CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 741

A Note on guilt aversion in the Battle of Sexes game

Giuseppe De Marco, Maria Romaniello, and Alba Roviello

December 2024

University of Naples Federico II
University of Salerno
Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

Working Paper no. 741

A Note on guilt aversion in the Battle of Sexes game

Giuseppe De Marco[*] , Maria Romaniello[†], and Alba Roviello[‡]

Abstract
We analyze the effects of guilt aversion in the Battle of Sexes game by exploiting the theory of psychological games and the concept of psychological Nash equilibrium. Then we examine the impact of ambiguity in the (second-order) beliefs by taking into account the theory of psychological games under ambiguity. Our results show that the sensitivity to guilt affects some equilibrium of the game since a player might be willing to accept a lower expected utility to compensate the other's disutility from guilt. Ambiguity, in turn, makes this effect more evident as it makes it greater the disutility from guilt.

Keywords: Battle of sexes, guilt aversion, psychological games, maxmin preferences.

[*] University of Napoli Parthenope and CSEF. Email: giuseppe.demarco@uniparthenope.it. *(Corresponding author)*

[†] University of Campania Vanvitelli.

[‡] University of Napoli Federico II.

1 Introduction

The *Battle of Sexes* is a classical model in the game theory literature. In its standard form, this game provides a coordination problem between two players who have conflicting preferences. Players must choose between two alternatives (often represented as "going to the Opera" vs. "going to the Boxing match"), each preferring one over the other, but both desiring coordination to avoid a complete loss of payoffs. The game has a key characteristic: it has multiple equilibria and (as the Battle of Sexes is a *generic game*) all the equilibria are stable with respect to perturbations, so that no refinement concept[2] can be effective. Moreover, joint deviations are not Pareto improving, whatever the equilibrium. As the multiplicity of equilibria makes coordination challenging, different approaches have been proposed as selection mechanism, such as pre-play communication, cheap talk and the focal point theory.

From another perspective, an interesting issue considered in the literature is the impact of psychological traits of players on the interaction described by the Battle of Sexes game: in Simon [2021], the Battle of Sexes game was used to model and test behaviors in individuals with personality disorders[3]. It emerged that participants' behavior was affected by personal traits[4]. In the paper by Zapata et al. [2018], weighted Rawlsian preferences were used and applied to the Battle of Sexes game in order to accommodate concerns that players may have about the payoff of the opponent and it emerges that Nash equilibria are strongly affected by this kind of players' concerns[5].

The present paper aims to provide an alternative insight on the influence of feelings on the Battle of Sexes game by considering the consequences of the sentiment of *guilt aversion*. This psychological trait arises when agents' utility is negatively affected by harming other individuals and seems to be particularly suited for the Battle of Sexes, as in this game players have incentives to coordinate their choices with their opponents. In particular, we refer to the formal model of guilt aversion by Battigalli, Dufwenberg [2007], Battigalli, Dufwenberg [2009], Attanasi, Nagel [2008], in which a guilt averse agent (say Bob) has a disutility if he believes that his opponent (say Ann) is disappointed by his play, as she receives a lower payoff than the one she originally expected, given her beliefs. The approach previously described is one of the possible models in which players' utilities may be affected by a psychological term (see [Attanasi, Nagel, 2008], [Chang, Smith, 2015], [Hauge, 2016], [Bellemare et al., 2017], [Attanasi et al., 2019]). Standard

[2]We refer to the classical refinements based on properties of stability with respect to perturbations, such as *trembling hand perfection* [Selten, 1975] or *properness* [Myerson, 1978] (see also [Van Damme, 1989] for an extensive survey on Nash equilibrium refinements).

[3]The test concerns dominant (i.e. prioritizing personal status) and submissive (i.e. prioritizing the relationship above their own individual desires) behaviors in individuals with narcissistic and dependent traits.

[4]The behavior may be influenced by an increase in basic individuals' fears or in the accuracy of underlying beliefs.

[5]When one of the players is pro-social, the Nash equilibrium in mixed strategies of the original Battle of Sexes game is not an equilibrium anymore, regardless of the other player attitude.

game theory models cannot be exploited to cover this type of situation, as highlighted by the pioneering paper by Geanakoplos et al. [1989]. Indeed, a wide range of phenomena can be better explained by letting payoff functions depend explicitly on hierarchies of beliefs, that is, utilities depend not only on what every player does but also on what he thinks every player believes, on what he thinks every player believes the others believe, and so on. Therefore, psychological game theory must be taken into account and the classical Nash equilibrium concept must be replaced by the *Psychological Nash Equilibrium*[6].

In this work, we construct a psychological Battle of Sexes game in which a player is sensitive to guilt aversion and we characterize the pyschological Nash equilibria of the game. Then, we investigate whether results are robust when we allow beliefs to be ambiguous. The so-called *strategic ambiguity* arises when players have ambiguous (first-order) beliefs about their opponents' play, and it consistently affects players' behavior and equilibria[7]. Several papers have already investigated the effects of strategic ambiguity in the Battle of Sexes (see [Kelsey, Le Roux, 2018] [Riedel, Sass, 2013] [Marinacci, 2000])[8]. Here we look at ambiguity in the psychological game and apply the general tools from De Marco et al. [2022][9] where the psychological Nash equilibrium concept has been generalized to the cases in which the entire hierarchy of beliefs might be ambiguous. In this paper, we characterize such *psychological Nash equilibria under ambiguity* in the Battle of Sexes game.

The results that we get, both from the non-ambiguous and the ambiguous model, are that the equilibria are confirmed to be extremely robust also in the case of guilt aversion. In particular, the two pure strategy equilibria are not affected at all by the presence of a guilt averse player, while in the mixed strategy equilibrium, only the strategy of the player that has no psychological utility (the non-psychological player Ann) is affected by guilt aversion. In the non-ambiguous case, the equilibrium strategy of the non-psychological player prescribes that she will play the less preferred strategy with a probability higher than the one she would have played without any psychological effects. Moreover, the former probability increases as the sensitivity to guilt of the psychological player (Bob) increases; in other words, Ann is willing to accept a lower expected utility to compensate Bob's disutility due to guilt aversion. Nevertheless, in equilibrium, the non-psychological player Ann will always play her most preferred strategy with a probability higher

[6]Different theoretical frameworks like static or dynamic psychological games were examined ([Geanakoplos et al., 1989], [Battigalli, Siniscalchi, 1999], [Attanasi, Nagel, 2008], [Battigalli, Dufwenberg, 2009], [Battigalli et al., 2019], [Battigalli, Dufwenberg, 2020]), as well as specific models for different kinds of feelings ([Rabin, 1993], [Guerra, Zizzo, 2004], [Battigalli, Dufwenberg, 2007], [Dufwenberg, Kirchsteiger, 2019], just to quote a few).

[7]The effects of strategic ambiguity in games is well established, both theoretically ([Dow, Werlang, 1994], [Lo, 1996], [Marinacci, 2000], [Eichberger, Kelsey, 2000], [Stinchcombe, 2003], [Riedel, Sass, 2013]) and empirically ([Pulford, Colman, 2007], [Eichberger et al., 2008], [Eichberger, Kelsey, 2011], [Kelsey, Le Roux, 2018],[De Marco, Romaniello, 2015]).

[8]More details are given in Remark 4.2.

[9]De Marco et al. [2022] is a first step in which ambiguity theory and psychological game theory have been merged into a unique model.

than the one of the less preferred one. Similarly, the presence of ambiguity affects only the mixed strategy equilibrium: Ann will play the less preferred strategy with a probability higher than the one she would have played without ambiguity, whatever is the sensitivity to guilt of Bob. However, in this case, Ann will play her most preferred strategy with a probability lower than the one of the less preferred one, if Bob's sensitivity to guilt is sufficiently high.

