GAMES AND ECONOMIC BEHAVIOR **20,** 31–50 (1997)
ARTICLE NO. GA970535

Dynamic Consistency and Imperfect Recall

Pierpaolo Battigalli*

Department of Economics, Princeton University, Princeton, New Jersey 08544-1021

Received December 15, 1995

We argue that a notion of constrained time consistency is appropriate to evaluate the interim incentives to deviate from a plan in a decision problem with imperfect recall. Under perfect recall, constrained time consistency is equivalent to the standard notion of time consistency. It turns out that a behavioral strategy β is constrained time consistent if and only if every realization equivalent strategy β' is a modified multiselves equilibrium and this implies that every optimal strategy is constrained time consistent. Furthermore, every constrained time consistent strategy is equivalent to a modified multiselves sequential equilibrium. *Journal of Economic Literature* Classification Numbers: C72, D81.

1. INTRODUCTION

A player in an extensive game (possibly against nature) has perfect recall if he or she can always remember all the information previously acquired and all her past choices (Kuhn, 1953). In their seminal book on game theory Von Neumann and Morgenstern allowed for the possibility of *im*perfect recall in an extensive game, arguing that strategic situations involving teams of imperfectly communicating individuals with identical interests are best modeled by regarding the team as a single forgetful player. They illustrated this point with Bridge, describing the elaborate system of conventional signals between, say, *N*orth and *S*outh as part of an overall strategy of the team *NS* (Von Neumann and Morgenstern, 1944, pp. 52–53). Yet, most of the theory of extensive games relies on the assumption of perfect recall, especially after Selten's (1975) article on

* I thank Paolo Ghirardato, Joseph Halpern, Ebbe Hendon, Lorand Ambrus-Lakatos, Michele Piccione, Ariel Rubinstein, and an anonymous referee for helpful comments on earlier versions of this paper. The usual disclaimer applies.

0899-8256/97 $25.00

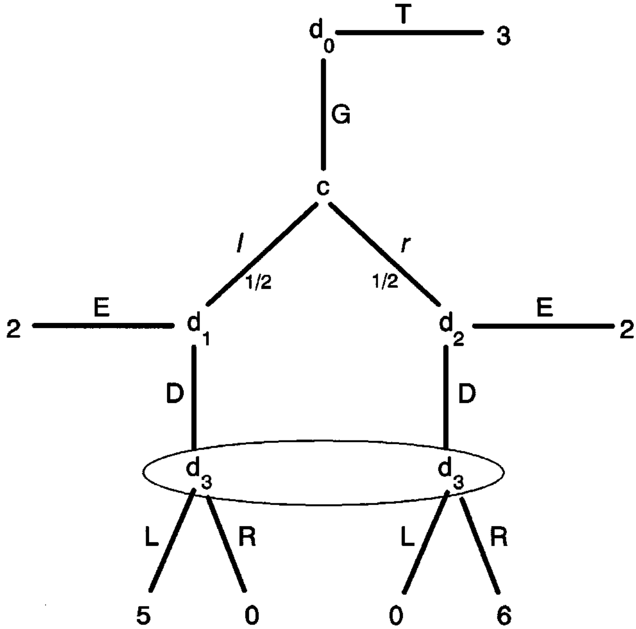

FIGURE 1

perfect Nash equilibria, as perfect recall is regarded as a rationality assumption.[1]

Not surprisingly, the recent research on bounded rationality has prompted some effort to analyze games and decision problems with imperfect recall. In a recent, thought-provoking paper, Piccione and Rubinstein (1997) (henceforth P & R) discuss some ambiguities in interpretation and paradoxes which arise when the perfect recall assumption is dropped. As a logical first step, they focus on one-person extensive games, i.e., dynamic decision problems.

One of their points is that, although the decision maker's preferences over outcomes do not change as the problem unfolds (unlike Strotz, 1956), an *ex ante* optimal plan need not be time consistent. Consider the decision problem depicted in Fig. 1 (two nodes in the same information set are enclosed in an ellipse).[2] The decision maker (by convention, a female) can either terminate the game immediately (T) and get 3 utils or play a game of chance (G). If she chooses G she gets 2 utils. Then, after having

[1] Selten (1975, p. 27) wrote: "Since game theory is concerned with the behavior of absolutely rational decision makers whose capabilities of reasoning and memorizing are unlimited, a game, where the players are individuals rather than teams, must have perfect recall."

[2] This is a modification of Example 2 in P & R.

observed the outcome of a random move (l or r), she can either stop (Exit) or go Down. If she goes down she can get a high reward or lose the 2 utils, but she forgets the outcome of the random move, which is crucial to determine the value of the reward (5 or 6) and whether Left or Right is the high-reward action. The optimal plan is to play the subgame and choose E after l (information set d_1), D after r (d_2), and R after D (d_3). This yields 4 utils in expectation. According to P & R, who simply extend "as is" the standard definition of time consistency given for situations with perfect recall, the optimal plan (actually, every plan in this example) is not time consistent: at node/information set d_1 the decision maker would reassess her plan and decide to go Down and Left.

We find it hard to make sense of this conclusion. How can the decision maker at information set d_1 plan to change her behavior at d_3 where a deviation from the original plan cannot be observed? It is plausible to suppose that if a decision maker does not observe any deviation from the original plan, she believes that no deviation has occurred. Indeed this seems to be part of the notion of having a plan. If the decision maker at d_3 thinks that no deviation has occurred, the best response is R. Thus a successful deviation at d_1 is possible only if the decision maker can find a way to change her future beliefs at an information set which is consistent with the original plan. But such a possibility should be explicitly modeled and the decision problem would be different.

This seems even more compelling when the decision maker is a team. Suppose that Player II has the move at d_3 and Player I at the other decision nodes. There is no reason for Player I to deviate from the optimal plan at d_1 as she has no way to "signal" that the outcome is l. On the contrary, Down would be interpreted by Player II as a signal that the outcome is r.

