THE RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

Working Paper No. 577, 2002

Uniqueness in Infinitely Repeated Decision Problems

by Nicolas Vieille and Jörgen W. Weibull

IUI, The Research Institute of Industrial Economics
P.O. Box 5501
SE-114 85 Stockholm
Sweden

Uniqueness in infinitely repeated decision problems

Nicolas Vieille
Laboratoire d'Econométrie, Ecole Polytechnique, and
Département Economie et Finance, HEC, 78 351 Jouy en Josas

Jörgen W. Weibull*
Department of Economics, Stockholm School of Economics
P.O. Box 6501, SE 113 83 Stockholm, and
Research Institute of Industrial Economics
P.O. Box 5501, SE 114 85 Stockholm

April 16, 2002

Abstract. Dynamic decision-making without commitment is usually modelled as a game between the current and future selves of the decision maker. It has been observed that if the time-horizon is infinite, then such games may have multiple subgame-perfect equilibrium solutions. We provide a sufficient condition for uniqueness in a class of such games, namely infinitely repeated decision problems with discounting. The condition is two-fold: the range of possible utility levels in the decision problem should be bounded from below, and the discount function should exhibit weakly increasing patience, that is, the ratio between the discount factors attached to periods $t+1$ and t should be non-decreasing in t, a condition met by exponential, quasi-exponential and hyperbolic discounting.

JEL codes: C61, C72, C73, D90
Keywords: game theory, time preference, hyperbolic discounting, repeated decision problems.

*Both authors thank Maria Saez-Marti and Bill Sandholm for comments that helped us improve the presentation, and Weibull thanks the Laboratoire d'Econométrie at Ecole Polytechnique for its hospitality during part of this research.

1. Introduction

A recent strand in economic theory concerns non-exponential time preferences as a source of dynamic inconsistency in intertemporal decision making.[1] Such dynamic decision problems are usually viewed as a sequential game between the decision maker's different selves.[2] A decision rule then takes the form of a strategy profile between these selves. In the absence of precommitment possibilities, a "consistent plan" for the decision maker is a subgame perfect equilibrium in this game. We follow this approach. Moreover, the mentioned studies assume certain parametric representations of quasi-exponential discount functions which approximate the hyperbolic functions identified in experimental psychology.[3] As compared with exponential functions, all these functions have decreasing, rather than constant, discount rates - that is, they exhibit relatively more patience in the distant future than in the near future. It hence seems desirable to establish results for all discount functions with this qualitative property, which we will call *weakly increasing patience*. This is the route followed here.

It is well-known that infinitely repeated games between patient players have a very large set of subgame perfect equilibria - with outcomes spanning all feasible and individually strictly rational outcomes. This is true even when the stage game has a unique Nash equilibrium and hence every finite repetition has a unique subgame perfect equilibrium. Under what conditions can such lack of (lower hemi-) continuity of the solution correspondence at infinity arise in games between different selves without precommitment possibilities? It should be noted that while the effect is similar to that of the Folk theorem, the reason is different. The key mechanism behind the Folk theorem is credible punishments among players who can choose whole strategies, that is, who can precommit to future actions. By contrast, the key mechanism here is the lack of such precommitment possibilities.

Examples of infinite-horizon decision problems with multiple, even infinitely many, subgame perfect equilibria have been given by Asilis et al (1991), Laibson (1994), Kocherlakota (1996) and Asheim (1997). These authors view this multiplicity and the ensuing indeterminacy of the outcome as a pathology that asks for remedy. They have therefore suggested refinements of the subgame perfection criterion for such classes of games.[4] The main point of the present study is not to take a position

[1]See, for example, Laibson (1997), Barro (1999), Krusell and Smith (1999) and Laibson and Harris (2001).

[2]See references in the preceding footnote, and also Strotz (1956), Pollak (1968), Phelps and Pollak (1968), Peleg and Yaari (1973), Elster (1979), Goldman (1980), Asheim (1997) and Bernheim et al. (1999).

[3]See, for example, Mazur (1981), Herrnstein (1987) and Ainslie (1992).

[4]Laibson (1994) requires that a certain set of continuation payoffs be bounded, while Asheim

for or against those or other refinements. Instead, our goal is to identify a class of dynamic decision problems in which there is no issue of multiplicity to begin with, that is, a class of decision problems, each of which has a unique subgame perfect equilibrium outcome. More exactly, we provide a sufficient condition for infinitely repeated decision problems (technically speaking, programs) to have a unique such solution.

Consider a decision maker who repeatedly, over an infinite sequence of discrete time periods t, faces the same decision problem over and over, namely to choose some action x_t from a given set X, the same in all periods. The decision maker is presumed to have stationary and additively separable time preferences represented by a sum of discounted instantaneous utilities (or period payoffs). The condition is two-fold:

(A) *Lower bound*: the set of feasible instantaneous utility levels should be bounded from below,

and

(B) *Weakly increasing patience*: The ratio between the discount factors attached to periods $t+1$ and t should be non-decreasing in t.