Paper is organized as follows: Section 2 illustrates the main concepts of equilibrium in psychological games that are useful for our analysis. In Section 3, the Battle of Sexes game under guilt aversion is introduced and the characterization of equilibria is presented. Section 4 is devoted to ambiguity and equilibria. Section 5 concludes.

2 Preliminaries

In order to include psychological aspects and ambiguity in the Battle of Sexes game, we take into account the model for static ambiguous psychological games introduced in [De Marco et al., 2022].

Summarily, consider a finite set of players $I = \{1, \dots n\}$ and, for each player $i \in I$, denote with $A_i = \{a_i^1, \dots, a_i^{k(i)}\}$ the finite set of pure strategies of player i. Following the standard notation, $A = A_1 \times \cdots \times A_n = \prod_{i \in I} A_i$ represents the set of strategy profiles and $A_{-i} = A_1 \times \cdots \times A_{i-1} \times A_{i+1} \times \cdots \times A_n = \prod_{j \neq i} A_j$ represents the set of i's opponents strategy profiles. The set Σ_i denotes the set of mixed strategies of player i and $\Sigma = \prod_{i \in I} \Sigma_i$, $\Sigma_{-i} = \prod_{j \neq i} \Sigma_j$ denote respectively the set of mixed strategies profiles and the set of i's opponents mixed strategies. We will often use the notation (σ_i, σ_{-i}) with $\sigma_i \in \Sigma_i$ and $\sigma_{-i} \in \Sigma_{-i}$ to indicate the mixed strategies profile $\sigma \in \Sigma$.

Beliefs structure

The hierarchical structure of beliefs is constructed in the following way: denoting with $\Delta(X)$ the set of probability measures on X, then

$$B_i^1 := \Delta(\Sigma_{-i}) \text{ is the set of first order beliefs of player } i,$$
$$B_i^2 := \Delta(\Sigma_{-i} \times B_{-i}^1) \text{ is the set of second order beliefs of player } i,$$
$$\vdots$$
$$B_i^k(\Sigma_{-i} \times B_{-i}^1 \times \cdots \times B_{-i}^{k-1}) \text{ is the set of } k\text{-th order beliefs of player } i,$$

and so on, where

$$B_{-i}^k := \prod_{j \neq i} B_j^k \text{ is the set of } k\text{-th order beliefs of } i\text{'s opponents},$$

for every player i and for every $k \in \mathbb{N}$

Therefore, k-th order beliefs of player i is represented by a probability measure over others' mixed strategies and other's beliefs up to the $(k-1)$-th order. In the end, the set of hierarchies of beliefs of player i is

$$B_i = \prod_{k=1}^{\infty} B_i^k,$$

whose elements are infinite hierarchies of beliefs $b_i = (b_i^1, b_i^2, \cdots, b_i^k, \cdots)$. Therefore, player i's beliefs represent (via probability measures) what player i believes the others will play, what player i thinks the others believe their opponents will play, and so on. Since, for every $k \in \mathbb{N}$, B_i^k is compact and can be metrized as a separable metric space, the set B_i is, in turn, metrizable and separable. Moreover, it results to be compact under the topology induced by this metric (see [De Marco et al., 2022] for further details).

We will restrict our attention to the subset of collectively coherent beliefs $\overline{B}_i \subset B_i$, which is the set of beliefs of player i in which he is sure (i.e. with probability equal to 1) that it is common knowledge that beliefs are coherent. Specifically, a belief $b_i \in B_i$ is said to be coherent if, for every $k \in \mathbb{N}$, the following holds:

$$\operatorname{marg}(b_i^{k+1}, \Sigma_{-i} \times B_{-i}^1, \times \cdots \times B_{-i}^{k-1}) = b_i^k. \tag{1}$$

The set $\overline{B}_i$ is compact as well (the detailed construction of the set of collectively coherent beliefs can be found in [Geanakoplos et al., 1989] and the proof of its compactness in [De Marco et al., 2022]).

A *standard psychological game* is described by $G^{GPS} = \{A_1, \cdots, A_n, u_1, \cdots, u_n\}$ where, for every $i \in I$, the utility functions u_i have the form $u_i : \overline{B}_i \times \Sigma \to \mathbb{R}$ ([Geanakoplos et al., 1989]). In the present work, we will follow the more general model by De Marco et al. [2022], in which it is allowed for ambiguity in the beliefs, i.e. in this case, the belief of player i is represented by a compact subset $K_i \subseteq \overline{B}_i$. This choice enables to cover the case in which agents are uncertain about other players' actions and beliefs; therefore agent i does not have a precise belief b_i as in [Geanakoplos et al., 1989] but knows that the belief can be any $b_i \in K_i$. In order to compare ambiguous alternatives, *maxmin* preferences are considered[10], i.e. each agent i is endowed with an utility function of the form $U_i : \mathscr{K}_i \times \Sigma \to \mathbb{R}$, defined by

$$U_i(K_i, \sigma) = \inf_{b_i \in K_i} u_i(b_i, \sigma) \qquad \forall (K_i, \sigma) \in \mathscr{K}_i \times \Sigma,$$

where $\mathscr{K}_i$ denotes the set of all compact subsets of $\overline{B}_i$. Therefore, a *psychological game under ambiguity* is described by $G = \{A_1, \cdots, A_n, U_1, \cdots, U_n\}$.

Equilibrium notion for G^{GPS} and G

The notion of psychological Nash equilibria introduced in [Geanakoplos et al., 1989] takes into account correctness in equilibrium for the beliefs, i.e. each player is equipped with a function $\beta_i : \Sigma \to \overline{B}_i$ which selects, for every $\sigma \in \Sigma$, the beliefs $\beta_i(\sigma) = (b_i^1, b_i^2, \cdots, b_i^k, \cdots)$ in which player

[10] This model was generalized in [De Marco et al., 2024] for different kind of preferences.

i believes (with probability 1) that his opponents follow the mixed strategies profile σ_{-i} and that each opponent $j \neq i$ believes that his opponents follow σ_{-j}, and so on. In this framework, a *psychological Nash equilibrium* is defined as a pair (b^*, σ^*) where $b^* = (b_1^*, \cdots, b_n^*)$ with $b_i^* \in \overline{B}_i$ and $\sigma^* \in \Sigma$ such that, for every player i,

i) $b_i^* = \beta_i(\sigma^*)$;

ii) $u_i(b_i^*, \sigma^*) \geqslant u_i(b_i^*, (\sigma_i, \sigma_{-i}^*))$ for every $\sigma_i \in \Sigma_i$.

We can also say that $(\beta(\sigma^*), \sigma^*)$ is a psychological Nash equilibrium of the game G^{GPS}.

