Discussion Paper No. 1128

Individual and Collective Time-Consistency

by Geir B. Asheim* **

June 1995

Abstract

This paper reconsiders the Strotz-Pollak problem of consistent planning and argues that a solution to this problem requires a refinement of subgame-perfectness. Such a refinement is offered through an analysis based on Greenberg's 'theory of social situations'. A unifying framework is presented whereby consistent one-person planning as a problem of individual time-consistency and renegotiation-proofness as a problem of collective time-consistency are captured through the same general concept.

* I thank Nabil Al-Najjar, Joseph Greenberg, Narayana Kocherlakota, Torsten Persson, Gaute Torsvik, Torben Tranæs and Eric van Damme for valuable discussion and correspondence, seminar participants at ESEM'92 and McGill, Oslo and Queen's Universities for helpful comments, and referees for detailed suggestions and constructive criticism. The usual disclaimer applies. This paper was initiated while visiting CentER at Tilburg University and revised while visiting MEDS Department, KGSM at Northwestern University. I gratefully acknowledge the hospitality of these institutions as well as financial support from the Norwegian Research Council.
** Department of Economics, University of Oslo, Box 1095 Blindern, N-0317 Oslo, Norway (Internet: gasheim@econ.uio.no).

1. INTRODUCTION

Consider a single person facing an intertemporal decision problem. At the initial node, the person would like to realize a path that maximizes the person's payoff as evaluated at the initial node. Such a path is said to be *optimal*. Likewise, a decision rule (defined by the property that it at every decision node determines an action) is said to be *optimal* if it generates an optimal path. A optimal decision rule is *time-consistent* if at each decision node reachable by the optimal path, the decision rule is still optimal in the sense of maximizing the person's payoff as evaluated at the reached node. Strotz (1955-56) and Pollak (1968) are concerned with the case where there is no optimal and time-consistent decision rule because the person has time-inconsistent preferences. Time-inconsistency is a real-life phenomenon, sometimes of economic relevance. The following illustrations are included in order to support the claim that seemingly rational individual decision makers may in fact have time-inconsistent preferences.

Procrastination. It is a common experience that people tend to postpone unpleasent tasks, preferring to have them done in the next period (day, week, ...). Yet, when the next period comes along, still further postponement seems preferable. Such "... [p]rocrastination occurs when the present costs are unduly salient in comparison with future costs ..." in the words of Akerlof (1991, p.1), who gives the subject an interesting treatment filled with economically relevant real-life examples. Hence, at any time, the time preference between the present and the first future period is greater than between a future period and its successor. These are the kind of *inconsistent time preferences* which are explicitly analyzed by Strotz (1955-56).

Intoxication. The following situation may also seem realistic: After work, some would prefer to go by the local pub and have *one* beer instead of going straight home. At the pub, *after* the first beer, it may, however, seem preferable to consume another three beers. These preferences are time-inconsistent if, when leaving work, going straight home is preferable to consuming four beers at the pub. Such *endogenous preferences* are treated by e.g. Hammond (1976), who argues that there is no need in a formal analysis to distinguish between inconsistent time preferences (changing exogenously due to the passing of time) and endogenous preferences (changing endogenously due to the person's choice of actions; e.g., the consumption of alcohol and other intoxicating substances). Both types of time-inconsistency are thoroughly reviewed by Elster (1979), whose terminology I have adopted.

Addiction may be considered as a combination of both types of time-inconsistency. When a person becomes addicted to a substance (or unproductive habit), his preferences changes endogenously due to actions taken: Past abuse increases the craving for the substance. When the person is addicted and tries to drop the habit, however, it is a common phenomenon that the person—while acknowledging his health damaging (or unproductive) behavior—wants to indulge for yet another period before quitting. This latter perspective on addiction as procrastination is stressed by Akerlof (1991).

If there is no optimal and time consistent decision rule, Strotz (1955-56) suggests two possibilities: *Precommitment* or *Consistent planning*. Precommitment is a special case of consistent planning if the original decision problem allows for it; else it amounts to changing the decision problem. Precommitment as such will not be discussed here. The problem of consistent planning is according to Strotz (1955-56, p.173) for the person "to find the best plan among those that he will actually follow". If preferences are time-inconsistent, the problem of consistent one-person planning can be analyzed by turning the intertemporal decision problem into an extensive game where the decisions at different nodes are taken by different selves of the decision maker. These selves are agents of the same person, but corresponds formally to separate players. The payoff of an agent depends on his decision node and the path—originating at his node—that the decision rule generates. The strategy of an agent is the action that the decision rule determines at his decision node. Thus, a decision rule corresponds to a strategy profile. Peleg and Yaari (1973) and Goldman (1980) analyze the notion of consistent one-person planning in such a game-theoretic context. They claim that a plan is the best that will actually be followed ("optimal in the Strotz-Pollak sense", Peleg and Yaari (1973, p.345); "a Strotz-Pollak equilibrium", Goldman (1980, p.534)) if and only if it is a SPE of the corresponding extensive game.

As long as the decision problem has a finite time horizon and the agents are not indifferent between any two outcomes, this solution concept uniquely determines the consistent plan. If, however, the time horizon is infinite, or if there are indifferences, then multiple SPEa may exist. Furthermore, as the following examples will make clear, not all of these SPEa are solutions to the problem that Strotz (1955-56) posed.

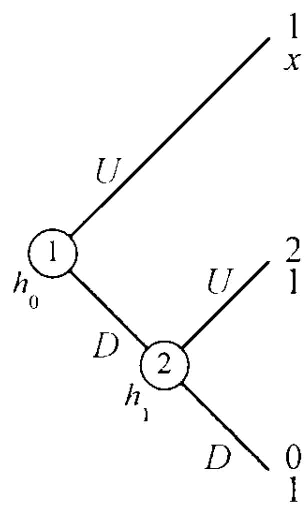

FIGURE 1

EXAMPLE 1 (see Figure 1) considers a person who lives at times 1, 2, and 3. At time 1 he has to decide whether to *U*ndertake or *D*elay an unpleasant task. If he chooses U at time 1, there is no decision at time 2. If he chooses D at time 1, he has to decide whether to *U*ndertake or *D*elay at time 2. With D at time 1, the person is—at time 2—indifferent between U and D. Note that (D,U) is the *unique optimal and time-consistent decision rule* since this is the only decision rule that is optimal at all nodes that is reached if the rule is followed. Using (D,U) enables the person to delay the unpleasant task from period 1 to period 2. However, both (D,U) and (U,D) are SPEa. Hence, there exists a SPE in which the person at the initial node does not choose the best plan among those that he will actually follow. Rather, he receives a lower payoff by undertaking the task immediately. He does so fearing that if he delayed the task at time 1, he would delay the task at time 2 as well and be worse off as evaluated at the initial node.

EXAMPLE 2—which is inspired by an example by Asilis et al. (1991)—has the person choosing $a_t \in \{0,1\}$ at each time $t \in \aleph$ with payoff at time t being given as $min_{\tau \geq t} a_\tau$. Also in this example there exists a *unique optimal and time consistent decision rule*, viz., "always choosing 1". Note, however, that "always choosing 0" is a SPE since no single agent can profit by a unilateral deviation. Then, by using "always choosing 0" as a punishment, any feasible action path can be supported as a SPE path. Hence, even though the problem of consistent planning is trivial and has a unique solution, the concept of a SPE has no bite what-so-ever in this example.

EXAMPLE 3 (presented in detail in Section 3) is an infinite horizon capital accumulation problem where the person has inconsistent time preferences of the kind that Strotz (1955-56) explicitly analyzed. Here, there is no optimal and time-consistent decision rule. However, the decision rules "always maintaining the capital stock intact" and "always consuming the entire capital stock" are both SPEa even though the person prefers the former decision rule to the latter. This is an economically relevant example, where—in contrast to Example 1—the person is not indifferent between two different actions, and where—in constrast to Example 2—actions taken in the distant future do not influence the person's payoff significantly (i.e., there is 'continuity at infinity').