In this paper we try to analyze the *interim* incentives of a decision maker in a one-person, finite extensive game with (possibly) imperfect recall. In Section 2 the formal set up is introduced. Section 3 puts forward a notion of *constrained time consistency* which is shown to be equivalent to the usual definition of time consistency in decision problems with perfect recall. The modified definition simply assumes that, when reassessing her plan at an information set X, the decision maker expects to behave as prescribed by the original plan s at all the information sets which can follow X and are consistent with s (reachable with positive probability if s is implemented). It turns out, perhaps not surprisingly, that every optimal plan is time consistent in this sense. But the converse is not true (for example, the inefficient plan $G.D.E.L.$ is constrained time consistent).

Section 4 considers modified multiselves equilibria and sequential rationality. In decision problems without absentmindedness—that is, if no information set contains two nodes on the same path—a modified multiselves equilibrium is simply a Nash equilibrium of the agent form, and a modified multiselves sequential equilibrium is a sequential equilibrium of the agent form. The adjective "modified" refers to a particular way to define the conditional expected value of one-shot deviations at information sets which exhibit absentmindedness (see, e.g., Section 7 of P & R and Aumann *et al*. (1997)). It turns out that a behavioral strategy β is constrained time consistent if and only if every realization equivalent strategy β' is a modified multiselves equilibrium. Furthermore, it is shown that every constrained time consistent strategy is equivalent to a modified multiselves sequential equilibrium. It is argued that constrained time consistency may be interpreted as a forward induction refinement of multiselves sequential equilibrium outcomes. An appendix contains some of the proofs and related results.

2. MAIN CONCEPTS

We consider dynamic decision problems representable as finite one-person games in extensive form. This section collects definitions and notation about some concepts used in the literature on extensive games.

2.1. *Decision Problems in Extensive Form*

It is convenient here to follow P & R and adopt the formal definition given in Osborne and Rubinstein's (1994) textbook. The definition is given below for completeness, but the reader should consult the textbook above for further details and comments.

A *decision problem* (or *one-person game*) is a five-tuple $\Gamma = \langle H, u, C, \rho, I \rangle$, where

- H is a finite set of sequences of elements of some set A (interpreted as *histories* or feasible sequences of actions) including the empty sequence ϕ and such that every prefix[3] (initial subhistory) of an element of H is also in H.

—We write $h' \preceq h''$ to denote that h' is a prefix of h''. We say that h' (strictly) *precedes* h'' (or h'' *follows* h'), written $h' \prec h''$, if h' is a prefix of h'' different from h'' itself. (Thus $\prec$ is the asymmetric part of $\preceq$). The pair $(H, \prec)$ can be regarded as a tree with root ϕ. This interpretation is used in the diagrammatic representation of a decision problem.

[3] A *prefix* of a sequence h is a subsequence given by any number of leading elements in h. The empty sequence ϕ and h itself are prefixes of h.

—$Z := \{h \in H: \forall a \in A, (h, a) \notin H\}$ is the set of *terminal* histories and, for $X \subset H$, $Z(X) := \{z \in Z: \exists h \in X, h \preceq z\}$ is the set of terminal histories (weakly) following some history in X.

—For each $h \in H \setminus Z$, $A(h) := \{a \in A: (h, a) \in H\}$ is the set of actions available at h.

- $u: Z \to R$ is the *payoff* (VNM utility) function.
- $C \subset H \setminus Z$ corresponds to the set of *chance* moves.
- $\rho \in \Pi_{h \in C} \Delta(A(h))$ assigns the probabilities of chance actions and is assumed strictly positive.
- $\mathcal{I}$ is a partition of the set $H \setminus (C \cup Z)$ (the histories after which the decision maker chooses an action) into information sets. Generic elements of $\mathcal{I}$ will be denoted by X and Y. It is assumed that the decision maker always knows her available actions: for all $X \in \mathcal{I}$, all $h', h'' \in X$, $A(h') = A(h'')$. The set of available actions at information set X is denoted $A(X)$. Information sets are represented diagrammatically by enclosing the nodes/histories in the same information set in an ellipse.

2.2. *Imperfect Recall and Absentmindedness*

The *experience* of the decision maker at a history h is the sequence of the decision maker's information sets and actions along history h.[4] A decision problem Γ exhibits *perfect recall* if the decision maker has the same experience at all histories in the same information set and *imperfect recall* otherwise.

P & R say that a decision problem Γ exhibits *absentmindedness* if there is an information set $X \in \mathcal{I}$ containing two histories h' and h'' such that h' precedes h''. P & R analyze absentmindedness as a special case of imperfect recall. Other authors (e.g., Kuhn, 1953; Selten, 1975) exclude this case from the definition of an extensive game. When an information set X exhibits absentmindedness it is not obvious how to define updated probabilities and the expected value of one-shot deviations conditional on X. P & R consider different definitions. Here we adopt the one which we find most compelling. But since we do not focus on the specific problems

[4] Let ex(h) denote the experience at h; ex(h) can be defined by induction on the length of h:

—ex(ϕ) = ϕ, if $\phi \in C$;
—ex(ϕ) = (X), if $\phi \in X \in \mathcal{I}$;
—ex(h, a) = ex(h), if $h, (h, a) \in C \cup Z$;
—ex(h, a) = (ex(h), a), if $h \in H \setminus (C \cup Z), (h, a) \in C \cup Z$;
—ex(h, a) = (ex(h), X), if $h \in C, (h, a) \in X \in \mathcal{I}$;
—ex(h, a) = (ex(h), a, X), if $h \in H \setminus C \cup Z, (h, a) \in X \in \mathcal{I}$.