In other words, there should be a lower bound on how much damage the decision maker can cause him- or herself in a period, and the decision maker's patience should not be greater concerning events in the near future than concerning events in the distant future. Condition A is met in many, if not most, of the decision problems analyzed in the economics literature. Laibson (1994) notes that without this condition, any feasible path, in a certain class of dynamic decision problems, is subgame perfect. Condition B is met under traditional exponential discounting as well as under quasi-exponential and hyperbolic discounting. This condition has also turned out to be useful for another, related, purpose, see Saez-Marti and Weibull (2002).

We proceed by first, in Section 2, giving three examples in order to high-light the multiplicity issue and to provide a key to the basic logic behind the subsequent analysis. The first two examples are close in spirit to those in Laibson (1994), although his setting is slightly different, and our third example is close in spirit to Example 2 in Asheim (1997).[5] Section 3 provides formal definitions and establishes the uniqueness result.

(1997) requires that the strategy profile be immune to future favorable revisions.

[5]According to Ashiem (1997), his example is in turn inspired by an example in an unpublished manuscript by Asislis et al. (1991).

2. Examples

2.1. Cobb-Douglas preferences. Assume that the decision maker in each period τ has the following Cobb-Douglas preferences over consumption profiles $\bar{x} = (x_0, x_1, x_2, ...)$:

$$U_\tau(\bar{x}) = \prod_{t=0}^{\infty} (x_{\tau+t})^{\lambda_t} \qquad \forall \tau \in \mathbb{N},$$

where $x_{\tau+t}$ is consumption in period $\tau + t$, and all parameters λ_t are non-negative and add up to 1, with at least λ_0 and λ_1 positive.[6]

First, suppose that the good is non-storable and that the decision maker faces the same compact set of alternatives in each period: $x_t \in X = [0, 1]$ for all t. At first sight, the decision problem may seem trivial: choose $x_t = 1$ in every period. Indeed, this is a subgame perfect equilibrium profile. For if future consumption levels are positive and independent of current consumption, then it is a best reply to choose maximal consumption in the current period. However, there are infinitely many subgame perfect equilibria, and the resulting utility levels span the whole range $[0, 1]$ of possible values. For example, to consume 0 in each period, irrespective of history, is clearly subgame perfect. For if the decision maker expects to consume nothing in some future period, then he or she is indifferent as to current consumption, and can thus just as well consume nothing in the current period too.[7] Moreover, for any positive consumption level $x = a > 0$ in X, the following stationary strategy profile is subgame perfect: consume a if consumption in all earlier periods was a, otherwise consume 0. To see that this indeed is subgame perfect, it suffices to verify two conditions: deviations from the "norm" a should be sufficiently punished, and neglect of punishment of deviations from the norm should be sufficiently punished. The temptation to consume more than a in the current period, if a was consumed in all preceding periods, just results in an expected total utility of zero, since then zero consumption is expected in the next period. Second, if some preceding consumption level differed from a, then current consumption does not matter, since zero consumption is anyhow expected in the next period.

Secondly, suppose instead that the good is storable and that the set of alternative consumption streams has to meet a life-time budget constraint. More exactly, suppose that consumption has to be non-negative in all periods t, and sum up to some given positive number, which we normalize to unity: $\sum_{t=0}^{\infty} x_t \leq 1$. If the consumer in

[6]For example, these weights may decline exponentially over time; $\lambda_t = (1 - \delta)\,\delta^t$ for some $\delta \in (0, 1)$.

[7]Note, however, that this strategy is weakly dominated.

the initial period could pre-commit to a consumption profile, then he or she would evidently choose $x_t = \lambda_t$ for all periods t - this choice would maximize $U_0(\bar{x})$. However, this dynamic decision problem has infinitely many subgame perfect equilibria as well. First, to consume 0 in each period, irrespective of history, is clearly subgame perfect. For if the decision maker expects zero consumption in the next period, then he or she is again indifferent as to current consumption. Moreover, for any stream $\bar{x}$ of positive consumption levels x_t, summing up to 1 or less, the following strategy profile is subgame perfect: consume x_t in each period t if consumption in all earlier periods followed $\bar{x}$, otherwise consume 0. Again the temptation to consume more than x_t in period t, if the prescribed stream $\bar{x}$ was followed so far, results in an expected total utility of zero and likewise if in some preceding period consumption deviated from $\bar{x}$.

In this example, there is no real incentive to punish: the punisher is indifferent between punishing and not punishing. Moreover, while most models of intertemporal decision making in economics are additively separable, the present model is not. We may, of course, just take the logarithm of the present utility function U_τ in order to obtain additive separability. However, while in the present example the lower bound on consumption, zero, belongs to the domain of the subutility function, this is not the case after the logarithmic transformation. Though seemingly insubstantial, this transformation does matter, since then the threat of zero consumption as a punishment is no longer available. We study this case in the following example, where there is a strict incentive to punish.

