Scienza e Pace
Science & Peace

ISSN 2039-1749 VOL. XV, N. 1 (2024)

Expanding the educational message conveyed by the "Battle of the Sexes" game

Tommaso Luzzati

Pietro Guarnieri

Francesca Rosa

Rivista online del Centro Interdisciplinare
"Scienze per la Pace" - Università di Pisa

Paper soggetto a **double-blind peer review**

Ricevuto: 12 maggio 2025.
Accettato: 29 luglio 2025.
Pubblicato: 31 luglio 2025.

Per citare il paper:

Luzzati, T., Guarnieri, P., Rosa, F. (2024), “Expanding the educational message conveyed by the ‘Battle of the Sexes’ game”, *Scienza e Pace*, XV (1), pp. 113-134.

Expanding the educational message conveyed by the "Battle of the Sexes" game[+]

Tommaso Luzzati[*], Pietro Guarnieri[**], Francesca Rosa[***]

Abstract

The "Battle of the Sexes" (BoS) game is often used to illustrate the challenges of coordination in competitive situations. However, when reading Luce and Raiffa (1957), who introduced this game, one notices that the narrative they propose does not perfectly align with the payoffs they uses. What happens if the payoffs are modified to better reflect their narrative? How can the game be modified to yield other equilibria, thereby allowing to interpret a wider range of real-life coordination problems? By explicitly introducing an utility function, we propose a general framework that addresses these questions and broaden the educational message that is conveyed.

Keywords: Teaching game theory, Coordination problems, Compromise

Abstract

Il gioco "Battle of the Sexes" (BoS) viene spesso utilizzato per illustrare le sfide del coordinamento in situazioni competitive. Tuttavia, leggendo Luce e Raiffa (1957), che hanno introdotto questo gioco, si nota che la narrazione che propongono non è perfettamente allineata con i payoff che utilizzano. Cosa succede se i guadagni vengono modificati per riflettere meglio la loro narrazione? Come può essere modificato il gioco per ottenere altri equilibri, consentendo così di interpretare una gamma più ampia di problemi di coordinamento nella vita reale? Introducendo esplicitamente una funzione di utilità, proponiamo un quadro generale che affronta queste domande e amplia il messaggio educativo che viene trasmesso.

Parole chiave: Insegnamento della teoria dei giochi, Problemi di coordinamento, Compromesso

---

[+] We are very grateful to the referees for their comments; in particular, one of them pointed out a conceptual error in our original manuscript. This allowed us to thoroughly revise the work and frame it from a new perspective. The paper is the result of a joint effort by the authors. Nonetheless, the different sections can be attributed as follows: §1 to F. Rosa; §2, §3.2, and §5 to P. Guarnieri; and §3.1, §4, and §6 by T. Luzzati. All authors approved the final manuscript.

[*] Department of Economics and Management, University of Pisa. E-mail: tommaso.luzzati@unipi.it.

[**] Department of Economics and Management, University of Pisa.

[***] Oxford Department of International Development (ODID), University of Oxford.

Introduction

In the Preface to his *Lecture notes in microeconomic theory*, Ariel Rubinstein writes "The word 'model' sounds more scientific than 'fable' or 'fairy tale', but I don't see much difference between them" (Rubinstein, 2006, p. X). This sentence expresses a powerful epistemological statement concerning the role of models in economics. In Rubinstein's view, they are not meant neither to represent, nor to predict reality, but to provide a convincing narrative of economic facts. Moreover, "an economic model differs substantially from a purely mathematical model in that it is a combination of a mathematical model and its interpretation." (Rubinstein, 2006, p. IX).

Inconsistency is sometimes observed between the narrative and the mathematical formulation of economic models, even for famous ones. This is the case for the well-known *Battle of the Sexes* game (from here onwards BoS) (Luce and Raiffa, 1957). The BoS game exhibits two Nash equilibria with unequally distributed payoffs, benefitting one player more than the other depending on the equilibrium. Textbooks use it as an example of the challenges posed by coordination in competitive situations, and to discuss some general difficulties with solution concepts in game theory.