Since in psychological games under ambiguity beliefs might be ambiguous (or imprecise) in equilibrium, we consider the concept of psychological Nash equilibrium under ambiguity, introduced by De Marco et al. [2022]. Precisely, the function β_i that maps strategy profiles to correct beliefs is replaced by a set-valued map (called ambiguous belief correspondence), that maps strategy profiles to the subsets of those beliefs that players perceive to be consistent with the corresponding strategy profile. In detail, each agent i is now endowed with a set-valued map $\gamma_i : \Sigma \rightsquigarrow \overline{B}_i$, where each set $\gamma_i(\sigma)$ is not empty and compact, i.e. for every $i \in I$:

$$\emptyset \neq \gamma_i(\sigma) \in \mathscr{K}_i \quad \forall \sigma \in \Sigma.$$

Each subset $\gamma_i(\sigma) \subseteq \overline{B}_i$ provides the set of beliefs that player i perceives to be consistent given the strategy profile σ. In particular, a *psychological Nash equilibrium under ambiguity* of the game G with belief correspondences $\gamma = (\gamma_1, \ldots, \gamma_n)$ is a pair (K^*, σ^*), where $K^* = (K_1^*, \ldots, K_n^*)$ with $K_i^* \subseteq \overline{B}_i$ and $\sigma^* \in \Sigma$ such that, for every player i:

i) $K_i^* = \gamma_i(\sigma^*)$;

ii) $U_i(K_i^*, \sigma^*) \geqslant U_i(K_i^*, (\sigma_i, \sigma_{-i}^*))$ for every $\sigma_i \in \Sigma_i$.

In this case, we can also say that $(\gamma(\sigma^*), \sigma^*)$ is a psychological Nash equilibrium under ambiguity. It is necessary to underline that the concept of psychological Nash equilibrium under ambiguity is γ-dependent, in the sense that same games but with different γ correspondences may have different equilibria.

Summary utility functions and best replies

One can characterize psychological Nash equilibria under ambiguity as classical Nash equilibria via the functions $w_i : \Sigma \times \Sigma \to \mathbb{R}$ called *summary utility functions* defined by:

$$w_i(\sigma, \tau) = U_i(\gamma_i(\sigma), \tau) = \inf_{b_i \in \gamma_i(\sigma)} u_i(b_i, \tau), \qquad \forall (\sigma, \tau) \in \Sigma \times \Sigma. \tag{2}$$

Indeed, as shown in Lemma 2.8 in [De Marco et al., 2022], $(\gamma(\sigma^*), \sigma^*)$ is a psychological Nash equilibrium under ambiguity if and only if, for every $i \in I$,

$$w_i((\sigma_i^*, \sigma_{-i}^*), (\sigma_i^*, \sigma_{-i}^*)) \geqslant w_i((\sigma_i^*, \sigma_{-i}^*), (y_i, \sigma_{-i}^*)), \qquad \forall y_i \in \Sigma_i, \tag{3}$$

or equivalently $\sigma_i^* \in BR_i(\sigma_{-i}^*)$ with

$$BR_i(\sigma_{-i}^*) = \left\{\sigma_i \in \Sigma_i \,|\, w_i((\sigma_i, \sigma_{-i}^*), (\sigma_i, \sigma_{-i}^*)) \geqslant w_i((\sigma_i, \sigma_{-i}^*), (y_i, \sigma_{-i}^*)),\ \forall y_i \in \Sigma_i\right\}. \tag{4}$$

A similar formulation was firstly obtained by Geanakoplos et al. [1989] for standard psychological Nash equilibria; in that case, the summary utility functions are defined as follows:

$$w_i^{GPS}(\sigma, \tau) = u_i(\beta_i(\sigma), \tau), \qquad \forall (\sigma, \tau) \in \Sigma \times \Sigma. \tag{5}$$

Then $(\beta(\sigma^*), \sigma^*)$ is a psychological Nash equilibrium if and only if, for every $i \in I$,

$$w_i^{GPS}((\sigma_i^*, \sigma_{-i}^*), (\sigma_i^*, \sigma_{-i}^*)) \geqslant w_i^{GPS}((\sigma_i^*, \sigma_{-i}^*), (y_i, \sigma_{-i}^*)), \qquad \forall y_i \in \Sigma_i, \tag{6}$$

or equivalently $\sigma_i^* \in BR_i^{GPS}(\sigma_{-i}^*)$ with

$$BR_i^{GPS}(\sigma_{-i}^*) = \left\{\sigma_i \in \Sigma_i \,|\, w_i^{GPS}((\sigma_i, \sigma_{-i}^*), (\sigma_i, \sigma_{-i}^*)) \geqslant w_i^{GPS}((\sigma_i, \sigma_{-i}^*), (y_i, \sigma_{-i}^*)),\ \forall y_i \in \Sigma_i\right\}. \tag{7}$$

Finally, note that in case of no ambiguity and correct beliefs, i.e. when γ_i reduces to β_i, the correspondence BR_i coincides with BR_i^{GPS}. If, additionally, there are no psychological terms in the utility function of player i (which then coincides with the classical expected utility), BR_i gives the classical best reply correspondence.

3 Battle of Sexes game with guilt aversion

In this section we analyze the Battle of Sexes game, introducing the sentiment of guilt in the payoff of Player 2 (Bob), who feels guilty to let Player 1 (Ann) down. Our result shed light on the fact that the sentiment of guilt generated by the disappointment of others' expectation has a significant impact on the completely mixed equilibrium strategy of the standard game, while the two pure strategy equilibria are robust with respect to this kind of preference's perturbation. Moreover, it turns out that, in the former case, the equilibrium mixed strategy played by Ann, whose payoff is not affected by any psychological aspect, actually depends on Bob's sensitivity to guilt.

The standard game

The Battle of Sexes game in its standard form consists in the following: Player 1 (Ann) and Player 2 (Bob) have to choose which event to partecipate in the evening, going to the Opera or to a Boxing match, so the strategy sets are $A_1 = A_2 = \{\text{Opera, Boxing}\}$. Bob prefers the Boxing match while Ann prefers the Opera, but they both prefer to go out together. The matrix form is the following:

Player 2 / Player 1	Boxing	Opera
Boxing	$1, 2$	$0, 0$
Opera	$0, 0$	$2, 1$

A generic mixed strategy of Player 1 (Ann) is $p \in [0,1]$, where as usual (with an abuse of notation), $p = \text{prob}(Boxing)$ and $1-p = \text{prob}(Opera)$. Similarly, a generic mixed strategy of Player 2 (Bob) is $q \in [0,1]$, where (with an abuse of notation), $q = \text{prob}(Boxing)$ and $1-q = \text{prob}(Opera)$. Therefore, (p,q) denotes a generic strategy profile. Note that the set of mixed strategies reduces to $\Sigma_1 = [0,1]$ and $\Sigma_2 = [0,1]$. The term $\pi_1(p,q)$ (resp. $\pi_2(p,q)$) denotes the classical expected payoff of Player 1 (resp. Player 2) for every $(p,q) \in \Sigma_1 \times \Sigma_2$. For the sake of simplicity, we denote with a_1 the generic pure strategy of Ann where $a_1 = 1$ means that Ann plays *Boxing* and $a_1 = 0$ means that Ann plays *Opera*. Similarly, we denote with a_2 the generic pure strategy of Bob where $a_2 = 1$ means that Bob plays *Boxing* and $a_2 = 0$ means that Bob plays *Opera*.

In this form the game has three equilibria:

- $(p,q) = (1,1)$: both players choose the Boxing Match;
- $(p,q) = (0,0)$: both players choose the Opera;
- $(p,q) = \left(\frac{1}{3}, \frac{2}{3}\right)$: both players randomize.