These examples show that the notion of subgame-perfectness is not conceptually valid as a solution to the problem of consistent planning.[1] The reason is the following: The notion of consistent planning is based on the premise that the agents (representing the selves of the person at different decision nodes) are symmetric with respect to their ability to influence later agents. In a SPE, each agent can reconsider only his own action; thus, by the symmetry requirement the agent at the initial node is not the one suggesting or coordinating on a particular SPE. Hence, by playing according to a SPE the person is not at any decision node doing any planning; instead each of his selves chooses a best response given that his other selves play according to an exogenously given decision rule. This discussion—as well as the examples above—suggest that a

---

[1] In a slightly different context, this is also illustrated by the model of Asheim (1988a).

refinement of subgame-perfectness is required in order to conceptually solve the problem of consistent one-person planning. Such a refinement will be offered in Section 2 through a concept which is based on the following premise: *The person can—given that one of his decision nodes has been reached—select any decision rule for the remaining problem, taking into account that he can reselect at any later decision node.* This concept of a *revision-proof decision rule* respects the symmetry between the different selves of the person, while allowing him to engage in planning by choosing "the best plan among those that he will actually follow". It will be shown that the suggested concept solves the examples above in an attractive manner. In particular, Section 3 analyzes its implications in the economically relevant Example 3.

The subsequent sections are organized as follows. In Section 4 it is pointed out that the definition of a revision-proof decision rule may be considered as a special case of Bernheim et al.'s (1987) definition of a Perfectly coalition-proof equilibrium, extended to an infinite time horizon. This in turn permits relating consistent one-person planning as a problem of individual time-consistency to renegotiation as a problem of collective time-consistency. *Stationary* one-person decision problems allow for the adaptation of concepts of renegotiation-proofness in repeated games. Kocherlakota (1995) exploits this to suggest an alternative refinement which he terms "reconsideration-proofness". An evaluation of his concept is contained in Section 5. For this purpose, the time-consistency of a pay-as-you-go pension system is considered as Example 4.

The present paper entails that it matters for the analysis whether the agents of a game are interpreted as different selves of the same person, or as different persons. Section 6 addresses this question. Furthermore, the present definition of a revision-proof decision rule is based on Greenberg's (1990) theory of social situations. In Section 6 it is pointed out that the concept of a revision-proof decision rule is related to Greenberg's (1990) notion of optimistic stability in the tree situation. Finally, the assumptions underlying the definition are chosen in order to keep the analysis simple and thereby make the conceptual arguments more transparent. Section 6 contains a brief discussion of possible extensions. All proofs are relegated to Section 7.

2. REVISION-PROOF DECISION RULES

Consider a one-person T-stage decision problem. The sets of *histories* are defined inductively: The set of histories at the beginning of the first stage 1 is $H_0 = \{h_0\}$. Let H_{t-1} denote the set of histories at the beginning of stage t. At $h \in H_{t-1}$, the person's action set is A_i^h. Define the set of histories at the beginning of stage t+1 as follows: $H_t := \{(h,a_i)|h \in H_{t-1}$ and $a_i \in A_i^h\}$. This concludes the induction. Then $H := \bigcup_{t=0}^{T-1} H_t$ is set of *decision nodes*, and (if $T < \infty$) H_T is the set of *terminal nodes*. Assume that, $\forall h \in H$, A_i^h is non-empty, but possibly singular. The latter case means that a trivial decision is taken at h. The person's *decision rule* s_i determines for each $h \in H$ an action $s_i(h)$ in A_i^h, with S_i being the set of decision rules. Given h, s_i determines a path of decision nodes, $H^h(s) := \{g \in H|\ g$ is reached given h and $s\}$, originating at h. Write $H^h := \bigcup_{s_i \in S_i} H^h(s_i)$. The person's *payoff* given h and s_i equals $u_i^h(s_i)$, where $u_i^h(s_i)$ depends only on h and the properties of s_i on $H^h(s_i)$. If payoff functions can be normalized such that, $\forall h \in H$ and $\forall s_i \in S_i$, $u_i^h(s_i) = u_i^{(h,s_i(h))}(s_i)$, then the person's preferences are said to be *time-consistent*.[2] A decision rule s_i is *optimal* at h if $s_i \in \arg\max_{r_i \in S_i} u_i^h(r_i)$. A decision rule s_i is *optimal* and *time-consistent* at h if, $\forall g \in H^h(s_i)$, $s_i \in \arg\max_{r_i \in S_i} u_i^g(r_i)$. If preferences are not time-consistent, an optimal and time-consistent decision rule may not exist.

Assume that, at each decision node, the person can select any decision rule. However, he takes into account that he may revise the decision rule at the current or any later decision node. In order to capture this *one-person planning situation*, let a *standard of behavior* (SB) be a correspondence σ_i assigning to each decision node $h \in H$ a subset $\sigma_i(h)$ of S_i. A SB σ_i in the one-person planning situation is *optimistic internally stable* if

(IS) $\forall h \in H$, $\forall s_i \in \sigma_i(h)$, there do not exist $g \in H^h$ and $r_i \in \sigma_i(g)$ such that $u_i^g(r_i) > u_i^g(s_i)$.

A SB σ_i in the one-person planning situation is *optimistic externally stable* if

(ES) $\forall h \in H$, $\forall s_i \in S \backslash \sigma_i(h)$, there exist $g \in H^h$ and $r_i \in \sigma_i(g)$ such that $u_i^g(r_i) > u_i^g(s_i)$.

---

[2] See Blackorby et al. (1973, Theorem 3).

A SB σ_i in the one-person planning situation is said to be *optimistic stable* if it is both optimistic internally and externally stable.[3] Consistent one-person planning can now be defined.

DEFINITION 1. A decision rule s_i is *revision-proof* at h if there exists an optimistic stable SB σ_i in the one-person planning situation, with $s_i \in \sigma_i(h)$.

The rationale underlying this definition is the following: Interpret $\sigma_i(h)$ as the set of plans that will be followed at h. (IS) means that if a plan is followed, then there does not exist, now or later, a strictly preferred plan that will be followed. This explains why a plan in $\sigma_i(h)$ will be followed. Conversely, (ES) means that if a plan is not followed, then there exists, now or later, a strictly preferred plan that will be followed. This explains why a plan outside $\sigma_i(h)$ will not be followed.

EXAMPLE 1 *(Continued)*. As indicated in Figure 1, refer to the decision nodes where non-trivial decisions are made as h_0 and h_1, with $S_i = \{(U,U),(U,D),(D,U),(D,D)\}$. Let σ_i be any optimistic stable SB. Then $\sigma_i(h_1) = S_i$ by (ES) since no domination can occur at or subsequent to h_1. Furthermore, $(D,U) \in \sigma_i(h_0)$ since otherwise (ES) would have been violated, while $\{(U,U),(U,D),(D,D)\} \subseteq S_i \backslash \sigma_i(h_0)$ since otherwise (IS) would have been violated. Note that σ_i so determined satisfies (IS) and (ES). By Definition 1, (D,U) is the unique revision-proof decision rule at h_0.

This analysis shows that in Example 1, Definition 1 picks out the unique optimal and time-consistent decision rule. The following proposition establishes general results on the relation between revision-proofness and the set of optimal and time-consistent plans. Part (i) shows that if an optimal and time-consistent decision rule exists, then a revision-proof decision rule gives the person the same payoff. By part (ii), "always choosing 1" is the only revision-proof decision rule in Example 2. Hence, also in this example does Definition 1 yield the satisfactory solution.

---

[3] This is an application of the general notion of optimistic stability, which is closely related to von Neumann and Morgenstern (1953) stability, and which is a central solution concept in Greenberg's (1990) *theory of social situation*. The term 'optimistic stability'—coined by Greenberg (1990)—refers here to the optimistic attitude of the person at h in the sense of believing that he can select any viable decision rule at h. Greenberg's (1990) alternative notion of *conservative* stability, which does not correspond to vN-M stability, will not be treated here.

