)\right) + \beta\ \left(|a_2|\ \hat{J}_\theta(i) + a_1\ \hat{J}_\theta(i+1)\right)$$

so that,

$$|a_2|\ J_\theta(i) + a_1\ J_\theta(i+1) = \left(\delta + \frac{\beta}{1-\beta}\right)\left(|a_2|\ u(i,a_1) + a_1\ u(i+1,a_2)\right)$$

By comparison, one has

$$|a_2|\ J_0(i) + a_1\ J_0(i+1) = \left(\delta + \frac{\beta}{1-\beta}\right)\left(|a_2|\ u(i,0) + a_1\ u(i+1,0)\right)$$

To prove the lemma, it is then sufficient to prove that, for any state i and any consumptions a_1 and a_2 with $a_1 > 0 > a_2$,

$$\left(|a_2|\ u(i,a_1) + a_1\ u(i+1,a_2)\right) < \left(|a_2|\ u(i,0) + a_1\ u(i+1,0)\right) \qquad (*)$$

By concavity of $u(i,.)$, for any a_1, a_2 with $a_1 > 0 > a_2$,

$$u(i,a_1) < u(i,0) + a_1 u_2(i,0)$$

$$u(i+1,a_2) < u(i+1,0) - |a_2| u_2(i+1,0)$$

Therefore :

$$\left(|a_2|\ u(i,a_1) + a_1\ u(i+1,a_2)\right) < \left(|a_2|\ u(i,0) + a_1\ u(i+1,0)\right) + a_1|a_2|\ \left(u_2(i,0) - u_2(i+1,0)\right)$$

Assumption A (u_2(i,0) is constant) then clearly implies (*).

QED