Modeling guilt aversion

Guilt aversion arises when harming others, and could be explained by many reasons. However, it is often a consequence of letting others down. This is the perspective mainly considered in the theoretical papers devoted to this issue. We refer to the models by [Battigalli, Dufwenberg, 2007], [Battigalli, Dufwenberg, 2009], [Attanasi, Nagel, 2008] in which it is said that player j is let down if his actual material payoff $\hat{\pi}_j$, received after the play, is lower than the payoff he initially expected to get $E_j[\pi_j^e]$. Therefore, player j disappointment is given by

$$\max\{0, E_j[\pi_j^e] - \hat{\pi}_j\}.$$

Given the strategy profile σ, player i's beliefs b_i and player i's guilt-sensitivity parameter $\theta_i > 0$, the guilt-dependent utility of Player i can be constructed as follows:

$$u_i(b_i, \theta_i, \sigma) = \hat{\pi}_i(\sigma) - \theta_i \max\{0, E_j[\pi_j^e(\sigma), b_i] - \hat{\pi}_j(\sigma)\}, \tag{8}$$

where $\max\{0, E_j[\pi_j^e(\sigma), b_i] - \hat{\pi}_j(\sigma)\}$ represents player i's expectation of player j disappointment. In particular, $E_j[\pi_j^e(\sigma), b_i]$ represents what player i believes is the payoff that player j initially expects to get and $\hat{\pi}_j(\sigma)$ is what player j actually gets.

We build on this guilt aversion model and consider a variation of the Battle of Sexes game in which Bob feels guilty about letting Ann down. Alternatively, Ann is not affected by psychological factors. Let $(p, q) \in \Sigma_1 \times \Sigma_2$ denote the mixed strategy of the players, $\tilde{p}$ is the expectation of Ann's first-order beliefs about Bob's strategy choice q and $\tilde{q}$ is the expectation of Bob's second-order beliefs about the first-order belief of Ann $\tilde{p}$. The term θ_B denotes Bob's sensitivity to guilt.

Now, recall that π_1, π_2 are the classical expected utility functions. If a pure strategy profile, say (a_1, a_2), with $a_1, a_2 \in \{0, 1\}$, is actually played, Bob believes that Ann initially expected a payoff $E_1[\pi_1^e(a_1, a_2), \tilde{q}]$ which corresponds to Ann choosing a_1 and Bob randomizing with probabilities $\tilde{q}$ and $1 - \tilde{q}$; more precisely

$$E_1[\pi_1^e(a_1, a_2), \tilde{q}] = \pi_1(a_1, \tilde{q}) = \tilde{q}\pi_1(a_1, 1) + (1 - \tilde{q})\pi_1(a_1, 0),$$

Then, for every pure strategy profile $(a_1, a_2) \in \{0, 1\}^2$, Bob's utility is

$$u_2(\tilde{q}, \theta_B, (a_1, a_2)) = \pi_2(a_1, a_2) - \theta_B \max\{0, E_1[\pi_1^e(a_1, a_2), \tilde{q}] - \pi_1(a_1, a_2))\},$$

which leads to the following psychological game:

Player 1 \ Player 2	Boxing	Opera
Boxing	$1, 2$	$0, -\theta_B\tilde{q}$
Opera	$0, -2\theta_B(1 - \tilde{q})$	$2, 1$

Psychological Nash equilibria

We now analize the psychological Nash equilibria of the game introduced above. Ann's expected utility does not depend on any psychological term so her best reply correspondence coincides with the one of the standard game:

$$BR_1(q) = \begin{cases} 0 & \text{if } q < 2/3, \\ [0, 1] & \text{if } q = 2/3, \\ 1 & \text{if } q > 2/3. \end{cases} \tag{9}$$

Bob's expected utility, instead, is computed as follows from the psychological game:

$$u_2(\tilde{q}, \theta_B, (p, y)) = 2py - \theta_B \tilde{q}(1-y)p - 2\theta_B(1-\tilde{q})(1-p)y + (1-p)(1-y) =$$
$$y(3p - 1 - 2\theta_B + 2\theta_B p + 2\theta_B \tilde{q} - \theta_B p \tilde{q}) - \theta_B p \tilde{q} + 1 - p. \tag{10}$$

Given a mixed strategy profile $(p, q) \in \Sigma_1 \times \Sigma_2$, the correct belief function $\beta_2(p, q)$ imposes that[11] the corresponding second-order belief $\tilde{q}$ of Bob must coincide with q, i.e. $\tilde{q} = q$. It follows that the summary utility function defined in (5) is:

$$w_2^{GPS}((p', q), (p, y)) = u_2(q, \theta_B, (p, y)) \tag{11}$$

for all $(p', q), (p, y) \in \Sigma_1 \times \Sigma_2$.
Therefore, Bob's best reply, constructed as in (7), takes the form of the correspondence

$$BR_2^{GPS} : [0, 1] \rightsquigarrow [0, 1]$$

defined by

$$BR_2^{GPS}(p) = \{q \in [0, 1] \mid w_2^{GPS}((p, q), (p, q)) \geqslant w_2^{GPS}((p, q), (p, y)) \ \forall y \in [0, 1]\} =$$
$$= \{q \in [0, 1] \mid u_2(q, \theta_B, (p, q)) \geqslant u_2(q, \theta_B, (p, y)) \ \forall y \in [0, 1]\} \quad \forall p \in [0, 1]. \tag{12}$$

The next proposition characterizes Bob's best reply and the psychological Nash equilibria of the game.

Proposition 3.1: *For every $\theta_B > 0$, Bob's best reply correspondence is given by the following:*

$$BR_2^{GPS}(p) = \begin{cases} 0 & \text{if } 0 \leqslant p < \frac{1}{3+\theta_B}, \\ \{0, 1\} & \text{if } p = \frac{1}{3+\theta_B}, \\ \{0, 1, Q(p, \theta_B)\} & \text{if } \frac{1}{3+\theta_B} < p < \frac{1+2\theta_B}{3+2\theta_B}, \\ \{0, 1\} & \text{if } p = \frac{1+2\theta_B}{3+2\theta_B}, \\ 1 & \text{if } \frac{1+2\theta_B}{3+2\theta_B} < p \leqslant 1, \end{cases} \tag{13}$$

where

$$Q(p, \theta_B) = \frac{1 - 3p + 2\theta_B - 2\theta_B p}{(2-p)\theta_B} \qquad \forall p \in \left[\frac{1}{3+\theta_B}, \frac{1+2\theta_B}{3+2\theta_B}\right],$$

and the function $Q(\cdot, \theta_B) : \left[\frac{1}{3+\theta_B}, \frac{1+2\theta_B}{3+2\theta_B}\right] \to \mathbb{R}$ is strictly decreasing with

$$Q\left(\frac{1}{3+\theta_B}, \theta_B\right) = 1 \quad \text{and} \quad Q\left(\frac{1+2\theta_B}{3+2\theta_B}\right) = 0. \tag{14}$$

Therefore, the Battle of Sexes game with guilt aversion has the following PNE:

[11] A precise condition for Bob's correct belief would require much more details that, however, are irrelevant for the present model. The idea is that Bob's hierarchy of belief belonging to $\beta_2(p, q)$ should have a second-order belief whose expectation $\tilde{q}$ corresponds to q. As the expectation $\tilde{q}$ is the unique psychological term that affects Bob's utility, we don't need to characterize Bob's entire hierarchy of beliefs.

i) $(p, q) = (0, 0)$;

ii) $(p, q) = (1, 1)$;

iii) $(p, q) = \left(\frac{3+2\theta_B}{9+4\theta_B}, 2/3\right)$.

Proof. From equation (10), we know that, given any mixed strategy p of Player 1 (Ann), Bob's expected utility is:

$$u_2(\tilde{q}, \theta_B, (p, y)) = y(3p - 1 - 2\theta_B + 2\theta_B p + 2\theta_B \tilde{q} - \theta_B p\tilde{q}) - \theta_B p\tilde{q} + 1 - p.$$

Equation (12) ensures that $q \in BR_2^{GPS}(p)$ if and only if

$$u_2(q, \theta_B, (p, q)) = w_2^{GPS}((p, q), (p, q)) \geqslant w_2^{GPS}((p, q), (p, y)) = u_2(q, \theta_B, (p, y)) \quad \forall y \in [0, 1].$$

Consequently, we need to maximize $u_2(\tilde{q}, \theta_B, (p, y))$ with respect to y, under the condition that $\tilde{q}$ must be consistent with the maximum point.

