

# Loss Aversion in Repeated Games <sup>\*†</sup>

Jonathan Shalev <sup>‡</sup>

February 16, 1998

## Abstract

The Nash equilibrium solution concept for strategic form games is based on the assumption of expected utility maximization. Reference dependent utility functions (in which utility is determined not only by an outcome, but also by the relationship of the outcome to a reference point) are a better predictor of behavior than expected utility.

In a repeated situation, the value of the previous payoff is a natural reference point for evaluating each period's payoff, and loss aversion implies that decreases are treated more severely than increases. We characterize the equilibria of infinitely repeated games for the case of extreme loss aversion, and show how these are related to the equilibria of stochastic games with state-independent transitions.

**Keywords:** loss aversion, reference dependence, repeated games.

JEL Classification: **C72**.

## 1 Introduction

Expected utility dominates the analysis of game-theoretic situations, despite overwhelming evidence that it fails to adequately describe or predict human behavior. Kahneman and Tversky's (1979) prospect theory proposes an alternative to expected utility in which outcomes are evaluated with respect to a reference point. Such *reference dependent* utility functions are successful in explaining many systematic deviations from the maximization of expected utility.

---

\*Version 0.16, 06/02/98 (First version - 5/97).

†I am grateful to Gaetano Bloise, Fabrizio Germano and Eilon Solan for helpful discussions.

‡CORE, 34 voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium. Fax: +32-10-474301, Phone: +32-10-478186, E-mail: SHALEV@CORE.UCL.AC.BE.

Rabin (1996) writes that “reference dependence deserves to be, and is gradually becoming, an important part of economic modeling.”

The most striking result of the investigation of reference-dependent utility functions is the demonstration of existence of loss aversion. Experimental works in both the psychological and the economic literature suggest that people are motivated to minimize losses (relative to a reference point) much more than they are motivated to maximize gain. For example, Fishburn and Kochenberger (1979) empirically assessed utility functions over changes in wealth. They found that the slope of the utility function below the reference point was on average almost five times as steep as the slope above the reference point. Other examples emphasizing the different treatment of losses and gains (and implicitly or explicitly implying reference dependence) are De Dreu, Emans and Van de Vliert (1992), Kahneman and Tversky (1979), Kahneman, Knetsch and Thaler (1990, 1991), Kramer (1989), Taylor (1991), and Tversky and Kahneman (1992)). Tversky and Kahneman (1991) state that “much experimental evidence indicates that choice depends on the status quo or reference level: changes of reference point often lead to reversals of preference.” The choices dealt with by Tversky and Kahneman were single person decisions. Decision makers in their model evaluate situations while taking into account how the situation was arrived at (reference dependence) and giving different treatment to increases and decreases (loss aversion).

We present a model which implements the consequences of their model for the case of repeated interaction between decision makers. Many interactions between decision makers (persons, corporations, countries or other entities) are repeated continuously over a period of time. It is reasonable to expect the actions taken in each encounter to depend on the outcomes of previous contacts. This mode of repeated encounters is a good candidate for modeling with repeated games. In a repeated game, the same interaction is repeated a number of times, with payoffs at each stage depending on the actions taken by the players (the participants) at that stage. For a survey on repeated games, see Mertens, Sorin and Zamir (1994a, 1994b and 1994c).

The methods generally used in the mathematics and economics literature for evaluating streams of payoffs received in repeated games are (i) the sum of the stage payoffs, (ii) the average payoff, or (iii) a discounted sum of the payoffs. Assuming that loss aversion and reference dependence are relevant factors in evaluating outcomes, one should take into account not

only the stage payoffs themselves (the material payoffs), but also the differences between pairs of payoffs. Thus, the stage payoffs, in addition to being carriers of utility, are also reference points for future payoffs. For instance, an increasing stream of payoffs might be preferred to one with decreases, if the sum of the material payoffs is the same in both cases. As in Tversky and Kahneman (1991), “The basic intuition concerning loss aversion is that losses (outcomes below the reference state) loom larger than corresponding gains (outcomes above the reference state).”

In the model presented here, we assume that the stream of payoffs received in a repeated game is evaluated with the utilities of the outcomes depending on the outcomes of previous rounds. Losses will be regarded more severely than corresponding gains. Depending on the importance of the changes in payoffs relative to the importance of the material payoffs themselves, this change in evaluation can lead to equilibria totally different from those found in models with classical methods of evaluating streams of payoffs.

Shalev (1997a) axiomatized loss aversion in a multi-period model, in which a single decision maker evaluated payoff streams. Preferences over streams of payoffs were characterized with a set of axioms, which include a weaker version of von-Neuman and Morgenstern’s independence axiom, that together accommodate preferences violating temporal monotonicity<sup>1</sup>. Examples have been given both in the economic literature and in that of psychology, where subjects have expressed preferences for streams in a way that violates temporal monotonicity. For example, subjects preferred increasing streams of money to decreasing ones so strongly that even dominated streams were preferred in some cases to those dominated by them. References to such cases are given in Shalev (1997a). The axioms in Shalev’s model are similar to those in Gilboa (1989), which characterize variation aversion. This is no coincidence, and in Section 5 we show an equivalence between variation and losses in a repeated game with infinitely many periods, when the differences between pairs of payoffs (losses and gains) are assumed to have more importance than the actual material payoffs.

Ferreira, Gilboa and Maschler (1995) deal with games in which the utilities of the players may change during the play of the game. They derive an extension of the concept of Nash equilibrium for these games which they call credible equilibrium. Similarly to the situation

---

<sup>1</sup>Temporal monotonicity states that if the outcome of stream  $v_1$  is preferred to the outcome of stream  $v_2$  in every single period, then stream  $v_1$  should be preferred to stream  $v_2$ .

they present, in the model proposed in this paper the utilities of the players from outcomes received in the present can depend on the actions chosen in the past. However, in contrast to their paper, where the changes in utilities are exogenous and given as part of the description of the game, in the model we present here the changes in utilities are endogenous and are derived specifically from differences between pairs of consecutive payoffs.