(The information set containing h is included in ex(h) as a matter of convention.)

due to absentmindedness, we refer to P & R, Aumann *et al*. (1997), Dekel and Gul (1996), and Gilboa (1997) for a thorough discussion. Otherwise, the reader may simply assume that the decision problems analyzed here do not exhibit absentmindedness.[5]

2.3. *Strategies, Plans, and Beliefs*

A (pure) *strategy* is a function s which assigns an available action to each information set. A *behavioral strategy* is a function β which assigns a probability measure on available actions for each information set. It is convenient to regard pure and behavioral strategies as elements of the Cartesian sets $S := \Pi_{X \in \mathbf{I}} A(X)$ and $B := \Pi_{X \in \mathbf{I}} \Delta(A(X))$, respectively. Pure strategies are identified with the corresponding degenerate behavioral strategies whenever convenient. In probabilistic computations, it is understood that the probability assigned by β to $a \in A(X)$ is the probability that a is chosen conditional on any $h \in X$. This conditional probability is denoted by $\beta(a \mid X)$.

The probability of history h'' conditional on h' given β is denoted by $p(h'' \mid h'; \beta)$.[6] Thus $p(h \mid \phi; \beta)$ is the prior probability of history h given β. Let

$$p(X \mid \phi; \beta) := \sum_{h \in X} p(h \mid \phi; \beta).$$

In decision problems *without* absentmindedness $p(X \mid \phi; \beta)$ is simply the prior probability of reaching information set X. More generally, $p(X \mid \phi; \beta) > 0$ if and only if strategy β does not prevent X from being reached. We say that an information set is *relevant* for β, or β-relevant, if $p(X \mid \phi; \beta) > 0$.

Two behavioral strategies β' and β'' are *equivalent*, written $\beta' \cong \beta''$, if they induce the same prior probability for every history (or, equivalently, for every terminal history). A necessary and sufficient condition for $\beta' \cong \beta''$ is that β' and β'' induce the same class of relevant information sets and prescribe the same behavior at such relevant information sets (see, e.g., Kuhn, 1953). The equivalence class containing strategy β is denoted by $[\beta]$.

[5] An earlier version of this paper (Battigalli, 1995) considered only decision problems without absentmindedness.

[6] Formally, $p(h'' \mid h'; \beta) = 1$ if $h' = h''$; $p(h'' \mid h'; \beta) = 0$ if h'' does not (weakly) follow h'; for $h' \prec h'' = (h, a)$:

—$p((h, a) \mid h'; \beta) = p(h \mid h'; \beta)\beta(a \mid X)$, if $h \in X \in \mathbf{I}$;

—$p((h, a) \mid h'; \beta) = p(h \mid h'; \beta)\rho(a \mid h)$, if $h \in C$.

A strategy β^* is (*ex ante*) *optimal* if it maximizes the expected utility

$$E(u \mid \phi; \beta) := \sum_{z \in Z} p(z \mid \phi; \beta) u(z).$$

The notion of behavioral strategy may have two sorts of redundancies: (i) if there is no absentmindedness, the expected utility induced by β is multilinear in the probabilities $\beta(a \mid X)$ and the decision maker has no incentive to randomize her choices; (ii) a strategy β (pure or randomized) may prescribe behavior at information sets that β itself prevents from being reached and this behavior cannot affect the outcome. We will use the word "plan" to refer to instructions given by a strategy at the relevant information sets. A *plan* corresponds to an equivalence class of strategies $[\beta]$. If there is no absentmindedness, nothing is lost by assuming that a strategy β corresponds to a "pure plan of action" $[\beta]$; that is, β is deterministic at the β-relevant information sets. The value of β at irrelevant (unreachable) information sets may be interpreted as the decision maker's expectation about her own future behavior if she deviates from the plan (cf. Rubinstein, 1991). Even if β prescribes deterministic choices at the relevant information sets, it is conceivable that the decision maker is uncertain about her own reaction to a deviation and this uncertainty may be relevant for the analysis of a decision problem with imperfect recall (see Section 4). If there is absentmindedness, the expected utility induced by β need not be multilinear and there may be *ex ante* incentives to randomize (see, e.g., Section 4 of P & R).

A system of beliefs represents the decision maker's conditional probabilities over past histories at each information set. Formally, a *system of beliefs* is a collection of probability measures $\mu \in M := \Pi_{X \in I} \Delta(X)$. $\mu(h \mid X)$ denotes the probability of history h conditional on information set X. An *assessment* is a pair $(\mu, \beta) \in (M \times B)$. We say that an assessment (μ, β) is *weakly consistent* (or that μ is weakly consistent with β) if, for all $X \in I$, $h \in X$,

$$\mu(h \mid X) p(X \mid \phi; \beta) = p(h \mid \phi; \beta). \tag{2.1}$$

If there is no absentmindedness, Eq. (2.1) simply says that μ is derived from β (and the chance probabilities), according to the usual updating rule at each relevant information set.[7] The implicit assumption is that the decision maker believes that she has followed her plan of action if no deviation is observed. P & R (Section 5) and Aumann *et al.* (1997) motivate Eq. (2.1) for decision problems with absentmindedness.

[7] P & R say that when (2.1) is satisfied μ is "consistent" with β. Kreps and Wilson (1982) call "consistency" a much stronger condition. Here we follow Myerson's (1991) terminology.

3. CONSTRAINED TIME CONSISTENCY

A strategy (or plan of action) β is time consistent if there is no information set X relevant for β such that the decision maker has an incentive to deviate from β at X following a different "subplan" β'. For a given assessment (μ, β) let

$$E(u \mid X; \mu, \beta) := \sum_{h \in X} \mu(h \mid X) \sum_{z \in Z} p(z \mid h; \beta) u(z).$$

In decision problems with perfect recall, the notion of time consistency can be unambiguously formalized using the following definition.

DEFINITION 3.1. A behavior strategy β is *time consistent* if there is a system of beliefs μ weakly consistent with β such that for all information sets $X \in \mathbf{I}$ relevant for β

$$\forall \beta' \in B, \qquad E(u \mid X; \mu, \beta) \geq E(u \mid X; \mu, \beta'). \tag{3.1}$$

(Note that, by Eq. (2.1), the quantification "there exists μ weakly consistent with β" can be replaced by "for all μ weakly consistent with β," since $\mu(\cdot \mid X)$ is uniquely determined by β at each X relevant for β.) It is well known that in decision problems with perfect recall, a strategy is optimal if and only if it is time consistent.

