Driesen, B., Perea, A., and Peters, H. (2010) On loss aversion in bimatrix games. Theory and Decision, 68 (4). pp. 367-391. ISSN 0040-5833

Theory and Decision (2010) 68:367–391
DOI 10.1007/s11238-008-9102-6

BRAM DRIESEN, ANDRÉS PEREA, and HANS PETERS

ON LOSS AVERSION IN BIMATRIX GAMES

ABSTRACT. In this article three different types of loss aversion equilibria in bimatrix games are studied. Loss aversion equilibria are Nash equilibria of games where players are loss averse and where the reference points—points below which they consider payoffs to be losses—are endogenous to the equilibrium calculation. The first type is the fixed point loss aversion equilibrium, introduced in Shalev (2000; Int. J. Game Theory 29(2):269) under the name of 'myopic loss aversion equilibrium.' There, the players' reference points depend on the beliefs about their opponents' strategies. The second type, the maximin loss aversion equilibrium, differs from the fixed point loss aversion equilibrium in that the reference points are only based on the *carriers* of the strategies, not on the exact probabilities. In the third type, the safety level loss aversion equilibrium, the reference points depend on the values of the own payoff matrices. Finally, a comparative statics analysis is carried out of all three equilibrium concepts in 2×2 bimatrix games. It is established when a player benefits from his opponent falsely believing that he is loss averse.

KEY WORDS: bimatrix games, loss aversion, reference-dependence

JEL CLASSIFICATION: C72

1. INTRODUCTION

Since von Neumann and Morgenstern (1944) developed expected utility theory, it has been the dominant approach in decision-making under uncertainty. However, the use of expected utility in the economics of uncertainty has also been challenged on empirical grounds. One of the earliest examples is the Allais paradox (Allais, 1953).[1]

Several alternative models for decision-making under uncertainty have been proposed. One of the most successful non-expected utility theories is 'prospect theory,' developed by

Kahneman and Tversky (1979). This theory assumes that economic agents make choices between lotteries in two phases: an editing phase and an evaluation phase. In the editing phase, agents observe and interpret the options between which they must choose using several simple heuristics, one of which is the framing of payoffs as gains or as losses, using a reference point. In the evaluation phase, an agent modifies his utility function to a *reference-dependent* utility function, to account for the perception of the payoffs. That is, perceived losses are weighted downwards, a phenomenon frequently referred to as *loss aversion*. The agent then transforms the probabilities with which payoffs are realized by using a probability weighting function, and uses these modified probabilities to calculate the expected reference-dependent utility of the lottery. The latter aspect is ignored in this article.

Although expected utility theory remains the standard model for rational decision-making in mainstream economic theory, the non-expected utility theories—and prospect theory in particular—have proved to be successful challengers of the expected utility paradigm. Many of these theories have an equally solid mathematical basis as expected utility theory, making them acceptable alternatives for economists. More importantly, they tend to incorporate a number of behavioral patterns, documented in the psychology literature[2] that better explain the decisions of economic agents, and as a consequence, are better able to provide a theoretical basis for several empirically observed phenomena that do not fit with the standard theory of rational choice.[3]

Although a number of these behavioral aspects have been applied to the specific field of non-cooperative game theory,[4] the effects of loss aversion in non-cooperative games have not been extensively studied. There is some work in which the outcomes of certain well known examples of games are showed to be consistent with experimental or empirical observations, if the players are assumed to be loss averse.[5] However, this literature focuses on specific examples, and furthermore, usually assumes that the players' reference points are given by some exogenous status quo value. This does not fully reflect the idea of reference-dependence as it was originally intended: Tversky and Kahneman (1981) defined the framing of payoffs as 'the decision-maker's conception of the acts, outcomes, and contingencies associated with a particular choice' (p. 453). This implies that the reference points of players playing a non-cooperative game should not be fixed *ex ante*, but must be based on their own strategies (the *acts*), their payoffs (the *outcomes*), and the strategies of their opponents (the *contingencies*). Thus, game theory adds another dimension to the issue of framing payoffs, and to loss aversion in general, that is often ignored.

One paper in which reference-dependence is treated consistently with Tversky and Kahneman's definition is Shalev (2000). There, an equilibrium concept is developed in which each player transforms his basic utility payoffs with a reference point such that his expected reference-dependent equilibrium payoff is exactly equal to that reference point. Thus, the players' reference points can be interpreted as their expected payoffs in equilibrium. In line with Tversky and Kahneman's definition, the reference points thus depend on the equilibrium strategies and on the players' individual basic utility payoff matrices.

We develop two other equilibrium concepts that take into account the players' loss aversion in a way that is consistent with Kahneman and Tversky's definition. We restrict ourselves to bimatrix games, i.e., two-player games in which each player has finitely many pure strategies. The first new concept, called 'maximin loss aversion equilibrium,' assumes that each player's reference point is equal to his pure maximin value, taking into account only those pure strategies of the opponent that are played with positive probability. This differs from Shalev's equilibrium concept in two significant ways: first, a player's reference point depends on the *carrier* of the opponent's strategy. In addition, it assumes that players are cautious, in the sense that they base their reference points on 'worst-case' values. Since a player's reference point depends on the carrier of his opponent's strategy, it can exhibit discontinuities when the opponent's carrier changes. Indeed, maximin loss aversion equilibrium may fail to exist. Nonetheless, we show existence if at least one player has at most two pure strategies.

For the second new concept, the idea of loss aversion safety level is central. The loss aversion safety level of a player is the value of the matrix game, derived from the basic payoff matrix with that value as reference point. A safety level loss aversion equilibrium is an equilibrium in the bimatrix game obtained by transforming the basic payoffs with these loss aversion safety levels as reference points. This type of equilibrium shares the fixed point idea with Shalev's loss aversion equilibrium and the 'cautious player' property with the maximin loss aversion equilibrium. However, it is based on reference points that no longer depend on the opponent's equilibrium strategy, and that represent the payoff that a player can guarantee.

We conclude the article with a comparative statics analysis of the three equilibrium concepts in 2×2 bimatrix games. Specifically, we assume that both players are loss neutral, but that player 2 believes that player 1 is loss averse. We study the effect this has on the equilibrium payoff of player 1, and establish under which condition this is beneficial to player 1.

The article continues as follows. After preliminaries in Section 2, we discuss the 'myopic loss aversion equilibrium' of Shalev (2000) in Section 3. Section 4 discusses the maximin loss aversion equilibrium, and Section 5 the safety level loss aversion equilibrium. In Section 6 we present the comparative statics results mentioned above.

2. PRELIMINARIES

Before introducing the different equilibrium concepts we define bimatrix games and Nash equilibrium, and indicate how loss aversion of the players can be incorporated.

2.1. *Bimatrix games and nash equilibria*

Players 1 and 2 have sets of pure strategies $I=\{1,\dots,m\}$ and $J=\{1,\dots,n\}$, respectively. If player 1 plays i and player 2 plays j, then player 1 (2) receives a_{ij} (b_{ij}), the number in the payoff matrix A (B) in row i and column j. The pair (A,B) is called a *bimatrix game*.