In Rabin (1993) assumptions about fairness are used to modify the evaluation of outcomes in a game. Similarly to what we do here, the modification of the evaluation of outcomes is endogenous. In both the model presented by Rabin and in the model presented here, the exogenous specification of the game initially includes only the material payoffs, and the psychological assumptions are used to modify the evaluations of the outcomes in a consistent manner. The modifications lead to the set of equilibria being different from the set of equilibria obtained when only the material payoffs are taken into consideration. The results obtained by Rabin have economic significance, reflecting certain stylized facts about fairness. Similarly, the results obtained from the model with loss aversion are significant in that they reflect certain conclusions found in experiments on loss aversion, such as those described in Tversky and Kahneman (1991).

The paper is organized as follows. In Section 2 we examine the equilibria of three well-known games when using loss aversion evaluation, and compare these to the case of regular evaluation. Section 3 formalizes the model. Section 4 discusses finitely repeated games, while Section 5 deals with infinitely repeated games, concentrating on situations with extreme loss aversion. Section 6 contains some final remarks and directions for future research.

## 2 Basic Examples

The payoff evaluation used in the examples in this section is according to the following definitions. A more detailed discussion including the motivation of the definitions is given in Section 3. A (one-shot) game  $G$  is given by  $G = \langle N, S, h \rangle$ , where  $N = \{1, 2\}$  is the set of players<sup>2</sup>,  $S_i, i = 1, 2$  the finite sets of pure strategies of the players;  $S = S_1 \times S_2$ , and  $h : S \rightarrow \mathbb{R}^2$  the payoff function. The game  $G$  is repeated  $T$  times. At each stage  $1 \leq t \leq T$  the players each choose

---

<sup>2</sup>We deal with more than two players in the general model.

simultaneously one of their pure strategies  $s_i^t \in S_i$  (they are allowed to use private randomization devices to implement mixed actions). After each stage, the players are informed of the actions chosen at that stage. Denote the material payoff to player  $i$  in period  $t$  by  $v_i^t = h_i(s_1^t, s_2^t)$ . The loss aversion payoff is defined as  $f_i^t = \min\{0, \lambda_i(v_i^t - v_i^{t-1})\}$ , for  $t > 1$ , where  $\lambda_i \geq 0$  is player  $i$ 's *loss aversion coefficient*, and reflects her degree of loss aversion. Thus, a decrease in payoffs gives a negative loss aversion payoff, while an increase or a repetition of the same payoff gives a loss aversion payoff of zero.

We define the loss aversion utility evaluation of a stream of payoffs by

$$u_i(v_i^1, \dots, v_i^T) = \frac{1}{T} \left[ \alpha \sum_{t=1}^T v_i^t + (1 - \alpha) \sum_{t=2}^T f_i^t \right] \quad (1)$$

The constant  $\alpha \in [0, 1]$  is situation dependent and reflects how relevant loss aversion is to the situation. At the extreme points, if  $\alpha = 0$ , only decreases matter (and are to be avoided for higher utility), and if  $\alpha = 1$  only the material payoffs matter. The case  $\alpha = 1$  is the standard situation of repeated games without taking loss aversion into account and without discounting. A deeper discussion of the payoff evaluation is given in Section 3.

If mixed actions are used, the expectation should be taken over streams, i.e. each possible stream is evaluated according to (1), and an appropriately weighted average of these evaluations gives the final utility. For example, if a strategy profile in a two-stage game gave player  $i$  the stream  $(1, -1)$  with probability  $\frac{1}{3}$  and the stream  $(-1, 1)$  with probability  $\frac{2}{3}$ , then her utility from this lottery over streams is

$$\frac{1}{3} \left( \frac{1}{2} [\alpha(1 - 1) + (1 - \alpha)(-2\lambda_i)] \right) + \frac{2}{3} \left( \frac{1}{2} [\alpha(-1 + 1) + (1 - \alpha)0] \right) = \frac{-\lambda_i(1 - \alpha)}{3}.$$

Note that the payoffs for player  $i$  are first defined on the streams of outcomes, and then defined for mixtures. This sequence is important.

Now that we have defined the payoff evaluation, we give some examples of games, and examine some equilibria obtainable in these games with loss-aversion evaluation.

## 2.1 Finitely Repeated Prisoners' Dilemma

In the first example, loss-aversion evaluation can induce cooperation in the first periods of the finitely repeated prisoners' dilemma. With regular evaluation, defection at all stages is the only equilibrium outcome. The material payoffs at each stage are given in the following table:

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

C and D represent cooperation and defection respectively. The stage game is repeated  $T$  times. For  $\alpha$  close to one, the only equilibrium is the well-known one with both players defecting at every stage. However, for  $\alpha$  sufficiently close to zero (a situation where extreme loss-aversion is assumed) and strictly positive loss-aversion coefficients  $\lambda_i$ , we can find equilibrium paths with cooperation continuing until the stage before last. An example of such a strategy pair that is in equilibrium is the following. Both players have the same strategy, which doesn't depend on the history, and consists of playing C in the first  $T - 1$  periods, and D in the last period. The payoff for each player is  $4\alpha - (1 - \alpha)\frac{3\lambda_i}{T}$ . For  $\alpha$  close enough to zero this cannot be improved by any deviation, since the first term is negligible, and given the strategy of the other player, with any deviation the multiplier of  $(1 - \alpha)\frac{\lambda_i}{T}$  is strictly less than  $-3$  (the sum of the decreases). For any  $T$ , the path with both players playing D at each stage can be supported in equilibrium, showing that loss-aversion evaluation enables cooperation, but does not necessarily induce it.