PROPOSITION 1. *(i) If s_i is revision-proof at h and r_i is optimal and time-consistent at h, then* $u_i^h(s_i) = u_i^h(r_i)$. *(ii) If, $\forall h \in H$, there exists a non-empty set of decision rules that are optimal and time-consistent at h, then s_i is revision-proof at h iff s_i is optimal and time-consistent at $\forall g \in H^h$.*

Part (i) of the following proposition shows that a revision-proof decision rule exists at a decision node iff there exists an optimistic stable SB in the one-person planning situation. Hence, part (ii) establishes existence, given a finite horizon and a restriction to finite action.

PROPOSITION 2. *(i) If there exists an optimistic stable SB σ_i in the one-person planning situation, then, $\forall h \in H$, $\sigma_i(h) \neq \emptyset$. (ii) Assume that $T < \infty$, and that, $\forall h \in H$, A_i^h is finite. Then there exists a unique optimistic stable SB σ_i in the one-person planning situation.*

Even though—under the assumptions of part (ii)—there exists a unique optimistic stable SB, there may still exist multiple revision-proof decision rules at some $h \in H$. Example 1 illustrates this possibility since $\sigma_i(h_1) = S_i$. Under different assumptions, the existence of a unique optimistic SB cannot generally be established: With infinite and compact action, there are finite horizon examples where no optimistic stable SB exists (see Hellwig & Leininger (1987; Section III); this example is included in Asheim (1991b, Example 6)). With finite action, but infinite horizon, no counter example to existence is available; in this case, there are examples (see Example 4 of Section 5) with multiple optimistic SBs.

Turn the one-person T-stage decision problem into a perfect information game Γ, where the decision maker is regarded as having different selves at different decision nodes. These selves are agents of the same person, but have different preferences and are formally separate players. I.e., H is the set of players, and, $\forall h \in H$, $s_i(h)$ is player h's strategy, and $u_i^h(\cdot)$ is player h's payoff function. Then s_i is a *multi-self subgame-perfect equilibrium* (SPE) of Γ at h, if there do not exist $g \in H^h$ and $r_i(g) \in A_i^g$ such that $u_i^g(r_i(g), s_i \setminus s_i(g)) > u_i^g(s_i)$. The concept of a revision-proof decision rule is in general a refinement of the concept of a multi-self SPE of Γ.

PROPOSITION 3. *If the decision rule s_i is revision-proof at h, then s_i is a multi-self SPE of Γ at h.*

As Examples 1 and 2 of the introduction as well as Example 3 of the following section illustrate, this refinement can be strict.

Under the condition of the following proposition, the concepts coincide.

PROPOSITION 4. *If, $\forall h \in H$, (a) there exists a multi-self SPE of Γ at h, and (b) s_i' and s_i'' being SPEa of Γ at h implies $s_i' = s_i''$ on H^h, then s_i is a revision-proof decision rule at h iff s_i is a multi-self SPE of Γ at h.*

3. AN EXAMPLE OF INCONSISTENT TIME PREFERENCES

In this section, Definition 1 is applied to an economically relevant example of inconsistent time preferences, where the concept of revision-proofness yields a strict refinement of the concept of a multi-self SPE of Γ. For the example, the following definitions are useful. Let the *state space* K be a partition of H such that, with $h',h'' \in H$ and $k \in K$, $h',h'' \in k$ iff the problem faced at h' is isomorphic to the problem faced at h''. Say that a decision rule s_i is *Markovian* if $s_i(h') = s_i(h'')$ whenever $h',h'' \in k$.

EXAMPLE 3 *(Continued)*. Consider an infinite horizon capital accumulation problem where the capital stock can take values in $K := \{0, k^1, \ldots, k^{m-1}, k^m\}$, with $0 < k^1 < \ldots < k^{m-1} < k^m$. The discreteness of K reflects indivisible real capital or—in the case of financial capital—a smallest money unit. At each stage t, the person inherits a capital stock k_{t-1}, and his action is to choose a capital stock k_t which will be inherited at the next stage t+1. Let $A^h = \{k_t \in K | \, k_t \leq k_{t-1}/\rho\}$ if $h \in k_{t-1}$. Interpret $c_t := k_{t-1} - \rho k_t$ as consumption at stage t and $1/\rho$ as the gross rate of return. The person derives payoff from consumption. If the decision rule s_i generates the consumption path $(c_t, c_{t+1}, \ldots)$ given $h \in k_{t-1}$, then $u_i^h(s_i) = v(c_t) + \lambda \sum_{\tau=t+1}^{\infty} \delta^{\tau-t-1} v(c_\tau)$. Here, $v(c_\tau)$ is the utility derived from consumption in stage τ, with $v(\cdot)$ being strictly increasing and strictly concave. Normalize $v(\cdot)$ such that $v(0) = 0$. Assume inconsistent time preferences: λ is the discount factor between this stage and the next, and δ is the discount factor between a future stage and its successor, with $0 < \lambda < \delta < 1$. Also, assume that, $\forall k \in K$, $v(k^m) > v(k^m - \rho k) + \lambda v(k)$. This requires $\lambda < \rho$ and implies the person will choose to consume the entire capital stock at stage t if any remaining capital stock will be entirely consumed at stage t+1. Finally, assume that $(1 + \lambda + \lambda\delta + \lambda\delta^2 + \ldots =)\ 1 + \frac{\lambda}{1-\delta} = \frac{1}{1-\rho}\ (= 1 + \rho + \rho^2 + \rho^3 + \ldots)$. This requires that $\lambda < \rho < \delta$ (e.g., $\lambda = .4$, $\rho = .5$, and $\delta = .6$) and implies that the person will choose to maintain the capital stock intact at stage t $(k_t = k_{t-1})$ if the capital stock will be maintained intact at each future stage $(\forall \tau > t,\ k_\tau = k_{\tau-1})$.

Let $\widehat{s}_i$ denote the Markovian decision rule defined by, $\forall h \in H$, $\widehat{s}_i(h) = 0$ ("always consuming the entire capital stock"). If $h \in k$, then $u_t^h(\widehat{s}_i) = v(k)$. Let $\check{s}_i$ denote the Markovian decision rule defined by, $\forall h \in H$, $\check{s}_i(h) = k$ if $h \in k$ ("always maintaining the capital stock intact"). If $h \in k$, then $u_t^h(\check{s}_i) = (1 + \frac{\lambda}{1-\delta})v((1-\rho)k)$. Since $v(0) = 0$ and $1 + \frac{\lambda}{1-\delta} = \frac{1}{1-\rho}$, it follows from the strict concavity of $v(\cdot)$ that $u_t^h(\widehat{s}_i) < u_t^h(\check{s}_i)$ if $h \in k$ with $k > 0$. Hence, if $h \in k$ with $k > 0$, *the person strictly prefers* $\check{s}_i$ *to* $\widehat{s}_i$. Note that neither of these decision rules are optimal and time-consistent: If the grid K is sufficiently fine, at any stage t, the person's optimal, but time-inconsistent, decision rule would yield a consumption 'binge' at the present, followed by positive capital accumulation in the future.

Since the person strictly prefers $\check{s}_i$ to $\widehat{s}_i$, it would seem that consistent planning should lead the person to select $\check{s}_i$ rather than $\widehat{s}_i$. The following claim establishes that the concept of revision-proofness supports this intuition: While both $\widehat{s}_i$ and $\check{s}_i$ are multi-self SPEa of Γ, only $\check{s}_i$ is a revision-proof decision rule.

CLAIM 1. *Consider Example 3. (i)* $\forall h \in H$, *the Markovian decision rules* $\widehat{s}_i$ *("always consuming the entire capital stock") and* $\check{s}_i$ *("always maintaining the capital stock intact") are both multi-self SPEa of* Γ *at h. (ii)* $\forall h \in H$, *the Markovian decision rule* $\check{s}_i$ *("always maintaining the capital stock intact") is revision-proof at h. (iii) If a Markovian decision rule is revision-proof at h, then* $s_i = \check{s}_i$ *on* H^h.