Consider for simplicity the case in which the decision maker does not actually randomize; that is, β is equivalent to a pure strategy (see discussion in Section 2.3). The reason why this definition of time consistency is unambiguous under perfect recall is that, if β is time inconsistent, then there is an information set X where the decision maker can improve her conditional expected payoff simply by choosing an action a^* different from the one designated by β and by planning to behave differently at information sets that are irrelevant for β (i.e., unreachable under β) because they follow the deviation a^* and the decision maker remembers this deviation. In other words, the decision maker has to consider only the alternative plans which behave as β at all information sets relevant for β except X. This suggests an alternative notion of time consistency. In order to present it, we first have to define the conditional value of a one-shot deviation. Let

$$E(u \mid X, a; \mu, \beta) := \sum_{h \in X} \mu(h \mid X) \sum_{z \in Z} p(z \mid (h, a); \beta) u(z)$$

be the conditional expected value of an action $a \in A(X)$ at X given β. Note that if X exhibits absentmindedness and contains two histories h and $h' \succsim (h, a)$, the decision maker is assuming that if she were at h and chose a, the probability of choosing some action a' after h' would be $\beta(a' \mid X)$.

This particular definition of the conditional expected value of a one-shot deviation is put forward and discussed with some skepticism by P & R (Section 7). Aumann *et al.* (1997) and Dekel and Gul (1996) forcefully defend it. They argue that the (pure or mixed) actions taken by the decision maker at h and h' should be the same *in equilibrium*. But when the decision maker contemplates a deviation at X, conditional on any $h \in X$, she should take as given the (mixed) action controlled by her "twin selves" at other histories in X following h. Gilboa (1997) reaches similar conclusions, arguing that one should look for symmetric solutions of modified decision problems where the original information sets featuring absentmindedness are replaced by corresponding collections of "crossing information" sets. We do not have anything to add to this discussion. Although we do not claim that different approaches cannot be fruitful, we find the above definition of $E(u \mid X, a; \mu, \beta)$ consistent with the standard way to model deviations in game theory and we use it in the following analysis.

DEFINITION 3.2. A behavior strategy β is *constrained time consistent* if there is a system of beliefs μ weakly consistent with β such that, for all $X \in I$ relevant for β there is no alternative strategy β' which coincides with β at all β-relevant information sets except X, chooses some action $a \in A(X)$ with probability one, and yields a higher expected utility conditional on X. Equivalently, β is constrained time consistent if it satisfies $\forall X \in I$, $\forall a \in A(X)$, $\forall \beta' \in B$,

$$[(p(X \mid \phi; \beta) > 0) \,\&\, (\beta' \cong \beta)] \Rightarrow E(u \mid X; \mu, \beta) \geq E(u \mid X, a; \mu, \beta'). \tag{3.2}$$

Remark 1. The notions of optimality, time consistency, and constrained time consistency are properties of plans of actions: if a strategy is optimal (time consistent, constrained time consistent), then every equivalent strategy is optimal (time consistent, constrained time consistent).

Proof. If $\beta' \cong \beta$, β and β' induce the same prior probabilities, the same transition probabilities conditional on positive probability histories, the same class of relevant information sets, and the same conditional probabilities at relevant information sets. ∎

PROPOSITION 3.3. *Suppose that Γ exhibits perfect recall. Then a strategy β is time consistent if and only if it is constrained time consistent.*

COROLLARY 3.4. *In decision problems with perfect recall, a strategy is optimal if and only if it is constrained time consistent.*

PROPOSITION 3.5. *Every optimal strategy is constrained time consistent.*

Propositions 3.3 and 3.5 are proved in Section 4 where constrained time consistency is related to multiselves equilibria.

The example depicted in Fig. 1 shows that the equivalence stated in Proposition 3.3 and the converse of Proposition 3.5 do not hold in decision problems with imperfect recall: the optimal plan $G.E.D.R.$ is not time consistent and the constrained time consistent plan $G.D.E.L.$ is not optimal.

In that first example the decision maker forgets something she knew before. P & R (Proposition 2) show that if the decision maker always remembers what she knew about chance moves, but possibly forgets her own choices, a strategy is time consistent if and only if it is optimal. The simple example depicted in Fig. 2 shows that a constrained time consistent plan ($L'L''$ in Fig. 2) may be time inconsistent and suboptimal even in this restricted class of decision problems.

Now the question is: Which characterization of time consistency is appropriate for decision problems with imperfect recall? Of course, the answer depends on how we interpret the situation formally represented by a decision problem in extensive form. If we think that at the *ex ante stage* the decision maker devises a plan suggesting what to do and *what to expect* in every possible contingency, then the notion of constrained time consistency seems more appropriate. A plan (in particular, an optimal plan) works not only as a ‘‘book of instructions,’’ but also as a device to coordinate the decision maker’s conditional expectations at different information sets. This is the reason why, at $[\beta]$-relevant information sets, conditional beliefs about past moves are determined by the *ex ante* chosen plan $[\beta]$. Similarly, the decision maker should expect to follow $[\beta]$ at future $[\beta]$-relevant information sets. On the other hand, expectations about future moves are not determined for $[\beta]$-irrelevant information sets

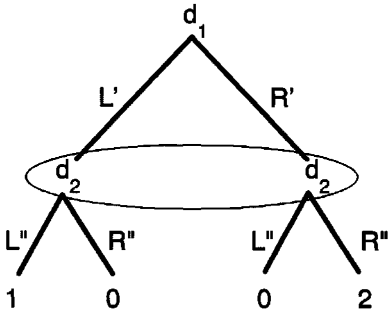

FIGURE 2

and the decision maker may at least hope to be able to signal to her "future selves" that a new plan $[\beta']$ is in place by causing such information sets to be reached. The possibility to signal strategic intent is further discussed in the next section, which relates constrained time consistency to multiselves equilibria.