## 2.2 Finitely Repeated Battle of the Sexes

The second example has the material payoffs of the stage game given by the battle of the sexes. These payoffs are:

	A	B
A	2, 1	0, 0
B	0, 0	1, 2

If the game is repeated twice, with  $\alpha$  sufficiently close to zero, then there are four pure strategy equilibrium outcomes. In two of these, one of the coordinated outcomes is repeated twice. In the other two, the equilibrium has the first round outcome of  $(B, A)$ , giving payoffs

of  $(0,0)$ , and then either  $(A,A)$  or  $(B,B)$ . A deviation by a player in the first round leads to the players playing that players least preferred coordinated outcome in the second round. There are other equilibria also (with mixed actions being played), such as the following, also with  $\alpha$  close to zero: both players play mixed actions in the first round, player 1 playing A with probability  $5/7$  and player 2 playing A with probability  $2/7$ . In the second round, if the outcome in the first round was  $(A,A)$  or  $(B,B)$ , this outcome is repeated. If the outcome was one of the non-coordinated ones, they play mixed actions, with player 1 playing A with probability  $2/3$  and player 2 playing A with probability  $1/3$ .

For more than 2 repetitions, the analysis of equilibria becomes more complex. For instance, it is not true that the players' actions in the last period (period T) always maximize their expected stage payoff, given the other players action. In some cases they may be willing to receive a lower expected payoff in round T if this gives a smaller chance of a loss relative to the outcome in round T-1. To see this, we now examine in detail the equilibrium actions in the last period of the thrice-repeated battle of the sexes, based on the outcome of the second period (the only relevant historical information).

1. Outcome of  $(0,0)$  in period 2

In this case, since there can be no losses when comparing the second and third periods, the only factor is maximization of expected payoffs for the third period, i.e. any of the 3 Nash equilibria of the stage game could be part of an equilibrium.

2. Outcome of  $(2,1)$  in period 2

$(A,A)$  and  $(B,B)$  are the only pure action profiles that could constitute part of an equilibrium, and they do so for any  $\lambda$ 's and any  $\alpha$ . Mixed actions that could constitute part of an equilibrium are those for which player 2 plays A with probability  $\frac{1}{3}$  and player 1 plays A with probability  $\frac{2\alpha+(1-\alpha)\lambda_2}{3\alpha+2(1-\alpha)\lambda_2}$ . When  $\alpha = 1$  this is equal to  $2/3$ , and when  $\alpha = 0$  this is equal to  $1/2$ . Thus, as losses become more important, player 1 places more weight on the action that could lead to player 2's favourable outcome. Note that player 2 is indifferent between her two actions (which is why she can play a mixed action), even though they give her different expected material payoffs.

3. Outcome of  $(1,2)$  in period 2

This case is analogous to the previous one, reversing the roles of the two players.

### 2.3 Finitely Repeated Matching Pennies

The last example is the game of matching pennies, with the following stage payoff matrix:

	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

At the last stage, both players will mix their actions with equal probabilities in any equilibrium, regardless of the values of the  $\lambda_i$ 's and  $\alpha$ . If the number of stages is large enough, then pure actions can be supported in equilibrium at earlier stages with threats of punishment after deviation from the equilibrium path. There also exists an equilibrium with both players playing  $(1/2, 1/2)$  unconditionally at all stages. We now investigate the possible equilibria in a two-stage game. At the second stage, both players prefer 1 to  $-1$ , for any values of  $\lambda_i$  and  $\alpha$ , after any first period outcome. Therefore, any equilibrium has both players mixing H and T with probability  $\frac{1}{2}$  each. Moving now to the first stage, and taking as given the second stage actions, an outcome of 1 at the first stage gives player  $i$  a (loss aversion) utility of  $\frac{\alpha - (1-\alpha)\lambda_i}{2}$ , and an outcome of  $-1$  gives  $\frac{\alpha}{2}$ . Thus, for  $\lambda_i > \frac{2\alpha}{1-\alpha}$ , player  $i$  prefers  $-1$  to 1 at the first stage. If  $\alpha$  is small enough, and  $\lambda_i > 0$  for  $i = 1, 2$ , then the only possible equilibrium has both players mixing equiprobably also at stage 1. A similar analysis shows that for a 3-stage game with  $\alpha$  close enough to zero, there are no equilibria with pure actions at the first stage. All equilibria start with both players playing each of their actions equiprobably. In the second stage of an equilibrium path there could either be pure actions, repeating the realized outcome of the first stage, or equiprobable mixtures by both players. The third and last stage of any equilibrium is with equiprobable mixtures by both players. When there are more than 3 stages, pure actions can be supported in equilibria at the beginning of the game, with threats of punishment (by mixing) for deviating at the first stage. Such punishments are not necessary for stages after the first one and before the last 2, as repeating the previous outcome is optimal for both players. To summarize, this simple example shows that when differences between pairs of outcomes are taken into consideration, the set of equilibria changes considerably. This same game, without loss aversion (either  $\lambda_i = 0 \quad \forall i$  or  $\alpha = 1$ ) admits only one equilibrium path - both players mixing equiprobably at each and every stage.

### 3 Notation and the Underlying Model

We use the following notation, some of which was given for the examples of the previous section, and is repeated here for convenience.

We start with the elements of a repeated game with loss aversion evaluation. The first three items define the stage game.

The set of players is  $N = \{1, 2, \dots, n\}$ .