The $(k-1)$-dimensional unit simplex Δ^k is defined as

$$\Delta^k := \left\{ \omega \in \mathbb{R}^k : \sum_{l=1}^{k} \omega_l = 1 \text{ and } \omega_i \geqslant 0 \text{ for all } i=1,\dots,k \right\}.$$

A *mixed strategy* for player 1 (2) is an element of Δ^m (Δ^n). A pure strategy l is identified with the mixed strategy e^l, where e^l is the vector with a one in position l and zeros otherwise. The *carrier* of a player's strategy is the set of pure strategies that a player plays with positive probability. That is,

$$\text{Car}(p) := \{i \in I : p_i > 0\} \text{ and } \text{Car}(q) := \{j \in J : q_j > 0\}.$$

A *Nash equilibrium* in an $m \times n$ bimatrix game (A,B) is a pair $(p^*,q^*) \in \Delta^m \times \Delta^n$ such that $p^*Aq^* \geqslant pAq^*$ for all $p \in \Delta^m$ and $p^*Bq^* \geqslant p^*Bq$ for all $q \in \Delta^n$.[6]

2.2. *Loss aversion*

Following Shalev (2000), we introduce loss aversion in two-player games (A,B) by specifying nonnegative *loss aversion coefficients* λ_1 and λ_2, respectively measuring player 1's and player 2's degrees of loss aversion. A *loss aversion bimatrix game* is an object of the form $((A,B),(\lambda_1,\lambda_2))$.

In addition to his loss aversion coefficient, player 1 (2) has a number r_1 (r_2) below which he considers the basic utility payoff entries of A (B) to be losses. These points, r_1 and r_2 are the players' respective *reference points*. The idea of loss aversion is captured by transforming the players' basic utility payoffs as follows:

$$a_{ij}^{\lambda_1,r_1} = a_{ij} - \lambda_1 \max\{r_1 - a_{ij}, 0\},$$
$$b_{ij}^{\lambda_2,r_2} = b_{ij} - \lambda_2 \max\{r_2 - b_{ij}, 0\}.$$

Observe that this transformation preserves the ordering over deterministic payoffs. That is, a decision-maker prefers basic utility payoff x to y if and only if he prefers $x^{\lambda,r}$ to $y^{\lambda,r}$ for all $\lambda \geqslant 0$ and $r \in \mathbb{R}$.

For each equilibrium concept to be considered below, we require that it is a Nash equilibrium in the bimatrix game $(A^{\lambda_1,r_1}, B^{\lambda_2,r_2})$, where the reference points r_1 and r_2 are endogenous. The three equilibrium concepts differ in the way the reference points are determined.

3. FIXED POINT LOSS AVERSION EQUILIBRIA

In Shalev (2000), a concept of loss aversion equilibrium is introduced where the players' reference points are found through a fixed point calculation. First, define

$$\underline{r} := \min \left\{ \min_{(i,j) \in I \times J} a_{ij}, \min_{(i,j) \in I \times J} b_{ij} \right\}$$

and

$$\overline{r} := \max \left\{ \max_{(i,j) \in I \times J} a_{ij}, \max_{(i,j) \in I \times J} b_{ij} \right\}.$$

In words, $\underline{r}$ and $\overline{r}$ are the lowest resp. the highest payoffs in A or B. Then, for a strategy profile (p, q) and a reference point $r_1 \in [\underline{r}, \overline{r}]$, player 1 has an expected payoff of $pA^{\lambda_1,r_1}q$. Observe that

$$pA^{\lambda_1,\underline{r}}q = pAq \geqslant \min_{(i,j) \in I \times J} a_{ij} \geqslant \underline{r}$$

and that

$$\overline{r} \geqslant \max_{(i,j) \in I \times J} a_{ij} \geqslant pAq \geqslant pA^{\lambda_1,\overline{r}}q.$$

This and the fact that $pA^{\lambda_1,r_1}q$ is a continuous function of r_1 implies that there is an $r_1^* \in [\underline{r}, \overline{r}]$ such that $r_1^* = pA^{\lambda_1,r_1^*}q$. Furthermore, r_1^* is unique because r_1 is strictly increasing on $[\underline{r}, \overline{r}]$, while $pA^{\lambda_1,r_1}q$ is non-increasing on $[\underline{r}, \overline{r}]$. Similarly, there is a unique $r_2^* \in [\underline{r}, \overline{r}]$ such that $r_2^* = pB^{\lambda_2, r_2^*}q$. Clearly, these 'fixed point' reference points can be interpreted as the utilities players expect to realize given the strategy profile (p, q).

Next, Shalev introduces a non-empty, compact- and convex-valued correspondence $\beta : \Delta^m \times \Delta^n \times [\underline{r}, \overline{r}]^2 \to \Delta^m \times \Delta^n \times [\underline{r}, \overline{r}]^2$ where

$$\begin{aligned} \beta(\hat{p}, \hat{q}, (\hat{r}_1, \hat{r}_2)) := \{(p, q, (r_1, r_2)) \in \Delta^m \times \Delta^n \times [\underline{r}, \overline{r}]^2 : \\ r_1 = pA^{\lambda_1, \hat{r}_1}\hat{q} \geqslant p'A^{\lambda_1, \hat{r}_1}\hat{q} \text{ for all } p' \in \Delta^m, \text{and} \\ r_2 = \hat{p}B^{\lambda_2, \hat{r}_2}q \geqslant \hat{p}B^{\lambda_2, \hat{r}_2}q' \text{ for all } q' \in \Delta^n\}. \end{aligned}$$

Since the (Nash) best reply-correspondence is upper semicontinuous and the players' payoff functions are continuous in their respective reference points, it follows that the correspondence β is also upper semicontinuous. Hence, by the Kakutani fixed point theorem there exists a fixed point $(p^*, q^*, (r_1^*, r_2^*))$. Note that the strategy pair (p^*, q^*) is a Nash equilibrium in the bimatrix game $(A^{\lambda_1, r_1^*}, B^{\lambda_2, r_2^*})$. Since the reference points are determined through a fixed point calculation, we refer to this equilibrium concept as *fixed point loss aversion equilibrium.*

DEFINITION 3.1. *A* fixed point loss aversion equilibrium $(p^*, q^*) \in \Delta^m \times \Delta^n$ *in a loss aversion bimatrix game* $((A, B), (\lambda_1, \lambda_2))$ *is a Nash equilibrium in the game* $(A^{\lambda_1, r_1^*}, B^{\lambda_2, r_2^*})$ *such that*

$$r_1^* = p^* A^{\lambda_1, r_1^*} q^* \text{ and } r_2^* = p^* B^{\lambda_2, r_2^*} q^*.$$

Although this is an effective way of dealing with loss aversion, it is certainly not the only possible approach. One of the less attractive features of this concept is that reference points are not unique: two loss aversion equilibria generally do not yield the same expected payoffs to the players. Furthermore, a player's reference point depends heavily on his own beliefs about the opponent's strategy. In what follows, we discuss two alternative equilibrium concepts which—to a greater or lesser extent—respond to these issues.