4. MULTISELVES EQUILIBRIA AND SEQUENTIAL RATIONALITY

An assessment (μ, β) is sequentially rational if β maximizes the conditional expected utility $E(u \mid X; \mu, \cdot)$ at each information set X. In decision problems with perfect recall a strategy β is optimal if and only if there is an assessment (μ', β') satisfying "full consistency" (a strengthening of weak consistency) and sequential rationality, whereby β' is equivalent to β. Furthermore, a necessary and sufficient condition for sequential rationality of a consistent assessment is the "no single improvement" property: for every information set X and every action $a \in A(X)$ with positive conditional probability, a must be a best response conditional on X (see, e.g., Hendon *et al.* (1996)). This means that to analyze sequential rationality we can regard the decision maker as a team of $\# I$ different agents,[8] or "selves," controlling the choices at different information sets.

These equivalence properties do not hold in decision problems with imperfect recall: time inconsistent optimal strategies cannot be equivalent to sequentially rational assessments (see Fig. 1) and strategies satisfying the "no single improvement" property need not be optimal (see Figs. 1, 2, and 3). However, we may still use "no single improvement" as a test for sequential rationality. Furthermore, it turns out that the multiselves approach can be used to characterize constrained time consistency.

We consider a notion of sequential equilibrium which strengthens a multiselves equilibrium concept proposed by P & R (Section 7).

DEFINITION 4.1. A weakly consistent assessment (μ, β) is a *modified multiselves equilibrium*[9] if for all β-relevant information sets X,

$$\forall a, a' \in A(X), \qquad \beta(a \mid X) > 0 \Rightarrow E(u \mid X, a; \mu, \beta) \geq E(u \mid X, a'; \mu, \beta). \tag{4.1}$$

An assessment (μ, β) is *fully consistent* if there exists a sequence of strictly positive weakly consistent assessments $((\mu_k, \beta_k))_{k \geq 1}$ such that

[8] $\# F$ denotes the cardinality of a finite set F.

[9] P & R use the phrase "modified multiselves consistent."

$\lim_{k \to \infty}(\mu_k, \beta_k) = (\mu, \beta)$. A fully consistent assessment (μ, β) is a *modified multiselves sequential equilibrium* if (4.1) holds for all information sets X.

The adjective "modified" refers to the particular definition of $E(u \mid X, a; \mu, \beta)$ given in the previous section, which yields a modified version of the "no single improvement" property in games with absentmindedness. Note that the "no single improvement property" holds only at *relevant* information sets in the definition of a modified multiselves equilibrium. Let $A(\Gamma)$ be the agent form of a finite extensive game Γ—that is, the $\# I$-person extensive game derived from Γ assigning a separate player with payoff function u to each information set. It is easy to see that if there is no absentmindedness:

(i) the set of modified multiselves equilibria of Γ corresponds to the set of Nash equilibria of $A(\Gamma)$;

(ii) the set of modified multiselves sequential equilibria of Γ coincides with the set of sequential equilibria of $A(\Gamma)$.

Aumann *et al*. (1997), Dekel and Gul (1996), and Gilboa (1997) defend the multiselves equilibrium approach to decision problems, but they do not distinguish between sequential and nonsequential equilibria as they focus on the "absentminded driver" example by P & R featuring only one information set.[10] However, even in decision problems with perfect information, if there are potentially irrelevant information sets, Nash equilibria of the agent form may correspond to completely irrational plans for the decision maker, as choices "on the equilibrium path" may be induced by the expectation of irrational choices "off the equilibrium path" (it is easy to find numerical examples). This is the reason why we consider the stronger sequential equilibrium concept.

The requirement of full consistency ensures that an assessment satisfies Bayes rule everywhere, while the weak consistency property only requires that Bayes rule is satisfied on the equilibrium path. But full consistency also incorporates the much stronger assumption that every unexpected event is interpreted as the result of a sequence of independent mistakes (cf. Selten, 1975; Kreps and Wilson, 1982). Another implicit assumption of the multiselves sequential equilibrium concept is that the decision maker is able to anticipate her own conditional expectations for every contingency and all the "selves" have common expectations. These assumptions may be too strong for decision problems with imperfect recall and are relaxed elsewhere (see Battigalli, 1995).

[10] Note that in any decision problem where all the information sets are β-relevant for all β (such as the "absentminded driver" problem in P & R) a strategy is constrained time consistent if and only if it is a modified multiselves (sequential) equilibrium.

We have emphasized that constrained time consistency is a property of plans whereas the multiselves equilibrium concepts apply to strategies (or strategy profiles of the agent form) and do not distinguish between the decision problem Γ and the multiplayer game of common interest $A(\Gamma)$. However, the multiselves equilibrium concept can be used to clarify the notion of constrained time consistency and its relation to optimality.

PROPOSITION 4.2. (i) (P & R) *If a strategy β is optimal, then for every system of beliefs μ weakly consistent with β, (μ, β) is a modified multiselves equilibrium.*

(ii) *A strategy β is constrained time consistent if and only if for every μ weakly consistent with β and every β' equivalent to β the assessment (μ, β') is a modified multiselves equilibrium.*

Proof. (i) This is Proposition 3 in P & R.

(ii) First realize that it follows from the definitions of $E(u \mid X; \mu, \beta)$ and $E(u \mid X, a; \mu, \beta)$ that

$$E(u \mid X; \mu, \beta) = \sum_{a \in A(X)} \beta(a \mid X) E(u \mid X, a; \mu, \beta).$$

Therefore if β is either constrained time consistent or (part of) a modified multiselves equilibrium for each β-relevant X and each action a such that $\beta(a \mid X) > 0$, $E(u \mid X; \mu, \beta) = E(u \mid X, a; \mu, \beta)$. It follows that each constrained time consistent strategy is (part of) a modified multiselves equilibrium.