Firstly observe that:

(1) If

$$(3p - 1 - 2\theta_B + 2\theta_B p + 2\theta_B \tilde{q} - \theta_B p\tilde{q}) > 0, \tag{15}$$

then $u_2(\tilde{q}, \theta_B, (p, \cdot))$ is strictly increasing, therefore it is maximized only for $y = 1$. The consistency condition with the maximum for correct beliefs implies that $\tilde{q} = 1$; it follows that (15) becomes

$$3p - 1 - 2\theta_B + 2\theta_B p + 2\theta_B - \theta_B p = p(3 + \theta_B) - 1 > 0.$$

Hence, when $p \in \left]\frac{1}{3+\theta_B}, 1\right]$,

$$w_2^{GPS}((p, 1), (p, 1)) \geqslant w_2^{GPS}((p, 1), (p, y)) \qquad \forall y \in [0, 1],$$

and $1 \in BR_2^{GPS}(p)$.

In this case, there cannot be any other best reply.

(2) If

$$(3p - 1 - 2\theta_B + 2\theta_B p + 2\theta_B \tilde{q} - \theta_B p\tilde{q}) < 0, \tag{16}$$

then $u_2(\tilde{q}, \theta_B, (p, \cdot))$ is strictly decreasing and it is maximized only for $y = 0$. In this case, the consistency condition with the maximum for correct beliefs implies that $\tilde{q} = 0$; it follows that (16) becomes $3p - 1 - 2\theta_B + 2\theta_B p = p(3 + 2\theta_B) - (1 + 2\theta_B) < 0$. Hence, for $p \in \left[0, \frac{1+2\theta_B}{3+2\theta_B}\right[$, it follows that

$$w_2^{GPS}((p, 0), (p, 0)) \geqslant w_2^{GPS}((p, 0), (p, y)) \qquad \forall y \in [0, 1],$$

and $0 \in BR_2^{GPS}(p)$.

In this case, there cannot be any other best reply.

(3) If

$$(3p - 1 - 2\theta_B + 2\theta_B p + 2\theta_B \tilde{q} - \theta_B p \tilde{q}) = 0, \tag{17}$$

then $u_2(\tilde{q}, \theta_B, (p, y)) = -\theta_B p \tilde{q} + 1 - p$, therefore every $y \in [0,1]$ maximizes $u_2(\tilde{q}, \theta_B, (p, \cdot))$. However, (17) implies that

$$\tilde{q} = \frac{1 - 3p + 2\theta_B - 2\theta_B p}{(2-p)\theta_B} := Q(p, \theta_B).$$

So $q = Q(p, \theta_B)$ is the unique Bob's best reply in this case, provided that $Q(p, \theta_B) \in [0,1]$. It can be easily checked that

$$\frac{1}{3+\theta_B} < \frac{1+2\theta_B}{3+2\theta_B} \qquad \text{for every } \theta_B > 0, \tag{18}$$

that (14) holds and that, in the interval $\left[\frac{1}{3+\theta_B}, \frac{1+2\theta_B}{3+2\theta_B}\right]$, $Q(\cdot, \theta_B)$ is strictly decreasing and attains only the values in $[0,1]$.

Summarising, for $p \in \left[\frac{1}{3+\theta_B}, \frac{1+2\theta_B}{3+2\theta_B}\right]$,

$$w_2^{GPS}\left((p, Q(p,\theta_B)), (p, Q(p,\theta_B))\right) \geqslant w_2^{GPS}\left((p, Q(p,\theta_B)), (p, y)\right) \qquad \forall y \in [0,1],$$

and $Q(p, \theta_B) \in BR_2^{GPS}(p)$.

In this case, there cannot be any other best reply.

From the previous arguments it follows directly the characterization of $BR_2^{GPS}(p)$ given in (13).

Finally, from Ann's best reply, described in (9), it follows that $(0,0)$ and $(1,1)$ are PNE. Moreover there exists only another equilibrium that is in completely mixed strategies. In this equilibrium, Bob must play $q = 2/3$ as a best reply to Ann's mixed strategy $\overline{p}$. We find $\overline{p}$ as follows:

$$\frac{2}{3} = Q(\overline{p}, \theta_B) \implies \overline{p} = \frac{3+2\theta}{9+4\theta}$$

Hence the third equilibrium is $(p, q) = \left(\frac{3+2\theta_B}{9+4\theta_B}, 2/3\right)$. □

Equilibrium analysis

Previous results substantially confirm that the equilibria of the game are stable with respect to perturbations. It is well known that the Battle of Sexes is a generic game; as such, all the equilibria are stable with respect to classical perturbations and no refinement concept, such as *trembling hand perfection* [Selten, 1975] or *properness* [Myerson, 1978][12], can be effective. Moreover, joint deviations are not Pareto improving, whatever the equilibrium. In our case, the presence of guilt

[12] See also [Van Damme, 1989] for an extensive survey on Nash equilibrium refinements.

aversion does not affect the two pure strategy equilibria of the standard game, while the completely mixed strategy equilibrium $(1/3, 2/3)$ of the standard game is the limit point of the equilibrium $\left(\frac{3+2\theta_B}{9+4\theta_B}, 2/3\right)$ of the game when guilt aversion vanishes (that is, θ_B converges to 0).

Nevertheless, there are interesting insights arising from the mixed strategy equilibrium. First of all, Bob's sensitivity to guilt affects only the equilibrium strategy of Ann. While this result might seem surprising, it is instead consistent with previous literature: the paper by Attanasi, Nagel [2008] analyzes a sequential Trust game with guilt aversion and finds that θ_B affects only the equilibrium strategy equilibrium of the player with no guilty feelings. Moreover, as the function $\frac{3+2\theta_B}{9+4\theta_B}$ is strictly increasing, in equilibrium Ann will play Boxing with an higher probability with respect to the case $\theta_B = 0$, meaning that she will have a lower expected utility to compensate Bob's disutility of the guilty feelings. However, when $\theta_B \to +\infty$, $\frac{3+2\theta_B}{9+4\theta_B}$ increases to $1/2$ so that she will never play *Boxing* with a probability greater then the one of *Opera*.

4 Equilibria under Ambiguity

In this section, we consider the case in which Bob has ambiguous second-order beliefs, analyzing the corresponding effect on equilibria. In particular, we consider the event of full ignorance: Bob ignores completely what Ann believes his choice will be. From the mathematical point of view, Bob is endowed with a belief correspondence γ_2 that does not impose any restriction on his second-order beliefs so that $\tilde{q}$ can be any element in the interval $[0, 1]$. Therefore, the summary utility function of Bob, defined in (2), takes the following form:

$$w_2((p', q), (p, y)) = \min_{\tilde{q}\in[0,1]} u_2(\tilde{q}, \theta_B, (p, y)), \tag{19}$$

for all $(p', q), (p, y) \in \Sigma_1 \times \Sigma_2$.
In this case, Bob's best reply (defined in (4)) takes the form of the correspondence $BR_2 : [0, 1] \rightsquigarrow [0, 1]$ defined, for every $p \in [0, 1]$, by

$$\begin{aligned} BR_2(p) &= \{q \in [0,1] \mid w_2((p, q), (p, q)) \geqslant w_2((p, q), (p, y)) \; \forall y \in [0, 1]\} = \\ &= \{q \in [0,1] \mid \min_{\tilde{q}\in[0,1]} u_2(\tilde{q}, \theta_B, (p, q)) \geqslant \min_{\tilde{q}\in[0,1]} u_2(\tilde{q}, \theta_B, (p, y)) \quad \forall y \in [0, 1]\}. \end{aligned} \tag{20}$$

The next proposition gives the characterization of Bob's best reply correspondence and of the equilibria in case of ambiguity.