The (finite) set of pure actions of player  $i$  in the stage game is  $S_i$ . The set of action  $n$ -tuples is  $S = S_1 \times S_2 \times \dots \times S_n$ .

The payoff function for player  $i$  in the stage game (material payoffs) is given by  $h_i : S \rightarrow \mathbb{R}$ . The vector of payoffs for all the players is denoted  $h = h_1 \times h_2 \times \dots \times h_n$ .

This defines the stage game with the material payoffs,  $\langle N, S, h \rangle$ .

The number of stages in the game is  $T$ .  $T \in \{1, 2, \dots\} \cup \infty$ . The case of finite  $T$  is treated in Section 4, and infinitely repeated games are addressed in Section 5.

The loss aversion coefficients of the players are  $\lambda = (\lambda_i)_{i \in N}$ , with  $\lambda_i \geq 0$  for each  $i$ . Higher values of  $\lambda_i$  indicate greater loss aversion of player  $i$ .  $\lambda_i = 0$  indicates that player  $i$  is not loss averse, and cares only about material payoffs.

The constant  $\alpha \in [0, 1]$  is situation dependent and reflects how relevant loss aversion is to the situation. The closer  $\alpha$  is to 0, the more relevant are the losses relative to the material payoffs. This is analogous to the scaling factor  $X$  used in Rabin (1993), where higher values of  $X$  give more importance to the material payoffs. We assume here for convenience that  $\alpha$  is common for all the players, but a model could be developed where  $\alpha_i$  gave the the relevance of the losses for player  $i$ . This cannot be done simply by modifying the  $\lambda_i$ 's.

The above definitions are the data of the game. The following items give the strategies, and the evaluation of the payoffs.

The set of possible histories until (not including) stage  $t$  is denoted by  $H^t = S^{t-1}$ .

A (behavioral) strategy of player  $i$  in the repeated game is given by  $\sigma_i = (\sigma_i^t)_{t=1, \dots, T}$ , where  $\sigma_i^t : H^t \rightarrow \Delta(S_i)$  is the single period (mixed) action of player  $i$  for period  $t$ .

We denote by  $\Sigma_i$  the set of player  $i$ 's strategies in the repeated game.

A profile of actions in period  $t$  is  $s^t = (s_i^t)_{i \in N}$ .

$v_i^t = h_i(s^t)$  denotes the material payoff to player  $i$  in period  $t$ . The material payoffs for the set of players is given by  $v^t = (v_i^t)_{i \in N}$ .

The loss aversion payoff for player  $i$  at period  $t$  is defined as  $f_i^t = \min\{0, \lambda_i(v_i^t - v_i^{t-1})\}$ , for  $t \geq 2$ . The function  $f$  gives the disutility obtained from losses (relative to the previous period). In a similar fashion one could add utility from gains (which loss aversion implies would be less than the absolute value of the disutility of comparable losses). One of the simplifying assumptions we make is ignoring the effect of gains and focusing only on the effect of losses. When the game is repeated infinitely, this is without loss of generality, as we show in Lemma 1 that the average increase in payoffs (between adjacent stages) is equal to the average decrease in payoffs. Shalev (1997a) gives a more general model (in a decision theoretical setup, i.e. with a single player) in which both gains and losses are taken into account.

The utility evaluation of a stream, as used in the examples of Section 2, is given by

$$u_i(v_i^1, \dots, v_i^T) = \frac{1}{T} \left[ \alpha \sum_{t=1}^T v_i^t + (1 - \alpha) \sum_{t=2}^T f_i^t \right] \quad (2)$$

This is player  $i$ 's  $T$ -period utility evaluation.

The function  $u$  is based on a multiperiod representation of the value function used by Tversky and Kahneman (1991), incorporating loss aversion and reference dependence. The reference point for evaluating the payoff at stage  $t$  (for  $t > 1$ ) is  $v_{t-1}$ . Thus, we assume that the reference points fully adjust to new payoffs, the moment they are received. For a discussion on the speed of adjustment of reference points and the implications for different equilibrium concepts, see Shalev (1997b).

A strategy profile  $\sigma$  induces a distribution over the sequences of outcomes. As described in Section 2, each stream is evaluated according to (2), and the expectation is taken over these evaluations. As usual,  $\sigma$  is a *loss-aversion (Nash) equilibrium* if no player  $i$  can strictly gain from a unilateral deviation to a strategy  $\sigma'_i$ .

## 4 Remarks on Equilibria of the Finitely Repeated Game

### 1. Repetition of a Nash equilibrium of the stage game

Without taking loss aversion into account, the repetition of a Nash equilibrium of the stage game always represents an equilibrium path of a finitely repeated supergame (e.g. with strategies independent of the history). This is true also when loss aversion is used in the evaluation of the payoff streams<sup>3</sup>. If the number of repetitions is large enough, and  $\alpha$  is close enough to zero, *any* pair of actions repeated at each stage (and not just those representing a Nash equilibrium of the stage game) can be supported in equilibrium for most of the periods of the repeated game (the last few may have to include mixed actions).

### 2. Zero-sum stage games

If one is only interested in the material payoffs (and not the loss aversion payoffs), then for a zero-sum stage game the equilibria are with each player's expected stage payoffs being equal to her value in the stage game. However, with loss aversion evaluation, the repeated game is not generally zero-sum, as the evaluation of differences is not symmetrical for positive and negative payoffs. Since one player's gain is the other player's loss, and losses are evaluated more severely than gains, the repeated zero-sum stage game with loss aversion evaluation becomes a game with non-positive payoffs when loss-aversion is highly relevant ( $\alpha$  near zero).