(Only if) Suppose that (μ, β) is weakly consistent and β' is equivalent to β. Then (μ, β') is weakly consistent. If β is constrained time consistent, then also β' is constrained time consistent and—by the argument above—(μ, β') must be a modified multiselves equilibrium.

(If) Suppose that, for each μ weakly consistent with β and every β' equivalent to β, (μ, β') is a modified multiselves equilibrium. Then for each β'-relevant information set X, each action $a \in A(X)$ such that $\beta'(a \mid X) > 0$ and each $b \in A(X)$

$$E(u \mid X; \mu, \beta') = E(u \mid X, a; \mu, \beta') \geq E(u \mid X, b; \mu, \beta').$$

Since β' is equivalent to β, this implies that for all β-relevant X and all $b \in A(X)$

$$E(u \mid X; \mu, \beta) = E(u \mid X; \mu, \beta') \geq E(u \mid X, b; \mu, \beta').$$

Thus (3.2) is satisfied and β is constrained time consistent. ∎

Proof of Proposition 3.5. We must show that optimality implies constrained time consistency. By Proposition 4.2(i) every optimal strategy β is part of a modified multiselves equilibrium. Since every β' equivalent to β is optimal as well, Proposition 4.2(i), (ii) imply that β must be constrained time consistent. ■

Constrained time consistent strategies (or plans) pass the multiselves sequential equilibrium test in the following sense.

PROPOSITION 4.3. *For every constrained time consistent behavioral strategy β there is a modified multiselves sequential equilibrium (μ', β') such that β' is equivalent to β.*

Proof. See the Appendix. ■

Proof of Proposition 3.3. We must show that in decision problems with perfect recall constrained time consistency is equivalent to time consistency. Time consistency trivially implies constrained time consistency. To prove the converse, consider that a constrained time consistent strategy β is equivalent to some β' which is part of a modified multiselves sequential equilibrium (Proposition 4.3). If the decision problem has perfect recall, β' must be optimal and time consistent. Therefore, also β must be time consistent. ■

Note that Proposition 4.3 along with 3.5 provides a stronger statement than 4.2(i): not only optimal strategies are modified multiselves equilibria; they also induce modified multiselves *sequential* equilibrium outcomes. Optimal strategies satisfy even stronger properties. For example, it is shown in the Appendix that for every optimal strategy β there is a modified multiselves sequential equilibrium (μ', β') with $\beta' \cong \beta$, which satisfies optimality in every subform.[11]

We have just shown that constrained time consistency is a refinement of multiselves sequential equilibrium *outcomes*. The parametrized example depicted in Fig. 3 (with $x \leq 9$) shows that the refinement can be strict and has a "forward induction" flavor.[12] Strategy (profile) Tc is a multiselves sequential equilibrium. In this equilibrium the first agent (at d_1) is afraid that the second agent (at d_2) would be uncertain about the actual deviation from T. But T is not a constrained time consistent plan. Suppose that, for some unexplained reason, plan T is chosen *ex ante*. According to constrained time consistency, since d_2 is inconsistent with T, the first agent is confident of being able to signal to the second agent that T has

[11] A subform is a sort of subgame with possibly multiple initial nodes (see Kreps and Wilson, 1982).

[12] Note that most forward induction refinements are applied to equilibrium *outcomes* (see, e.g., Kohlberg and Mertens, 1986, and Kohlberg, 1989).

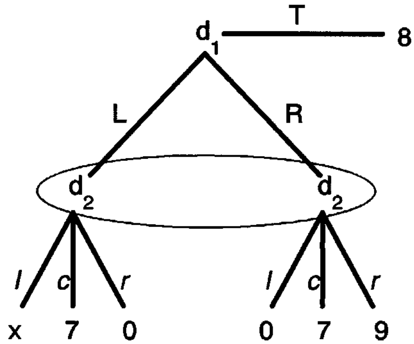

FIGURE 3

been replaced by a better plan, such as Rr.[13] Is this confidence justified? It is certainly more plausible to assume that the second agent is able to correctly interpret the fact that T has not been chosen when both agents are different "selves" of the same decision maker rather than distinct individuals. But the plausibility of this forward induction argument also depends on the precise payoffs.

If $x < 8$, L is a strictly dominated action and any forward induction refinement would yield Rr as the unique solution independently of the interpretation of Fig. 3 as a decision problem or a two-player game with common interests. For $8 < x < 9$ we might have less confidence in the ability of the first agent to signal her strategic intent. Yet Rr is the unique optimal plan and this may make it salient. In the knife-edge case $x = 9$, the interpretation of Fig. 3 as a decision problem or a two-player game should affect our intuition about the solution. In this case Ll and Rr are equivalent strategies (profiles) and Tc seems a plausible equilibrium of the two-player game, but somewhat less so for the one-person decision problem.

Note that this analysis rests on the informal assumption that the decision maker remembers the *ex ante* chosen plan, which is used as a reference point in order to coordinate expectations at different relevant information sets and to interpret information about past actions. We argue elsewhere (Battigalli, 1995) that if we drop this assumption, a different approach related to extensive form rationalizability (Pearce, 1984) is more appropriate and in the knife-edge case above we should not exclude plan T.

[13] More generally, one can show the following: Suppose that (i) if the decision maker is at the root of the decision tree, she knows it (that is, $\{\phi\} \in I$); (ii) the decision maker has the possibility to terminate the game immediately, say choosing action T. Then a plan prescribing T is constrained time consistent if and only if it is optimal.