4. MAXIMIN LOSS AVERSION EQUILIBRIA

In maximin loss aversion equilibrium, each player chooses his reference point in such a way that his maximin payoff w.r.t. the strategies he believes his opponent plays with positive probability is exactly equal to that reference point. The Nash equilibria in the game that results from using these consistent reference points are maximin loss aversion equilibria.

Maximin loss aversion equilibria are similar to fixed point loss aversion equilibria, because in both concepts the players base their reference points on the carriers of their opponents' strategies. In a fixed point loss aversion equilibrium the reference points depend on the exact probabilities used in these strategies. In a maximin loss aversion equilibrium, the reference points depend only on the carriers of the strategies of the opponents. Each player considers the pure strategies of the opponents that can be realized with positive probability, and his reference point is the pure maximin value over those strategies.

4.1. *Definition of maximin loss aversion equilibria*

Formally, for a strategy combination $(p, q) \in \Delta^m \times \Delta^n$, the players' reference points are defined by[7]

$$r_1^* := \max_{i \in I} \min_{j \in \mathrm{Car}(q)} a_{ij} \text{ and } r_2^* := \max_{j \in J} \min_{i \in \mathrm{Car}(p)} b_{ij}. \tag{1}$$

Observe that this implies

$$r_1^* := \max_{i \in I} \min_{j \in \mathrm{Car}(q)} a_{ij}^{\lambda_1, r_1^*} \text{ and } r_2^* := \max_{j \in J} \min_{i \in \mathrm{Car}(p)} b_{ij}^{\lambda_2, r_2^*}.$$

In other words, these reference points do not change after the basic payoffs are transformed according to loss aversion with these reference points.

Note that these reference points are unique for each carrier played by the opponent.

Since a player's reference point only depends on the carrier of the strategy played by his opponent, rather than the strategy itself, reference points are more robust against wrong beliefs a player may have about his opponent.

DEFINITION 4.1. *A* maximin loss aversion equilibrium *in a loss aversion bimatrix game* $((A,B),(\lambda_1,\lambda_2))$ *is a Nash equilibrium* $(p^*,q^*)\in\Delta^m\times\Delta^n$ *in the bimatrix game* $(A^{\lambda_1,r_1^*},B^{\lambda_2,r_2^*})$ *such that* r_1^* *and* r_2^* *are the reference points for* (p^*,q^*) *defined by* (1).

Note that this concept of maximin loss aversion equilibrium does not solve the problem of multiple reference points. Furthermore, because reference points no longer depend continuously on the strategies played by the opponent, maximin loss aversion equilibria may fail to exist.

4.2. *Existence of maximin loss aversion equilibria*

We show by a counter example that maximin loss aversion equilibria may fail to exist. Next, we show existence if one of the players has no more than two pure strategies.

4.2.1. *An example showing non-existence*

Consider the following 3×3 bimatrix game:

$$A=\begin{bmatrix}8&1&0\\1&8&0\\4&4&-1\end{bmatrix}\qquad B=\begin{bmatrix}1&2&1\\2&1&1\\0&0&1\end{bmatrix}.$$

Let player 2 be loss neutral, i.e., $\lambda_2=0$, and assume $\lambda_1=1$. Due to player 2's loss neutrality, r_2 has no influence on the equilibrium. That is, $B=B^{\lambda_2,r_2}$ for all values of r_2. Observe that player 1's best reply against player 2 playing e^1 is e^1, and player 2's best reply against this is e^2. Hence, e^1 can never be an equilibrium strategy for player 2. Similarly, we can exclude e^2 as one of player 2's equilibrium strategies. This implies that player 1's equilibrium reference point is never equal to 8. This leaves two possibilities: $r_1=0$ or $r_1=4$.

- $r_1 = 0$: In this case, we have

$$A^{\lambda_1,0} = \begin{bmatrix} 8 & 1 & 0 \\ 1 & 8 & 0 \\ 4 & 4 & -2 \end{bmatrix}.$$

  The unique Nash equilibrium in $(A^{\lambda_1,0}, B)$ is $((.5, .5, 0), (.5, .5, 0))$, implying $r_1 = 4$.
- $r_1 = 4$: In this case, we have

$$A^{\lambda_1,4} = \begin{bmatrix} 8 & -2 & -4 \\ -2 & 8 & -4 \\ 4 & 4 & -6 \end{bmatrix}.$$

  The unique Nash equilibrium in $(A^{\lambda_1,4}, B)$ is $\left(\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right), \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)\right)$, implying $r_1 = 0$.

Each of player 1's possible reference points implies a loss aversion game in which the carrier of player 2's equilibrium strategy is such that another reference point should be chosen. Hence, there is no maximin loss aversion equilibrium.[8]

4.2.2. *Existence in $m \times 2$ and $2 \times n$ games*

Although maximin loss aversion equilibria do not exist in general, we do have existence in the case where one of the players has no more than two pure strategies.

PROPOSITION 4.2. *For all $\lambda_1, \lambda_2 \geqslant 0$ and all $m \times 2$ or $2 \times n$ matrices A and B, the loss aversion bimatrix game $((A, B), (\lambda_1, \lambda_2))$ has a maximin loss aversion equilibrium.*

Proof. Let $\lambda_1, \lambda_2 \geqslant 0$ and let A, B be $m \times 2$ matrices, and consider the game $((A, B), (\lambda_1, \lambda_2))$. Define

$$\tilde{r} := \max_{i \in I} \min_{j \in J} a_{ij},$$

and let (p^*, q) be a Nash equilibrium in the game $(A^{\lambda_1, \tilde{r}}, B)$. We distinguish three (exhaustive) cases:

i. $q = e^t$ where $t \in J$. This implies there is an $s \in \text{Car}(p^*)$ such that $b_{st} = \max_{j \in J} b_{sj}$, since otherwise e^t would not be a best reply of player 2 against p^*. Since loss aversion preserves the agent's preference ordering over pure alternatives, this implies $b_{st}^{\lambda_2, b_{st}} = \max_{j \in J} b_{sj}^{\lambda_2, b_{st}}$. From the fact that $s \in \text{Car}(p^*)$, it follows that $a_{st}^{\lambda_1, \tilde{r}} = \max_{i \in I} a_{it}^{\lambda_1, \tilde{r}}$, and thus also $a_{st}^{\lambda_1, a_{st}} = \max_{i \in I} a_{it}^{\lambda_1, a_{st}}$. Hence, (e^s, e^t) is a pure loss aversion equilibrium in $((A, B), (\lambda_1, \lambda_2))$.

ii. $(p^*, q) = (e^s, (\beta, 1 - \beta))$ where $s \in I$ and $\beta \in (0, 1)$. This implies

$$\beta a_{s1}^{\lambda_1, r_1} + (1 - \beta) a_{s2}^{\lambda_1, r_1} \geqslant \beta a_{i1}^{\lambda_1, r_1} + (1 - \beta) a_{i2}^{\lambda_1, r_1}$$

for all $i \in I$ with $r_1 = \tilde{r}$. Furthermore, it implies $b_{s1} = b_{s2} =: b$ from which it follows that $b_{s1}^{\lambda_2, r_2} = b_{s2}^{\lambda_2, r_2}$ with $r_2 = b$. Hence, $(e^s, (\beta, 1 - \beta))$ is a maximin loss aversion equilibrium in $((A, B), (\lambda_1, \lambda_2))$.

