Testing Threats in Repeated Games*

Ran Spiegler †

April 28, 2004

Abstract

Under most game-theoretic solution concepts, equilibrium beliefs are justified by off-equilibrium events. I propose an equilibrium concept for infinitely repeated games, called "*Nash Equilibrium with Tests*", according to which players can only justify their equilibrium beliefs with events that take place on the equilibrium path itself. In NEWT, players test every threat that rationalizes a future non-myopic action that they take. The tests are an integral part of equilibrium behavior. Characterization of equilibrium outcomes departs from the classical "folk theorems". The concept provides new insights into the impact of self-enforcement norms, such as reciprocity, on long-run cooperation.

KEY WORDS: Solution concepts, repeated games, threat testing, justifiability, reciprocity, trigger strategies

JEL Numbers: C73

---

*I am deeply grateful to Nuffield College and the Institute for Advanced Study at Princeton, for their stimulating hospitality. I am indebted to Eric Maskin for his fruitful suggestions concerning the folk theorem. I also thank Eddie Dekel, Kfir Eliaz, Glenn Ellison, Dana Heller, Johannes Horner, Shachar Kariv, Ariel Rubinstein, Larry Samuelson, Asher Wolinsky, an associate editor and a referee, as well as numerous seminar participants, for their helpful comments.

†School of Economics, Tel Aviv University, Tel Aviv, Israel 69978. E-mail: rani@post.tau.ac.il. URL: http://www.tau.ac.il/~rani.

1 Introduction

Game-theoretic equilibrium concepts are typically based on a criterion by which players justify their equilibrium beliefs of the opponents' strategies. For example, in subgame perfect equilibrium (SPE), the criterion is sequential rationality: player i justifies his belief in player j's threats with the argument that if player i had (counterfactually) tested any of these threats, it would have been optimal for player j to realize them. Thus, players validate their equilibrium beliefs by invoking counterfactual events that are never observed in equilibrium.

In this paper I construct an equilibrium concept that requires players to *justify their equilibrium beliefs by events that take place on the equilibrium path itself*. If a player rationalizes his equilibrium actions by attributing threats to his opponent, he seeks evidence for their existence in the actual equilibrium path. Threats must be seen, in order to be believed.

I explore this idea in the context of two-player, infinitely repeated games with discounting. The fundamental problem in repeated games is whether players can sustain non-myopic behavior. A player takes a myopically inferior action only if he can rationalize it with a threat of his opponent to punish him if he plays myopically. According to the equilibrium concept that I propose, called "*Nash Equilibrium with Tests*" (**NEWT**), this threat must be evident from the equilibrium path. Specifically, the player must test any threat that rationalizes a non-myopic action that he takes later in the course of the game. Both the tests and the optimal response are part of equilibrium behavior: the player's equilibrium strategy carries out both tasks.

The idea that players verify a threat by testing it first is intuitive. For example, imagine a leader of one country, who faces an aggressive rival country in an ongoing conflict over some territory. The leader believes that fighting, although costly in the short run, is necessary for making the rival country soften its policy in the long run. The leader validates this belief by trying to play dovishly first. This demonstrates the futility of being dovish and helps justifying the fight. The modeling innovation in this paper is to build threat testing into the definition of an equilibrium concept. A strategy profile is unstable (in the sense of NEWT) if it leaves untested any of the threats that rationalize the non-myopic actions taken along the play path.

As the example suggests, equilibrium threat testing fits situations in which players need to *justify* their behavior to some audience. The player may be a delegate who must account for his actions to a principal or a constituency. Alternatively, he may have an internal need to respond to outside criticism. The player's audience is not a player in the game, and therefore need not share the player's equilibrium knowledge of the opponent's off-path behavior. In order to persuade such an audience that a threat exists, "empirical" evidence from the actual play path is more convincing than counterfactual arguments about the opponent's off-path behavior.

Because the notion of threat testing is central to the equilibrium concept, we need a language to describe what it means for a player to test at one history the *same* threat that he encounters at another history. I employ the formalism of finite automata (see Ch. 8 in Osborne and Rubinstein (1994)). A finite automaton is a 'machine' that consists of a finite set of internal states (one of which is specified to be the initial state), an output function (which specifies the player's action when he is in a given state), and a transition function (which determines his state at the next period, given his current state and the opponent's current action).

For every machine state q in the automaton s_i, let $br_j(q)$ denote player j's stage-game ("myopic") best-reply to the action that player i takes when in the state q. Let $BR_j(q)$ be the action prescribed for player j by *long-run* optimization against s_i, given that player i is in q. if $BR_j(q) \neq br_j(q)$ - i.e., if the action prescribed by long-run optimization against q is myopically inferior - then the transition from q must be different when j plays $br_j(q)$ or $BR_j(q)$ against q. In other words, there must be a threat associated with q to punish j if he plays $br_j(q)$, rather than $BR_j(q)$, against q.

A profile of finite automata (s_1, s_2) is a **NEWT** if for every $i = 1, 2$, and every state q in player i's automaton: *(i)* if $BR_j(q) \neq br_j(q)$, then along the path induced by (s_1, s_2), player j plays $br_j(q)$ when he faces q for the first time, but past a certain period, he plays $BR_j(q)$ whenever he faces q; *(ii)* if $BR_j(q) = br_j(q)$, player j always plays this action against q in the path induced by (s_1, s_2).

NEWT thus captures the idea that players test threats that rationalize their future non-myopic actions, but do not depart from optimizing behavior for any other purpose. If $BR_j(q) \neq br_j(q)$, then playing $BR_j(q)$ against q can only be rationalized by associating a threat (technically, a non-constant transition) with q. NEWT requires player j to test this threat by playing myopically against q as soon as he faces this state. In the long run, however, player j will adhere to the action prescribed by long-run optimization against q. If $BR_j(q) = br_j(q)$, there is no need to associate a threat with q in order to rationalize $BR_j(q)$. Therefore, NEWT requires player j to play $BR_j(q)$

whenever he faces q. NEWT is *not* a refinement of Nash equilibrium (NE), because it may require players to depart from optimizing with respect to their equilibrium beliefs.

To make the concept of NEWT more concrete, imagine a scenario, in which each player has to justify his behavior to an outside "referee". As the repeated game unfolds, the referee may question the player's rationale for his behavior at any period t. If the player behaves non-myopically at period t, he answers to the query by stating his belief of the opponent's strategy, and rationalizing his move by the opponent's threat at that period. The referee will be satisfied with this reply only if the player actually tested the threat before: otherwise, there is no evidence at period t for the threat's reality. If the player behaves myopically at t, the referee does not initiate a debate about rationalizing threats, because the player does not need them to rationalize his move at t. In NEWT, each player can successfully reply to all of his referee's queries.[1]

My main task in the paper is to analyze the restrictions on repeated-game behavior implied by NEWT. I begin by characterizing equilibrium outcomes in a number of restricted domains of stage games. Consider a stage game with a unique NE. Suppose further that the NE also Pareto-dominates any other stage-game outcome. I show that in such a game, only infinite repetition of the stage-game NE is consistent with NEWT. (In contrast, SPE sustains non-myopic play patterns in this class of games.) I find this *departure from standard folk theorems* quite attractive: when individual and collective rationality imply the same outcome in the stage game, we do not expect non-myopic behavior to emerge in the repeated game. In contrast to this extreme selection result, a "NEWT folk theorem" holds for every $n \times n$ coordination game. The construction of equilibrium strategies illustrates how threat testing can be incorporated into equilibrium behavior, by "adding a wrinkle" to a simple structure, which sustains a Nash folk theorem.