5. CONCLUSION

We have analyzed the concepts of time consistency and sequential rationality for dynamic decision problems with imperfect recall. Under the assumption of perfect recall, there is an essentially unique appropriate formalization of these concepts, and optimality is equivalent to time consistency and realization equivalent to sequential rationality. But in decision problems with imperfect recall it is not obvious how these dynamic properties should be formalized. We have argued that a notion of constrained time consistency of plans is appropriate in this more general case: if the decision maker decides at the outset to follow a plan s and then evaluates the expected utility of a deviation from s at an information set X (consistent with s), she expects to behave as prescribed by s at each future information set Y consistent with the original plan s because at such information sets she cannot remember whether she has deviated to a new plan.

Under perfect recall, constrained time consistency coincides with the standard notion of time consistency and, hence, with optimality. In general, optimality, constrained time consistency, sequential rationality, and multiselves equilibria are related as follows:

(a) Every optimal behavioral strategy is constrained time consistent.

(b) A behavioral strategy is constrained time consistent if and only if every equivalent strategy is a modified multiselves equilibrium.

(c) Every constrained time consistent behavioral strategy is equivalent to some modified multiselves sequential equilibrium.

Simple examples show that the converses of (a) and (c) do not hold.

The analysis can be extended to multiperson games, considering Nash equilibria (in behavioral strategies), instead of optimal strategies and taking as fixed the opponents' strategies when checking for constrained time consistency of a player's strategy in a given profile. The definitions of multiselves equilibria are extended in an obvious way to multiperson games. Since for each player i and each fixed profile of opponents' strategies one can define an associated decision problem for i, one can use some of the decision theoretic arguments of this paper to show that

(a′) Every Nash equilibrium in behavioral strategies is constrained time consistent.

(b′) A constrained time consistent profile of behavioral strategies is a modified multiselves equilibrium.

APPENDIX

For any $\epsilon > 0$, let B_ϵ denote the set of behavioral strategies which assign probability at least ϵ to every action. B_ϵ is compact and is not empty if $\epsilon \in (0, \epsilon_\Gamma)$, where $\epsilon_\Gamma = (\max_{X \in I} \#A(X))^{-1}$.

DEFINITION 6.1. Fix $\epsilon > 0$. A weakly consistent assessment (μ, β) is an ϵ-modified multiselves equilibrium if $\beta \in B_\epsilon$ and for all information sets X

$$\forall a, a' \in A(X), \qquad \beta(a \mid X) > \epsilon \Rightarrow E(u \mid X, a; \mu, \beta) \geq E(u \mid X, a'; \mu, \beta).$$

By continuity of $E(u \mid X, a; \mu, \beta)$ in (μ, β), a limit of ϵ-modified multiselves equilibria as $\epsilon \searrow 0$ is a modified sequential equilibrium. ϵ-modified equilibria are considered here for technical reasons. In games without absentmindedness a limit of ϵ-modified multiselves equilibria of Γ as $\epsilon \searrow 0$ is a trembling hand perfect Nash equilibrium of $\mathcal{A}(\Gamma)$ (cf. Selten, 1975; Harsanyi and Selten, 1988, pp. 63–64).

We say that a strategy β is *ϵ-constrained optimal* if it maximizes the *ex ante* expected utility in the restricted set B_ϵ. For any strictly positive behavioral strategy β, μ_β denotes the unique system of beliefs (weakly and fully) consistent with β.

LEMMA 6.2. (i) *If β is ϵ-constrained optimal, then (μ_β, β) is an ϵ-modified equilibrium.*

(ii) *For every $\epsilon \in (0, \epsilon_\Gamma)$, there exists an ϵ-modified multiselves equilibrium (thus modified multiselves (sequential) equilibria always exist).*

Proof. (i) The same argument as in the proof of Proposition 3 in P & R can be applied here.

(ii) An ϵ-constrained optimal strategy β exists by nonemptiness and compactness of B_ϵ and continuity of expected utility. By (i), (μ_β, β) is an ϵ-modified multiselves equilibrium.[14] ■

Proof of Proposition 4.3. Fix a sequence $(\epsilon_k)_{k \geq 1}$ in $(0, \epsilon_\Gamma)$, where $\lim_{k \to \infty} \epsilon_k = 0$. For each k, consider the modified game $\Gamma(\beta, \epsilon_k)$ derived

[14] The different definitions of multiselves equilibrium are extended to finite multiperson games in a straightforward way. Modified multiselves equilibria always exist, but a different, fixed-point argument has to be used. In fact, it is true that a (ϵ-constrained) Nash equilibrium in behavioral strategies of the agent form of a multiperson game is a (ϵ)-modified multiselves equilibrium; but simple numerical examples show that the agent form may not have Nash equilibria in behavioral strategies due to the nonconcavity of the *ex anto* expected utility of absentminded agents.

from Γ as follows: each personal move corresponding to a β-relevant information set X is replaced by a chance move with probabilities

$$\rho_k(a \mid h) = \epsilon_k \qquad \forall h \in X, \text{if } \beta(a \mid X) = 0,$$

$$\rho_k(a \mid h) = \beta(a \mid X) - c(\beta, X)\epsilon_k \qquad \forall h \in X, \text{if } \beta(a \mid X) > 0,$$

where $c(\beta, X) = [\#A(X) - \#\mathrm{Supp}(\beta(\cdot|X))]/\#\mathrm{Supp}(\beta(\cdot|X))$. For each k choose an ϵ-modified multiselves equilibrium σ_k of $\Gamma(\beta, \epsilon_k)$ (Lemma 6.2(ii)) and let (μ_k, β_k) be the corresponding assessment—that is, $\beta_k = (\sigma_k, \rho_k)$ and $\mu_k = \mu_{\beta_k}$. Let (μ', β') be a cluster point of the sequence (μ_k, β_k). Clearly (μ', β') is fully consistent. For each β-relevant information set X, $\lim_{k \to \infty} \beta_k(\cdot|X) = \beta(\cdot|X)$. Thus $\beta' \cong \beta$. By a standard continuity argument, the multiselves sequential rationality property (4.1) is satisfied at each β-irrelevant information set. Finally, constrained time consistency of β implies that (μ, β') is a modified multiselves equilibrium (Proposition 4.2(ii)). Therefore (4.1) is satisfied also at the relevant information sets. ■

A subset of histories $\hat{H}$ corresponds to a *subform* if it is closed under succession and preserves information sets—that is, for all $h \in \hat{H}$, $X \in I$, $h \prec h'$ implies $h' \in \hat{H}$, and $X \cap \hat{H} \neq \emptyset$ implies $X \subset \hat{H}$. The set of initial histories of a subform induced by $\hat{H}$ is $\hat{W} := \{h \in \hat{H} \mid \forall h' \in H, h' \prec h \Rightarrow h' \notin \hat{H}\}$. Given some probability measure $\hat{\rho} \in \Delta(\hat{W})$, one can define a corresponding game $\hat{\Gamma}(\hat{\rho})$.