2.2. Exponential time preferences. In this example, the decision maker has additively separable and exponential time preferences. Suppose, moreover, that the set of available actions is the same in each period, and that this set contains a unique point which maximizes instantaneous utility. Despite this "classical" setting of this example, all feasible utility levels are still consistent with subgame perfection.

More specifically, let the set of available actions in each period be $X = (0, 1]$, and let the instantaneous utility of consumption be logarithmic, $u(x) = \log x$. Hence, the instantaneous utility is maximized by the choice $x^* = 1$, and the associated utility level is $u(x^*) = 0$. Suppose that the decision-maker's preferences in each period τ are represented by the following utility function:

$$U_\tau(\bar{x}) = (1-\delta)\sum_{t=0}^{\infty} \delta^t \log(x_{\tau+t}),$$

for some discount factor $\delta \in (0, 1)$.[8]

[8] Hence, the current utility function is the logarithm of that in the preceding example, with $\lambda_t = (1-\delta)\delta^t$.

We argue that this game has infinitely many subgame perfect equilibria, and that the associated set of equilibrium utility levels for the decision maker, as viewed from the initial period, is the whole range $u(X) = (-\infty, 0]$ of possible utility levels. The strategy profiles which we construct below all have the feature that player 0 ("self 0") chooses a suboptimal action in terms of the one-shot decision problem, in the fear of otherwise invoking a "punishment" from the next player (self). The next player indeed does have an incentive to punish defections, since lack of punishment is expected to be punished etc., in an infinite hierarchy of ever harsher punishments.

More precisely, for any $a \in (0,1)$, we consider a certain strategy profile which induces the outcome $(a, 1, 1, 1, ..)$, that is, the suboptimal action a followed by an infinite strings of the optimal action 1, resulting in suboptimal utility to the decision maker, $U_0(\bar{x}) = \log a < 0$. Let player 0 choose $x_0 = a$. In this case, player 1 chooses $x_1 = x^* = 1$, otherwise he chooses $x_1 = a^2 < a$. Likewise, player 2 chooses $x_2 = x^* = 1$ if the preceding "history" is $h = (a, 1)$ or if it is $h' = (x_0, x_1)$ with $x_1 = a^2$. Otherwise, player 2 chooses $x_2 = a^4$. More generally, for any integer $\tau \geq 1$, player τ chooses 1 if player $\tau - 1$ punished when she should, otherwise player τ chooses $a^{p(\tau)}$, where $p(\tau) = 2^\tau$. Formally, the choice of player τ depends on the identity i_τ of the player who last deviated. This identity is defined recursively as follows: set $i_0 = -1$. For $\tau \geq 1$, set $i_\tau = \tau - 1$ (the preceding player), if $i_{\tau-1} = \tau - 2$ (if the preceding player should punish), and $x_{\tau-1} \neq a^{p(\tau-1)}$ (but did not punish), otherwise, set $i_\tau = i_{\tau-1}$ (the player who last deviated before the preceding player). The strategy of player τ, for any $\tau \geq 0$, is to choose $x_\tau = a^{p(\tau)}$ if $i_\tau = \tau - 1$ and $x_\tau = 1$ otherwise.

Is this strategy profile subgame perfect? The answer is affirmative if $\delta > 1/2$.[9] To see this, consider any player $\tau \geq 0$. If $\tau > 0$ and she deviates when $i_\tau < \tau - 1$, i.e. when she should play 1, she can only loose. Secondly, if she deviates when $i_\tau = \tau - 1$, i.e. when she should play $a^{p(\tau)}$, then she obtains

$$0 + \delta p(\tau + 1) \log a + 0 + 0 + ...$$

instead of

$$p(\tau) \log a + 0 + 0 + ...$$

The former payoff falls short of the latter iff $2\delta \log a < \log a$, or, equivalently (since $a < 1$), iff $\delta > 1/2$. Consequently, for any $a \in (0,1)$, the given strategy profile is subgame perfect.

[9] Our claim that the sequence $(a, 1, 1, ...,)$ as a a subgame perfect outcome is in fact valid as soon as $\delta > 0$. If $\delta \leq 1/2$, punishments should be more stringent.

2.3. Non-exponential time preferences. In the preceding example, punishers had a strict incentive to punish, but the range of instantaneous utilities was unbounded from below, thus allowing for an infinite hierarchy of ever harsher punishments. The first example had a lower bound on instantaneous utilities, but punishers had only a weak incentive to punish. In the following example, the range of instantaneous utilities is bounded from below and punishers have a strict incentive to punish.