4. PERFECTLY COALITION-PROOF EQUILIBRIA

Consider a n-person T-stage decision problem. The sets of *histories* are defined inductively: The set of histories at the beginning of the first stage 1 is $H_0 = \{h_0\}$. Let H_{t-1} denote the set of histories at the beginning of stage t. At $h \in H_{t-1}$, each person's action set is A_i^h, with $i \in N := \{1, \ldots, n\}$. Write $A^h := A_1^h \times \cdots \times A_n^h$. Define the set of histories at the beginning of stage t+1 as follows: $H_t := \{(h,a)|h \in H_{t-1}$ and $a \in A^h\}$. This concludes the induction. Then $H := \bigcup_{t=0}^{T-1} H_t$ is the set of *subgames* and (if $T < \infty$) H_T is the set of *terminal nodes*. Person i's *strategy* s_i determines for each $h \in H$ an action $s_i(h)$ in A_i^h, with S_i being the set of strategies. Write, $\forall J \subseteq N$, $S := S_1 \times \ldots \times S_n = S_J \times S_{-J}$. Given h, $s \in S$ determines $H^h(s) := \{g \in H|\ g$ is reached given h and $s\}$. Write $H^h := \bigcup_{s \in S} H^h(s)$. Person i's *payoff* given h and s equals $u_i^h(s)$, where $u_i^h(s)$ depends only on h and the properties of s on $H^h(s)$. If payoff functions can be normalized such that, $\forall h \in H$ and $\forall s \in S$, $u_i^h(s) = u_i^{(h,s(h))}(s)$, then person i's preferences are said to *time-consistent*.

Assume that, in each subgame, any coalition can coordinate on any strategy profile, taking into account that, in the current and each later subgame, any weakly included coalition can do so in turn. In order to analyze this n-person decision problem, consider the following *perfectly coalition-proof situation*: A *standard of behavior* (SB) is a vector of correspondences σ assigning to each coalition $(\emptyset\neq)\ J \subseteq N$ and each subgame $h \in H$ a subset $\sigma_J(h)$ of S. A SB σ in the perfectly coalition-proof situation is *optimistic internally stable* if

(IS) $\forall (\emptyset\neq)\ J \subseteq N$, $\forall h \in H$, $\forall s=(s_J, s_{-J}) \in \sigma_J(h)$, there do not exist $(\emptyset\neq)\ I \subseteq J$, $g \in H^h$, and $r=(r_I, s_{-I}) \in \sigma_I(g)$ such that, $\forall i \in I$, $u_i^g(r) > u_i^g(s)$.

A SB σ in the perfectly coalition-proof situation is *optimistic externally stable* if

(ES) $\forall (\emptyset\neq)\ J \subseteq N$, $\forall h \in H$, $\forall s=(s_J, s_{-J}) \in S \backslash \sigma_J(h)$, there exist $(\emptyset\neq)\ I \subseteq J$, $g \in H^h$, and $r=(r_I, s_{-I}) \in \sigma_I(g)$ such that, $\forall i \in I$, $u_i^g(r) > u_i^g(s)$.

A SB σ in the perfectly coalition-proof situation is said to be *optimistic stable* if it is both optimistic internally and externally stable. Note that if $N = \{i\}$, the perfectly coalition-proof situation reduces to the one-person planning situation.

With time-consistent preferences and a finite horizon, Bernheim et al. (1987) define a *perfectly coalition-proof equilibrium* inductively as follows: Let ρ be a vector of correspondences assigning to each coalition ($\emptyset \neq$) $J \subseteq N$ and each subgame $h \in H$ a subset $\rho_J(h)$ of S. Assume, $\forall (I, g) \neq (J, h)$ with $\emptyset \neq I \subseteq J$ and $g \in H^h$, $\rho_I(g)$ has been defined. Let $\overline{\rho}_J(h) := \{s \in S|$ there does not exist $(I, g) \neq (J, h)$ with $\emptyset \neq I \subseteq J$ and $g \in H^h$ such that $s \in S \backslash \rho_I(g)\}$.[4] Then $\rho_J(h)$ is defined by $s \in \rho_J(h)$ iff $s \in \overline{\rho}_J(h)$ and there does not exist $r=(r_J, s_{-J}) \in \overline{\rho}_J(h)$ such that, $\forall i \in J$, $u_i^h(r) > u_i^h(s)$. A strategy profile s is a *perfectly coalition-proof equilibrium* at h if $s \in \rho_N(h)$. Note that the assumption of time-consistent preferences plays no role in this definition. Hence, it can in a straight-forward manner be generalized to (possibly) time-inconsistent preferences. Using the perfectly coalition-proof situation, the following proposition characterizes Bernheim et al.'s (1987) concept given a restriction to finite action.

PROPOSITION 5.[5] *Assume that $T < \infty$, and that, $\forall h \in H$, A^h is finite. Then there exists a unique optimistic stable SB σ in the perfectly coalition-proof situation. Furthermore, a strategy profile s is a perfectly coalition-proof equilibrium at h iff $s \in \sigma_N(h)$.*

Using this characterization, Bernheim et al.'s (1987) concept can be extended to infinite horizon decision problems.

DEFINITION 2. A strategy profile s is a *perfectly coalition-proof equilibrium* at h if there exists an optimistic stable SB σ in the perfectly coalition-proof situation, with $s \in \sigma_N(h)$.

Note that it follows from Definition 2 that a perfectly coalition-proof equilibrium is a revision-proof decision rule in the special case where $N = \{i\}$. Hence, *the extended concept of a perfectly coalition-proof equilibrium yields the definition of one-person consistent planning offered in Definition 1 of this paper.* Bernheim et al. (1987) do not, however, discuss time-inconsistent preferences and do not consider their concept as a solution to a one-person decision problem.

[4] Note that $\overline{\rho}_J(h) = S$ if $J = \{i\}$ and $h \in H_{T-1}$.

[5] This characterization was noted by Asheim (1988b). It is discussed by Asilis and Kahn (1992).

In a repeated game, a SPE can be supported by a threat which—if called—is Pareto-inferior to the original SPE. Hence, it the players can coordinate before each stage of the game, they prefer renegotiating back to the original SPE rather than undertaking the threat. However, this undermines the credibility of the threat and questions the viability of the original SPE. The literature on renegotiation-proof equilibria[6] seeks to answer the following question: What SPEa are not prone to this kind of criticism? A SPE that is not renegotiation-proof may be considered as a strategy profile that is not collectively time-consistent (see e.g. Bernheim and Ray, 1989): There are subgames where the grand coalition as a collective gains by revising the strategy profile. In these terms there is an obvious similarity between the notion of consistent one-person planning as a problem of individual time-consistency and the notion of renegotiation-proofness as a problem of collective time-consistency. This similarity is brought forward below by noting that *the extended concept of a perfectly coalition-proof equilibrium yields also a definition of renegotiation-proofness*: Let the n-person T-stage decision problem be a repeated game with n = 2. Then, with T finite, Bernheim et al.'s (1987) perfectly coalition-proof equilibrium coincides with Bernheim and Ray's (1985, 1989) *Pareto-perfect* (or *Consistent*) *equilibrium*. Furthermore, with T finite or infinite, a perfectly coalition-proof equilibrium as defined in Definition 2 coincides with a Pareto-perfect equilibrium as defined in Definition 1 of Asheim (1991a) (provided that $\forall i \in N$, $\forall h \in H$, $\forall s_{-i} \in S_{-i}$, $\arg\max_{s_i \in S_i} u_i^h(s_i, s_{-i}) \neq \varnothing$). Hence the concept of a perfectly coalition-proof equilibrium yields the usual and uncontroversial definition of renegotiation-proofness in finitely repeated games as well as the extension that I have suggested for infinitely repeated games.

The reason why problems of individual and collective time-consistency are related is that both types of problems arise because the person (in consistent one-person planning) or the grand coalition (in a repeated game) is at any stage able to select a decision rule or a SPE. Taking into account the opportunity for future selection, the individual decision maker or the grand coalition is forced to select a decision rule or a SPE that will actually be followed at later stages.