DEFINITION 6.3 (Kreps and Wilson, 1982). A behavioral strategy β is *extended subgame perfect* if there exists a mapping $\hat{H} \mapsto \hat{\rho} \in \Delta(\hat{W})$ which assigns to each $\hat{H}$ corresponding to a subform a probability measure on initial histories such that (a) for every $\hat{H}$ (inducing a subform), β induces a Nash equilibrium of $\hat{\Gamma}(\hat{\rho})$; (b) for every pair $\hat{H}', \hat{H}''$ (inducing subforms), if $\hat{H}' \subset \hat{H}''$ and $\sum_{h' \in \hat{W}'} \hat{\rho}'(h') \sum_{h'' \in \hat{W}''} p(h'' \mid h'; \beta) > 0$, then $\hat{\rho}''$ is derived from $\hat{\rho}'$ and β via Bayes' rule. (See Kreps and Wilson, 1982, p. 877.)

PROPOSITION 6.4. *A behavioral strategy β is optimal if and only if there is a modified multiselves equilibrium (μ', β') such that β' is extended subgame perfect and is equivalent to β.*

Proof. Construct a sequence of perturbed games $\Gamma(\beta, \epsilon_k)$ as in the proof above and for each k choose an ϵ_k-constrained optimal strategy σ_k of $\Gamma(\beta, \epsilon_k)$. By Lemma 6.2(i) σ_k is also an ϵ-modified multiselves equilibrium of $\Gamma(\beta, \epsilon_k)$. Let (μ', β') be a cluster point of the corresponding sequence of assessments (μ_k, β_k). We have already shown in the proof above that β' is equivalent to β and (μ', β') is a modified multiselves sequential equilibrium. To show that β' is extended subgame perfect, for each set of initial histories $\hat{W}$ of a subform, let $\hat{\rho}$ be obtained as the limit (possibly along a subsequence) of the probabilities $[\sum_{h' \in \hat{W}} p(h \mid \phi; \beta_k)]^{-1} \cdot p(h \mid \phi; \beta_k)$ (cf. Kreps and Wilson, 1982, p. 877). ■

The proof above suggests that we can formulate a generalized notion of sequential rationality as the limit of ϵ-constrained optimality and show that an optimal strategy is always equivalent to a sequentially rational one (cf. Kreps and Wilson, 1982, Proposition 6).

REFERENCES

1. Aumann, R., Hart, S., and Perry, M. (1996). "The Absent-Minded Driver," in *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Sixth Conference* (*TARK 96*) (Y. Shoham, Ed.), pp. 97–116. San Francisco, CA: Morgan Kaufmann.
2. Battigalli, P. (1995). "Time Consistency, Sequential Rationality and Rational Inferences in Decision Problems with Imperfect Strategy Recall," mimeo, Department of Economics, Princeton University.
3. Dekel, E., and Gul, F. (1996). "Rationality and Knowledge in Game Theory," in *Advances in Economics and Econometrics* (D. Kreps and K. Wallis, Eds.). Cambridge, UK: Cambridge Univ. Press.
4. Gilboa, I. (1996). "A Comment on the Absent Minded Driver Paradox," *Games Econ. Behav.*, to appear.
5. Harsanyi, J. C., and Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.
6. Hendon, E., Jacobsen, J., and Sloth, B. (1996). "The One-Deviation Principle for Sequential Rationality," *Games Econ. Behav.* **12,** 274–282.
7. Kohlberg, E. (1989). "Refinements of Nash Equilibrium: the Main Ideas," in *Game Theory and Applications* (T. Ichiishi, A. Neyman, and Y. Tauman, Eds.), pp. 3–45. San Diego: Academic Press.
8. Kohlberg, E., and Mertens, J. F. (1986). "On the Strategic Stability of Equilibria," *Econometrica* **54,** 1003–1037.
9. Kreps, D., and Wilson, R. (1982). "Sequential Equilibria," *Econometrica* **50,** 863–894.
10. Kuhn, H. W. (1953). "Extensive Games and the Problem of Information," in *Contributions to the Theory of Games II* (H. W. Kuhn and A. W. Tucker, Eds.), pp. 193–216. Princeton, NJ: Princeton Univ. Press.
11. Myerson, R. (1991). *Game Theory. Analysis of Conflict*. Cambridge, MA: Harvard Univ. Press.
12. Osborne, M., and Rubinstein, A. (1994). *A Course in Game Theory*. Cambridge, MA: MIT Press.
13. Pearce, D. (1984). "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica* **52,** 1029–1050.
14. Piccione, M., and Rubinstein, A. (1995). "On the Interpretation of Decision Problems with Imperfect Recall," *Games Econ. Behav.* **14** (1997).

15. Rubinstein, A. (1991). "Comments on the Interpretation of Game Theory," *Econometrica* **59,** 909–924.

16. Selten, R. (1975). "Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games," *Int. J. Game Theory* **4,** 25–55.

17. Strotz, R. H. (1956). "Myopia and Inconsistency in Dynamic Utility Maximization," *Rev. Econ. Stud.* **23,** 165–180.

18. Von Neumann, J., and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton Univ. Press (3rd ed.).