Next, I turn from equilibrium outcomes to the *structure of equilibrium strategies*. A structural restriction on equilibrium automata that is shared by all players in a symmetric game can be viewed as a "norm of punishment". I analyze two natural norms of punishment: a class of automata that display an element of reciprocal behavior, and a class of "forgiving" trigger strategies.

[1] Note that in this scenario, nothing prevents player i from communicating to his referee a false belief of player j's strategy, as long as it is *consistent* with the play path. In order to accommodate this possibility, we should interpret s_j as a consistent belief of player j's strategy, rather than as his true strategy. I shall ignore this distinction in the sequel.

These norms do not constrain long-run cooperation in NE, and are mildly restrictive under SPE. Under NEWT, however, these norms impose severe restrictions on long-run cooperation. These results shed some light on the way in which certain norms can facilitate or hamper long-run cooperation.

2 An Example

Let us introduce the equilibrium concept by way of a simple example. Let the stage game be the following instance of "Chicken" and suppose that the discount factor is close to one:

	C	D
C	$4,4$	$1,5$
D	$5,1$	$0,0$

In order to sustain infinite repetition of (C,C) in the long-run, we must construct strategies that contain threats to punish deviations. Consider the strategy for player 1, given a diagrammatic finite-automata representation as s_1 in Figure 1. Because C is a myopically inferior action against itself, playing C against q^0 is optimal for player 2 *only* because of player 1's threat to punish player 2 if he plays D against q^0. This threat is represented by the D-transition from q^0. If player 2's automaton were identical to s_1, we would have a NE that sustains infinite repetition of (C,C). The threats that rationalize the players' cooperation would never be realized along the equilibrium path, and therefore, this automata profile is not a NEWT.

Consider the automata profile (s_1, s_2) given by Figure 1. Given the stage game's payoff structure, $br_2(q^0) = br_2(q^2) = D$ and $br_2(q^1) = C$. Given the discount factor, playing always D against q^0 yields a maximal discounted payoff of $\frac{11}{3}$, whereas playing always C against q^0 yields a discounted payoff of 4. Therefore, $BR_2(q^0) = C$. Note that q^1 and q^2 have a constant transition - i.e., player 1's continuation is independent of the action that player 2 takes against q^1 or q^2. Therefore, myopic and long-run optimization prescribe the same action for player 2 against each of these states: $BR_2(q^1) = br_2(q^1)$ and $BR_2(q^2) = br_2(q^2)$. Now turn to player 1's br and BR actions. First, $br_1(r^0) = br_1(r^2) = C$ and $br_1(r^1) = D$. Given the discount factor, playing always D against r^1 yields a maximal discounted payoff of 3, whereas playing always C against r^1 yields a discounted payoff of 4. Therefore, $BR_2(r^1) = C$. Because r^0 and r^2 have a constant transition, $BR_1(r^0) = br_1(r^0)$ and $BR_1(r^2) = br_1(r^2)$.

[Insert Figure 1]

We can now verify that (s_1, s_2) is a NEWT. Consider player 2's reaction to s_1. First, in the path induced by (s_1, s_2), player 2 plays myopically at period 1 - the first time that he faces q^0 - whereas in the long run (from period 4 onwards), he always plays $BR_2(q^0)$ against q^0. Second, player 2 always plays myopically against q^1 and q^2. Now turn to player 1's reaction to s_2. First, he plays myopically at period 2 - the first time that he faces r^1 - whereas in the long run, he always plays $BR_1(r^1)$ against r^1. Second, player 1 always plays myopically against r^0 and r^2.

Thus, both players depart from optimizing behavior only for the sake of testing the threats that rationalize the non-myopic actions that they take in equilibrium. Starting at period 4, the players enter a cyclic phase, in which they carry out a repeated-game NE. We can see that NEWT captures a "build-up" phase that precedes Nash equilibria which sustain non-myopic behavior. All the threat testing takes place during the build-up phase.

Note that player 2's automaton "displays reciprocity" at the C-state r^1: if player 1 plays $a \in \{C, D\}$ against r^1, player 2 plays a in the next period. Player 1's automaton violates this property: even if player 2 plays D against the C-state q^2, player 1 plays C in the next period. In Section 5, we will see that this is a general feature: in order to sustain infinite long-run repetition of (C, C) in repeated Chicken, at least one player must refrain from displaying reciprocity at some C-state. According to NEWT, there is a tension between long-run cooperation and reciprocal behavior in this game.

3 The Equilibrium Concept

I define the solution concept for two-player, infinitely repeated games with perfect information and discounting. Let A_i denote the stage-game action set for player i. Let u_i denote player i's stage-game payoff function. Assume that $u_i(a_i, a_j) \neq u_i(a_i', a_j)$ for every $a_j \in A_j$ and $a_i, a_i' \in A_i$. In particular, there is a unique stage-game best-reply for player i against any action taken by player j. Define $a_i^* = \arg\min_{a_i} \max_{a_j} u_j(a_i, a_j)$. That is, a_i^* is the action that "*minimaxes*" player j.

As to repeated-game strategies, I consider only pure strategies with a finite automata representation. An automaton is a quadruple $(Q_i, q_i^0, f_i, \tau_i)$, where: Q_i is a finite set of machine states; q_i^0 is the initial state; $f_i : Q_i \rightarrow A_i$ is an output function, which specifies the action taken by player i when he is in state $q \in Q_i$; and $\tau_i : Q_i \times A_j \rightarrow Q_i$ is a transition function, where $\tau_i(q, a_j)$ specifies the state to which the automaton switches from state $q \in Q_i$ when the opponent plays a_j against q. We will say that $q \in Q_i$ is an *a-state* if $f_i(q) = a$; and we will say that $q \in Q_i$ has a *constant transition* if $\tau_i(q, a_j)$ is the same for all $a_j \in A_j$. Henceforth, s_i denotes an automaton for player i.[2]

The path induced by (s_1, s_2) is $z(s_1, s_2) = (a_1^t, a_2^t)_{t=0,1,...}$, where $a_i^t \in A_i$ is player i's action at period t, such that $a_i^0 = f_i(q_i^0)$, $a_i^1 = f_i[\tau_i(q_i^0, a_j^0)]$, etc. Player i's discounted payoff is $(1-\delta)\sum_{t=0}^{\infty} \delta^t u_i(a_i^t, a_j^t)$. Given (s_1, s_2), define a function $p_i : \{0, 1, ...\} \rightarrow Q_i$, which depicts the evolution of player i's machine state along $z\,(s_1, s_2)$, where $p_i(t) = q$ signifies that s_i is in the state q at period t. We will say that player j *faces a state* $q \in Q_i$ at period t if $p_i(t) = q$. A state $q \in Q_i$ is *visited* in $z(s_1, s_2)$ if $p_i(t) = q$ for some t in $z(s_1, s_2)$.

It is known that a profile of finite automata induces a path that eventually enters a *cycle*. That is, there exist a period t^* and an integer L, such that $p_i(t) = p_i(t+L)$ for every $i = 1, 2$, $t \geq t^*$. In the sequel, t^* always denotes the period in which the cyclic phase of $z(s_1, s_2)$ begins. The phase that precedes t^* is referred to as the "introductory phase" of $z(s_1, s_2)$.