Luce and Raiffa state that each player has a preferred event, at the same time the payoff matrix shows that if the other partner does not attend that event, each player is indifferent between going to the disliked event or the preferred one. For instance, the man who goes to the ballet without his partner receives the same payoff as going alone to the prize fight, despite Luce and Raiffa stating that he prefers the fight over the ballet. True, game theorists would argue that the mathematical structure of a game is far more important than the story used to justify it. Indeed, game theory textbooks often present payoffs structures without telling any story[1], treating games as pedagogical tools to teach students the basics. However, telling a specific story is not irrelevant, as it encourages students to attribute epistemic value to the game, that is, to use it as a tool for interpreting reality. If games are to be assigned this epistemic value, narratives play a crucial role in mediating between models and facts, as

[1] The following are some comments we received to a preliminary version of this paper: 'the story that goes with the structure of the game is a pedagogical tool and not the relevant thing', 'I am one of those game theorists who think that the story that accompanies the Battle of the Sexes is an optional extra. The point is not to model the story better, but to analyse the game as traditionally given.'

highlighted in the economic methodology literature (e.g., Morgan, 2001; Grüne-Yanoff & Schweinzer 2008).

Which is then the message that is conveyed by the BoS game? Not only that coordination is difficult, as intended by game theory teachers, but also that either one partner or the other gets higher payoffs in equilibrium, excluding both compromise outcomes and the possibility of going alone. By explicitly modelling how the BoS payoffs can be obtained, we propose a general framework showing first what happens when the inconsistency is eliminated, second, how to get other realistic outcomes. This framework generates different outcomes depending on the relative intensity of the intrinsic preference intensity and will make possible to model when the partners go alone to the preferred event. Moreover, by adding a new option, a compromise between the prize fight and the ballet becomes possible in equilibrium.

In the next section, we recall the origin of the BoS and its use in game theory textbooks. The third section first shows that to make the payoff structure more adhering to the narrative will not change the Luce and Raiffa's outcome, then under what conditions other outcomes become possible as pure strategy equilibria. The fourth section illustrates our framework in the continuous space. Finally, we show how can be used to apply it to an unexpected government agreement that occurred in Italy between *Lega* and *Movimento 5 Stelle* after the 2018 general election. We hope that readers will appreciate our simple proposal not only for strengthening the pedagogical power of the BoS, but also for stimulating the reflection on the interpretation of economic models, their development through the history of economic thought, and their actual relevance.

1. The battle of the sexes

To the best of our knowledge, the first appearance of the BoS game was in the seminal book *Games and Decisions: Introduction and Critical Survey* by Luce and Raiffa (1957). They used this example mainly to show that "the analysis of non-zero-sum games is so much wilder and [...] so much more interesting than is the zero-sum case” (Luce and Raiffa, 1957, p. 92). Unfortunately, textbooks in game theory very rarely mention the important authorship of this game.

As it is well known, the story involves a couple deciding how to spend the evening. The man and the woman have two options: going to a prize fight or to a ballet. The man prefers the fight, and the woman prefers the ballet; however, both would rather go out together than attend their preferred event separately. This story ends up with two Nash equilibria in pure strategies, namely (*Fight*, *Fight*) and (*Ballet, Ballet*), the first of which gives higher payoffs to the man and the second one to the woman. Under the payoff specification given by Luce and Raiffa, and shown in Figure 1, the equilibrium in mixed strategies is [$p^M(F)$=3/5 , $p^W(F)$=2/5], which gives an expected payoff of 1/5 to each player and is Pareto dominated by the equilibria in pure strategies.

		W	
		Fight	*Ballet*
M	*Fight*	**2 1**	-1 -1
	Ballet	-1 -1	**1 2**

Figure 1. The payoff matrix of the BoS in Luce and Raiffa

The BoS belongs to the category of non-cooperative games because of the assumption that players cannot communicate and make binding agreements before playing the game. Each player decides without knowing what the other is opting for – as if the couple forgot where they decided to meet. Luce and Raiffa also discuss what would happen if a player disclosed his/her strategy first. If, for example, player *M* reveals that he opts for *Fight*, and player *W* believes in this announcement, then it is in her interest to choose *Fight* and go to the boxing event. The same holds for an analogous announcement by player *W*. Moreover, the authors emphasise that pre-play communication may have the effect of changing players’ preferences, and hence the payoffs matrix and the game itself. For example, an arrogant approach to the discussion by one of the two may elicit resentment in the other.