Proposition 4.1: *Bob's best reply correspondence is given by the following:*

$$BR_2(p) = \begin{cases} 0 & \text{if } 0 \leqslant p < \frac{1}{3+\theta_B}, \\ \left[0, \frac{p}{2-p}\right] & \text{if } p = \frac{1}{3+\theta_B}, \\ \frac{p}{2-p} & \text{if } \frac{1}{3+\theta_B} < p < \frac{1+2\theta_B}{3+2\theta_B}, \\ \left[\frac{p}{2-p}, 1\right] & \text{if } p = \frac{1+2\theta_B}{3+2\theta_B}, \\ 1 & \text{if } \frac{1+2\theta_B}{3+2\theta_B} < p \leqslant 1. \end{cases} \tag{21}$$

Therefore, the Battle of Sexes game with guilt aversion gives the following PNEUA:

- *If* $\theta_B > \frac{7}{2}$*:*
  - *i)* $(p, q) = (0, 0)$*;*
  - *ii)* $(p, q) = (1, 1)$*;*
  - *iii)* $(p, q) = \left(\frac{4}{5}, \frac{2}{3}\right)$.
- *If* $\theta_B \leqslant \frac{7}{2}$*:*
  - *i)* $(p, q) = (0, 0)$*;*
  - *ii)* $(p, q) = (1, 1)$*;*
  - *iii)* $(p, q) = \left(\frac{1+2\theta_B}{3+2\theta_B}, \frac{2}{3}\right)$.

Proof. From the characterization (19) of the summary utility functions and from the formula of Bob's expected utility function (10), we get

$$w_2((p, q), (p, y)) = \min_{\tilde{q} \in [0,1]} \theta_B \tilde{q}(2y - p - py) + y[3p - 1 - 2\theta_B(1-p)] + 1 - p.$$

Then, for every $\theta_B > 0$, we get:

$$\arg\min_{\tilde{q} \in [0,1]} u_2(\tilde{q}, \theta_B, (p, y)) = \begin{cases} 0 & \text{if } 2y - p - py > 0, \\ 1 & \text{if } 2y - p - py < 0, \\ [0, 1] & \text{if } 2y - p - py = 0. \end{cases}$$

Consequently, for every (p, y) we have that:

$$w_2((p, q), (p, y)) = \min_{\tilde{q} \in [0,1]} u_2(\tilde{q}, \theta_B, (p, y)) = \begin{cases} y[3p - 1 - 2\theta_B(1-p)] + 1 - p & \text{if } 2y - p(1+y) \geqslant 0, \\ y(3p - 1 + \theta_B p) + 1 - p - \theta_B p & \text{if } 2y - p(1+y) < 0. \end{cases}$$

To compute the best reply BR_2, recall that

$$q \in BR_2(p) \iff w_2((p, q), (p, q)) \geqslant w_2((p, q), (p, y)) \qquad \forall y \in [0, 1].$$

Now:

(1) If

$$2y - p(1+y) \geqslant 0, \quad \text{i.e. } y \geqslant \frac{p}{2-p},$$

then

$$w_2((p,q),(p,y)) = y[3p - 1 - 2\theta_B(1-p)] + 1 - p.$$

Therefore:

- for $3p - 1 - 2\theta_B(1-p) > 0$, that is $p > \frac{1+2\theta_B}{3+2\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is strictly increasing in $[\frac{p}{2-p}, 1]$;
- for $3p - 1 - 2\theta_B(1-p) < 0$, that is $p < \frac{1+2\theta_B}{3+2\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is strictly decreasing in $[\frac{p}{2-p}, 1]$;
- for $3p - 1 - 2\theta_B(1-p) = 0$, that is $p = \frac{1+2\theta_B}{3+2\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is constant in $[\frac{p}{2-p}, 1]$.

(2) If

$$2y - p(1+y) < 0, \quad \text{i.e. } y < \frac{p}{2-p},$$

then

$$w_2((p,q),(p,y)) = y(3p - 1 + \theta_B p) + 1 - p - \theta_B p.$$

Therefore:

- for $3p-1+\theta_B p > 0$, that is $p > \frac{1}{3+\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is strictly increasing in $[0, \frac{p}{2-p}[$;
- for $3p-1+\theta_B p < 0$, that is $p < \frac{1}{3+\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is strictly decreasing in $[0, \frac{p}{2-p}[$;
- for $3p-1+\theta_B p = 0$, that is $p = \frac{1}{3+\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is constant in $[0, \frac{p}{2-p}[$.

Recall that (18) holds, that is

$$\frac{1}{3+\theta_B} < \frac{1+2\theta_B}{3+2\theta_B} \qquad \text{for every } \theta_B > 0.$$

It follows that:

a) If $0 \leqslant p < \frac{1}{3+\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is strictly decreasing in the interval $[0,1]$ and attains its maximum in $y = 0$. Therefore:

$$w_2((p,0),(p,0)) \geqslant w_2((p,0),(p,y)) \qquad \forall y \in [0,1].$$

b) If $p = \frac{1}{3+\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is constant and equal to $1 - p - \theta_B p$ in the interval $[0, \frac{p}{2-p}]$ and strictly decreasing in the interval $[\frac{p}{2-p}, 1]$. Then every $q \in [0, \frac{p}{2-p}]$ is a maximum point. So, for every $q \in [0, \frac{p}{2-p}]$, we have

$$w_2\left((p,q),(p,q)\right) \geqslant w_2\left((p,q),(p,y)\right) \qquad \forall y \in [0,1].$$

c) If $\frac{1}{3+\theta_B} < p < \frac{1+2\theta_B}{3+2\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is strictly increasing in $[0, \frac{p}{2-p}]$ and strictly decreasing in $[\frac{p}{2-p}, 1]$. Hence the unique maximum point is attained at $y = \frac{p}{2-p}$. Therefore :

$$w_2\left(\left(p, \frac{p}{2-p}\right), \left(p, \frac{p}{2-p}\right)\right) \geqslant w_2\left(\left(p, \frac{p}{2-p}\right), (p,y)\right) \quad \forall y \in [0,1].$$

d) If $p = \frac{1+2\theta_B}{3+2\theta_B}$, the function $w_2((p,q),(p,\cdot))$ is increasing for $y \in [0, \frac{p}{2-p}]$, while it is constant and equal to $1-p$ in the interval $[\frac{p}{2-p}, 1]$. This means that every $q \in [\frac{p}{2-p}, 1]$ is a maximum point. So, for every $q \in [\frac{p}{2-p}, 1]$:

$$w_2\left((p,q),(p,q)\right) \geqslant w_2\left((p,q),(p,y)\right) \qquad \forall y \in [0,1].$$

e) If $\frac{1+2\theta_B}{3+2\theta_B} < p \leqslant 1$, the function $w_2((p,q),(p,y))$ is strictly increasing for $y \in [0,1]$ and attains its maximum in $y = 1$. Therefore:

$$w_2((p,1),(p,1)) \geqslant w_2((p,1),(p,y)) \qquad \forall y \in [0,1].$$

The previous arguments immediately imply that the best reply BR_2 satisfies (21).