For every $q \in Q_i$, denote $br_j(q) = \arg\max_{a_j \in A_j} u_j[f_i(q), a_j]$. In words, $br_j(q)$ is player j's *myopic best-reply action* against q. In contrast, define $BR_j(q)$ as player j's *long-run best-reply action* against q - i.e., the action that long-run optimization prescribes for player j whenever player i is in the state q. Let us ensure that $BR_j(q)$ is a well-defined, single-valued function. By the assumption of discounting, finding a best-reply to s_i is a Markovian decision problem with a well-defined set of solutions. For every $q \in Q_i$, there is a non-empty subset of actions $A_j^*(q) \subseteq A_j$, which long-run optimization against s_i prescribes for player j at any history in which s_i is in the state q.[3]

[2] In Figure 1, s_1 consists of: $Q_1 = \{q^0, q^1, q^2\}$, $f_1(q^0) = f_1(q^2) = C$, $f_1(q^1) = D$, $\tau_1(q^0, C) = q^0$, $\tau_1(q^0, D) = q^1$, $\tau_1(q^1, \cdot) = q^2$ and $\tau_1(q^2, \cdot) = q^0$. In the same figure, s_2 consists of: $Q_2 = \{r^0, r^1, r^2\}$, $f_2(r^0) = f_2(r^2) = D$, $f_2(r^1) = C$, $\tau_2(r^0, \cdot) = \tau_2(r^2, \cdot) = r^1$, $\tau_2(r^1, C) = r^1$ and $\tau_2(r^1, D) = r^2$.

[3] See Osborne and Rubinstein (1994, Ch. 9) for more details.

Let $BR_j(q) = \arg\max_{a_j \in A_j^*(q)} u_j[f_i(q), a_j]$. That is, $BR_j(q)$ is the myopically optimal action against q in $A_j^*(q)$.

Both $BR_j(q)$ and $br_j(q)$ are functions from Q_i to A_j. Of course, both functions are defined relative to a given s_i. Note that $BR_j(q) \neq br_j(q)$ only if the transition $\tau_i(q, \cdot)$ is not constant. In this case, the *threat* associated with q is represented by the transitions $\tau_i[q, a_j]$, $a_j \neq BR_j(q)$.

We are now ready to formulate an equilibrium concept, which requires players to test the threats that rationalize their non-myopic actions.

Definition 1 *An automata profile* (s_1, s_2) *is a Nash Equilibrium with Tests (**NEWT**) if for every player* $i \in \{1, 2\}$, $j \neq i$, *and every state* $q \in Q_i$*:*

1. If $BR_j(q) \neq br_j(q)$, then:
    (a) $a_j^t = br_j(q)$ at the earliest t in $z(s_1, s_2)$ for which $p_i(t) = q$.
    (b) $a_j^t = BR_j(q)$ at every $t \geq t^*$ in $z(s_1, s_2)$ for which $p_i(t) = q$.
2. If $BR_j(q) = br_j(q)$, then $a_j^t = BR_j(q)$ at every t in $z(s_1, s_2)$ for which $p_i(t) = q$.

The first condition in Definition 1 says that if the long-run best-reply action against $q \in Q_i$ is non-myopic, then player j should play myopically when he faces q for the first time, and play the long-run best-reply action against q in the long run. The second condition says that if the myopic and long-run best-reply actions against q coincide, player j should play this action whenever he faces q.

NEWT is an equilibrium concept, in which players validate their equilibrium beliefs with events that take place on the equilibrium path itself. Playing non-myopically can only be justified by a belief in a threat by the opponent to punish the myopic move. The belief-selection criterion underlying NEWT is that this threat must be evident from *actual* behavior. Therefore, the threat that justifies the non-myopic move must be tested in equilibrium.

NEWT is *not* a refinement of NE. Every NEWT that involves non-myopic behavior is sustained by threats that are tested in equilibrium. Testing a threat is costly, hence players do not optimize with respect to their beliefs.

Note, however, that the *cyclic* phase of NEWT *does* constitute a repeated-game NE. In fact, every NEWT (s_1, s_2) can be transformed into a NE (s_1', s_2'), where s_j' is identical to s_j except that the initial state is $p_j(t^*)$ instead of q_j^0.

[Insert Figure 2]

Example. In Section 2, we have seen an example of NEWT in repeated Chicken. The automata profile given by Figure 2 is another example. Consider player 1's reaction to s_2. His long-run best-reply actions are: $BR_1(r^0) = C$, $BR_1(r^1) = D$ and $BR_1(r^2) = D$. Only in the case of r^2 do player 1's myopic and long-run best-reply actions coincide. And indeed, player 1 plays D against r^0 at period 1 (the first time that he faces r^0), and he plays only C against r^0 in the cyclic phase. He plays D against r^2 whenever he faces it. He plays C against r^1 at period 2 (the first time that he faces r^1). Since player 1 never faces r^1 in the long run, condition 1(b) in Definition 1 is vacuously satisfied with respect to r^1.

Turning to player 2's reaction to s_1, his best-reply actions are: $BR_2(q^0) = C$, $BR_2(q^1) = C$ and $BR_2(q^2) = D$. Only in the case of q^0 do his myopic and long-run best-reply actions coincide. And indeed, player 2 plays C against q^0 whenever he faces it. He plays D at period 2 (the first time that he faces q^1), and he plays only C against q^1 in the cyclic phase. He plays C at period 3 (the first time that he faces q^2). Since player 2 never faces q^2 in the long run, condition 1(b) in Definition 1 is vacuously satisfied with respect to q^2.

Although in NEWT players depart from optimizing behavior, they do not experiment arbitrarily with sub-optimal actions. Instead, NEWT assigns a special status to the myopic best-reply actions. This is consistent with the justifiability-based interpretation of NEWT expounded in the introduction. According to this interpretation, NEWT captures a situation in which each player attempts to justify his repeated-game behavior to some audience. Myopic moves can be rationalized without attributing any threat to the opponent, whereas non-myopic moves can only be rationalized by a threat. In this case, since the threat is essential to the player's justification, he must provide evidence for the threat from past behavior, in order to convince his audience that the justification is sound.

Tests as "single deviations". I have identified a threat with a transition from a machine state. Testing a threat amounted to playing myopically against a state, without specifying the continuation. In other words, a test is like a "single deviation". Alternatively, one could identify a threat with a sequence of transitions. In this case, testing a threat would consist of "multiple deviations" that extend over a number of periods. I find the notion of a test as a single deviation more natural.

High-order threats. In the NEWT given by Figure 2, q^2 is visited only because player 2 tests the threat associated with q^1. Similarly, r^1 is visited only because player 1 tests the threat associated with r^0. If player 2 always played $BR_2(q^1)$ against q^1, q^2 would not be reached and player 2 would not face the threat associated with q^2. Nevertheless, since $BR_2(q^2) \neq br_2(q^2)$, NEWT requires player 2 to test the threat associated with q^2 by playing $br_2(q^2)$ when he faces q^2 for the first time. We can see that in NEWT, players have to test *every* threat that rationalizes a non-myopic action that they take, even if the threat is faced only as a result of testing another threat. I refer to such a threat as a "high-order threat".

Exact vs. full automata. Definition 1 relies on the formalism of *exact* automata - i.e., τ_i conditions only on player i's current state and player j's current action. By contrast, a *full* automaton would also condition on player i's current *action*. The restriction to exact automata entails no loss of generality, because NEWT does not deal with behavior off the equilibrium path. One could re-define NEWT in terms of full automata, but this would only complicate notation. Later in the paper, when we compare the predictions of NEWT and SPE, the comparison is well-defined only if both concepts are defined in terms of full automata.

4 Analysis

Let us begin our analysis of NEWT with a simple existence result.

Proposition 1 *A NEWT exists if and only if the stage game possesses a pure-strategy NE.*

Proof. Suppose that the stage game possesses a pure-strategy NE (a_1, a_2). The following automata profile is a NEWT in the repeated game. Player i's automaton consists of a single state q_i satisfying $f_i(q_i) = a_i$. Since

$BR_j(q_i) = br_j(q_i)$, only condition 2 in the definition of NEWT is relevant. And indeed, $a_j^t = br_j[p_i(t)]$ for every player $i = 1, 2$, $j \neq i$, and every period t along $z(s_1, s_2)$.