### 3. Mixed actions and loss aversion evaluation

With (strictly) mixed actions at stage  $t - 1$  and loss aversion evaluation of payoffs, the players' actions for stage  $t$  ( $2 \leq t \leq T$ ) will in general depend on the *outcome* of stage  $t - 1$ , which is known only after stage  $t - 1$ . This is in contrast with the situation where the evaluation is simply the sum (or weighted sum, discounted sum, or average) of the material payoffs, in which case, along the equilibrium path, the actions need not depend upon the outcomes of previous randomizations. The expectation of the payoff for the en-

---

<sup>3</sup>It is interesting to note that if we include a premium for gains (in addition to the penalty for losses), then this is not necessarily true. For this method of evaluation, with  $\alpha$  close to zero, the players prefer to have as low a payoff as possible in the first period. In this case, some single-stage Nash equilibria cannot arise as the first stage of equilibria of the repeated game. An example is the action profile  $(A,A)$  for the twice-repeated battle of the sexes, which cannot appear as the first stage of an equilibrium.

tire game with loss-aversion evaluation is not just a function of the expected one-round payoffs, but a function of the expected one-round payoffs and the expected differences (and the direction of these differences) between pairs of adjacent stage payoffs. Therefore, to calculate the expected future payoff from any point in the game, one must take into consideration the realization of the previous outcome.

#### 4. Existence of loss-aversion equilibria

For every finitely repeated game with loss-aversion evaluation of the payoffs there exists a loss aversion equilibrium. One can regard the repeated game as an extensive game in which the payoffs at each terminal node are calculated according to the loss aversion formulae given in the model. This is a finite game, and therefore, any Nash equilibrium of this game is a loss aversion equilibrium of the (equivalent) repeated game with loss aversion evaluation. Since a Nash equilibrium is known to exist for any such game, a loss aversion equilibrium exists for the equivalent game with loss aversion evaluation.

#### 5. Dominating actions

Strictly dominating actions in the stage game need not be played in a loss aversion equilibrium (except in the last stage). This can be seen in the example of the prisoners' dilemma in the previous section. With regular evaluation (utility is the sum of the stage payoffs), backwards induction leads us to infer that the only equilibrium path has the players playing their dominating action at each stage.

## 5 Loss Aversion in Infinitely Repeated Games

Given an infinite stream of stage payoffs  $(v_i^1, v_i^2, \dots)$ , we use the following formula (limit of averages) to evaluate the loss-aversion payoff:

$$u_i(v_i^1, v_i^2, \dots) = \limsup_{T \rightarrow \infty} \frac{1}{T} \left[ \alpha \sum_{t=1}^T v_i^t + (1 - \alpha) \sum_{t=2}^T f_i^t \right]. \quad (3)$$

A strategy profile  $\sigma$  is a loss-aversion equilibrium for the infinite case if the expected payoff from  $\sigma$  cannot be improved upon by any player with a unilateral deviation from  $\sigma_i$ .

It is interesting to note that an infinitely repeated game with loss aversion payoffs can be modeled by an appropriate stochastic game. The set of states of the game is the set of action tuples, that we denoted by  $S$ , plus one initial state,  $s^0$ . The transition matrix is simply that from any state, playing actions  $s \in S$  causes a sure transfer to state  $s$ . Such transition matrices characterize the game as a stochastic game with state-independent transitions. The stage payoff of player  $i$ , given state  $s \in S$  and actions  $s' \in S$ , is  $\alpha h_i(s) + (1 - \alpha)(h_i(s) - h_i(s'))^-$ . The payoffs from actions  $s$  at state  $s^0$  are  $h(s)$ . Taking the limit of the averages evaluation, we arrive at the formula (3).

Recall that the purpose of  $\alpha$  is to represent the importance of the material payoffs relative to the importance of the changes in the material payoffs over time. When  $\alpha$  is close to one this represents low relevance of loss aversion, and when  $\alpha$  is close to zero this represents a situation where loss aversion is highly relevant. Two interesting cases are the extremes. The case  $\alpha = 1$  is when only the material payoffs count, which is equivalent to a regular repeated game without discounting. The case  $\alpha = 0$  is when the material payoffs are “immaterial”, and only losses count. This is the situation we investigate in this section. The results reflect fully the loss aversion aspects of the evaluation. We are not pretending that this is a realistic assumption for common situations, but the analysis of this extreme case can be compared to results for  $\alpha \neq 0$  and  $\alpha = 1$ . Similar propositions to the three given in this section hold for intermediate values of  $\alpha$ . A loss aversion equilibrium in an infinitely repeated game with  $\alpha = 0$  is characterized by each player minimizing the average decrease in her payoff stream. As shown in the following lemma, the utility of a stream of payoffs  $v = (v^1, v^2, \dots)$  to a player  $i$  is equal to  $-\lambda_i$  (the loss aversion coefficient of the player) multiplied by the average decrease between pairs of payoffs (where increases are treated as a decrease of zero).

Note that the utility of a payoff stream  $v = (v^1, v^2, \dots)$  when  $\alpha = 0$  is given by  $\lambda_i \lim_{T \rightarrow \infty} \sup \frac{1}{T} \sum_{t=2}^T (v_i^t - v_i^{t-1})^-$ , where  $x^-$  denotes  $\min(0, x)$ .