Suppose $X = [0, 1]$ and let $u(x) = x$. Hence, $x^* = 1$ and $u(x^*) = 1$. The utility functions U_τ are defined as Cobb-Douglas functions of, on the one hand, the minimal payoff in the nearest $T + 1$ periods (for any $T \geq 1$), and, on the other hand, the exponentially discounted (normalized) sum of future instantaneous utilities:[10]

$$U_\tau(\bar{x}) = \left[\min_{0 \leq t \leq T} x_{\tau+t}\right]^\lambda \cdot \left[(1-\delta)\sum_{t=0}^{\infty} \delta^t x_{\tau+t}\right]^{1-\lambda} \qquad \forall \tau \in \mathbb{N}, \tag{1}$$

for some $\delta, \lambda \in (0, 1)$, where λ is the weight attached to the minimal-payoff consideration.

We construct a strategy profile much along the lines of the preceding example. Player 0 again chooses a suboptimal action, in terms of the one-shot decision problem, for fear of triggering a punishment from the next player, who in turn carries out the punishment for fear of otherwise being punished, etc. in an infinite hierarchy of punishments. More precisely, we shall again prove that, for any action $a \in (0, 1)$, the outcome $\bar{x} = (a, 1, 1, ...)$ is subgame perfect. The resulting payoff to player 0, the *ex ante* expected utility to the decision maker, is accordingly $a^\lambda [(1-\delta)a + \delta]^{1-\lambda}$, a number below one that can be brought arbitrarily close to zero by a suitable choice of a.

One strategy profile which induces this outcome is the following. Player 0 chooses $x_0 = a \in (0, 1)$. In this case, player 1 chooses $x_1 = x^* = 1$, otherwise he chooses $x_1 = a/2$. Likewise, player 2 chooses $x_2 = x^* = 1$ if the preceding history is $h = (a, 1)$ or if it is $h' = (x_0, x_1)$ with $x_1 = a/2$. Otherwise, she chooses $x_2 = a/4$. More generally, for any integer $\tau \geq 1$, player τ chooses 1 if player $\tau - 1$ punished when she should, otherwise player τ chooses $a/p(\tau)$, where again $p(\tau) = 2^\tau$. Formally, the choice of player τ again depends on the identity i_τ of the player who last deviated, where $i_0 = -1$, and, for $\tau \geq 1$, $i_\tau = \tau - 1$ if $i_{\tau-1} = \tau - 2$ and $x_{\tau-1} \neq a/p(\tau - 1)$, and otherwise $i_\tau = i_{\tau-1}$ The strategy of player τ, for any $\tau \geq 0$, is to choose $x_\tau = a/p(\tau)$ if $i_\tau = \tau - 1$, and $x_\tau = 1$ otherwise.

Is this strategy profile subgame perfect? It suffices to show that U_0 assigns a higher

[10]The analysis is valid also if $T = +\infty$.

value to the stream $\bar{x} = (a', 1, 1, 1, ...)$ than to the stream $\bar{y} = (1, a'/2, 1, 1, ...)$, for all $a' \in (0, 1)$. We analyze the special case $\lambda = 1/2$.[11] Then

$$U_0(\bar{x}) > U_0(\bar{y}) \quad \Leftrightarrow \quad a'[(1-\delta)a' + \delta] > \frac{a'}{2}\left[(1-\delta)(1+\delta a'/2) + \delta^2\right] ,$$

and the latter inequality holds if and only if

$$(1-\delta)\left(1 - \frac{\delta}{4}\right)a' + \delta > \frac{1}{2}\left(1 - \delta + \delta^2\right),$$

an inequality which holds for all $a' \in (0, 1)$ if $\delta^2 - 3\delta + 1 \leq 0$. The latter inequality holds for all $\delta \geq 1/2$.[12] In sum: The studied strategy profile is subgame perfect for all $a \in (0, 1)$ and all $\delta, \lambda \geq 1/2$. By varying the initial action a, the *ex ante* utility to the decision maker can be any number in the interval $(0, 1)$.

3. Infinitely repeated decision problems

Having considered examples of non-uniqueness, we now address the task of identifying sufficient conditions for uniqueness in infinitely repeated decision problems with discounting.[13] More exactly, consider a decision maker who repeatedly faces the same decision problem, namely to choose an action x_t from some fixed set X, in each period $t = 0, 1, 2, ...$ An *outcome* $\bar{x}$ is an infinite sequence of actions $x_t \in X$, and we write $\bar{x} = (x_0, x_1, ...) \in X^\infty$. The decision maker in each period τ has preferences which are represented by a utility function $U_\tau : X^\infty \to \mathbb{R}$, where

$$U_\tau(\bar{x}) = \sum_{t=0}^{\infty} f_t u(x_{\tau+t}) \tag{2}$$

for some function $u : X \to \mathbb{R}$ and some non-increasing, positive and summable sequence $f = (f_t)_{t\in\mathbb{N}}$.[14] We interpret $u(x)$ as the *instantaneous (sub)utility* from consuming $x \in X$, f as the *discount function*, and $f_t > 0$ as the *discount weight* attached to the instantaneous utility from consumption t periods later. Without loss of generality, we normalize the sum of the discount weights to one: $\sum_t f_t = 1$.