Now suppose that the stage game does not possess a pure-strategy NE. Then, $a_j^1 \neq br_j[p_i(1)]$ for some player j. Either $br_j[p_i(1)] = BR_j[p_i(1)]$, in which case condition 2 in Definition 1 is violated, or $br_j[p_i(1)] \neq BR_j[p_i(1)]$, in which case condition 1(a) in Definition 1 is violated. ■

An immediate corollary of Proposition 1 is that in every NEWT, (a_1^1, a_2^1) is a stage-game NE. Players play myopically at the beginning of the game, and they may develop a non-myopic play pattern later.

Let us turn to the task of characterizing equilibrium outcomes. The standard folk theorems will provide a natural benchmark. Consider, for example, the following stage game:

	A	B
A	$3, 3$	$1, 1$
B	$1, 1$	$0, 0$

Every discounted-payoff profile above $(1, 1)$ is SPE-sustainable.[4] This is a particularly unattractive prediction, given that (A, A) is implied by individual as well as collective rationality in the stage game.

Proposition 2 *Suppose that the stage game has a unique NE (a, a). Suppose further that (a, a) Pareto-dominates any other stage-game outcome. Then, only infinite repetition of (a, a) is consistent with NEWT.*

Proof. Suppose, contrary to the claim, that there exists a NEWT that induces non-myopic behavior. Because (a, a) is the unique stage-game NE, Proposition 1 implies that $a_1^1 = a_2^1 = a$. Let t denote the earliest period in $z(s_1, s_2)$, for which (a_1^t, a_2^t) is not a stage-game NE. Then, $t > 1$ and $a_1^k = a_2^k = a$ for every $k < t$. Because $a_j^t \neq br_j[p_i(t)]$ for at least one player j, by NEWT there is a period $k < t$, such that $p_i(k) = p_i(t)$. If $a_j^t \neq a$ for both $j = 1, 2$, then $a_i^k \neq a$ for some player i, a contradiction.

[4] The restriction to finite automata implies that generically, real-valued payoffs can only be approximated, though arbitrarily closely.

It follows that exactly one player, say player 2, plays a at period t. Moreover, there exists a period $k < t$, such that $p_2(k) = p_2(t)$ and $a_1^k = a$. By definition of k and t, $a_1^l = a_2^l = a$ whenever $k < l < t$. Thus, player 1 plays a against $p_2(t)$ at some $k < t$; both players subsequently play a repeatedly, until player 2's automaton returns to $p_2(t)$ at period t. Since (a, a) is the best outcome in the game for player 1, $BR_1[p_2(t)] = a$. But this means that player 1's move at period t violates condition 2 in Definition 1. ■

Thus, *NEWT effects a departure from the standard folk theorems.* The departure is in an intuitive direction: when individual and collective rationality unequivocally favor the same outcome in the stage game, we should not expect non-myopic behavior to emerge in the repeated game.

To see why NEWT confirms this intuition, recall the "referee scenario" described in the introduction. In the early stages of the game, players play (a, a) until some period t, in which player 1, say, switches to some other action b while player 2 continues to play a. At that period, player 1's referee may ask him to explain his move. Player 1 faces the following predicament. Because b is not myopically optimal against a, he must rationalize it by attributing a threat to player 2. On one hand, if player 1 did not test the threat prior to period t, the referee is not convinced by his story. On the other hand, if he tested the threat, he merely showed that a is the optimal action in that situation, because the "punishment" was a repeated play of (a, a), which is the best outcome in the game for player 1. Either way, player 1 is unable to convince the referee that his move at period t was sound.

I am unable to provide a complete characterization of equilibrium outcomes for arbitrary stage games. The next result, however, establishes a "folk theorem" for a restricted domain. A stage game is a *coordination game* if for every player i and every $a \in A_i$, $(a_i, br_j(a_i))$ is a NE. Our next result shows that a "NEWT folk theorem" holds for every coordination game.

Proposition 3 *Every individually rational discounted payoff profile in a repeated coordination game can be approximated (arbitrarily closely) by some NEWT, provided that the discount factor is sufficiently close to one.*

Proof. Consider a cycle (of length L) of action profiles $(a_1^l, a_2^l)_{l=1,...,L}$, which sustains a profile of discounted payoffs that exceed each player's min-max level, when the discount factor is sufficiently close to one. Within the cycle, let $((a_1^{k_1}, a_2^{k_1}), ..., (a_1^{k_m}, a_2^{k_m}))$, $1 \leq m \leq L$, be a sub-sequence of action profiles which do not constitute a stage-game NE. Because the stage game is a coordination game, neither player plays myopically at any of these m periods. It will be notationally convenient to assume that $k_i = i$ for every $i = 1, ..., m$. - i.e., that (a_1^k, a_2^k) is a stage-game NE for every $k = m+1, ..., L$. If the discount factor is sufficiently close to one, this re-ordering of the cycle does not affect any of the arguments in the proof.

Let us construct an automata profile (s_1, s_2) that induces this cycle in the cyclic phase of $z(s_1, s_2)$. For every player $i = 1, 2$, partition Q_i into three non-empty subsets: $Q_i = QC_i \cup QP_i \cup QT_i$. We will design the automata, such that: QC_i will be the set of states that are visited in the cyclic phase of $z(s_1, s_2)$; QP_i will be the set of states that carry out punishments for player j's deviations from his cyclic play pattern; and QT_i will be the set of states that are responsible for testing player j's threats. Let us now describe the outputs and transitions of these different classes of states. We will begin by specifying the elements that are common to both players. We will then complete the description by specifying the elements that are idiosyncratic to each player. In the sequel, let qc_i, qp_i and qc_i denote a typical "cyclic", "punishment" or "testing" state, respectively.

The "cyclic" states (QC_i). Let $QC_i = \{qc_i^1, ..., qc_i^L\}$. For every $k = 1, ..., L$, let $f_i(qc_i^k) = a_i^k$. For every $k = 1, ..., m$, $\tau_i(qc_i^k, a_j^k) = qc_i^{k+1 \bmod L}$, and $\tau_i(qc_i^k, a_j) = qp_i(k, 1)$ for every $a_j \neq a_j^k$. (The state $qp_i(k, 1)$ belongs to QP_i, and its precise definition is given below.) For every $k = m+1, ..., L$, let $\tau_i(qc_i^k, \cdot) = qc_i^{k+1 \bmod L}$.

The interpretation of this part of the construction is as follows. The set of states QC_i is responsible for carrying out player i's cyclic play pattern. He cycles through QC_i as long as player j adheres to his own cyclic play pattern. If player j deviates at one of the first m periods in the cycle, player i's automaton switches to a "punishment" state. Player i does not punish player j for deviations during the last $L - m$ periods of the cycle.

The "punishment" states (QP_i). Let us partition QP_i into m subsets: $QP_i = QP_i^1 \cup \cdots \cup QP_i^m$. Let $QP_i^k = \{qp_i(k, 1), ..., qp_i(k, N_i^k)\}$, $N_i^k \geq 1$, for every $k = 1, ..., m$. For every $l = 1, ..., N_i^k$, let $f_i[qp_i(k, l)] = a_i^*$. For every $l = 1, ..., N_i^k - 1$, let $\tau_i[qp_i(k, l), \cdot] = qp_i(k, l+1)$. For every $k < m$, let $\tau_i[qp_i(k, N_i^k), \cdot] = qc_i^{k+1}$. (The specification of $\tau_i[qp_i(k, N_i^m), \cdot]$ is different for each player, and will be given below.)

The interpretation of this part of the construction is as follows. The set of states QP_i is responsible for punishing player j for deviating from his cyclic play pattern. The number N_i^k is the duration of the punishment that player i inflicts on player j for having failed to play a_j^k against one of the first $m-1$ "cyclic" states qc_i^k When the punishment ends, player i's automaton proceeds to the "next" cyclic state, qc_i^{k+1}. As we shall see below, the case of $k=m$ requires some qualification.