Now we look at the equilibria of the game; from the previous characterization of BR_2 and from the characterization of BR_1 given in (9), it immediately follows that $(p,q) = (0,0)$ and $(p,q) = (1,1)$ are PNEUA. Moreover there are no other equilibria in which one of the two players chooses a pure strategy as $BR_i(0)$ and $BR_i(1)$ are singletons for $i = 1,2$.

Here we show that there is another equilibrium in completely mixed strategies. In this kind of equilibrium, Bob's strategy can only be $q = 2/3$ because it is the unique Bob's completely mixed strategy contained in Ann's best reply set. Note also that every Ann's mixed strategy is a best reply to $q = \frac{2}{3}$. So we have only to find some strategy $\overline{p}$ of Ann whose Bob's best reply contains $q = 2/3$. Now,

(i) $\overline{p} \notin [0, \frac{1}{3+\theta_B}[$ and $\overline{p} \notin]\frac{1+2\theta_B}{3+2\theta_B}, 1]$ because in these cases $BR_2(\overline{p}) = \{0\}$ and $BR_2(\overline{p}) = \{1\}$ respectively and $q = \frac{2}{3}$ cannot be a best reply to $\overline{p}$.

(ii) $\overline{p} \neq \frac{1}{3+\theta_B}$ because in this case $BR_2(\overline{p}) = [0, \frac{1}{5+2\theta}]$ while $\frac{1}{5+2\theta_B} < \frac{2}{3}$ for every $\theta_B > 0$. So $q = \frac{2}{3}$ cannot be a best reply to $\overline{p}$.

(iii) If $\frac{1}{3+\theta_B} < \overline{p} < \frac{1+2\theta_B}{3+2\theta_B}$, then (21) implies that $q = \frac{2}{3}$ is a best reply to $\overline{p}$ if and only if $\frac{\overline{p}}{2-\overline{p}} = \frac{2}{3}$ that is equivalent to $\overline{p} = \frac{4}{5}$. However

$$\frac{1}{3+\theta_B} < \frac{4}{5} < \frac{1+2\theta_B}{3+2\theta_B} \quad \text{if and only if } \theta_B > \frac{7}{2}.$$

So $(p,q) = \left(\frac{4}{5}, \frac{2}{3}\right)$ is a PNEUA if and only if $\theta_B > \frac{7}{2}$.

(iv) If $\overline{p} = \frac{1+2\theta_B}{3+2\theta_B}$, then $q = \frac{2}{3}$ is a best reply to $\overline{p}$ if and only if $\frac{2}{3} \in BR_2(\overline{p}) = [\frac{1+2\theta_B}{5+2\theta_B}, 1]$. However,

$$\frac{1+2\theta_B}{5+2\theta_B} \leqslant \frac{2}{3} \text{ if and only if } \theta_B \leqslant \frac{7}{2}.$$

Hence $(p,q) = \left(\frac{1+2\theta_B}{3+2\theta_B}, \frac{2}{3}\right)$ is a PNEUA if and only if $\theta_B \leqslant \frac{7}{2}$.

From the previous arguments, the assertion immediately follows. □

Equilibrium Analysis

Figure 1 shows how Bob's best reply correspondence varies depending on the parameter θ_B. In particular, the previous proposition shows that ambiguity does not affect the two pure strategy equilibria. However, it can be easily checked that $\frac{1+2\theta_B}{3+2\theta_B} > \frac{3+2\theta_B}{9+4\theta_B}$ for every $\theta_B > 0$. This implies that, in the mixed strategy equilibrium, Ann must play her less preferred alternative (*Boxing*) with a probability that is larger than the one in the non-ambiguous case; it follows that ambiguity imposes Ann to reduce her expected utility more since Bob's disutility from guilt is larger in the ambiguous case. Moreover, while in the non-ambiguous case Ann will never play *Boxing* with a probability larger than the one of *Opera*, in the ambiguous one, Ann will play *Boxing* with a probability larger than 1/2 when $\theta_B > 1/2$. This probability is strictly increasing in θ_B as until θ_B reaches the value 7/2, then it becomes constant and equal to 4/5, as illustrated in Figure 2. This implies that Ann will always have an expected utility that is larger than the one she gets playing *Boxing* with probability 1. Summarizing, as for the non-ambiguous case, the equilibria of the Battle of Sexes game are rather robust with respect to the effects of guilt aversion. Nevertheless, the equilibrium in mixed strategy is substantially perturbed when θ_B is sufficiently large and the perturbation is always greater than the corresponding one in the non-ambiguous case.

Remark 4.2: From another perspective, several papers consider the impact of ambiguity on the classical Battle of Sexes model (with no psychological utilities): [Marinacci, 2000] highlights that players' ambiguous beliefs may appear as a consequence of different common culture or history. In his work, a parametric approach is used, leading to the introduction of ambiguous games where non-additive probabilities are exploited as a tool to manage ambiguity. In the context of Ellsberg games, again with the use of capacities, Kelsey, Le Roux [2018] analyze a modified version of the Battle of Sexes game from an experimental point of view. Applying the concept of equilibrium under ambiguity (EUA), previously introduced by Eichberger et al. [2009], they