Since Luce and Raiffa’s contribution, the BoS has become a classic in game theory textbooks, and gender-neutral versions have been proposed. For instance, *Fight* and *Ballet* have been substituted by concerts where the music played is either by Bach or by Stravinsky (Osborne-Rubinstein, 1994, p. 15). Usually, the game serves the purpose of illustrating the notion of Nash equilibrium and its potential multiplicity. It is often presented alongside other coordination games, such as the *Hawk-Dove* game, also named *Chicken* game (e.g., in Binmore 2007a, pp. 11-12), or the *Stag Hunt* game (e.g., in Fudenberg and Tirole, 1991, p. 18). The BoS game also serves as a valuable didactic tool for explaining mixed-strategy equilibria (e.g. Gintis, 2009, pp. 46-48) and related concepts, such as “pre-play randomization” and “cooperative payoff regions” (Binmore 2007b, p. 199). Furthermore, it is used to introduce key ideas including focal points and Pareto optimality (Fudenberg and Tirole, 1991, p. 18), the effects of restricting alternatives and issuing threats (Davis 2000, pp. 85-89), and the concept of rationalizability in discussions of multiple-equilibria selection (e.g. Tadelis, 2013, pp. 68-79).

Aumann (1974, p. 70) explains the notion of “correlated equilibrium” by using the game shown in Figure 2 with generic labels for strategies.

	Left	*Right*
Top	**2 1**	0 0
Bottom	0 0	**1 2**

Figure 2. The payoff matrix in Aumann (1974)

As in the BoS, the game has three Nash equilibria, two in pure strategies, (*Top*, *Left*) and (*Bottom*, *Right*), and one in mixed strategies, [$p(T)$=2/3, $q(R)$=2/3], implying an expected payoff equal to 2/3 for each player. As noted by Aumann, each player can gain an expected payoff of 3/2 if all players jointly randomise, rather than independently as in mixed strategies. Hence, he formalised the concept of correlated equilibrium based on *correlated strategies*, namely, strategies that are played according to correlated randomization determined by the observation of the same public signal.

This kind of equilibrium was already discussed in Luce and Raiffa (1957, p. 94 and pp. 115-118) who argued that, if pre-play communication was allowed, the couple could toss a coin and agree to jointly play either *Fight* or *Ballet* depending on the toss. Given the signal of the coin, each player would not gain from deviating unilaterally from the signal and, hence, this is an equilibrium involving an expected payoff equal to 3/2 each. An unfair coin would also help because “any point in the convex hull of the Nash equilibrium payoffs of any game can be achieved in a similar fashion and will also be in equilibrium” (Aumann 1974, p. 70[2]).

The notion of correlated equilibrium is illustrated by Fudenberg and Tirole (1991, pp. 53-54) using the payoff matrix shown in Figure 3, a matrix that is also used for the same purpose by Gintis (2009, p. 151). It is important to note that this payoff matrix has the same structure as the Battle of the Sexes, except for the lack of symmetry between (*Left, Down*) and (*Up, Right*). Aumann, Fudenberg and Tirole, and Gintis, while discussing correlated equilibria, do not mention Luce and Raiffa or this similarity.

	Left	*Right*
Up	**5 1**	0 0
Down	4 4	**1 5**

Figure 3. The matrix used in Fudenberg and Tirole (1991, p. 54) to explain the correlated equilibrium.

This game has two equilibria in pure strategies and one in mixed strategies, where each pure strategy is played with probability of yielding a payoff of 2.5. If the players could jointly observe a fair coin toss before playing, they could reach a correlated equilibrium, in which each receives an expected payoff of 3. In the words of Fudenberg and Tirole (1991, p. 69) “[…] players who may engage in pre-play discussion, […] go off to isolated rooms to choose their strategies […] might gain if they could build a ‘signalling device’ that sent signals to the separate rooms”.