[6] See e.g. Farrell and Maskin (1989), Bernheim and Ray (1989), Asheim (1991a) as well as Abreu et al. (1993) and Bergin and MacLeod (1993).

5. STATIONARY DECISION PROBLEMS

An infinite horizon decision problem is stationary if, $\forall h', h'' \in H$, the problem faced at h' is isomorphic to the problem faced at h''. An infinitely repeated game is a stationary decision problem. Given the relation between renegotiation-proofness and consistent one-person planning, concepts of renegotiation-proofness as developed for infinitely repeated games can be applied to stationary one-person decision problems. Kocherlakota (1995) successfully applies Farrell and Maskin's (1989) concept of *weakly renegotiation-proof* and *strongly renegotiation-proof* equilibria, the analogs being called *symmetric* and *reconsideration-proof* equilibria, respectively. By normalizing the payoff-function such that $u_i^h(s_i)$ does not depend on h but only on the actions taken along $H^h(s_i)$, a symmetric equilibrium is defined as a multi-self SPE of Γ yielding the person the same payoff at any decision node. Kocherlakota (1995) considers such symmetry a necessary condition for consistent planning.[7] A symmetric equilibrium is reconsideration-proof if there is no symmetric equilibrium yielding the person a higher payoff, a concept for which Kocherlakota (1995) establishes general existence. Kocherlakota (1995) also shows that if an optimal and time-consistent decision rule exists, then this rule is reconsideration-proof, a property that—by Proposition 1(ii)—is shared by revision-proofness. The following proposition gives a result on the relation between revision-proof decision rules and reconsideration-proof equilibria.

PROPOSITION 6. (Kocherlakota, private communication). *In a stationary one-person decision problem, a revision-proof decision rule is a reconsideration-proof equilibrium iff it is symmetric.*

However, there are examples where revision-proof decision rules are not symmetric. In fact, Example 4 below illustrates the case where the set of revision-proof decision rules and the set of reconsideration-proof equilibria are disjoint.

To evaluate the concept of reconsideration-proofness, two issues must be addressed:

---

[7] In a slightly different context, this view is supported by Dekel and Farrell (1990).

- What is the economic relevance of stationary one-person decision problems with time-inconsistent preferences?
- Given a stationary one-person decision problem with time-inconsistent preferences, what is the appropriate solution concept?

Section 1 lists three empirically relevant examples of time-inconsistent preferences: Procrastination, intoxication, and addiction. These examples have in common that the problems faced at different decision nodes are *not* isomorphic. The problem facing a person that still needs to undertake a task is different from the problem facing a person who has already done the task. The problem facing an intoxicated person is not the same as the problem facing a sober person. The problem facing a person addicted to some substance or unproductive habit, is not the same as the problem facing a person that is not addicted. Even though these examples are far from exhaustive, it seems restrictive to analyze consistent planning in a stationary environment.[8]

For the second issue it is instructive to consider the following example of a stationary one-person decision problem. Actually, since the stationarity makes it hard to interpret the example as a *one-person* problem in a literal sense, I choose to present it as a decision problem of a *dynasty*.

EXAMPLE 4. Let a dynasty be described by an overlapping generations model where—leaving reproductive issues aside—there at each stage are two individuals: an agent of the dynasty and his parent. The agent can choose whether to pay a pension to the parent ($a = 1$) or not ($a = 0$). The parent does not act. Hence, $\forall h \in H$, $A^h = \{0, 1\}$. The cost of paying a pension is 1; the benefit of receiving a pension is 2. Hence, $\forall h \in H$ and $\forall s_t \in S_t$, $u_t^h(s_t) = -s_t(h) + 2s_t(h, s_t(h))$. At each stage, the agent of the dynasty can select any decision rule to be followed by the present and future agents of the dynasty, taking into account that future agents can do so in turn.

[8] As Example 4 of Kocherlakota (1995) illustrates, a reconsideration-proof equilibrium need not be a Markovian decision rule. Hence, it is not possible to generalize the concept of a reconsideration-proof equilibrium to non-stationary one-person decision problems by imposing that the decision rules be Markovian, and then choosing a most preferred Markovian decision rule.

Here, the unique symmetric equilibrium is $\hat{s}_i$ defined by, $\forall h \in H$, $\hat{s}_i(h) = 0$. Hence, $\hat{s}_i$ ("never paying a pension") is the unique reconsideration-proof equilibrium, yielding, $\forall h \in H$, $u_i^h(\hat{s}_i) = 0$. However, following Hammond (1975) and Kotlikoff et al. (1988), it is straightforward to construct decision rules where a pay-as-you-go pension system is established. For this purpose, determine inductively a partition of H as follows: Let it be given whether the agent at h_0 has a 'cooperative' or a 'non-cooperative' parent. Consider the agent at $h \in H$. If his parent was 'cooperative', then choosing 1 is 'cooperative' and choosing 0 is 'non-cooperative'. If his parent was 'non-cooperative', then choosing 0 is 'cooperative' and choosing 1 is 'non-cooperative'. Let s_i^1 (s_i^0) be the decision rule "always choosing cooperatively" given that the agent at h_0 has a 'cooperative' ('non-cooperative') parent. Note that, with $a = 1$ or $a = 0$, $u_i^h(s_i^a) = 1$ (= 2) if the agent at h has a 'cooperative' ('non-cooperative') parent. Hence, $\forall h \in H$, s_i^a yields a higher payoff than the unique reconsideration-proof equilibrium $\hat{s}_i$. The following claim establishes that s_i^1 and s_i^0 are revision-proof, while $\hat{s}_i$ is not.

CLAIM 2. *Consider Example 4. (i) $\forall h \in H$, s_i^1 and s_i^0 are revision-proof at h. (ii) If s_i is a reconsideration-proof equilibrium, then, $\forall h \in H$, s_i is not revision-proof at h.*

The problem with the concept of reconsideration-proofness in the context of this example is that the imposition of symmetry as a necessary condition implies that plans yielding all generations a higher payoff in any contingency are not considered. In particular, the concept does not comply with the view that a plan will not be followed only if there exists a strictly preferred plan that will be followed.

Finally, note that the proof of Claim 2 shows that in this example there are multiple optimistic stable SBs in the one-person planning situation.

6. CONCLUDING REMARKS

The possibility of communication is a distinctive feature of a one-person decision problem: With perfect recall, any agent of the person is able to communicate to later agents of the person, while the converse is not true.[9] This enables each agent to select a decision rule for the remaining decision problem. Without communication or with bi-directional communication, the refinement of Definition 1 need not be applicable. This can be seen by considering Example 1 as a two-person game with $x = 2$. Then, in line with the argument of Bennett and van Damme (1991) and Tranæs (1995) (see also Nalebuff and Shubik (1988)), it seems natural to expect that 2 will communicate his intention to punish 1 for not playing U, by himself playing D. If 1 believes that this is a credible threat, she will indeed be induced to choose U. Furthermore, without communication this still seems like a plausible outcome. Even with communication from 1 to 2 only, 1 cannot rule out that 2, if asked to play, will behave out of vengeance. Hence, Definition 1 relies on uni-directional communication as well as no agent behaving out of vengeance. This seems to be appropriate in the case of a one-person decision problem. It also implies that Definition 1 appears to be applicable to any entity that exists over time and where agents of the entity do not act vengefully. Examples here include dynasties (like Example 4 above) and organizations (e.g. firms). It is less applicable to a government in a democratic society where the different governing parties are antagonistic.