The analysis is focused on the case when the decision maker cannot precommit to actions in future periods. Following the literature mentioned in the introduction, we accordingly model the decision maker in each period τ as a separate player, "self τ."

[11] It is easily verified that the inequalities below in fact hold for all $\lambda \geq 1/2$.

[12] More precisely, it holds for all $\delta \geq \left(3 - \sqrt{5}\right)/2 \approx 0.38$.

[13] In particular, the second case considered in the first example - that of a life-time budget constraint - is excluded.

[14] We use $\mathbb{R}$ to denote the reals, $\mathbb{R}_+$ the non-negative reals, and $\mathbb{N}$ the non-negative integers.

Any "decision rule" of self τ takes the form of a *pure behavior strategy* $s_\tau : H_\tau \to X$, where H_τ is the set of "histories" up to, but not including, period τ. More exactly, $H_\tau = X^\tau$ for τ positive and $H_0 = \{h_0\}$, where h_0 is the "null history" in period zero.[15] The set of players is thus $\mathbb{N}$. A sequence $s = (s_\tau)_{\tau \in \mathbb{N}}$ of behavior strategies will be called a *strategy profile*, and the set of pure strategy profiles is denoted S. Clearly every pure strategy profile generates a unique outcome $\bar{x}$. The payoff to each player $\tau \in \mathbb{N}$ is the utility $U_\tau(\bar{x})$ to the decision maker in that period. Player τ's payoff function $\pi_\tau : S \to \mathbb{R}$ is thus defined by $\pi_\tau(s) = U_\tau(\bar{x})$, where $\bar{x}$ is the outcome induced by s. The *ex-ante utility* to the decision maker is defined as $U_0(\bar{x})$. This defines an infinite extensive-form game of perfect and complete information, to be denoted Γ. The solution concept we will use is that of subgame perfection; each player's strategy is required to be a best reply to the others' strategies, given any history leading up to the player's decision period.

Let v^* be the (possibly infinite) least upper bound on the decision maker's instantaneous (as well as total) utility, $v^* = \sup_{x \in X} u(x)$, and let X^* be the (possibly empty) set where u attains its supremum value:

$$X^* = \{x \in X : u(x) = v^*\} \ . \tag{3}$$

In standard microeconomic settings, X is compact and u continuous, implying that u attains its supremum value, or formally, $X^* \neq \emptyset$. This case is covered by, but not presumed, in the subsequent analysis.

Had the decision problem been repeated a finite number of times, then the set of subgame perfect equilibria is self-evident: (a) if $X^* = \emptyset$, then no equilibrium exists, (b) if X^* contains exactly one action, x^*, then the unique equilibrium is to choose x^* in each period, irrespective of history, resulting in utility v^* to the decision maker, and (c), if X^* contains more than one action, then every equilibrium prescribes some action from X^* in each period, where the choice may be conditioned on the date (τ) and/or on the history (h_τ), again resulting in utility v^* to the decision-maker. However, as the initial examples show, uniqueness may be lost if the decision problem is repeated infinitely many times.

It is natural to allow for the possibility that a player τ randomizes when choosing an action. Hence, while the main analysis will be restricted to pure behavioral strategies, as outlined above, we will also show how the analysis can be extended to general behavior strategies $\sigma_\tau : H_\tau \to \Delta(X)$, where $\Delta(X)$ is the set of probability distributions over X, and where all randomizations are statistically independent. Each such general behavior strategy profile σ induces a unique probability distribution over the set of outcomes, see Remark 2 below.

[15]See remarks below concerning randomized behavior strategies.

3.1. Uniqueness. We have seen that uniqueness may fail if the range of possible payoffs in the one-shot decision problem is not bounded from below, or if time preferences differ much from exponential discounting. We here establish uniqueness under the two-fold condition (A)-(B) mentioned in the introduction, namely, that the payoff range be bounded from below, and that the discount function exhibit weakly increasing patience. More precisely, the set $u(X) \subset \mathbb{R}$ should have a lower bound, and the sequence $g_t = f_{t+1}/f_t$ should be non-decreasing. This last condition, introduced by Saez-Marti and Weibull (2002), is clearly met by standard exponential discounting, in which case $g_t = \delta$ for all t. The condition is also met by the quasi-exponential discount functions used in the macroeconomics literature (see, for example, Phelps and Pollak (1968), Laibson (1997), Barro (1999), Krusell and Smith (1999), and Laibson and Harris (2001)): then $g_0 = \beta\delta \leq g_t = \delta$ for all positive t, where $\beta, \delta \in (0,1)$. The condition is also met by the hyperbolic discount functions used in the psychology literature (see, for example, Mazur (1981), Herrnstein (1987) and Ainslie (1992)). There, $g_t = [(1+at)/(1+a+at)]^b$, for $a, b > 0$.