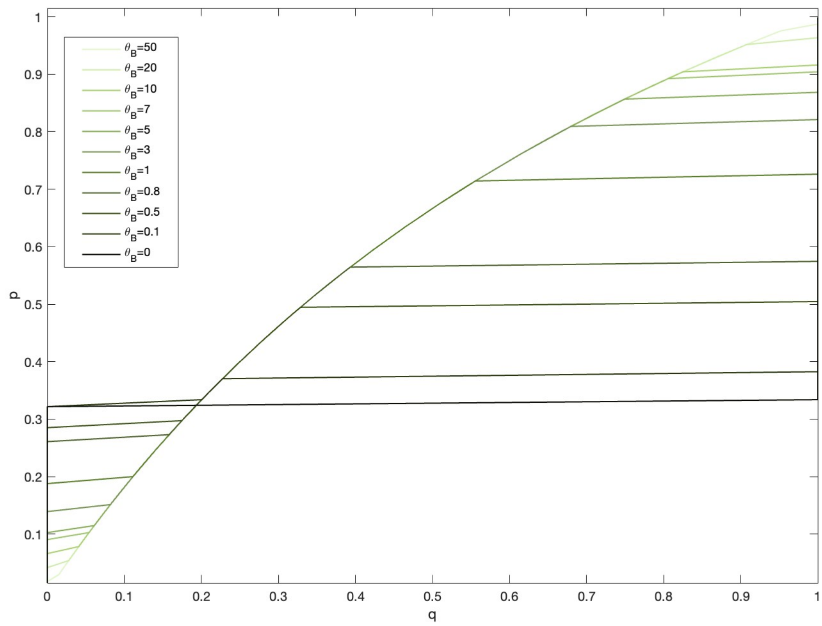

Figure 1: The best reply correspondence BR_2 for different values of θ_B

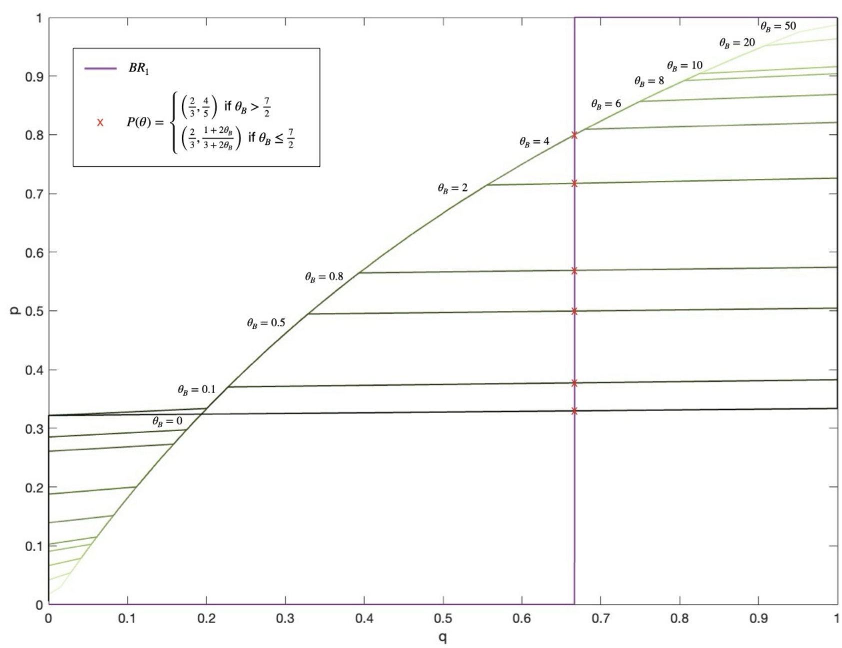

Figure 2: Mixed strategies equilibria for different values of θ_B

conclude that ambiguity aversion strongly influences the behavior of the players, since individuals tend to choose an ambiguity-safe option which results to be off the table by applying the standard Nash equilibrium concept. In the paper [Riedel, Sass, 2013], the Battle of Sexes game is studied as a special case of 2×2 coordination games. They apply the concept of Ellsberg equilibrium (based on an imprecise probabilistic approach) and compute the Ellsberg equilibria in a general 2×2 coordination game. The set of equilibria turns to be given by intervals of probability distributions such that Nash equilibria and max-min strategies provide the boundaries of the largest one.

5 Conclusion

The Battle of Sexes is a very simple game and it turns to have many natural applications. Moreover, it is characterized by the multiplicity of equilibria and by the fact that it is a generic game, so that coordination on a single equilibrium is challenging on the basis of classical game theory. For all these reasons, the model has always attracted the attention of new research.

In this paper, we provide a different perspective since we look at the Battle of Sexes game in case of guilt aversion and study the corresponding psychological game. It turns out that the pure strategy equilibria are confirmed to be extremely robust with respect to *perturbations* of agents' preferences while the mixed strategy equilibrium is sensitive in such a way that a non-psychological (or *material*) player is willing to accept a lower expected utility to compensate others' disutility from guilt. In the ambiguous case the sensitiveness is more effective as it is higher the disutility from guilt. In both cases, the mixed strategy equilibrium remains stable as it converges to the classical one when guilt aversion vanishes. The present paper shows that guilt aversion and ambiguity do not break the *structure* of the equilibria, but they turn to have a clear effect whose interpretation is rather natural but not obvious. Finally, this paper also contributes to the line of research concerning the issue of ambiguous higher order beliefs in psychological games and shows that such issue can be relevant even in the simplest class of games.

References

Attanasi G., Battigalli P., Manzoni E., Nagel R. Belief-dependent preferences and reputation: Experimental analysis of a repeated trust game // Journal of Economic Behavior & Organization. 2019. 167. 341–360.

Attanasi G., Nagel R. A survey of psychological games: theoretical findings and experimental evidence // Games, Rationality and Behavior. Essays on Behavioral Game Theory and Experiments. 2008. 204–232.

Battigalli P., Corrao R., Dufwenberg M. Incorporating belief-dependent motivation in games // Journal of Economic Behavior & Organization. 2019. 167. 185–218.

Battigalli P., Dufwenberg M. Guilt in games // American Economic Review. 2007. 97, 2. 170–176.

Battigalli P., Dufwenberg M. Dynamic psychological games // Journal of Economic Theory. 2009. 144, 1. 1–35.

Battigalli P., Dufwenberg M. Belief-dependent motivations and psychological game theory // CESifo Working Paper. 2020.

Battigalli P., Siniscalchi M. Hierarchies of conditional beliefs and interactive epistemology in dynamic games // Journal of Economic Theory. 1999. 88, 1. 188–230.

Bellemare C., Sebald A., Suetens S. A note on testing guilt aversion // Games and Economic Behavior. 2017. 102. 233–239.

Chang L.J., Smith A. Social emotions and psychological games // Current Opinion in Behavioral Sciences. 2015. 5. 133–140.

De Marco G., Romaniello M. Variational preferences and equilibria in games under ambiguous belief correspondences // International Journal of Approximate Reasoning. 2015. 60. 8–22.

De Marco G., Romaniello M., Roviello A. Psychological Nash equilibria under ambiguity // Mathematical Social Sciences. 2022. 120. 92–106.

De Marco G., Romaniello M., Roviello A. On Hurwicz Preferences in Psychological Games // Games. 2024. 15, 4. 27.

Dow J., Werlang S.R.C. Nash Equilibrium under Uncertainty: Breaking Down Backward Induction // Journal of Economic Theory. 1994. 64. 305–324.

Dufwenberg M., Kirchsteiger G. Modelling kindness // Journal of Economic Behavior & Organization. 2019. 167. 228–234.

Eichberger J., Kelsey D. Non-Additive Beliefs and Strategic Equilibria // Games and Economic Behavior. 2000. 30. 183–215.

Eichberger J., Kelsey D. Are the treasures of game theory ambiguous? // Economic Theory. 2011. 48, 2. 313–339.

Eichberger J., Kelsey D., Schipper B. C. Granny versus game theorist: Ambiguity in experimental games // Theory and Decision. 2008. 64. 333–362.

Eichberger J., Kelsey D., Schipper B. C. Ambiguity and social interaction // Oxford Economic Papers. 2009. 61, 2. 355–379.

Geanakoplos J., Pearce D., Stacchetti E. Psychological games and sequential rationality // Games and Economic Behavior. 1989. 1, 1. 60–79.

Guerra G., Zizzo D. J. Trust responsiveness and beliefs // Journal of Economic Behavior & Organization. 2004. 55, 1. 25–30.

Hauge K.E. Generosity and guilt: The role of beliefs and moral standards of others // Journal of Economic Psychology. 2016. 54. 35–43.

Kelsey D., Le Roux S. Strategic ambiguity and decision-making: An experimental study // Theory and Decision. 2018. 84. 387–404.

Lo K. C. Equilibrium in Beliefs under Uncertainty // Journal of Economic Theory. 1996. 71. 443–484.

Marinacci M. Ambiguous Games // Games and Economic Behavior. 2000. 31. 191–219.

Myerson R.B. Refinements of the Nash equilibrium concept // International Journal of Game Theory. 1978. 7. 73–80.

Pulford B. D., Colman A. M. Ambiguous games: Evidence for strategic ambiguity aversion // Quarterly Journal of Experimental Psychology. 2007. 60, 8. 1083–1100.

Rabin M. Incorporating fairness into game theory and economics // The American economic review. 1993. 1281–1302.

Riedel F., Sass L. Ellsberg Games // Theory and Decision. 2013. 76. 1–41.

Selten R. Reexamination of the perfectness concept for equilibrium points in extensive games // International Journal of Game Theory. 1975. 4. 24–55.

Simon L. I. L.and Reed. Narcissistic and dependent behaviors in the Battle of the Sexes game. // Personality Disorders: Theory, Research, and Treatment. 2021. 12, 3. 286.

Stinchcombe M. Choice with ambiguity as sets of probabilities // Unpublished manuscript. 2003.

Van Damme E. Stability and Perfection of Nash Equilibria. 1989.

Zapata A., Mármol A. M., Monroy L., Caraballo M. A. When the other matters. The battle of the sexes revisited // New Trends in Emerging Complex Real Life Problems. 2018. 501–509.