The "testing" states (QT_i). Let us partition QT_i into m subsets: $QT_i = QT_i^1 \cup \cdots \cup QT_i^m$. Let $QT_i^k = \{qt_i(k,0), qt_i(k,1), ..., qt_i(k, N_j^k)\}$, for every $k = 1, ..., m$. Let $f_i[qt_i(k,0)] = br_i(a_j^k)$. For every $l = 1, ..., N_j^k$, let $f_i[qt_i(k,l)] = br_i(a_j^*)$. For every $l = 1, ..., N_j^k - 1$, let $\tau_i[qt_i(k,l), \cdot] = qt_i(k, l+1)$. For every $k = 1, ..., m-1$, let $\tau_i[qt_i(k, N_j^k), \cdot] = qt_i(k+1, 0)$. As to the case of $k = m$, let $\tau_i[qt_i(m, N_j^m), \cdot] = qc_i^1$.

The interpretation of this part of the construction is as follows. The set of states QT_i carries out the threat testing. For every $k = 1, ..., m$, the state $qt_i(k,0)$ is responsible for playing myopically against the "cyclic" state qc_j^k in player j's machine. The next N_j^k states are responsible for playing myopically against player j's min-max action during the N_j^k periods of punishment. Subsequently, player i plays myopically against qc_j^{k+1}, and so forth. After all the testing states are visited, player i's machine switches to its first "cyclic" state, qc_i^1.

The description so far holds for both players. The difference between s_1 and s_2 consists in the initial state q_i^0, as well as in the state that precedes the first testing state $qt_i(1,0)$. Complete the description of s_1 by setting $q_1^0 = qc_1^1$ and $\tau_1[qp_1(m, N_1^m), \cdot] = qt_1(1,0)$. Complete the description of s_2 by setting $q_2^0 = qt_2(1,0)$ and $\tau_2[qp_2(m, N_2^m), \cdot] = qc_2^1$. In words, the testing states in player 1's machine appear right after the last sequence of punishment states. In contrast, the testing states in player 2's machine appear right at the beginning.

In order to conclude the construction, it remains to determine the N_i^k's. For every $k = 1, ..., m-1$ - and in the case of $i = 2$, also for $k = m$ - if player j plays $a_j \neq a_j^k$ against qc_i^k, he is punished, in the form of being "min-maxed" for N_i^k periods. After the punishment is over, player i's machine switches to a "cyclic" state. Because the cycle generates a payoff that lies strictly above player j's min-max level, there exists a number N^*, such that if $N_i^k > N^*$, the punishment is deterring. It is only in the case of $i = 2$ and $k = m$, that player i's machine does not switch to a "cyclic" state immediately after the punishment is over. However, if N_1^m is sufficiently high, there are enough a_1^*-states in this punishment phase to ensure that the punishment is deterring.

In this way, we have guaranteed that for every $k = 1, ..., m$, $a_j^k = BR_j(qc_i^k)$. For every $k = m + 1, ..., L$, $a_j^k = br_j(a_i^k)$ and qc_i^k has a constant transition, hence $a_j^k = BR_j(qc_i^k)$.

Finally, let us verify that (s_1, s_2) is a NEWT. By construction, each player i cycles through QC_i in the cyclic phase of $z(s_1, s_2)$. As we have seen, $a_j^k = BR_j(qc_i^k)$ for every $k = 1, ..., L$. Thus, players stick to the long-run best-reply actions in the cyclic phase. Now consider the introductory phase of $z(s_1, s_2)$. For every $k = m + 1, ..., L$, qc_i^k remains unvisited in this phase. For every $k = 1, ..., m$, qc_i^k is visited exactly once in the introductory phase, and player j plays $br_j(qc_i^k)$ on that occasion. Finally, all the states in QP_i and QT_i have a constant transition, and by construction, player j always plays myopically against them. It follows that (s_1, s_2) is a NEWT. ■

Figure 3 illustrates the construction in the context of a symmetric stage game. It displays a NEWT that sustains a two-period cycle $((a_1^1, a_2^1), (a_1^2, a_2^2))$, where neither outcome is a stage-game NE. The structure of equilibrium strategies "adds a wrinkle" to a simple construction, which sustains a Nash folk theorem. In the basic construction, s_i punishes deviations from the cyclic play pattern, by playing the min-max action a^* for some finite number of periods (which is independent of player j's behavior while being punished). When the punishment is over, s_i switches to some "cyclic" state.[5]

[Insert Figure 3]

The variation on this basic construction consists in adding a sequence of constant-transition states to each player's automaton. These states are responsible for testing the opponent's threats. The states are placed in each player's machine in such a way that in equilibrium, player 1 starts testing player 2's threats only after player 2 has finished all his tests. When player 1 finishes his tests, the players proceed to the cyclic phase, in which they carry out a repeated-game NE.[6]

[5] In Figure 3, a box with $K \cdot a$ written inside stands for a sequence of K consecutive a-states with a constant transition.

[6] Figure 1 provides another, particularly simple example for the construction. Using the notation of the proof of Proposition 3: $m = L = 1$; $QC_1 = \{q^0\}$, $QP_1 = \emptyset$, $QT_1 = \{q^1, q^2\}$; $QC_2 = \{r^1\}$, $QP_2 = \{r^2\}$, $QT_2 = \{r^0\}$; and $M_1^1 = M_2^1 = 1$. (The only difference from the construction given in the proof is that QP_1 is empty.)

Proposition 3 demonstrates that the strong selection result given by Proposition 2 hinges on the uniqueness of the stage-game NE. For example, in a common-interest coordination game, there is a NE which Pareto-dominates all other stage-game outcomes. However, because there are other NE in the stage game, it is possible to sustain individually rational, inefficient outcomes.

The relative simplicity of the construction of Proposition 3 owes to our focus on coordination games. Suppose that as part of the construction, we need player j to test a threat associated with some "cyclic" state $q \in Q_i^C$. Then, at some period t, player j plays $br_j(q)$ against q. Because we are dealing with a coordination game, (a_1^t, a_2^t) is necessarily a stage-game NE. Therefore, it is possible to incorporate player j's test into the play path without adding a new threat to player j's machine. Such a construction becomes infeasible when the stage game is *not* a coordination game.

However, in the context of *symmetric* 2×2 stage games, it can be shown that a "NEWT folk theorem" holds for the class of games not covered by Propositions 2 and 3. (This is essentially PD, except that the dominant-strategy equilibrium need not be inefficient.) The construction is quite similar to the one given in Proposition 3, only somewhat more complicated because it now involves high-order threats (non-cyclic states with a non-constant transition), in order to overcome the difficulty described in the previous paragraph. Because of the limited scope of this "folk theorem" and for the sake of brevity, I omit the proof. At any rate, this finding completes the characterization for symmetric 2×2 games: except for the games covered by Proposition 2, a folk theorem holds for every stage game in this class.

5 Long-Run Cooperation: The Role of Norms

Section 4 examined the set of outcomes that emerge under NEWT. The focus of this section is on the *structure* of equilibrium strategies. Specifically, I examine how certain restrictions on the structure of equilibrium strategies affect the sustainability of mutually beneficial outcomes. I use the term "*norm of punishment*" to describe a structural restriction on strategies that is shared by all players in a repeated game. These structural restrictions are meant to capture, in a stylized way, self-enforcement that we observe in real-life repeated interactions.