Farrell and Rabin (1996, pp. 114-116) explicitly refer to the BoS and propose a “battle of the sexes with one round of talk” to account for the effect of cheap talk on coordination. They underline the lack of realism of the original BoS in the following quotation:

> While [the BoS] is usually analysed as a simple simultaneous-move game, we suggest that this assumes the couple is rather odd. Normal couples, in

[2] Aumann attributes this observation to Harsanyi and Selten (1972), while already Luce and Raiffa (1954, p. 117-118) formulated it implicitly.

> situations like this, talk and try to agree on where to go. Although one can forcibly assume that cheap talk is impossible (they must decide right away, are at separate workplaces and can't phone), doing so strips the game of most of what happens in most strategic interactions whose eventual payoffs look like this. Thus, while the (directly) payoff-relevant moves in the battle of the sexes can be represented as this bimatrix, the strategic interaction consists of a negotiation: in other words, this game is preceded by talking about what they will do.

In other words, the authors assume that the BoS decision is preceded by a round of cheap talk, in which each player simultaneously announces a plan concerning the decision at stake in the second round. Such a game has several equilibria. Farrell and Rabin (1996) underline that one salient equilibrium obtained through cheap talk could be the above-mentioned correlated equilibrium where the outcome is (2,1) or (1,2) with equal probability. In this respect, they notice that "Even if somehow the players cannot toss a coin to decide which of these Nash equilibria to play, they could do something equivalent: for instance, each player can name 'odd' or 'even', and if both name the same they play (2,1) and if they differ they play (1,2); of course, there are many other equilibria with the same distribution of outcomes" (p. 115, note 16). A crucial point, however, is that it is also very likely that the two players, instead of using this 'fair' decision method, will argue in favour of his/her preferred equilibrium. For this reason, this modified BoS can be considered a 'negotiation game' in which cheap talk might lead to an agreement.

2. Small variations on the BoS

In this section, we will highlight that the story told by Luce and Raiffa implies neither the payoff structure they assumed nor the restriction of the game to only two options. Accordingly, we will discuss the implications of relaxing these two assumptions.

2.1. *Breaking payoff symmetry through preference intensity*

Luce and Raiffa assume that each player receives a payoff of -1 when they go out separately, implying that if the partner does not join, the player is indifferent between attending the disliked or the preferred event. The implicit assumption here is that going out separately cancels any enjoyment from the event itself. Such indifference is rather odd and is in contrast with what Luce and Raiffa wrote to describe the players' preferences in the BoS story, that is,

> following the usual cultural stereotype the man *much prefers* the fight and the woman the ballet; however, to both it is *more* important that they go out together than that each see the preferred entertainment (Luce and Raiffa 1957, p. 91, our emphasis).

By using the comparative form 'more' they explicitly consider the relative importance of going out together relative to the preference toward the event itself. One could then change the narrative to make it consistent with Luce and Raiffa's payoffs and suggest, for instance, that «for both partners, attending the same event together is of utmost importance. Ending up at separate events would make them so unhappy that they would forgo the man's preference for the fight and the woman's for the ballet».

What if one adjusts the payoffs to align the game with the terms used by Luce and Raiffa, allowing individuals to still derive some enjoyment when each attends their favourite event alone? To model this, let's assume that the utility functions are additively separable into two components, one determined by the personal attitude towards the entertainment event and the other by 'going out together'. Denote the latter component with T and U_T the utility going together at the same event; U_P the utility of attending the preferred entertainment, and U_D the utility of attending the disliked one; for the sake of simplicity let us assume that $U_D=0$. The term 'preferred entertainment' is used by Luce and Raiffa to indicate the personal taste of each partner towards the event itself. We use it interchangeably with 'favourite'. In any case, it does not refer to the technical meaning of preference ordering (over the outcomes).

To remain adherent to Luce and Raiffa's narrative that "to both it is more important that they go out together than that each see the preferred entertainment" (Luce and Raiffa 1957, p. 91, our emphasis), one must assume that $U_T > U_P$. If this is the case, the implied payoffs are such that the equilibrium in pure strategies is the same as in the original BoS, as shown by the example in Figure 4a[3], where $U_T=5$ and $U_P=4$.