**Lemma 1** *For any payoff stream  $v = (v^1, v^2, \dots)$  from a repeated game with  $\alpha = 0$ , the following equations hold:*

$$u_i(v) = \lambda_i \lim_{T \rightarrow \infty} \sup \frac{1}{T} \sum_{t=2}^T (v_i^t - v_i^{t-1})^-$$

$$\begin{aligned}
&= \lambda_i \limsup_{T \rightarrow \infty} \frac{-1}{T} \sum_{t=2}^T (v_i^t - v_i^{t-1})^+ \\
&= \lambda_i \limsup_{T \rightarrow \infty} \frac{-1}{2T} \sum_{t=2}^T |v_i^t - v_i^{t-1}|
\end{aligned}$$

**Proof:** For any  $T \geq 2$  it is true that

$$\sum_{t=2}^T (v_i^t - v_i^{t-1}) = v_i^T - v_i^1$$

and that

$$\sum_{t=2}^T (v_i^t - v_i^{t-1}) = \sum_{t=2}^T (v_i^t - v_i^{t-1})^+ + \sum_{t=2}^T (v_i^t - v_i^{t-1})^-$$

Therefore, since

$$u_i(v) = \lambda_i \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=2}^T (v_i^t - v_i^{t-1})^-,$$

we have

$$\begin{aligned}
u_i(v) &= \lambda_i \limsup_{T \rightarrow \infty} \frac{1}{T} \left( (v_i^T - v_i^1) - \sum_{t=2}^T (v_i^t - v_i^{t-1})^+ \right) \\
&= \lambda_i \limsup_{T \rightarrow \infty} \frac{-1}{T} \sum_{t=2}^T (v_i^t - v_i^{t-1})^+
\end{aligned}$$

since  $v_i^T$  and  $v_i^1$  are both bounded, as game payoffs. The last equality now follows from the fact that

$$\sum_{t=2}^T |v_i^t - v_i^{t-1}| = \sum_{t=2}^T (v_i^t - v_i^{t-1})^+ - \sum_{t=2}^T (v_i^t - v_i^{t-1})^-$$

for all  $T \geq 2$ .

■(Lemma 1)

Since any payoff stream will give a non-positive payoff evaluation, a payoff of zero cannot be improved upon. Thus, any strategy path where the players play fixed actions is in equilibrium, and is also efficient. To fully characterize the equilibria in the infinitely repeated game, we will use a modification of the Folk Theorem. Any equilibrium is characterized by the payoffs being feasible and individually rational for all the players. We now investigate what are the feasible and individually rational payoffs for loss-aversion evaluation. First we investigate the feasible payoffs.

## 5.1 Feasibility

Given a stage game  $G = (N, S, h)$ , the feasible payoffs at each stage belong to the set  $V = \{(h_1(s), h_2(s), \dots, h_n(s)) \mid s \in S\}$ . The following proposition characterizes the nature of the payoff evaluations that are feasible given a stage game  $G$ . First, some notation. Define a cycle of length  $k$  ( $k \geq 1$ ) as a  $(k+1)$ -tuple  $c = (v^1, v^2, \dots, v^{k+1} = v^1)$ , where  $v^s \neq v^t$  for  $s \neq t$ ,  $s, t \in \{1, 2, \dots, k\}$ , and each  $v^t$  belongs to  $V$ . Since the number of outcomes is finite, so is the number of possible cycles. For a cycle  $c$  of length  $k$ , define the average cycle losses for the players, denoted by  $\delta(c) = (\delta_1(c), \delta_2(c), \dots, \delta_n(c))$ , by  $\delta_i(c) = \frac{1}{k} \sum_{t=1}^k (v_i^{t+1} - v_i^t)^-$ ,  $i \in N$ . Define  $\delta_G$  as the set of average cycle losses from cycles over outcomes in  $G$ . This set is also finite.

**Proposition 1** *The set of feasible outcomes in the infinitely repeated game  $G_\infty$  with loss aversion evaluation and  $\alpha = 0$  is equal to the set of convex combinations of elements of  $\delta_G$ , multiplied (coordinatewise) by  $\lambda$ .*

**Proof:** Take an infinite payoff stream  $v = (v^1, v^2, \dots)$ . We show that the utilities of the players from this stream is equal to a convex combination of elements of  $\delta_G$  multiplied (coordinate by coordinate by)  $\lambda$ .

Fix  $T \geq 2$ , and take  $\tilde{v}^T = (v^1, v^2, \dots, v^T)$ . Find the first cycle in  $\tilde{v}^T$ . Denote it  $c_1$  and its length  $k_1$ . Remove the first  $k_1$  elements of this cycle (all except the last one) from  $\tilde{v}^T$ , and we are left with a payoff stream  $k_1$  elements shorter than  $\tilde{v}^T$ . Denote it  $\tilde{v}^{T_1}$ . From the definition of

$u_i$ ,

$$u_i(\tilde{v}^T) = \frac{\lambda_i}{T} \sum_{t=2}^T (v^t - v^{t-1})^- = \frac{\lambda_i k_1}{T} \delta_i(c) + \left( \frac{T - k_1}{T} \right) u_i(\tilde{v}^{T_1}).$$

Continue with the first cycle remaining in  $\tilde{v}^{T_1}$ , denoting it  $c_2$  and its length  $k_2$ . Continue until no cycles remain. Denote the number of cycles by  $r$  and the length of the remaining stream  $\tilde{v}^{T_r}$  by  $k_{r+1}$ . Since  $\tilde{v}^{T_r}$  contains no cycles, we have  $k_{r+1} \leq |S|$ . Thus,

$$u_i(\tilde{v}^T) = \lambda_i \sum_{j=1}^r \frac{k_j}{T} \delta_i(c_j) + \lambda_i \frac{k_{r+1}}{T} u_i(\tilde{v}^{T_r}).$$

When  $T$  approaches infinity, the last term is negligible, and

$$u_i(v) = \limsup_{T \rightarrow \infty} \lambda_i \sum_{j=1}^{r(T)} \frac{k_j}{T} \delta_i(c_j)$$

for all  $i \in N$ , so  $u(v)$  is a convex combination of average cycle losses, multiplied coordinatewise by  $\lambda$ .