iii. p^* and q satisfy $|\text{Car}(p^*)| \geqslant 2$ and $|\text{Car}(q)| = 2$. Then there exist pure strategies s and s' in $\text{Car}(p^*)$ such that $[b_{s1} > b_{s2}$ and $b_{s'1} < b_{s'2}]$ or $[b_{s1} = b_{s2}$ and $b_{s'1} = b_{s'2}]$. In both cases, there exists an $\alpha \in (0, 1)$ such that both player 2's pure strategies are best replies against player 1's strategy $(\alpha e^s + (1 - \alpha) e^{s'})$, with payoffs given by B^{λ_2, r_2} where $r_2 = \max_{j \in J} \min_{i \in \{s, s'\}} b_{ij}$. But then $((\alpha, 1 - \alpha), q)$ is a maximin loss aversion equilibrium in the game $((A, B), (\lambda_1, \lambda_2))$.

This completes the proof of the $m \times 2$ case. The $2 \times n$ case is analogous. □

Note that the proof of this proposition provides a method to calculate maximin loss aversion equilibria. One can modify player 1's payoff matrix under the assumption that player 2 plays both his strategies with positive probability, and calculate the Nash equilibria in the resulting game. These Nash equilibria can then be transformed into maximin loss aversion equilibria.

5. SAFETY LEVEL LOSS AVERSION EQUILIBRIA

In a safety level loss aversion equilibrium, a player's reference point is the value of the matrix game which is obtained by adapting his basic payoff matrix to account for loss aversion. If this basic payoff matrix is C and the loss aversion coefficient is λ, then this reference point is equal to the number r if r is the value of the matrix game $C^{\lambda,r}$. Details are spelled out below.

This reference point does not depend on a player's belief about the strategy of the opponent in equilibrium. Instead, a player computes what he can obtain for sure and considers payoffs below this number as losses. Moreover, these reference points are unique, and safety level loss aversion equilibria always exist.

5.1. *Definition and existence of safety level loss aversion equilibria*

The safety level is a concept that dates back to von Neumann's (1928) analysis of zero-sum games.

Formally, for a bimatrix game (A, B), the players' safety levels are defined by

$$v_1(A) := \max_{p\in\Delta^m} \min_{q\in\Delta^n} pAq, \qquad \text{and} \qquad v_2(B) := \max_{q\in\Delta^n} \min_{p\in\Delta^m} pBq.$$

Since a player can guarantee his safety level, it would make an intuitively appealing reference point in a loss aversion bimatrix game $((A, B), (\lambda_1, \lambda_2))$. However, in the payoff matrices adapted by loss aversion the safety levels may change. Therefore, we look for reference points which in the transformed matrices are equal to the safety levels. That is, we wish to find r_1^* and r_2^* such that $r_1^* = v_1(A^{\lambda_1,r_1^*})$ and $r_2^* = v_2(B^{\lambda_2,r_2^*})$. Such reference points are called *loss aversion safety levels*.

In order to show that there is a unique r_1^* such that $r_1^* = v_1(A^{\lambda_1,r_1^*})$, it is sufficient to show that $v_1(A^{\lambda_1,r})$ is a continuous, non-increasing function of r on the interval $[\underline{r}, \overline{r}]$, and that

$$v_1(A^{\lambda_1,\underline{r}}) \geqslant \underline{r}, \quad \text{and} \quad v_1(A^{\lambda_1,\overline{r}}) \leqslant \overline{r}.$$

It is obvious that $v_1(A^{\lambda_1,r})$ is continuous in r. In order to show that it is non-increasing in r on the interval $[\underline{r},\overline{r}]$, let r_1 and s_1 be reference points in $[\underline{r},\overline{r}]$ with $r_1 > s_1$, and let

$$p^* \in \arg\max_{p\in\Delta^m}\min_{q\in\Delta^n} pA^{\lambda_1,r_1}q, \text{ and}$$
$$q^* \in \arg\min_{q\in\Delta^n} p^*A^{\lambda_1,s_1}q.$$

Then

$$\begin{aligned} v_1(A^{\lambda_1,s_1}) &= \max_{p\in\Delta^m}\min_{q\in\Delta^n} pA^{\lambda_1,s_1}q \\ &\geqslant \min_{q\in\Delta^n} p^*A^{\lambda_1,s_1}q \\ &= p^*A^{\lambda_1,s_1}q^*. \end{aligned}$$

Note that $pA^{\lambda_1,r}q$ is non-increasing in r for any given strategy pair (p,q). Therefore, $p^*A^{\lambda_1,s_1}q^* \geqslant p^*A^{\lambda_1,r_1}q^*$. Now observe that

$$\begin{aligned} p^*A^{\lambda_1,r_1}q^* &\geqslant \min_{q\in\Delta^n} p^*A^{\lambda_1,r_1}q \\ &= \max_{p\in\Delta^m}\min_{q\in\Delta^n} pA^{\lambda_1,r_1}q \\ &= v_1(A^{\lambda_1,r_1}). \end{aligned}$$

This shows that $v_1(A^{\lambda_1,r})$ is non-increasing in r. Furthermore,

$$v_1(A^{\lambda_1,\underline{r}}) = v_1(A) \geqslant \min_{(i,j)\in I\times J} a_{ij} \geqslant \underline{r},$$

and since $v_1(A^{\lambda_1,r})$ is non-increasing in r, we also have

$$v_1(A^{\lambda_1,\overline{r}}) \leqslant v_1(A^{\lambda_1,\underline{r}}) = v_1(A) \leqslant \max_{(i,j)\in I\times J} a_{ij} \leqslant \overline{r}.$$

Thus, there must be a unique $r_1^* \in [\underline{r},\overline{r}]$ such that $r_1^* = v_1(A^{\lambda_1,r_1^*})$. Similarly, there exists a unique $r_2^* \in [\underline{r},\overline{r}]$ such that $r_2^* = v_2(B^{\lambda_2,r_2^*})$.

The players transform their payoff matrices using their degrees of loss aversion and their loss aversion safety levels. The *safety level loss aversion equilibria* are the Nash equilibria in the transformed game, i.e., the bimatrix game $(A^{\lambda_1,r_1^*}, B^{\lambda_2,r_2^*})$.