The following result establishes that the uniqueness properties of the static decision program are inherited by the infinitely repeated program, under the mentioned two conditions.

Proposition 1. *Suppose $u(X)$ is bounded from below and (g_t) is non-decreasing. If X^* is a singleton set, then there exists a unique subgame perfect equilibrium. In this equilibrium, every player chooses $x^* \in X^*$, irrespective of history. If $X^* \neq \emptyset$, then every player chooses some action from X^* in each period, in all subgame perfect equilibria.*

Proof: We focus on pure strategy profiles. The case of general strategy profiles is dealt with in Remark 2 below. Moreover, we focus on the case $X^* = \{x^*\}$. A minor and straightforward modification of the proof establishes the claim for the case when X^* is non-empty but not necessarily a singleton.

Suppose, thus, that $X^* = \{x^*\}$, and let s^* assign $x_\tau = x^*$ to all periods τ, irrespective of history. Evidently s^* is a subgame perfect equilibrium (SPE), and $\pi_\tau(s^*) = v^*$ for all τ. This establishes existence. For uniqueness, suppose there is a pure SPE $s \neq s^*$. Then there is a period τ and an history $h_\tau \in H_\tau$ at which player τ chooses $x_\tau \neq x^*$. Consider the subgame Γ' which begins at the node in the game tree which corresponds to this h_τ. Let s' be the profile induced by s in Γ'. By the choice of h_τ, this profile has player τ choose a suboptimal action in Γ': $\pi_\tau(s') < v^*$. Since all subgames are "isomorphic" to the original game Γ, by relabeling of players and periods, we may view s' as a profile of the original game Γ. Hence, there exists a subgame perfect profile s' in Γ such that $\pi_0(s') < v^*$.

Suppose furthermore that, given any such suboptimal SPE s', there exists an (even more suboptimal) SPE s'', such that

$$v^* - \pi_0(s'') \geq \frac{1}{1-f_0}(v^* - \pi_0(s')). \tag{4}$$

If this is true, then it follows, by using (4) repetitively, that there are SPE's s'' for which the payoff difference $v^* - \pi_0(s'')$ is arbitrarily large (since $0 < f_0 < 1$). Thus, $u(X)$ is then not bounded from below, contradicting the hypothesis in the proposition. Hence, it suffices to prove that, given any SPE s', there exists a SPE s'' satisfying (4). We deal first with the easier case of exponential discounting.

Exponential discounting (g_t constant): Define $V_1 : X^\infty \to \mathbb{R}$ by shifting back discount factors by one period, $V_1(\bar{x}) = \sum_{t=0}^{\infty} f_{t+1}u(x_t)$. Thus, writing $\bar{x} = (x_0, \bar{y})$, with $\bar{y} \in X^\infty$, we have $U_0(\bar{x}) = f_0 u(x_0) + V_1(\bar{y})$. Under exponential discounting, $V_1(\bar{y}) = (1 - f_0)U_0(\bar{y})$. Let s' be an arbitrary SPE and let s'' be the profile induced by s' in the subgame Γ'' starting in period 1, after player 0 has chosen x^*. Since s' is a SPE, so is s'', and, the fact that s' is a SPE implies[16]

$$\pi_0(s') \geq f_0 v^* + (1 - f_0)\pi_0(s''), \tag{5}$$

an inequality which is equivalent to (4).

Generalized hyperbolic discounting (g_t non-decreasing): Define s'' as in the previous paragraph. By the equilibrium condition for s' we now have

$$\pi_0(s') \geq f_0 u(x^*) + V_1(\bar{y}), \tag{6}$$

where $\bar{y} = \bar{x}^1 = (x_1, x_2, ...)$ is the outcome induced by s''. Recall that $V_1(\bar{y})$ is how player 0 evaluates the sequence of decisions taken in periods $1, 2, ...$

We re-normalize the discount factors that appear in V_1 by setting $\beta = \sum_{t=1}^{\infty} f_t = 1 - f_0$, and writing $\widetilde{U}_1(\bar{x}) = V_1(\bar{x})/\beta$ for all $\bar{x} \in X^\infty$. For $\overline{x} \in X^\infty$ and $\tau > 0$, let $\bar{x}^\tau = (x_\tau, x_{\tau+1}, ...) \in X^\infty$ be the outcome obtained from the choices in periods $\tau, \tau+1, ...$ We first prove that $\widetilde{U}_1(\bar{x}^1)$ is a convex combination of $\{U_0(\bar{x}^\tau) : \tau \geq 1\}$. In other words, the re-normalized utility to self 0 from the choices made by his future selves is a convex combination of their utilities (see also Proposition 2 in Saez-Marti and Weibull (2002)). This is a consequence of (g_t) being non-decreasing.