[3] Figure 4a has the same structure of the examples used by Auman, Fudenberg and Tirole and Gintis mentioned in the previous section. Here, we reconnect these examples to the original contribution by

However, the equilibrium outcome changes if the relative importance assigned to going out together (*T*) is not sufficiently high, for example if $U_T=3$. In this case, illustrated in Figure 4b, the game becomes a "Deadlock" game (Buchanan, 1977), where the strategy 'preferred event' is dominant so that the Nash equilibrium becomes (*Fight*, *Ballet*) and each partner ends up attending their preferred event.

	Fight		*Ballet*	
Fight	**9**	**5**	4	4
Ballet	0	0	**5**	**9**

Figure 4a: symmetrical preferences and 'Together' more important than the preferred event ($U_T=5 > U_P=4$)

	Fight		*Ballet*	
Fight	**7**	3	**4**	**4**
Ballet	0	0	3	**7**

Figure 4b: symmetrical preferences and 'Together' less important than the preferred event ($U_T=3 < U_P=4$)

The same preference framework is also useful to illustrate the effects of non-symmetrical preferences between individuals. To this purpose, let us label the individuals with *j* and *k* – which guarantee a gender-neutral narrative – and assume that only for one partner *T* is more important than his/her intrinsic preference for the event, for instance $U^j_T>U^j_P$ and $U^k_T<U^k_P$. This generates the payoff matrix of the "Samaritan game" (Helbing et al., 2005; Hwang, 2018), and implies that the unique pure-strategy equilibrium is going together to the preferred event of the partner who assigns relatively higher utility to it. This outcome is shown in Figure 5.

	Fight		*Ballet*	
Fight	**7**	**3**	**4**	2
Ballet	0	0	3	**5**

Figure 5. Only for partner *j* 'together' is stronger than the preference for the preferred event ($U_T=3$, $U^k_P=4$ $U^j_P=2$)

To sum up, by explicitly modelling the utility function into the two components imagined by Luce and Raiffa, one gets a unified framework for interpreting different real-life situations.

Luce and Raiffa and allow for changes in preferences.

2.2 *Getting a compromise by expanding the options*

Another peculiarity of the BoS is that mixed strategies are the only possible compromise when the game is played non-cooperatively. This is a consequence of limiting the number of options to two. In real life, however, considering a broader range of options is a common way to resolve 'conflicts'. For this reason, it is reasonable to introduce an intermediate option, *M* (e.g., going to a movie), with U_M=2 for both individuals. Figures 6a and 6b present two possible games, whose outcomes depend on the relative strength of the two components of the preferences. If 'going out together' is more important than the intermediate option but less important than each individual's favourite choice, two Nash equilibria in pure strategies exist: 'going out separately' and 'going to the movies together' (Figure 6a). However, when the preference for spending the evening together is strong enough for both individuals, attending any event together becomes a Nash equilibrium, as in the BoS game (Figure 6b)[4].

	Fight	*Movie*	*Ballet*
Fight	**7.5** 3.5	4 2	**4 4**
Movie	2 0	**5.5 5.5**	2 4
Ballet	0 0	0 2	3.5 **7.5**

Figure 6a: symmetrical preferences and 'together' rather important (U_T=3.5; U_P=4; U_M=2)

	Fight	*Movie*	*Ballet*
Fight	**9 5**	4 2	4 4
Movie	2 0	**7 7**	2 4
Ballet	0 0	0 2	**5 9**

Figure 6b: symmetrical preferences and 'together' very important (U_T=5; U_P=4; U_M=2)

The changes that we suggested above do not alter the nature of the BoS narrative, which remains a game that illustrates the difficulties of coordination in the presence of conflicting preferences. However, adding options may introduce more reasonable ways to play the game and assist in selecting among equilibria–particularly in the case depicted in Figure 6b. As Schelling (1960) suggested, people are often able to coordinate by relying on a *focal point*, determined by salient characteristics of the strategic options that are not captured by the formal game structure–such as the names attached to the options or the order in which "equilibrium" (*Movie*, *Movie*) may serve as a focal point, as it represents an intermediate option that avoids direct conflict between the players and provides a Pareto-efficient solution.

[4] Obviously, when the intensity of the preference of going together is lower than the preference for the intermediate options themselves, both would go separately to his/her own favourite event.