DEFINITION 5.1. *A* safety level loss aversion equilibrium *in a loss aversion bimatrix game* $((A, B), (\lambda_1, \lambda_2))$ *is a Nash equilibrium* $(p^*, q^*) \in \Delta^m \times \Delta^n$ *in the bimatrix game* $(A^{\lambda_1, r_1^*}, B^{\lambda_2, r_2^*})$ *such that*

$$r_1^* = v_1(A^{\lambda_1, r_1^*}) \text{ and } r_2^* = v_2(B^{\lambda_2, r_2^*}).$$

5.2. *Strict dominance when players are loss averse*

In a safety level loss aversion equilibrium the determination of the equilibrium strategies is not related to the determination of the reference points. Hence, in contrast to the previous equilibrium concepts, reference points may also depend on strategies that are not played in equilibrium. Suppose for instance that player 2 has a strictly dominated column in his payoff matrix B, say $b_{i,n-1} > b_{i,n}$ for all $i = 1, \dots, m$. Then, whatever the reference level and the adapted payoff matrix are going to be, player 2 will not put any probability on column n in an equilibrium. One could argue that player 1, in determining his loss aversion safety level, should take into account that player 2 is not going to play column n. More generally, one could argue that, before actually computing loss aversion safety levels, first strictly dominated strategies should be iteratively eliminated. This raises the question which strategies are strictly dominated in the loss aversion context. The remainder of this section is devoted to studying this question.

We say that a pure strategy $i \in I$ is *strictly dominated* in A if there is a strategy $p \in \Delta^m$ with $p_i = 0$, such that $pAe^j > e^i Ae^j$ for all $j \in J$. A pure strategy $i \in I$ is said to be *strictly dominated* in (A, λ_1), if it is strictly dominated in $A^{\lambda_1, \rho}$ for all $\rho \in [\underline{r}, \overline{r}]$. Then, to eliminate a pure strategy $i \in I$ from the game, it is no longer sufficient that it is strictly dominated in A. In order to see this, consider the following example:

$$A = \begin{bmatrix} 5 & 0 \\ 0 & 5 \\ 2 & 2 \end{bmatrix}.$$

Observe that player 1's pure strategy e^3 is strictly dominated by the mixed strategy $(0.5, 0.5, 0)$. Now let $\lambda_1 = 1$ and $r_1 = 2$. Then the transformed payoff matrix is

$$A^{1,2} = \begin{bmatrix} 5 & -2 \\ -2 & 5 \\ 2 & 2 \end{bmatrix}.$$

Clearly, strategy e^3 is no longer strictly dominated in $A^{1,2}$, even though it was in A.

Since $A = A^{\lambda_1, \underline{r}}$, e^i being strictly dominated in A is still a necessary condition for e^i to be strictly dominated in (A, λ_1). The following proposition gives a necessary and sufficient condition.

PROPOSITION 5.2. *In the game* $((A, B), (\lambda_1, \lambda_2))$ *where* $\lambda_1 > 0$, *a strategy* $p \in \Delta^m$ *strictly dominates the pure strategy* $i \in I$ *in* (A, λ_1) *if and only if* p *strictly dominates* i *in* $A^{\lambda_1, \rho}$ *for all*

$$\rho \in \left[\min_{j \in J} a_{ij}, \max_{(i', j) \in \mathrm{Car}(p) \times J} a_{i'j}\right].$$

Proof. Let a pure strategy i be strictly dominated by a strategy p in the payoff matrix A. Then define $\underline{\rho} := \min_{j \in J} a_{ij}$ and $\overline{\rho} := \max_{(i', j) \in \mathrm{Car}(p) \times J} a_{i'j}$.

$\Leftarrow$: Let the pure strategy i be strictly dominated in (A, λ_1). Then it is strictly dominated in $A^{\lambda_1, \rho}$ for all $\rho \in [\underline{r}, \overline{r}]$. Since $[\underline{\rho}, \overline{\rho}] \subseteq [\underline{r}, \overline{r}]$, strategy i is strictly dominated in $A^{\lambda_1, \rho}$ for all $\rho \in [\underline{\rho}, \overline{\rho}]$.

$\Rightarrow$: Let the pure strategy i be strictly dominated by p in $A^{\lambda_1, \rho}$ for all $\rho \in [\underline{\rho}, \overline{\rho}]$. Then p strictly dominates i in $A^{\lambda_1, \underline{\rho}}$. That is,

$$pA^{\lambda_1, \underline{\rho}} e^j > e^i A^{\lambda_1, \underline{\rho}} e^j$$

for all $j \in J$. Let $\rho \in [\underline{r}, \underline{\rho}]$. Then $e^i A^{\lambda_1, \underline{\rho}} e^j = e^i A^{\lambda_1, \rho} e^j$. Furthermore, $pA^{\lambda_1, \underline{\rho}} e^j \leqslant pA^{\lambda_1, \rho} e^j$ since $\underline{\rho} \geqslant \rho$. Hence, $pA^{\lambda_1, \rho} e^j > e^i A^{\lambda_1, \rho} e^j$ as well. So p strictly dominates i in $A^{\lambda_1, \rho}$ for all $\rho \in [\underline{r}, \underline{\rho}]$.

Similarly, p strictly dominating i in $A^{\lambda_1,\rho}$ for all $\rho \in [\underline{\rho}, \overline{\rho}]$ implies that p strictly dominates i in $A^{\lambda_1,\overline{\rho}}$. That is, $pA^{\lambda_1,\overline{\rho}}e^j > e^i A^{\lambda_1,\overline{\rho}}e^j$ for all $j \in J$. Observe that for all $\rho \in [\overline{\rho}, \overline{r}]$, we have that

$$pA^{\lambda_1,\rho}e^j = (1+\lambda_1)pAe^j - \lambda_1\rho.$$

Note that for all $j \in J$ there is an $i' \in \mathrm{Car}(p)$ such that $e^i A^{\lambda_1,\rho}e^j \leqslant e^{i'} A^{\lambda_1,\rho}e^j$. Hence, for all $j \in J$ we have that $e^i A^{\lambda_1,\rho}e^j \leqslant \rho$ for all $\rho \in [\overline{\rho}, \overline{r}]$, implying

$$e^i A^{\lambda_1,\rho}e^j = (1+\lambda_1)e^i Ae^j - \lambda_1\rho$$

for all $j \in J$ and $\rho \in [\overline{\rho}, \overline{r}]$. Hence, $(1+\lambda_1)pAe^j - \lambda_1\overline{\rho} > (1+\lambda_1)e^i Ae^j - \lambda_1\overline{\rho}$ for all $j \in J$, and the inequality is preserved if we replace $\overline{\rho}$ by any $\rho \in [\overline{\rho}, \overline{r}]$. But then $pA^{\lambda_1,\rho}e^j > e^i A^{\lambda_1,\rho}e^j$ for all $j \in J$ and $\rho \in [\overline{\rho}, \overline{r}]$. That is, p strictly dominates i in $A^{\lambda_1,\rho}$ for all $\rho \in [\overline{\rho}, \overline{r}]$.

Thus, p strictly dominates i in $A^{\lambda_1,\rho}$ for all $\rho \in [\underline{r}, \overline{r}]$, which means that p strictly dominates i in (A, λ_1). □

REMARK 5.3. It could happen that none of the l strategies $p^1, \ldots, p^l \in \Delta^m$ dominates a pure strategy i in (A, λ_1), while taken together they do. Suppose none of the strategies $p^1, \ldots, p^l$ dominates i in (A, λ_1). Then for each p^k there is an interval $R^k \subseteq [\underline{r}, \overline{r}]$ such that p^k does not strictly dominate i in $A^{\lambda_1,\rho}$ for any $\rho \in R^k$. Then, as long as $\bigcap_{k=1}^{l} R^k = \emptyset$, there is always some strategy p^k strictly better than pure strategy i.