For the other direction, we are given a convex combination,

$$u_i = \lambda_i \sum_{j=1}^r \alpha_j \delta_i(c_j)$$

with  $\sum_{j=1}^r \alpha_j = 1$ ,  $\alpha_j > 0 \ \forall j$ , and  $\delta(c_j) \in \delta_G \ \forall j$ . Take  $\mu_j = \lfloor 2^s \alpha_j \rfloor$ , for fixed  $s \geq 1$ . Create a stream of outcomes giving each cycle  $\mu_j$  times, excluding the last element of the cycle each time except the last. The number of cycles is no greater than  $2^s$ , so we have a finite length stream. Repeat this stream infinitely many times, and we have a payoff stream whose payoff is close to that of the finite stream. As we increase  $s$ ,  $\frac{\mu_j}{\sum_{k=1}^r \mu_k}$  gets closer to  $\alpha_j$ , and the payoff valuation approaches the desired one. As  $s$  tends to infinity, we receive the desired result.

■(Proposition 1)

We have thus characterized the set of feasible payoffs that can be achieved by the players. We now examine which of these payoffs are individually rational, and therefore can be reached in an equilibrium.

## 5.2 Individual Rationality

A payoff is individually rational for a player if it is at least as large as the minimal payoff that she can ensure for herself (against all the others acting together, but not with correlated mixtures). Thus,  $x_i$  is an individually rational payoff for player  $i$  if

$$x_i \geq \min_{\sigma_{-i} \in \Sigma_{-i}} \max_{\sigma_i \in \Sigma_i} U_i(\sigma)$$

A strategy  $\sigma_{-i}$  achieving this value will be denoted an optimal punishment strategy against player  $i$ . Our first step is to characterize optimal punishment strategies. First we look at a number of examples, with two players. The punisher in each case is player 2, the column player.

Example 1:

	L	R
T	1, 1	0, 0
B	0, 0	1, 1

The severest punishment that player 2 can inflict is by mixing her two pure actions with equal probabilities each round. Whatever strategy player 1 follows, her expected payoff change per period is  $\frac{1}{2}$ , giving an average loss of  $\frac{1}{4}$  per period.

Example 2:

	L	R
T	1, 1	0, 0
B	1, 1	0, 0

For this game the severest punishment player 2 can inflict on player 1 is by playing  $L$  at even stages and  $R$  at odd stages. This gives a fixed change of 1 per period for player 1, which is an average loss of  $\frac{1}{2}$  per period.

Example 3:

	L	M	R
T	6, 0	0, 0	4, 0
B	0, 0	6, 0	4, 0

If player 2 mixes between L and M

with equal probabilities at each stage, player 1 has an average change of 3 per period. However, player 2 can punish even more severely. An optimal punishment strategy of player 2 is: play  $R$  in stage 1. For stages after the first, if the outcome gave player 1 a payoff of 0, then play  $R$ . Otherwise mix  $L$  and  $M$  with equal probabilities. The payoff stream will contain approximately  $\frac{1}{3}$  each of the three outcomes 6, 4 and 0 for player 1. Every 0 is followed by a 4, every 4 is

followed by either a 6 or a 0 (equal probabilities), and every 6 is followed by either a 6 or a 0 (equal probabilities). Taking the averages, the expected change per period in player 1's payoffs is  $\frac{10}{3}$ , which is greater than 3.

From these three examples, it can be seen that optimal punishment strategies must sometimes include almost only mixed actions (i.e. the set of stages at which a pure action is taken is finite), must sometimes include almost only pure actions, must sometimes include an infinite number of both, and may depend on the history. The following proposition establishes the existence of an optimal punishment strategy which is stationary (i.e. it depends only on the previous outcome).

**Proposition 2** *An optimal punishment strategy against player  $i$  exists, in which the actions of the other players at each stage are a function only of the actions in the last stage of the history, and this function is the same at every stage (i.e. it is stationary).*

**Proof:** The problem of finding an optimal punishment strategy is equivalent to finding an optimal punishment strategy in the following stochastic game. The states are the possible action  $n$ -tuples of the stage game, previously denoted by  $S$ . The payoff to player  $i$  at stage  $t + 1$  from the actions at this stage is a function of her payoff in the stage game at stage  $t$  (which is a function of the state) and her payoff in the stage game at stage  $t + 1$ . This function is exactly  $(v_{t+1} - v_t)^- = (h_i(s_{t+1}) - h_i(s_t))^-$ , where  $s_{t+1}$  is the action tuple at stage  $t + 1$  and  $s_t$  is the state at stage  $t + 1$ , which is the action tuple at stage  $t$ . The transition rule is that (from any state at stage  $t$ ), the new state at stage  $t + 1$  is the action tuple played at stage  $t$ , i.e.  $s_t$ . This stochastic game has the state-independent-transition property, and therefore according to Solan (1998), the other players  $N \setminus \{i\}$  have stationary optimal punishment strategies when the evaluation of the payoffs is according to the limit of the averages, as it is in our game. ■(Proposition 2)

In calculating the optimal punishment strategy (according to the proposition it is sufficient to find the optimal action after each outcome), one must take into account both the expected one-round differences, and also the expected distribution over the outcomes before the next stage. As the following proposition shows, it is not enough to maximize the one-round differences in a “greedy” manner.

**Proposition 3** *Maximizing the expected one-round differences in material payoffs for a player (playing “greedy”) is not necessarily an optimal punishment.*

**Proof:** To prove this statement it suffices to find a game in which “greedy” is not an optimal punishment. We take the stage game to be the battle of the sexes (Section 2.2). Player 2 is the punisher.