The definition of a revision-proof decision rule is based on the general framework of Greenberg's (1990) 'theory of social situations'. As I have explored in Asheim (1991b, Proposition 1), his specific concept of optimistic stability in the tree situation is formally closely related to the present concept of revision-proofness. However, Greenberg (1990) does not interpret optimistic stability in the tree situation in terms of consistent planning. In fact, it is hard to give an interpretation that directly corresponds to the one-person problem of finding "the best

[9] For an analysis of one-person decision problems without perfect recall, see Piccione and Rubinstein (1994).

plan among those that he will actually follow" (Strotz, 1955-56, p.173).[10] Furthermore, the concept of optimistic stability in the tree situation relies on paths, not decision rules. This makes it hard to extend the analysis to incorporate exogenous uncertainty. The present analysis can in a straight-forward manner be extended to exogenous uncertainty by introducing moves by nature.

Even though Proposition 2 ensures that a revision-proof decision rule exists if there exists an optimistic stable SB in the one-person planning situation, existence of the latter is established only under restrictive assumptions. In order to ensure existence, Asilis and Kahn (1992) argue in general that Roth (1976) *semi-stability* (or 'subsolution') should be substituted for von Neumann and Morgenstern (1953) stability. In the present specific context, this amounts to dividing—at each decision node—the set of decision rules into a *good*, a *bad*, and an *ugly* set. A decision rule is good if and only if there does not exist a good or ugly decision rule at the current or any later decision node that is strictly preferred at that node. A decision rule is bad if and only if there exists a good decision rule at the current or some later decision node that is strictly preferred at that node. This definition differs from the one given here by allowing there to be decision rules that are neither good nor bad; rather, they are ugly. Although it is conceptually hard to interpret the significance of a non-empty ugly set, it may still—if necessary—be worthwhile to pursue this variant of Definition 1 in applications.

---

[10] In the tree situation, an 'optimistic' agent acts at his node and suggests a path to be followed by later agents.

7. PROOFS

Since the one-person planning situation is a special case of the perfectly coalition-proof situation, this section first provides a useful observation for the latter as well as a proof of Proposition 5.

LEMMA 1. *Let* σ *be an optimistic stable SB in the perfectly coalition-proof situation. Then* $\sigma_I(g) \supseteq \sigma_J(g)$ *if* $\emptyset \neq I \subseteq J$ *and* $g \in H^h$.

Proof. Suppose $s \notin \sigma_I(g)$ for some $(\emptyset{\neq})\ I \subseteq J$ and $g \in H^h$. By (ES) of σ there exist $(\emptyset{\neq})\ I' \subseteq I$, $g' \in H^g$, and $r=(r_{I'},s_{-I'}) \in \sigma_{I'}(g')$ such that, $\forall i{\in}I'$, $u_i^{g'}(r) > u_i^{g'}(s)$. By (IS), $s \notin \sigma_J(h)$ since $(\emptyset{\neq})\ I' \subseteq J$ and $g' \in H^h$. □

Proof of Prop. 5. Let σ be any optimistic SB. Let $(\emptyset{\neq})\ J \subseteq N$ and $h \in H$. Assume, $\forall (I, g) \neq (J, h)$ with $\emptyset \neq I \subseteq J$ and $g \in H^h$, (i) $\rho_I(g) = \sigma_I(g)$ and (ii) $s \in \rho_I(g)$ iff there do not exist $(\emptyset{\neq})\ I' \subseteq I$, $g' \in H^g$, and $r=(r_{I'},s_{-I'}) \in \sigma_{I'}(g')$ such that, $\forall i{\in}I'$, $u_i^{g'}(r) > u_i^{g'}(s)$. By Lemma 1, $\overline{\rho}_J(h) \supseteq \sigma_J(h)$. If $s \in \rho_J(h)$, then — by the definition of ρ — there does not exist $r=(r_J,s_{-J}) \in \overline{\rho}_J(h) \supseteq \sigma_J(h)$ such that, $\forall i{\in}J$, $u_i^h(r) > u_i^h(s)$, implying by (ES) of σ that $\rho_J(h) \subseteq \sigma_J(h)$. Conversely, if $s \in \overline{\rho}_J(h)\backslash\rho_J(h)$, then — by the definition of ρ and the finiteness of $\overline{\rho}_J(h)$ — there exists $r=(r_J,s_{-J}) \in \rho_J(h) \subseteq \sigma_J(h)$ such that, $\forall i{\in}J$, $u_i^h(r) > u_i^h(s)$, implying by (IS) of σ that $\rho_J(h) \supseteq \sigma_J(h)$. By induction, it follows that ρ is an optimistic stable SB, and if σ is any optimistic stable SB in the perfectly coalition-proof situation, then $\rho = \sigma$. □

Proof of Prop. 1. (i) Since s_i is revision-proof at h, there exists an optimistic stable SB σ_i with $s_i \in \sigma_i(h)$. Consider q_i defined by $q_i = r_i$ on $H^h(r_i)$ and $q_i = s_i$ on $H\backslash H^h(r_i)$. By (ES), $q_i \in \sigma_i(h)$ since r_i is optimal and time-consistent and $s_i \in \sigma_i(h)$. By (IS), $u_i^h(s_i) = u_i^h(q_i) = u_i^h(r_i)$. *(ii)* Define the SB ρ_i by, $\forall h{\in}H$, $\rho_i(h) := \{s_i \in S_i | \forall g{\in}H^h,\ s_i \text{ is optimal and time-consistent at } g\}$. By assumption, $\forall h{\in}H$, $\rho_i(h) \neq \emptyset$. Let σ_i be any optimistic stable SB. Then, $\forall h{\in}H$, $\forall s_i \in \rho_i(h)$, there do not exist $g \in H^h$ and $r_i \in S_i \supseteq \sigma_i(g)$ such that $u_i^g(r_i) > u_i^g(s_i)$. Hence, ρ_i satisfies (IS), and by (ES) of σ_i, $\forall h \in H$, $\rho_i(h) \subseteq \sigma_i(h)$. Conversely, $\forall h \in H$, $\forall s_i \in S_i \backslash \rho_i(h)$, there exist $g \in H^h$ and $r_i \in \rho_i(g) \subseteq \sigma_i(g)$ such that $u_i^g(r_i) > u_i^g(s_i)$. Hence, ρ_i satisfies (ES), and by (IS) of σ_i, $\forall h \in H$, $\rho_i(h) \supseteq \sigma_i(h)$. □

Proof of Prop. 2. *(i)* Suppose there exists $h \in H$ with $\sigma_i(h) = \varnothing$. Let $G := \{g \in H^h | \sigma_i(g) \neq \varnothing\}$. By (ES) of σ_i, $G \neq \varnothing$. By Lemma 1, there exists $s_i \in S_i$ such that, $\forall g \in G$, $s_i \in \sigma_i(g)$. By (IS) of σ_i, there do not exist $g \in H^h$ and $r_i \in \sigma_i(g)$ such that $u_i^g(r_i) > u_i^g(s_i)$. By (ES) of σ_i, $s_i \in \sigma_i(h)$, which contradicts that $\sigma_i(h) = \varnothing$. *(ii)* is a corollary of Proposition 5. □

Proof of Prop. 3. Since s_i is revision-proof at h, there exists an optimistic stable SB in the one-person planning situation such that $s_i \in \sigma_i(h)$. Suppose s_i is not a multi-self SPE of Γ at h. Then there exist $g \in H^h$ and $r_i = (r_i(g), s_i \backslash s_i(g))$ with $r_i(g) \in A_i^g$ such that $u_i^g(r_i) > u_i^g(s_i)$. By (IS), $r_i \in S_i \backslash \sigma_i(g)$. By (ES), there exist $f \in H^g$ and $q_i \in \sigma_i(f)$ such that $u_i^f(q_i) > u_i^f(r_i)$. If $f = g$, then $u_i^f(q_i) > u_i^f(r_i) > u_i^f(s_i)$. If $f \in H^g \backslash \{g\}$, then $u_i^f(q_i) > u_i^f(r_i) = u_i^f(s_i)$. Since $s_i \in \sigma_i(h)$, this contradicts (IS). □