Thus, when considering safety level loss aversion equilibria in a loss aversion bimatrix game $((A, B), (\lambda_1, \lambda_2))$, we could assume that the payoff matrices A and B are the result of iterated elimination of strategies that are either strictly dominated in (A, λ_1) and (B, λ_2) respectively, or strictly dominated in the weaker sense, explained in the above remark.

6. COMPARATIVE STATICS

In this section we consider the effect of loss aversion on the equilibrium payoff of a player. Specifically, suppose that both players are loss neutral but player 2 thinks that player 1 is loss averse. This makes player 2 change his equilibrium strategy and we investigate when this is beneficial for player 1.

6.1. *Preliminaries*

For reasons of tractability we only consider 2×2 bimatrix games.[9] The basic utilities are represented by the following matrices:

$$A := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } B := \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Since pure Nash equilibria are equivalent to the pure versions of all three types of loss aversion equilibrium considered above, we can restrict ourselves to mixed equilibria. Following Berden and Peters (2006), we exclude the case where player 1 has a weakly dominant strategy, implying $a_{11} \neq a_{21}$ and $a_{12} \neq a_{22}$. W.l.o.g. assume $a_{11} > a_{21}$, $a_{12} < a_{22}$, and $a_{11} \geqslant a_{22}$. This leaves three exhaustive cases:

i. $a_{21} \geqslant a_{22}$;
ii. $a_{22} \geqslant a_{21} \geqslant a_{12}$;
iii. $a_{12} \geqslant a_{21}$.

In addition, we assume $b_{11} < b_{12}$ and $b_{21} > b_{22}$. Hence, also player 2 has no weakly dominant strategy. These assumptions imply that the game has a unique, completely mixed Nash equilibrium (p^*, q^*), with

$$p^* = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix} \quad \text{and} \quad q^* = \begin{bmatrix} \delta \\ 1-\delta \end{bmatrix},$$

where

$$\gamma = \frac{b_{21} - b_{22}}{b_{12} - b_{11} + b_{21} - b_{22}} \text{ and } \delta = \frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}}.$$

Assume that player 2 does not have any information about his opponent's attitude towards losses, allowing him to form a wrong belief about it. That is, player 2 may falsely believe that player 1 is loss averse, which causes player 2 to play differently in order to make player 1 indifferent between his pure strategies in equilibrium. Specifically, player 2 plays a strategy $\tilde{q} = (\tilde{\delta}, 1-\tilde{\delta})$ where $\tilde{\delta} \in (0,1)$. Player 1 knows the utility function of his opponent, so he keeps playing his previous strategy p^*. Thus, if player 2 misperceives λ_1, the mixed loss aversion equilibrium becomes $(p^*, \tilde{q})$.[10]

The equilibrium under player 2's wrong perception of λ_1 can be explained in two ways. First, player 1 could be naive in the sense that he does not know that player 2 does not perceive his degree of loss aversion correctly. Thus, player 1 plays his equilibrium strategy and is surprised by player 2's action. A second explanation would be that player 1—knowing that player 2 does not have any information about λ_1—intentionally misrepresents his degree of loss aversion, but is myopic in the sense that he is not able to determine the strategy he has to play in order to optimally exploit player 2's action.

Clearly, starting from a situation where both players are loss averse but player 2 overestimates player 1's loss aversion is not really different from the present case.

6.2. *Result*

We say that player 2's wrong perception of λ_1 *benefits* player 1 if $p^* A\tilde{q} \geqslant p^* A q^*$, and *hurts* him if $p^* A\tilde{q} \leqslant p^* A q^*$.[11] The following theorem presents the comparative statics result.

THEOREM 6.1. *In case i. player 1 benefits from player 2 misperceiving* λ_1. *In cases ii. and iii. player 1 benefits from player 2 misperceiving* λ_1 *if and only if*

$$\frac{b_{21}-b_{22}}{b_{12}-b_{11}+b_{21}-b_{22}} \geqslant \frac{a_{22}-a_{21}}{a_{11}-a_{12}+a_{22}-a_{21}}. \tag{2}$$

The proof can be found in the appendix.

Note that the left-hand side in (2) is equal to γ, i.e., the probability that player 1 puts on the first row in equilibrium. If non-negative, the right-hand side $\gamma' = \frac{a_{22}-a_{21}}{a_{11}-a_{12}+a_{22}-a_{21}}$ in (2) is the probability that player 1 would have to put on the first row in order to be indifferent between the actions of player 2. In that case, playing the mixed strategy $p' = (\gamma', 1-\gamma')$ would yield player 1 the Nash equilibrium payoff p^*Aq^*. Thus, as long as player 2 plays his Nash equilibrium strategy q^*, then player 1 is indifferent between playing p^* and p'. However, if player 2 plays $\tilde{q}$, i.e., erroneously believes player 1 is loss averse, then player 1 obtains the Nash equilibrium payoff by playing p', but could receive more or less by playing p^*.

If condition (2) on the payoffs of the players is satisfied, we thus obtain that pretending to be more loss averse makes a player better off. The comparative statics of the fixed point loss aversion equilibrium were also investigated in Shalev (2000), with different results however. Shalev investigated how a player's reference point moves with his own degree of loss aversion. As this reference point equals a player's equilibrium payoff by definition, this yields associated comparative statics results. It is not clear what the proper interpretation of these results is since payoffs for players with different utility functions are compared. For our approach this difficulty does not arise since we compare payoffs of one and the same player 1: only player 2's belief about player 1 changes, but not player 1 himself.

7. SUMMARY

In this article we have argued that in order to correctly incorporate the concept of loss aversion into non-cooperative game theory, it is necessary to let the reference points of the players depend on the strategies the players play. We have examined three different loss aversion equilibrium concepts that satisfy this requirement. Then we have established that in 2×2 bimatrix games, a simple condition on the payoffs is sufficient for a player to benefit from his opponent overestimating his (the player's) degree of loss aversion.

APPENDIX A: PROOF OF THEOREM 6.1

A player's reference point is in between the lowest and the highest payoff in his payoff matrix. In the 2×2 case, there are three intervals which could contain r_1: the upper, the middle and the lower interval. Let $\pi \in \mathbb{R}^4$ with

$$\pi = \begin{cases} (a_{12}, a_{22}, a_{21}, a_{11}) & \text{in case i.} \\ (a_{12}, a_{21}, a_{22}, a_{11}) & \text{in case ii.} \\ (a_{21}, a_{12}, a_{22}, a_{11}) & \text{in case iii.} \end{cases}$$

The following lemma says that the reference point in a fixed point loss aversion equilibrium lies in the middle interval.