To see this, set $\alpha_t = f_{t+1}/\beta$, so that $\widetilde{U}_1(\bar{x}^1) = \sum_{t=0}^{\infty} \alpha_t u(x_{t+1})$. Note that $U_0(\bar{x}^\tau) = \sum_{t=0}^{\infty} f_t u(x_{t+\tau})$. Hence it is sufficient to find nonnegative numbers $(\lambda_t)_{t\in\mathbb{N}}$ which sum

[16] As in the previous paragraph, we may also view s' as a profile in the original game Γ.

to 1, such that $\alpha_t = \sum_{k=0}^{t} \lambda_k f_{t-k}$ for each $t \in \mathbb{N}$, the λ_t being the coefficients in the convex combination. This condition on (λ_t) takes the form of a infinite triangular linear system. Since $f_0 \neq 0$, this system has a unique solution. By induction, it can be shown that $\lambda_t \geq 0$ for all t. Assume $\lambda_0, ..., \lambda_{t-1} \geq 0$. To prove that $\lambda_t \geq 0$, one has to prove that

$$\alpha_t \geq \sum_{k=0}^{t-1} \lambda_k f_{t-k}. \tag{7}$$

Note that $\alpha_t / \alpha_{t-1} = f_{t+1}/f_t = g_t$ for all $t > 0$. Since (g_t) is non-decreasing,

$$\frac{f_{t-k}}{f_{t-1-k}} \leq \frac{f_{t+1}}{f_t} = \frac{\alpha_t}{\alpha_{t-1}} \qquad \forall k \leq t-1.$$

Therefore,

$$\frac{\sum_{k=1}^{t-1} \lambda_k f_{t-k}}{\sum_{k=0}^{t-1} \lambda_k f_{t-1-k}} \leq \frac{\alpha_t}{\alpha_{t-1}}. \tag{8}$$

Since the denominators on both sides of (8) coincide, (7) follows. The fact that the sequence $(\lambda_t)_{t\in\mathbb{N}}$ sums up to one is a consequence of the fact that both sequences $(f_t)_{t\in\mathbb{N}}$ and $(\alpha_t)_{t\in\mathbb{N}}$ sum up to one. Indeed,

$$\sum_{t=0}^{\infty} \alpha_t = \sum_{t=0}^{\infty} \sum_{k=0}^{t} \lambda_k f_{t-k} = \left(\sum_{k=0}^{\infty} \lambda_k \right) \left(\sum_{t=0}^{\infty} f_t \right).$$

Having proved that the utility to self 0 is a convex combination of the utilities to his future selves, we note that, by (6),

$$\pi_0(s) \geq f_0 v^* + (1 - f_0) \widetilde{U}_1(\bar{x}^1).$$

It follows from the above that $\widetilde{U}_1(\bar{x}^1) \geq U_0(\bar{x}^\tau)$ for some $\tau \geq 1$. Therefore,

$$v^* - U_0(\bar{x}^\tau) \geq \frac{1}{1 - f_0} \left[v^* - \pi_0(s) \right].$$

Let $\hat{s}_\tau$ denote the profile induced by s'' in the subgame starting at stage τ, after the history $h_\tau = (x_1, ..., x_{\tau-1})$, the path induced by s'' up to period τ. The profile $\hat{s}_\tau$, viewed as a profile in the full game Γ, is a subgame perfect equilibrium. Since $\pi_0(\hat{s}_\tau) = U_0(\bar{x}^\tau)$, this establishes the inequality (4), and thus the claimed uniqueness.

End of proof.

Remark 1: If $X^* = \emptyset$ and $u(X)$ is bounded from below, the dynamic game has no SPE. This can be proved by noting that if s would be a subgame perfect equilibrium strategy, then it would have to satisfy $\pi_0(s) < v^*$. A contradiction can then be obtained along the same lines as in the above proof.

Remark 2: Randomized behavior strategies can be ruled out in the case when X^* is a singleton. In order to establish this claim, the above proof may be amended as follows. Given a general behavior-strategy profile σ, let $\mathbf{P}_\sigma$ denote the induced probability distribution over outcomes $\bar{x} \in X^\infty$, and denote by $\mathbf{E}_\sigma$ the expectation w.r.t. $\mathbf{P}_\sigma$. Arguing as in the proof for pure-strategy profiles, the existence of a SPE $\sigma = (\sigma_\tau)_{\tau\in\mathbb{N}}$ which differs from s^* implies the existence of a SPE $\sigma' = (\sigma'_\tau)_{\tau\in\mathbb{N}}$ such that $\pi_0(\sigma') < v^*$ (σ' is obtained as the profile induced by σ in a specific subgame). Since player 0 then makes suboptimal choices $x \in X$ with positive probability, there is such a choice x_0 such that the profile σ'' induced by σ' after $h_1 = (x_0)$ satisfies

$$v^* - \mathbf{E}_{\sigma''}\left[\widetilde{U}_1(\overline{x}^1)\right] \geq \frac{1}{1-f_0}\left[v^* - \pi_0(\sigma')\right]. \tag{9}$$