Let us restrict attention to symmetric stage games (i.e., $A_1 = A_2 = A$ and $u_1(a, b) = u_2(b, a)$). Recall that a^* denotes the *min-max* action in the game. Assume that there exists an action $a^0 \in A$, $a^0 \neq a^*$, such that $br(a^0) \neq a^0$ (hence, (a^0, a^0) is not a stage-game NE), yet $u(a^0, a^0)$ exceeds the min-max payoff. Infinite repetition of (a^0, a^0) thus has some amount of social desirability. For example, in PD and Chicken, $a^* = D$ and $a^0 = C$. Let $u(a^0, a^0) = M$ and normalize $u(br(a^0), a^0) = M + 1$. If a NEWT sustains infinite repetition of (a^0, a^0) in the cyclic phase, I will say that it "sustains long-run cooperation". The problem I study in this section is whether certain norms of punishment are consistent with long-run cooperation, under NEWT.

5.1 Reciprocity

Reciprocity is a familiar and intuitive norm of punishment in repeated games. Game theorists' interest in reciprocal behavior intensified in the wake of Axelrod's famous PD tournaments (Axelrod (1984)). In a tournament designed to mimic an evolutionary environment, the Tit-for-Tat automaton emerged as a "winner".[7] Theoretical works on reputation effects in repeated PD (Kreps et. al. (1982), Watson (1994), etc.) relied on Tit-for-Tat as a "crazy type", whose presence facilitates cooperation. This sub-section examines a weaker notion of reciprocity, which requires players to display reciprocal behavior only in some situations.

Definition 2 *In a repeated symmetric game, an automaton s displays* ***reciprocity at a^0-states*** *if whenever $f(q) = a^0$, $f[\tau(q, a)] = a$ for every $a \in A$.*

Reciprocity at a^0-states is a norm that requires players to display reciprocal behavior only when they are in a state whose output is a^0. Player 2's automaton in Figure 1 and player 1's automaton in Figure 2 display reciprocity at C-states, whereas player 1's automaton in Figure 1 and player 2's automaton in Figure 2 violate this property.

In every symmetric game that falls into the category presented at the beginning of this section, it is possible to sustain infinite repetition of (a^0, a^0) in NE, using automata that display reciprocity at a^0-states. The same holds under SPE, if $u(a^0, a^0) = M$ is sufficiently large (provided that we employ full, rather than exact automata).

[7] The Tit-for-Tat automaton is: $Q = \{q^0, q^1\}$; $f(q^0) = C$, $f(q^1) = D$; and for every $q \in Q$, $\tau(q, a) = q$ if and only if $a = C$.

In contrast, we shall see that the implications of reciprocity at a^0-states on the NEWT-sustainability of long-run cooperation depends on whether $(br(a^0), br(a^0))$ or $(br(a^0), a^0)$ is a stage-game NE.[8] The NEWT's given by Figures 1 and 2 sustain long-run cooperation in repeated Chicken. In Figure 1 (2), only s_2 (s_1) displays reciprocity at C-states. The violation of this notion of reciprocal behavior turns out to be necessary for sustaining long-run cooperation in a class of stage games, of which Chicken is an instance.

Proposition 4 *Consider a symmetric stage game, in which $(a^0, br(a^0))$ and $(br(a^0), a^0)$ are NE. Fix $u(a, b)$ for every $a, b \neq a^0, br(a^0)$. There exists M^*, such that for every $M > M^*$, if a NEWT (s_1, s_2) induces $a_1^t = a_2^t = a^0$ for every $t \geq t^*$, then s_1 or s_2 do not display reciprocity at a^0-states.*

Proof. By assumption, $a_1^t = a_2^t = a^0$ for every $t \geq t^*$ along $z(s_1, s_2)$. Denote $br(a^0) = b$. Because $b \neq a^0$, there must be a period $t < t^*$ for every player $j = 1, 2$, in which $p_i(t) = p_i(t^*)$ and $a_j^t = b$. Because s_i displays reciprocity at $p_i(t)$, $a_i^{t+1} = b$. Let k be the latest period along $z(s_1, s_2)$, at which any of the players (suppose that it is player 1, without loss of generality) plays b. Suppose that $a_2^k = a^0$. Then, because s_2 displays reciprocity at $p_2(k)$, $a_2^{k+1} = b$, contradicting the definition of k. It follows that $a_2^k \neq a^0$. We have established the existence of a period in the play path, in which one player plays b while the other player plays $a \neq a^0$.

Let t denote the earliest period in which any of the players (suppose that it is player 1, without loss of generality) plays b, while his opponent plays $a \neq a^0$. By our assumption on the stage game's payoff structure, neither player plays myopically at period t. By NEWT, there must exist a period $h < t$, for which $p_1(h) = p_1(t)$ and $a_2^h = a^0$. If $a_2^t = b$, then there also exists a period $l < t$, $l \neq h$, for which $p_2(l) = p_2(t)$ and $a_1^l = a^0$. In this case, let $h < l$, without loss of generality.

Because player 2 displays reciprocity at $p_2(h)$, $a_2^{h+1} = b$. If $h + 1 = t$, then $a_2^t = b$, and this contradicts the existence of a period l between h and t, in which $p_2(l) = p_2(t)$ and $a_1^l = a^0$. Therefore, $h + 1 < t$. By the definition of t, $a_1^{h+1} = a^0$. Because s_1 displays reciprocity at a^0-states, it follows that player 2 could play a^0 against $p_1(t)$, and guarantee a continuation consisting of infinite repetition of (a^0, a^0). If M is sufficiently large, this is the best

[8] Recall that by our assumptions on payoffs, every stage-game NE is strict. Hence, no more than one of these outcomes can be a stage-game NE.

continuation that is feasible for player 2, given that s_1 displays reciprocity at a^0-states. (The reason is that the maximal stage-game payoff against a^0 is $M + 1$, and the maximal stage-game payoff against any $a \neq a^0$ is lower than $M - 1$, when M is sufficiently large.) Therefore, because $a^0 = br_2[p_1(t)]$, $a^0 = BR_2[p_1(t)]$ if M is sufficiently large. By NEWT, player 2 should always play a^0 against $p_1(t)$. His behavior at period t violates NEWT. ■

Thus, the norm of reciprocity at a^0-states rules out the long-run sustainability of (a^0, a^0), when $(a^0, br(a^0))$ is a stage-game NE and $u(a^0, a^0)$ is sufficiently high (relative to $u(a, b)$, for all $a, b \neq a^0$). To see the intuition for this result, consider the problem of sustaining long-run repetition of (C, C) in repeated Chicken. Reciprocity at C-states implies that the build-up phase must contain a "war": a period t in which both players play (D, D). In Chicken, (D, D) is not a NE. Therefore, each player has to justify his behavior in t with a test. However, these tests turn out to generate another war, which triggers further tests, and so forth. Reciprocity at C-states is a norm that generates "too many" tests in the build-up phase, and thus obstructs cooperation altogether.[9]

By contrast, consider the following example. Let the stage game be PD, with $2 \cdot u_1(C, C) > u_1(C, D) + u_1(D, D)$. The automata profile given by Figure 4 is a NEWT. First, consider player 2's reaction to s_1: *(i)* $BR_2(q^0) = BR_2(q^2) = C$, a myopically dominated action; player 2 should play D when he faces q^i for the first time, and play only C whenever he faces q^i in the cycle, which is indeed the case; *(ii)* $BR_2(q^1) = D$, a myopically dominant action; player 2 should play D whenever he faces q^1, which is indeed the case. Now turn to player 1's reaction to s_2: *(i)* $BR_1(r^1) = BR_1(r^2) = C$; player 1 should play D when he faces r^i for the first time, and play only C whenever he faces r^i in the long run, which is indeed the case; *(ii)* $BR_1(r^0) = D$; player 1 should play D whenever he faces r^0, which is indeed the case.