Proof of Prop. 4. Consider σ_i defined by, $\forall h \in H$, $\sigma_i(h) := \{s_i \in S_i | s_i$ is a multi-self SPE of Γ at $h\}$. By Proposition 3, it suffices to show that σ_i is an optimistic stable SB. (IS) is satisfied since, $\forall h \in H$, $s_i', s_i'' \in \sigma_i(h)$ implies, $\forall g \in H^h$, $s_i'(g) = s_i''(g)$. To show (ES), suppose $\tilde{s}_i \in S_i$ and there do not exist $\tilde{g} \in H^h$ and $\tilde{r}_i \in \sigma_i(\tilde{g})$ such that $u_i^{\tilde{g}}(\tilde{r}_i) > u_i^{\tilde{g}}(\tilde{s}_i)$. If $\tilde{s}_i \in S_i \backslash \sigma_i(h)$, there exist $g \in H^h$ and $q_i = (q_i(g), \tilde{s}_i \backslash \tilde{s}_i(g))$ with $q_i(g) \in A_i^g$ such that $u_i^g(q_i) > u_i^g(\tilde{s}_i)$. Let $s_i \in \sigma_i(g)$. Then r_i defined by $r_i = q_i$ on $H^g(q_i)$ and $r_i = s_i$ on $H \backslash H^g(q_i)$ is a multi-self SPE of Γ at g. This leads to a contradiction since $g \in H^h$, $r_i \in \sigma_i(g)$ and $u_i^g(r_i) = u_i^g(q_i) > u_i^g(\tilde{s}_i)$. Hence, $\tilde{s}_i \in \sigma_i(h)$ □

Proof of Claim 1. *(i) (Always consuming the entire capital stock is a multi-self SPE of Γ.)* Suppose $r_i(g) = \tilde{k} > 0$ for some g with $g \in k$ and $k > 0$. Then, $u_i^g(\tilde{s}_i) = v(k)$ and $u_i^g(r_i(g), \tilde{s}_i \backslash \tilde{s}_i(g))$ $= v(k - \rho\tilde{k}) + \lambda v(\tilde{k})$. Hence, $u_i^g(\tilde{s}_i) - u_i^g(r_i(g), \tilde{s}_i \backslash \tilde{s}_i(g))$ $= v(k) - v(k - \rho\tilde{k}) - \lambda v(\tilde{k}) \geq v(k^m) - v(k^m - \rho\tilde{k}) - \lambda v(\tilde{k}) \geq 0$ by assumption, exploiting the concavity of $v(\cdot)$. *(Always maintaining* *the capital stock intact is a multi-self SPE of* Γ.) Suppose $r_i(g) = \widetilde{k} \neq k$ for some g with $g \in k$ and $k > 0$. Then, $u_i^h(\breve{s}_i) = (1+\frac{\lambda}{1-\delta})v((1-\rho)k)$ and $u_i^g(r_i(g), \breve{s}_i \setminus \breve{s}_i(g))$ $=$ $v(k-\rho\widetilde{k})$ $+$ $\frac{\lambda}{1-\delta}v((1-\rho)\widetilde{k})$. Hence, $u_i^h(\breve{s}_i) - u_i^g(r_i(g), \breve{s}_i \setminus \breve{s}_i(g))$ $=$ $v((1-\rho)k)$ $-$ $v(k-\rho\widetilde{k})$ $+$ $\frac{\lambda}{1-\delta}(v((1-\rho)k) - v((1-\rho)\widetilde{k}))$ $>$ $v'(k)(\rho(\widetilde{k}-k) + \frac{\lambda}{1-\delta}(1-\rho)(k-\widetilde{k}))$ $=$ $v'(k)(1-\rho)(k-\widetilde{k})$ $(1+\frac{\lambda}{1-\delta}-\frac{1}{1-\rho})$ $= 0$ by assumption, exploiting the strict concavity of $v(\cdot)$. *(ii)* Consider $\breve{\sigma}_i$ defined by, $\forall h \in H$, $\breve{\sigma}_i(h) := \{s_i \in S_i | \, s_i = \breve{s}_i \text{ on } H^h\}$. (IS) is trivially satisfied. To show (ES), suppose $r_i \in S_i$, and there does not exist $g \in H^h$ such that $u_i^g(\breve{s}_i) > u_i^g(r_i)$. Consider any $g \in H^h$. Let $\widetilde{k}$:= arg max$\{k \in K | h \in H^g(r_i) \setminus \{g\}$ and $h \in k\}$, and let $\widetilde{f}$ be the first decision node along $H^g(r_i)\setminus\{g\}$ at which $\widetilde{k}$ is reached. Let f be the predecessor of $\widetilde{f}$ along $H^g(r_i)$, with $f \in k_{t-1}$. Let $(k_t^r, k_{t+1}^r, \ldots)$ and $(c_t^r, c_{t+1}^r, \ldots)$ be the capital and consumption paths associated with $H^f(r_i)$. Let $q_i \in S_i$ satisfy $e \in \widetilde{k}$ whenever $e \in H^f(q_i)\setminus\{f\}$. Let $(k_t^q, k_{t+1}^q, \ldots)$ and $(c_t^q, c_{t+1}^q, \ldots)$ be the paths associated with $H^f(q_i)$. Since $\widetilde{k} = k_t^q = k_t^r$ and $c_t^q = c_t^r$, and, $\forall \tau > t$, $\widetilde{k} = k_\tau^q \geq k_\tau^r$ and $\widetilde{c} := (1-\rho)\widetilde{k} = c_\tau^q$, it follows that

$$
\begin{aligned}
& u_i^f(q_i) - u_i^f(r_i) = \lambda \textstyle\sum_{\tau=t+1}^{\infty} \delta^{\tau-t-1}(v(c_\tau^q) - v(c_\tau^r)) \\
& \geq \lambda v'(\widetilde{c}) \textstyle\sum_{\tau=t+1}^{\infty} \delta^{\tau-t-1}(c_\tau^q - c_\tau^r) \quad \text{(by the concavity of } v(\cdot)) \\
& = \lambda v'(\widetilde{c}) \textstyle\sum_{\tau=t+1}^{\infty} \delta^{\tau-t-1}(\delta-\rho)(k_\tau^q - k_\tau^r) \quad \text{(since } c_\tau = k_{\tau-1} - \rho k_\tau) \\
& \geq 0 \quad \text{(since } \delta > \rho \text{, with strict inequality if, for some } \tau > t,\ k_\tau^q > k_\tau^r).
\end{aligned}
$$

Hence, $u_i^f(q_i) \geq u_i^f(r_i)$ with strict inequality if $H^f(q_i) \neq H^f(r_i)$. Since $\breve{s}_i$ is a SPE of Γ at f (see part (i)), it follows that $u_i^f(\breve{s}_i) \geq u_i^f(q_i)$ with strict inequality if $H^f(\breve{s}_i) \neq H^f(q_i)$. Therefore, $u_i^f(\breve{s}_i) \leq u_i^f(r_i)$ only if $H^f(\breve{s}_i) = H^f(r_i)$. Note that $f = g$, since otherwise $\widetilde{f}$ is not the first decision node along $H^g(r_i)\setminus\{g\}$ at which $\widetilde{k}$ is reached. Hence, $r_i = \breve{s}_i$ on H^h and $r_i \in \breve{\sigma}_i(h)$. *(iii)* If s_i is a Markovian decision rule that revision-proof at h, then there exists an optimistic stable SB σ_i with $s_i \in \sigma_i(h)$. Suppose there exists $g \in H^h$ such that $u_i^g(\breve{s}_i) > u_i^g(s_i)$. Let $g \in k$. Since $\breve{s}_i$ and s_i are Markovian, $\breve{s}_i(g) \neq s_i(g)$. Let r_i satisfy $r_i(f) = \breve{s}_i(f)$ if $f \in H^g$ and $f \in k$, and $r_i(f) = s_i(f)$ if $f \in H^g$ and $f \notin k$. By (IS) of σ_i, $r_i \in S_i\setminus\sigma_i(h)$. By (ES) of σ_i, there exist $f \in H^g$ and $q_i \in \sigma_i(f)$ such that $u_i^f(q_i) > u_i^f(r_i)(\geq u_i^f(s_i))$. This contradicts (IS) of σ_i. Hence, if s_i is a Markovian decision rule that revision-proof at h, then there does not exist $g \in H^h$ such that $u_i^g(\breve{s}_i) > u_i^g(s_i)$. By the proof of part (ii), $s_i = \breve{s}_i$ on H^h. $\square$