LEMMA A.1. *If* $(p^*, q^*, (r_1, r_2))$ *is a fixed point loss aversion equilibrium in* $((A, B), (\lambda_1, \lambda_2))$*, then* $r_1 \in [\pi_2, \pi_3]$.

Proof. For any basic utility payoff x let $\hat{x}$ denote transformed payoff as a consequence of loss aversion. The expected payoff r_1 under the mixed fixed point loss aversion equilibrium is a convex combination of $\hat{a}_{21}$ and $\hat{a}_{22}$. Therefore, $r_1 \leqslant \max\{\hat{a}_{21}, \hat{a}_{22}\} \leqslant \max\{a_{21}, a_{22}\}$ and thus, $r_1 \leqslant \pi_3$.

In order to see that the reference point can never be in the lower interval, we consider each case separately.

Case i.: Here, $a_{21} \geqslant a_{22}$ and $a_{22} > a_{12}$. Assume $r_1 \in [a_{12}, a_{22})$. Note that $a_{22} - a_{21} \leqslant 0$ and $\hat{a}_{12} - a_{22} < 0$, and thus $(a_{22} - a_{21})(\hat{a}_{12} - a_{22}) \geqslant 0$. By an elementary computation it follows that

$$\frac{a_{11}a_{22} - a_{21}\hat{a}_{12}}{a_{11} - \hat{a}_{12} - a_{21} + a_{22}} \geqslant a_{22}.$$

In other words, player 1's expected payoff under loss aversion is larger or equal than a_{22}, contradicting $r_1 < a_{22}$. Thus, $r_1 \geqslant \pi_2$.

Case ii.: Here, $a_{11} > a_{21}$ and $a_{22} \geqslant a_{21}$. Assume $r_1 \in [a_{12}, a_{21})$. Now $(a_{11} - a_{21})(a_{22} - a_{21}) \geqslant 0$, which implies

$$\frac{a_{11}a_{22} - a_{21}\hat{a}_{12}}{a_{11} - \hat{a}_{12} - a_{21} + a_{22}} \geqslant a_{21},$$

contradicting $r_1 < a_{21}$. Hence, $r_1 \geqslant \pi_2$.

Case iii.: Repeating the argument from case ii., with a_{21} replaced by a_{12}, and $\hat{a}_{12}$ by $\hat{a}_{21}$, yields the desired contradiction for case iii.

It follows from the above that $r_1 \in [\pi_2, \pi_3]$. □

For the other loss aversion equilibrium concepts we have similar results.

LEMMA A.2. *If* $(p^*, q^*, (r_1, r_2))$ *is a maximin loss aversion equilibrium in* $((A, B), (\lambda_1, \lambda_2))$*, then* $r_1 = \pi_2$.

Proof. Since the equilibrium is completely mixed we have in Case i.: $\max_{i\in I} \min_{j\in J} a_{ij} = \max\{a_{12}, a_{22}\} = a_{22}$; and in Cases ii. and iii.: $\max_{i\in I} \min_{j\in J} a_{ij} = \max\{a_{12}, a_{21}\} = a_{21}$. Thus, $r_1 = \pi_2$. □

LEMMA A.3. *If* $(p^*, q^*, (r_1, r_2))$ *is a safety level loss aversion equilibrium in* $((A, B), (\lambda_1, \lambda_2))$*, then* $r_1 \in [\pi_2, \pi_3]$.

Proof. Recall that for any 2×2 matrix A, we have

$$v_1(A) = \max_{p\in\Delta^2} \min_{q\in\Delta^2} pAq \geqslant \max_{i\in I} \min_{q\in\Delta^2} e^i Aq = \max_{i\in I} \min_{j\in J} e^i Ae^j.$$

Assume $r_1 < \pi_2$. Then

$$\pi_2 = \max_{i\in I} \min_{j\in J} e^i A^{\lambda_1, r_1} e^j.$$

In safety level loss aversion equilibrium we have $r_1 = v_1(A^{\lambda_1, r_1})$. Then

$$v_1(A^{\lambda_1, r_1}) < \pi_2 = \max_{i\in I} \min_{j\in J} e^i A^{\lambda_1, r_1} e^j,$$

which is a contradiction. Hence, $r_1 \geqslant \pi_2$.

The safety level, $v_1(A^{\lambda_1, r_1})$, can be interpreted as player 1's Nash equilibrium payoff in the zero-sum game $(A^{\lambda_1, r_1}, -A^{\lambda_1, r_1})$. By a similar reasoning as above, we have that player 2's payoff, $-v_1(A^{\lambda_1, r_1})$, is above $-\pi_3$, implying $v_1(A^{\lambda_1, r_1}) \leqslant \pi_3$. Hence, $r_1 \in [\pi_2, \pi_3]$. □

Recall that $\tilde{q}=(\tilde{\delta},1-\tilde{\delta})$. We now compute $\tilde{\delta}$ for the three different cases.

Case i.: Here, we have $a_{11}>a_{21}\geqslant a_{22}>a_{12}$, and by Lemmas A.1–A.3, $r_1\in[a_{22},a_{21}]$. Thus,

$$\begin{aligned}\tilde{\delta}&=\frac{a_{22}-a_{12}-\lambda_1(r_1-a_{22})+\lambda_1(r_1-a_{12})}{a_{11}-a_{12}-a_{21}+a_{22}+\lambda_1(r_1-a_{12})-\lambda_1(r_1-a_{22})}\\&=\frac{(1+\lambda_1)(a_{22}-a_{12})}{a_{11}-a_{21}+(1+\lambda_1)(a_{22}-a_{12})}.\end{aligned}$$

Case ii.: Here $a_{11}>a_{22}\geqslant a_{21}>a_{12}$ with $a_{11}>a_{22}$, $a_{22}>a_{21}$, or both. By Lemmas A.1–A.3, we have $r_1\in[a_{21},a_{22}]$. Hence,

$$\begin{aligned}\tilde{\delta}&=\frac{a_{22}-a_{12}+\lambda_1(r_1-a_{12})}{a_{11}-a_{12}+\lambda_1(r_1-a_{12})-a_{21}+\lambda_1(r_1-a_{21})+a_{22}}\\&=\frac{a_{22}-a_{12}+\lambda_1(r_1-a_{12})}{a_{11}-a_{12}-a_{21}+a_{22}+\lambda_1(2r_1-a_{12}-a_{21})}.\end{aligned}$$

Case iii.: Here $a_{11}>a_{22}\geqslant a_{12}>a_{21}$. By Lemmas A.1–A.3, $r_1\in[a_{12},a_{22}]$, which implies that $\tilde{\delta}$ has the same value as in case ii. That is,

$$\tilde{\delta}=\frac{a_{22}-a_{12}+\lambda_1(r_1-a_{12})}{a_{11}-a_{12}-a_{21}+a_{22}+\lambda_1(2r_1-a_{12}-a_{21})}.$$

Having specified $\tilde{q}$, player 2's equilibrium strategy associated with a wrong belief about λ_1, for each case, we now investigate how it compares to q^*, player 2's equilibrium strategy associated with the correct belief about λ_1.