In this equilibrium, both s_1 and s_2 display reciprocity at C-states. Long-run cooperation in this equilibrium is "built up" over an introductory phase which lasts five periods.

[Insert Figure 4]

[9] In the case of Chicken, we can set $M^* = \frac{1}{2} \cdot [u(D, C) + u(C, D)]$. This is exactly the value of $u(C, C)$ that makes infinite repetition of (C, C) socially superior to any other symmetric repeated-game payoff profile.

The construction of Figure 4 can easily be adapted to any symmetric stage game in which $(br(a^0), br(a^0))$ is a NE, as long as M is sufficiently large. The output functions should be modified by replacing C with a^0 and D with $br(a^0)$. The transition functions should be modified by replacing C with a^0, D with $br(a^0)$, and for every action $a \neq a^0, br(a^0)$, the transitions $\tau_i(q, a)$ should be set to be equal to $\tau_i[q, br(a^0)]$. Indeed, the next result shows that this adapted construction yields the shortest cooperation build-up phase that is possible under NEWT in any symmetric stage game for which $(br(a^0), br(a^0))$ is a NE.

Proposition 5 *Consider a symmetric stage game, in which $(br(a^0), br(a^0))$ is a NE. Fix $u(a, b)$ for every $a, b \neq a^0, br(a^0)$. If NEWT (s_1, s_2) sustains cooperation and players are sufficiently patient, then $t^* \geq 6$.*

Proof. Denote $br(a^0) = b$. Because $b \neq a^0$, NEWT implies that for both $j = 1, 2$, $i \neq j$, $BR_j[p_i(t^*)] = a^0$ and there exists a period $t_j < t^*$, such that $p_i(t_j) = p_i(t^*)$ and $a_j^{t_j} = b$. Moreover, $t_j \neq t^* - 1$; otherwise, player j can guarantee, by playing $br(a^0)$ against $p_i(t^*)$, a continuation consisting of infinite repetition of (a^0, a^0), hence a^0 cannot be the best-reply action against $p_i(t^*)$, a contradiction.

Since $br(b) = b$, NEWT implies that for both $j = 1, 2$, there exists a period $k_j < t_j$, such that $p_j(k_j) = p_j(t_j)$ and $a_i^{k_j} = b$. It follows that $t^* \geq 5$. Suppose that $t^* = 5$. Then, $k_1 = k_2 = 1$ and the path must be as follows (assuming $t_1 > t_2$, without loss of generality):

Period	1	2	3	4	5	6	...
Player 1	b	a^0	b	b^1	a^0	a^0	...
Player 2	b	b	a^0	b^2	a^0	a^0	...

where $b^1, b^2 \neq a^0$. Note that $p_1(1) = p_1(3)$ and $p_1(2) = p_1(t^*)$. Suppose that $BR_2[p_1(1)] \neq b$. Then, if player 2 plays b against $p_1(t^*)$, there is a continuation action plan for him (playing $BR_2[p_1(1)]$ against $p_1(1)$, and then repeatedly playing $BR_2(q)$ against every $q \in Q_1$ subsequently reached), which yields a higher discounted payoff than M (as long as player 2 is sufficiently patient). But, this contradicts the condition that $a^0 = BR_2[p_1(t^*)]$. It follows that $BR_2[p_1(1)] = b$. Therefore, player 2's behavior at period 3 violates condition 2 in Definition 1. ∎

Proposition 5 implies that reciprocity at a^0-states is a "good norm" for sustaining long-run repetition of (a^0, a^0), for a class of stage games that includes PD. The norm is "good" in the sense that the build-up phase is the shortest possible under NEWT. Thus, according to NEWT, *the same norm that supports cooperation in "PD-like" games obstructs cooperation in "Chicken-like" games.* This is a distinction that emerges under NEWT, and not under standard solution concepts.

The intuition for this difference between "PD-like" and "Chicken-like" games is as follows. In both classes of games, reciprocity at a^0-states generates periods in which one player plays $br(a^0)$, while the other player plays some action $a \neq a^0$. In a stage game like PD, in which $(br(a^0), br(a^0))$ is a NE, such a period need not trigger any further tests. In contrast, in a stage game like Chicken, in which $(a^0, br(a^0))$ is a NE, such periods necessarily have to be justified by tests. Thus, reciprocity at a^0-states generates more threat testing in the latter class of games than in the former.

5.2 Forgiving Trigger Strategies

Trigger strategies are among the simplest norms of punishment in repeated games. A "trigger strategy" punishes a deviating player by "min-maxing" him for a number of periods. In a "grim" trigger strategy, the punishment phase is infinitely long. Clearly, grim strategies are incapable of sustaining cooperation in NEWT, because the players' tests trigger the grim threats as soon as they face them. Thus, in the context of NEWT, the trigger strategy must satisfy some form of "forgivingness" in order to sustain cooperation. In a "forgiving" trigger strategy, the punishment phase has finite duration, and when it is over, the punishing player resumes "business as usual".

Definition 3 *A automaton $s_i = (Q_i, q_i^0, f_i, \tau_i)$ in a repeated symmetric game is a "**forgiving trigger strategy**" if for every state $q \in Q_i$ for which $\tau_i(q, \cdot)$ is non-constant, there is a sequence of states $(q_0, ..., q_{m(q)+1})$, $m(q) \geq 0$, such that:*

1. $q_0 = q_{m(q)+1} = q$
2. $q_1 = \tau_i(q, a_j)$ *for every* $a_j \neq BR_j(q)$.

3. *For every n such that $0 < n < m(q)+1$, $f_i(q_n) = a^*$ and $\tau_i(q_n, a) = q_{n+1}$ for every $a \in A_j$.*

Forgiving trigger strategies consist of "normal" states and "punishment" states. When player j plays an action $a \neq BR_j(q)$ against a normal state $q \in Q_i$ with a non-constant transition $\tau_i(q, \cdot)$, player i plays a^* for $m(q) \geq 0$ periods. At the end of this punishment phase, player i's automaton returns to q - i.e., the same "situation" in which the original deviation took place.[10] For instance, the automaton s_2 in Figure 1 is a forgiving trigger strategy, with $m(r^1) = 1$. Likewise, the automaton s_1 in Figure 3 is a forgiving trigger strategy, with $m(q^0) = 0$ and $m(q^2) = 1$. Thus, the two main features of punishments in forgiving trigger strategies are: *(i)* the punishment purely consists of min-maxing the opponent, and (ii) the punishment's duration is independent of the deviant player's behavior while being punished.

Forgiving trigger strategies do not impose any restrictions on the ability to sustain cooperation under NE. (We only need to make sure that the $m(q)$'s are sufficiently high.) Under SPE, forgiving trigger strategies (when formulated as full automata) can sustain cooperation if (a_1^*, a_2^*) is a stage-game NE, as in PD. It turns out that under NEWT, this norm of punishment obstructs long-run cooperation in *any* game.

Proposition 6 *There exists no NEWT in forgiving trigger strategies that sustains long-run cooperation in any repeated symmetric game.*

Proof. Assume the contrary. Let (s_1, s_2) be a NEWT in forgiving trigger strategies, such that $a_1^t = a_2^t = a^0$ for every $t \geq t^*$. By assumption, $a^0 \neq br(a^0)$. Therefore, NEWT requires that for every player j and every period $t \geq t^*$: *(i)* $a^0 = BR_j[p_i(t)]$; *(ii)* there exists a period $t_j < t^*$ in $z(s_1, s_2)$, such that $p_i(t_j) = p_i(t)$ and $a_j^{t_j} \neq a^0$. For every player j, let k_j denote the latest such period. Clearly, $k_1 \neq k_2$. Without loss of generality, let $k_1 > k_2$. Observe that by our definition of a^* and a^0, $a^* \neq a^0$.