Proof of Prop. 6. If s_i is symmetric, but not reconsideration-proof, then there exists a symmetric equilibrium r_i such that, $\forall h \in H$, $u_i^h(r_i) > u_i^h(s_i)$. If s_i is revision-proof at h, then there exists an optimistic stable SB σ_i with $s_i \in \sigma_i(h)$. By (IS) of σ_i, $r_i \in S_i \backslash \sigma_i(h)$. By (ES) of σ_i, there exist $g \in H^h$ and $q_i \in \sigma_i(g)$ such that $u_i^g(q_i) > u_i^g(r_i)(> u_i^g(s_i))$. This contradicts (IS) of σ_i. Hence, if a revision-proof decision rule is a symmetric equilibrium, then it is reconsideration-proof. The converse is trivial. □

Proof of Claim 2. (i) Let $a = 1$ or $a = 0$. Consider σ_i^a defined by, $\forall h \in H$, $\sigma_i^a(h) := \{s_i \in S_i | s_i = s_i^a$ on $H^h\}$. (IS) is trivially satisfied. To show (ES), suppose $r_i \in S_i$, and there do not exist $g \in H^h$ such that $u_i^g(s_i^a) > u_i^g(r_i)$. If $s_i^a(g) = 0$, then $r_i(g) = 0$, because otherwise $u_i^g(s_i^a) = 2 > 1 \geq u_i^g(r_i)$. If $s_i^a(g) = 1$, then $r_i(g) = 1$, because otherwise $0 = s_i^a(g, r_i(g)) = r_i(g, r_i(g))$, implying that $u_i^g(s_i^a) = 1 > 0 = u_i^g(r_i)$. Hence, $r_i = s_i^a$ on H^h and $r_i \in \sigma_i^a(h)$. *(ii)* Let $h \in H$. Suppose there exists an optimistic stable SB σ_i with $\hat{s}_i \in \sigma_i(h)$. By (IS) of σ_i, $s_i^1 \in S_i \backslash \sigma_i(h)$. By (ES) of σ_i, there exists $g \in H^h$ and $r_i \in \sigma_i(g)$ such that $u_i^g(r_i) > u_i^g(s_i^1)(> u_i^g(\hat{s}_i))$. This contradicts (IS) of σ_i and establishes that no such optimistic stable SB exists. Hence, $\hat{s}_i$ is not revision-proof at h. □

REFERENCES

ABREU, D. PEARCE, D., and STACCHETTI, E. (1993), "Renegotiation and Symmetry in Repeated Games", *Journal of Economic Theory*, **60**, 217–240.

AKERLOF, G.A. (1991), "Procrastination and Obedience", *American Economic Review*, **81** (Papers and proceedings), 1–19.

ASILIS, C., KAHN, C., and MOOKHERJEE, D. (1991), "A Unified Approach to -Proof Equilibria" (University of Illinois).

ASILIS, C. and KAHN, C. (1992), "Semi-stability in game theory: A survey of ugliness", in *Game Theory and Economic Applications: Proceedings of the International Conference Held at the Indian Statistical Institute, New Delhi, India, December 18–22, 1990* (B. Dutta et al., Eds.; Berlin: Springer Verlag).

ASHEIM, G.B., (1988a), "Rawlsian Intergenerational Justice as a Markov-Perfect Equilibrium in a Resource Technology", *Review of Economic Studies*, **55**, 469–484.

ASHEIM, G.B., (1988b), "Renegotiation-Proofness in Finite and Infinite Stage Games through the Theory of Social Situations" (DP A-173, University of Bonn).

ASHEIM, G.B., (1991a), "Extending Renegotiation-Proofness to Infinite Horizon Games", *Games and Economic Behavior*, **3**, 278–294.

ASHEIM, G.B., (1991b), "Individual and Collective Time Consistency" (DP 9169, CentER, Tilburg University).

BENNETT, E. and VAN DAMME, E. (1991), "Demand Commitment Bargaining—The Case of Apex Games", in *Game Equilibrium Models III* (R. Selten, Ed.; Berlin: Springer Verlag).

BERGIN, J. and MACLEOD, W.B. (1993), "Efficiency and Renegotiation in Repeated Games", *Journal of Economic Theory*, **61**, 42–73.

BERNHEIM, D., PELEG, B., and WHINSTON, M.D. (1987), "Coalition-Proof Nash Equilibria; I. Concepts", *Journal of Economic Theory*, **42**, 1–12.

BERNHEIM, D. and RAY, D. (1985), "Pareto Perfect Nash Equilibria" (Stanford University).

BERNHEIM, D. and RAY, D. (1989), "Collective Dynamic Consistency in Repeated Games", *Games and Economic Behavior*, **1**, 295–326.

BLACKERBY, C., NISSEN, D., PRIMONT, D., and RUSSELL, R.R. (1973), "Consistent Intertemporal Decision Making", *Review of Economic Studies*, **40**, 239–248.

DEKEL, E. and FARRELL, J. (1990), "One-Sided Patience with One-Sided Communication Does Not Justify Stackelberg Equilibrium", *Games and Economic Behavior*, **2**, 299–303.

ELSTER, J. (1979) *Ulysses and the Sirens. Studies in Rationality and Irrationality* (Cambridge: Cambridge University Press).

FARRELL, J. and MASKIN, E. (1989), "Renegotiation in Repeated Games", *Games and Economic Behavior*, **1**, 327–360.

GOLDMAN, S. (1980), "Consistent Plans", *Review of Economic Studies*, **47**, 533–537.

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

HAMMOND, P. (1975), "Pure Exchange Equilibrium of Dynamic Economic Models", in *Charity: Altruism or Cooperative Egoism* (E.S. Phelps, Ed.; Russell Sage Foundation).

HAMMOND, P. (1976), "Changing Tastes with Coherent Dynamic Choice", *Review of Economic Studies*, **43**, 159–173.

HELLWIG, M and LEININGER, W. (1987), "On the Existence of Subgame-Perfect Equilibrium in Infinite-Action Games of Perfect Information", *Journal of Economic Theory*, **43**, 55–75.

KOCHERLAKOTA, N.R. (1995), "Reconsideration-Proofness: A Refinement for Infinite Horizon Time Inconsistency" (forthcoming in *Games and Economic Behavior*).

KOTLIKOFF, L.J., PERSSON, T. and SVENSSON, L.E.O. (1988), "Social Contracts as Assets: A Possible Solution to the Time-Consistency Problem", *American Economic Review*, **78**, 662–677.

NALEBUFF, B. and SHUBIK, M. (1988), "Revenge and Rational Play" (DP 138, Woodrow Wilson School, Princeton University).

PELEG, B. and YAARI, M.E. (1973), "On the Existence of a Consistent Course of Action when Tastes are Changing", *Review of Economic Studies*, **40**, 391–401.

PICCIONE, M. and RUBINSTEIN, A. (1994), "On the Interpretation of Decision Problems with Imperfect Recall" (WP 24-94, Sackler Institute, Tel Aviv University).

POLAK, R.A. (1968), "Consistent Planning", *Review of Economic Studies*, **35**, 201–208.

ROTH, A.E. (1976), "Subsolutions and the Supercore of Cooperative Games", *Mathematics of Operations Research*, **1**, 43–49.

STROTZ, R.H. (1955-1956), "Myopia and Inconsistency in Dynamic Utility Maximization", *Review of Economic Studies*, **23**, 165–180.

TRANÆS, T. (1995), "Tie-Breaking in Games of Perfect Information" (Copenhagen Business School).

VON NEUMANN, J. and MORGENSTERN, O. (1953) *Theory of Games and Economic Behavior* (3rd ed.; Princeton: Princeton University Press).