Next, we use the notation of the above proof: given an outcome $\overline{x} = (x_t)_{t\in\mathbb{N}} \in X^\infty$ and any $\tau \geq 1$, we let $\overline{x}^\tau$ be the outcome $(x_\tau, x_{\tau+1}, ...)$. Since $\widetilde{U}_1(\overline{x}^1) = \sum_{t=0}^\infty \lambda_t U_0(\overline{x}^{t+1})$, one obtains, by taking expectations,

$$\mathbf{E}_{\sigma''}\left[\widetilde{U}_1(\overline{x}^1)\right] = \sum_{\tau=0}^{\infty} \lambda_\tau \mathbf{E}_{\sigma''}\left[U_0(\overline{x}^{\tau+1})\right].$$

Letting $\mathcal{H}_\tau$ denote the information $\sigma-$algebra of player τ over outcomes, the preceding equality may be re-written as

$$\mathbf{E}_{\sigma''}\left[\widetilde{U}_1(\overline{x}^1)\right] = \sum_{\tau=0}^{\infty} \lambda_\tau \mathbf{E}_{\sigma''}\left[\mathbf{E}_{\sigma''}\left[U_0(\overline{x}^{\tau+1})|\mathcal{H}_{\tau+1}\right]\right].$$

Therefore, there exists a player $\tau \geq 1$ such that $\mathbf{E}_{\sigma''}\left[\widetilde{U}_1(\overline{x}^1)\right] \geq \mathbf{E}_{\sigma''}\left[U_0(\overline{x}^\tau)|\mathcal{H}_\tau\right]$ with positive $\mathbf{P}_{\sigma''}$-probability. Reinterpreting the right-hand side of this inequality, there is an history h_τ such that the profile $\widehat{\sigma}^\tau$ induced by σ'' in the subgame starting at h_τ satisfies

$$\pi_0(\widehat{\sigma}^\tau) \leq \mathbf{E}_{\sigma''}\left[\widetilde{U}_1(\overline{x}^1)\right].$$

By subgame perfection, also $\widehat{\sigma}^\tau$ is subgame perfect. Plugging the last inequality into (9) yields

$$v^* - \pi_0(\widehat{\sigma}^\tau) \geq \frac{1}{1-f_0}\left[v^* - \pi_0(\sigma')\right],$$

which is the exact analog of (4). The contradiction follows, just as in the proof for pure profiles.

References

[1] Ainslie G. W. (1992): *Picoeconomics*, Cambridge University Press (Cambridge UK).

[2] Asheim G. (1997): "Individual and collective time-consistency", *Review of Economic Studies* 64, 427-443.

[3] Asilis C., Kahn C. and D. Mookherjee (1991): "A unified approach to proof equilibria", mimeo. University of Illinois.

[4] Bernheim D., D. Ray and S. Yeltekin (1999): "Self-control, saving, and the low asset trap", mimeo.

[5] Elster J. (1979): *Ulysses and the Sirens. Studies in Rationality and Irrationality.* Cambridge University Press (Cambridge, UK).

[6] Goldman S. (1980): "Consistent plans", *Review of Economic Studies* 47, 533-537.

[7] Herrnstein R. (1981): "Self control and response strength", in *Quantification of Steady-State Operant Behavio*r, Christopher et al (eds.), North-Holland (Amsterdam).

[8] Kocherlakota N. R. (1996): "Reconsideration-proofness: A refinement for infinite-horizon time inconsistency", *Games and Economic Behavior* 15, 33- 54.

[9] Laibson D. (1994): "Self-control and saving", mimeo. Harvard University.

[10] Laibson D. and C. Harris (2001): "Dynamic choices of hyperbolic consumers", *Econometrica* 69, 935-957.

[11] Mazur J.E. (1987): "An adjustment procedure for studying delayed reinforcement", in *The Effect of Delay and Intervening Events on Reinforcement Value*, J. Commons et al (eds.), Erlbaum (NJ).

[12] Phelps E. and R. A. Pollak: "On second-best national saving and game-equilibrium growth", *Review of Economic Studies* 35, 185-199.

[13] Peleg B. and M. E. Yaari (1973): "On the existence of a consistent course of action when tastes are changing", *Review of Economic Studies* 40, 391-401.

[14] Pollak R. A. (1968): "Consistent planning", *Review of Economic Studies* 35, 201-208.

[15] Saez-Marti M. and J. Weibull (2002): "Discounting and future selves", IUI WP 575.

[16] Strotz R. H. (1956): "Myopia and inconsistency in dynamic utility maximization", *Review of Economic Studies* 23, 165-180.