LEMMA A.4. *Let* $(p^*,(\delta,1-\delta))$ *be the Nash equilibrium in* (A,B)*, and let* $(p^*,(\tilde{\delta},1-\tilde{\delta}))$ *be the Nash equilibrium in* (A^{λ_1,r_1},B)*, where* r_1 *is the equilibrium reference point associated with either of the three loss aversion equilibrium types. Then*

$$\tilde{\delta}\geqslant\delta$$

in all three cases.

Proof. Let $x := a_{22} - a_{12}$ and $y := a_{11} - a_{21} - a_{12} + a_{22}$. Note that x and y are strictly positive, and that $\delta = x/y$. Assume $\lambda_1 > 0$. Again we consider the different cases.
Case i.: Here we have

$$\tilde{\delta} = \frac{(1+\lambda_1)(a_{22}-a_{12})}{a_{11}-a_{21}+(1+\lambda_1)(a_{22}-a_{12})} = \frac{(1+\lambda_1)x}{y+\lambda_1 x}.$$

Since $y > x$ this implies $\tilde{\delta} > \delta$.
Cases ii. and iii.: Here we have

$$\begin{aligned}\tilde{\delta} &= \frac{a_{22}-a_{12}+\lambda_1(r_1-a_{12})}{a_{11}-a_{12}-a_{21}+a_{22}+\lambda_1(2r_1-a_{12}-a_{21})}\\ &= \frac{x+\lambda_1(r_1-a_{12})}{y+\lambda_1(2r_1-a_{12}-a_{21})}.\end{aligned}$$

Now $\tilde{\delta} \geqslant \delta$ follows by straightforward computation, using $r_1 \in [a_{21}, a_{22}]$. □

Proof of Theorem 6.1. We have

$$p^* A q^* = \left[\gamma a_{11} + (1-\gamma)a_{21} \;\; \gamma a_{12} + (1-\gamma)a_{22}\right] \begin{bmatrix} \delta \\ 1-\delta \end{bmatrix},$$

and

$$p^* A \tilde{q} = \left[\gamma a_{11} + (1-\gamma)a_{21} \;\; \gamma a_{12} + (1-\gamma)a_{22}\right] \begin{bmatrix} \tilde{\delta} \\ 1-\tilde{\delta} \end{bmatrix}.$$

From Lemma A.4, we have $\tilde{\delta} \geqslant \delta$. Then $p^* A \tilde{q} \geqslant p^* A q^*$ if and only if $\gamma a_{11} + (1-\gamma)a_{21} \geqslant \gamma a_{12} + (1-\gamma)a_{22}$, which is equivalent to (2). Observe that in case i. this condition is trivially satisfied. This concludes the proof. □

NOTES

1. For surveys of the literature on violations of expected utility theory, see Schoemaker (1982) or Machina (1987)
2. See for instance Kahneman and Tversky (1979), and Hershey et al. (1982), and others.
3. An overview of puzzles and the solutions proposed by prospect theory is provided in Camerer (2002).

4. Examples are Crawford (1990), Dekel et al. (1991), and Eichberger and Kelsey (1999).
5. For example, Fershtman (1996) studies an incumbency game, Berejikian (2002) a.o. the game of chicken and the prisoners' dilemma, and Butler (2007) an ultimatum game.
6. Note that we do not use the transposition notation for vectors and matrices if there is no confusion what is meant.
7. For simplicity of notation we do not use different symbols for the various reference point concepts in this article.
8. It can be checked—see Driesen et al. (2007)—that the appropriate best reply- correspondence is not upper semicontinuous, so that (indeed) the Kakutani fixed point theorem cannot be applied. In fact, it can be shown that the Kakutani fixed point theorem only applies in 2×2 games.
9. As an additional advantage there is no existence problem for maximin equilibria.
10. In this 2×2 framework, players keep playing strategies with full carriers for each type of loss aversion equilibrium.
11. Note that player 2's wrong belief about his opponent's loss attitude neither hurts nor benefits himself. That is, $p^*Bq^* = p^*B\tilde{q}$.

REFERENCES

Allais, M. (1953), Le comportement de l' homme rationnel devant le risque: critique des postulats et axiomes de l' école américaine, *Econometrica* 21, 503–546.

Berden, C. and Peters, H. (2005), On the effect of risk aversion in bimatrix games, *Theory and Decision* 60(4), 359–370.

Berejikian, J.D. (2002), Model building with prospect theory: A cognitive approach to international relations, *Political Psychology* 23(4), 759–786.

Butler, C.K. (2007), Prospect theory and coercive bargaining, *Journal of Conflict Resolution* 51(2), 227–250.

Camerer, C.F. (2002), Prospect theory in the wild: evidence from the field, in Tversky, A. and Kahneman, D. (eds.), *Choices, Values, and Frames*, chapter 16, Cambridge University Press, pp. 288–301.

Crawford, V.P. (1990), Equilibrium without independence, *Journal of Economic Theory*, 50, 127–154.

Dekel, E., Safra, Z., and Segal, U. (1991), Existence and dynamic consistency of Nash equilibrium with non-expected utility preferences, *Journal of Economic Theory*, 55(2), 229–246.

Driesen, B., Perea, A., and Peters, H. (2007), On loss aversion in bimatrix games. METEOR Research Memorandum 07/033, METEOR, Universtiy of Maastricht.

Eichberger, J. and Kelsey, D. (1999), Non-additive beliefs and strategic equilibria, *Games and Economic Behavior* 30(2), 183–215.

Fershtman, C. (1996), On the value of incumbency: Managerial reference points and loss aversion, *Journal of Economic Psychology* 17, 245–257.

Hershey, J.C., Kunreuther, H.C. and Schoemaker, P.J.H. (1982), Sources of bias in assessment procedures for utility functions, *Management Science* 28(8), 936–954.

Kahneman, D. and Tversky, A. (1979), Prospect theory: an analysis of decision under risk, *Econometrica* 47, 263–291.

Machina, M.J. (1987), Choice under uncertainty: Problems solved and unsolved, *Journal of Economic Perspectives* 1(1), 121–154.

Shalev, J. (2000), Loss aversion equilibrium, *International Journal of Game Theory* 29(2), 269–287.

Schoemaker, P.J.H. (1982), The expected utility model: Its variants, purposes, evidence and limitations, *Journal of Economic Literature*, 20(2), 529–563.

Tversky, A. and Kahneman, D. (1974), Judgment under uncertainty: Heuristics and biases, *Science* 185, 1124–1131.

Tversky, A. and Kahneman, D. (1981), The framing of decisions and the psychology of choice, *Science* 211, 453–458.

von Neumann, J. (1928), Zur theorie der gesellschaftsspiele, *Mathematische Annalen* 100(1), 295–320.

von Neumann, J. and Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton University Press.

Address for correspondence: Hans Peters, Department of Quantitative Economics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands. E-mail: H.Peters@ke.unimaas.nl

Bram Driesen, Andrés Perea, Department of Quantitative Economics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands E-mail: B.Driesen@ke.unimaas.nl, A.Perea@ke.unimaas.nl