We first calculate the “greedy” strategy for Player 2. There are three states, corresponding to outcomes 0, 1, and 2 for Player 1. For each state  $k \in \{0, 1, 2\}$ , the strategy specifies a probability  $q_k$  of playing the first action. Denote a strategy by the triplet  $(q_0, q_1, q_2)$ , with  $q_k$  being the probability that Player 1 plays A in round  $t + 1$  if the payoff in round  $t$  to Player 1 was  $k$ . Simple calculations show that the optimal strategy to maximize expected one-round differences is  $(\frac{1}{3}, 1, \frac{1}{3})$ .

Now assume that “greedy” is an optimal punishment strategy, i.e. it ensures that the punishment to Player 1 is at least  $v$ , with  $v$  being the value of the stochastic game (we will reach a contradiction to this assumption).

By the same arguments used in the previous proposition, Player 1 has an optimal stationary strategy ensuring that the punishment will not exceed  $v$ . In particular, this strategy is a best response against the “greedy” strategy played by Player 2, since we assumed that Player 1 cannot get a punishment less severe than  $v$  against it.

Calculation shows that any stationary best response against “greedy” is of the form  $\sigma_1 = (1, 0, r_2)$ , with  $r_2$  any probability. Since we assumed that “greedy” is optimal for Player 2, for some value of  $r_2$ ,  $\sigma_1$  is optimal for Player 1. However, there is a better strategy than “greedy” against *any* such  $\sigma_1$ , which is  $(0, 1, 1)$ .

Therefore, assuming that “greedy” is an optimal punishment strategy for Player 2 leads us to the fact that none of the  $\sigma_1$  strategies are optimal, whereas the same assumption led us to the fact that one of them is optimal. Therefore, “greedy” is not optimal. ■ Proposition 3

## 6 Concluding Remarks

In the previous sections, we have examined the effects of assuming reference dependence and loss aversion in repeated game situations. The exact assumptions we used focused on inter-

period losses, while disregarding the effect of interperiod gains. This kind of assumption has a drawback when the number of periods is finite, and there is scope to include the effect of gains in future research on finitely repeated games. Another restriction which could be relaxed is the assumption of extreme loss aversion (the utility of material payoffs is negligible relative to the utility of changes). However, the tradeoff between utility of material payoffs and utility of changes severely complicates the analysis. As the three examples at the beginning of Section 5.2 show, the construction of “optimal” punishments is not trivial, and while giving an existence proof, we leave open the question of how to construct an optimal punishment in general situations.

## Bibliography

- De Dreu, C. K. W., B. J. M. Emans and E. Van de Vliert (1992): “Frames of Reference and Cooperative Social Decision Making,” *European Journal of Social Psychology* **22** 297-302.
- Ferreira, J.-L., I. Gilboa and M. Maschler (1995): “Credible Equilibria in Games with Utilities Changing during the Play,” *Games and Economic Behavior* **10** 284-317.
- Fishburn, P. C. and G. A. Kochenberger (1979): “Two piece Von Neumann - Morgenstern Utility Functions,” *Decision Sciences* **10** 503-518.
- Gilboa, Itzhak (1989): “Expectation and Variation in Multi-Period Decisions,” *Econometrica* **57** 1153-1169.
- Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler (1990): “Experimental Tests of the Endowment Effect and the Coase Theorem,” *Journal of Political Economy* **98** (6) 1325-1348.
- Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler (1991): “The Endowment Effect, Loss Aversion and Status Quo Bias,” *Journal of Economic Perspectives* **5** (1) 193-206.
- Kahneman, Daniel and Amos Tversky (1979): “Prospect Theory: An Analysis of Decision Under Risk,” *Econometrica* **47** 263-291.

- Kramer, R. M. (1989): "Windows of Vulnerability or Cognitive Illusions: Cognitive Processes and the Nuclear Arms Race," *Journal of Experimental Social Psychology* **25** 79-100.
- Mertens, Jean-François, Sylvain Sorin and Shmuel Zamir (1994a): "Repeated Games, Part A: Background Material," CORE Discussion Paper 9420, Université Catholique de Louvain, Louvain la Neuve, Belgium.
- Mertens, Jean-François, Sylvain Sorin and Shmuel Zamir (1994b): "Repeated Games, Part B: The Central Results," CORE Discussion Paper 9421, Université Catholique de Louvain, Louvain la Neuve, Belgium.
- Mertens, Jean-François, Sylvain Sorin and Shmuel Zamir (1994c): "Repeated Games, Part C: Further Developments," CORE Discussion Paper 9422, Université Catholique de Louvain, Louvain la Neuve, Belgium.
- Rabin, Matthew (1993): "Incorporating Fairness into Game Theory and Economics," *American Economic Review* **83** (5) 1281-1302.
- Rabin, Matthew (1996): "Psychology and Economics," Mimeo, University of California, Berkeley.
- Shalev, Jonathan (1997a): "Loss Aversion in a Multi-Period Model," *Mathematical Social Sciences* **33** 203-226.
- Shalev, Jonathan (1997b): "Loss Aversion Equilibrium," CORE discussion paper 9774, Université Catholique de Louvain, Louvain la Neuve, Belgium.
- Solan, Eilon (1998): "Stochastic Games with State Independent Transitions," Mimeo, Hebrew University of Jerusalem.
- Taylor, S. E. (1991): "Asymmetrical Effects of Positive and Negative Events: The Mobilization-Minimization Hypothesis," *Psychological Bulletin* **110** 67-85.
- Tversky, Amos and Daniel Kahneman (1991): "Loss Aversion in Riskless Choice: a Reference Dependent Model," *Quarterly Journal of Economics* **106** (4) 1039-1061.
- Tversky, Amos and Daniel Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, **5** 297-323.