Let us show that $m[p_2(k_1)] > 0$. Assume the contrary and suppose that $m[p_2(k_1)] = 0$. Then, $p_2(k_1 + 1) = p_2(k_1)$. Recall that $p_2(k_1) = p_2(t)$ for

[10] The intersection between normal and punishment states is not necessarily empty. For instance, when $m(q) = 0$, q is both a normal state and a punishment state: the punishment for playing $a \neq BR_j(q)$ is that player i remains at q.

some $t \geq t^*$, and $BR_1[p_2(t)] = a^0$. But if player 1 can force, by playing $br(a^0)$ against $p_2(k_1)$, infinite repetition of $(br(a^0), a^0)$, then $a^0 \neq BR_1[p_2(t)]$, a contradiction. It follows that $m[p_2(k_1)] > 0$, hence $a_2^{k_1+1} = a^*$.

Let h denote the latest period $t' \in \{k_2, k_2+1, ..., k_1\}$, for which $p_1(t') = p_1(t)$ for some $t \geq t^*$. By the definition of k_2, there exists such a period h. Moreover, since $a_1^{k_1} = br(a^0) \neq a_0$, $h < k_1$. If $a_2^h = a^0$, then $p_1(h+1) = p_1(t+1)$, contradicting the definition of h. Therefore, $a_2^h \neq a^0$. If $m[p_1(h)] = 0$, then $p_1(h+1) = p_1(h)$. But once again, this contradicts the definition of h. Therefore, $m[p_1(h)] > 0$. By the structure of forgiving trigger strategies, player 1 plays a^* repeatedly immediately after period h for a number of periods, before returning to $p_1(h)$. By the definition of h, player 1 returns to $p_1(h)$ only *after* period k_1. If he returns to $p_1(h)$ at period k_1+1, we obtain a contradiction with the definition of k_1, because we have established that player 2 plays $a^* \neq a^0$ at period k_1+1.

It follows that both $p_1(k_1)$ and $p_1(k_1+1)$ are a^*-states with a constant transition. Therefore, $BR_2[p_1(k_1)] = br_2[p_1(k_1)]$ and $BR_2[p_1(k_1+1)] = br_2[p_1(k_1+1)]$. By NEWT, $a_2^{k_1} = a_2^{k_1+1} = br(a^*)$. However, we have shown that player 2 plays two different actions at k_1 and k_1+1, a contradiction. ■

The following example illustrates the reasoning involved in this result. Let the stage game be Chicken or PD. Suppose that (s_1, s_2) is a NEWT in forgiving trigger strategies, which induces $a_1^t = a_2^t = C$ for all $t \geq t^*$. Since $C \neq br(C)$ in either stage game, NEWT requires that for every player j: *(i)* $a_j^{t^*} = BR_j[p_i(t^*)]$; *(ii)* there exists a period $t_j < t^*$ in $z(s_1, s_2)$, such that $p_i(t_j) = p_i(t^*)$ and $a_j^{t_j} = D$. For each player j, let k_j denote the latest such period t_j. Without loss of generality, let $k_1 > k_2$. Then, the play pattern in the periods that immediately precede period t^* is:

Period	k_2	...	k_1	...	t^*
Player 1	C	...	D	...	C
Player 2	D	...	C	...	C

Player 1 is being punished by player 2 between k_2 and t^* for his failure to play C against $p_2(t^*)$ at period k_2. Similarly, player 2 is being punished by player 1 between k_1 and t^* for his failure to play C against $p_1(t^*)$ at period k_1. The proof established that $t^* > k_1 + 1$, such that the latter statement is not vacuous. The structure of forgiving trigger strategies has two implications on the players' behavior when one of them punishes the other. First, the punishing player plays the min-max action (D in both stage games). Second, the punished player plays myopically against the min-max action. But, this leads to a contradiction because player 2 plays C at period k_1 and D at period $k_1 + 1$. One of these actions is not the myopic best-reply to D, hence player 2's behavior violates NEWT.

Proposition 6 implies that forgiving trigger strategies are incapable of sustaining cooperation in any repeated symmetric game. The reason for their restrictiveness is that they force the tests that justify the players' cooperation to take place almost simultaneously. But this means that at least one player i takes a non-myopic action when he is being punished. Such a move would only be justified if player j's strategy contained a "high-order threat" - i.e., a "punishment state" with a non-constant transition. However, forgiving trigger strategies do not contain such punishment states. Thus, the simple punishment structure of forgiving trigger strategies does not allow the tests to be properly coordinated.

6 Concluding remarks

NEWT and experimentation

Under the justifiability-based interpretation of NEWT, each player knows his opponent's strategy; he performs tests not to resolve his own uncertainty, but to convince an outside observer. Let us now explore an alternative, experimentation-based interpretation. Suppose that player j enters the game with some uncertainty regarding the opponent's strategy. He has a prior belief that assigns a high probability to player i's true strategy s_i.[11] However, he is not sure whether the threats he ascribes to player i are real. His prior contains additional strategies, each dispensing with one of these threats.

Given this strategic uncertainty, it is optimal for player j to experiment. Threat testing is broadly consistent with optimal experimentation, if he is patient enough, depending on his exact prior, the stage game's payoffs and the structure of s_i. Calculating the optimal experimentation strategy is a complex task. A boundedly rational player may employ a simplifying rule of thumb: test every threat in s_i (the most probable automaton of the opponent)

[11] Kalai and Lehrer (1993) studied a model of how Bayesian rational players learn to play a repeated-game NE under such circumstances.

that rationalizes a non-myopic action, as soon as the threat is faced. If (s_1, s_2) is both a NEWT and an ε-NE, this means that the heuristic approximates the optimal experimentation strategy. In this case, we may interpret NEWT as a concept that describes how boundedly rational players play a repeated game when they have some initial uncertainty regarding the opponent's strategy.

Related literature

This paper is related to a strand in the literature, which formulates game-theoretic solution concepts that are based on individual decision procedures other than Bayesian rationality. See, for example, Osborne and Rubinstein (1998), Jehiel (2001), and Rabin and Eyster (2001). Spiegler (2002) proposes an alternative approach to justifiability in games. In that paper, I construct an equilibrium concept for extensive games, in which players do not choose strategies that maximize their utility, but strategies that they can justify to an imaginary "critic" in a post-game debate. A strategy is justifiable if for every objection of a certain form that the critic can raise, the player can come up with a "smashing" counter-objection.

References

[1] R. Axelrod, *The Evolution of Cooperation.* Basic Books, 1984.

[2] E. Eyster and M. Rabin, "Cursed Equilibrium", Mimeo, 2001.

[3] P. Jehiel, "Analogy-Based Expectation Equilibrium", *Journal of Economic Theory* (2003), forthcoming.

[4] E. Kalai and E. Lehrer, "Rational Learning Leads to Nash Equilibrium", *Econometrica* **61** (1993), 1019-1045.

[5] D. Kreps, R. Milgrom, J. Roberts and R. Wilson, "Rational Cooperation in the Finitely Repeated PD", *Journal of Economic Theory* **27** (1982), 245-252, 486-502.

[6] M. Osborne and A. Rubinstein, *A Course in Game Theory*, MIT Press, 1994.

[7] M. Osborne and A. Rubinstein, "Games with Procedurally Rational Players", *American Economic Review* **88** (1998), 834-847.

[8] R. Spiegler R, "Equilibrium in Justifiable Strategies: A Model of Reason-Based Choice in Extensive-Form Games", *Review of Economic Studies* **69** (2002), 691-706.

[9] J. Watson, "Cooperation in the Infinitely Repeated PD with Perturbations", *Games and Economic Behavior* **7** (1994), 260-285.