The case illustrated in Figure 6b is the most straightforward. The (*Movie*, *Movie*) outcome is a strong candidate for a focal point, as it is fairer than either (*Ballet, Ballet*) or *(Fight, Fight)*. Since (*Movie*, *Movie*) minimises the difference between the players' payoffs, it reinforces its role as an intermediate, compromise solution, while also being Pareto optimal. This outcome is analogous to a reasonable way of playing the original BoS–namely, tossing a coin and going together either to the ballet or to the fight. Although this solution is not an equilibrium in a non-cooperative framework, it is considered "equitable" by Luce and Raiffa (1957, p. 94).

In the case illustrated in Figure 6a, the fairness criterion does not discriminate among the multiple equilibria; however, another factor may influence the selection of a focal point. The game in Figure 6a closely resembles a Stag Hunt[5], except that the payoff-dominant equilibrium also represents an intermediate solution. This feature, combined with its payoff dominance, may lead players to perceive it as a focal point.

A possible counterargument is that two strategies (*Ballet* for the man and *Fight* for the woman) are dominated, effectively reducing the game in Figure 6a to a standard 2x2 Stag Hunt. Nonetheless, the mere presence of a third strategy–even if strictly dominated–may make (*Movie, Movie*) appear as a compromise solution and thus a plausible focal point.

In conclusion, coordination can emerge because the intermediate option offers the possibility of an implicit, self-enforcing agreement, even if payoff considerations alone do not compel the players to choose it.

[5] (Movie, Movie) would correspond to hunting a stag together, which is payoff dominant equilibrium, while (Fight, Ballet) would correspond to the risk dominant equilibrium in which each player hunts a hare alone. According to the definitions firstly proposed by Harsanyi and Selten (1988, 80–90, 355–359), a Nash equilibrium is payoff dominant if it is Pareto-superior to all the other Nash equilibria. On the other hand, a Nash equilibrium is risk dominant when it is the least risky for both players given the uncertainty concerning the other player's decision. In other words, in a symmetric 2x2 game, when the two players assign equal probability to the circumstances that the other player will choose one option or the other, and one of the two options results as strictly preferred for both, the strategy profile that they both opt for is the risk-dominant equilibrium.

3. BoS variations in the continuum

The BoS game can be interpreted as a Nash bargaining game, where individuals aim to share a good (e.g., an amount of money). If the total of their *bids* is equal to or less than the available amount, each receives their bid; otherwise, each receives an outside option. In the BoS, the value of the *outside* option is assumed to be less than or equal to zero, reflecting the idea that each partner strongly dislikes going out without the other. In its standard form, the BoS limits players to two strategies, each leading to an unequal distribution of payoffs in equilibrium. As shown above, introducing additional strategies allows for alternative "distribution of the total amount". Moreover, assuming that the *outside* option has a positive value may both turn 'going out separately' into an equilibrium and reduce the number of equilibria.

This section presents a continuous version of what was introduced in section 3. A continuum of multiple non-cooperative equilibria naturally arises in this setting. This multiplicity collapsed into a unique outcome under the cooperative solution proposed by Nash (1950). However, the analysis here remains within the framework of non-cooperative games.

We have already assumed that preferences are represented by an additively separable utility function with two components. One component depends on the individual's action, *x,* and the other on the discrete event *T* (*Together),* which occurs when both players make the same choice (*T*=1,0). The agents have opposing preferences regarding the action *x* itself.

Let us therefore assume that the utility functions are as follows[6]:

$$U^j=\alpha^j x^j+\varphi^j T x_M \qquad (1)$$

$$U^k=\alpha^k(x_M-x^k)+\varphi^k T x_M \qquad (2)$$

where T=1 iff ($x^i = x^j$), T=0 otherwise and $x\in[0,x_M]$

[6] The original BoS payoffs implicitly assume that the two components are not additively separable, for instance, $U^j=\alpha^j x^j T$.

The parameter α indicates the intensity of the preference for the action x, while φ represents the importance attached to coordination–that is, to choosing the same strategy (T). To allow for direct comparability between α and φ, we rescale T by the maximum value of x, denoted x_M. To replicate the payoffs in the discrete cases presented above, we set x_M=4.

Note that this formalisation allows for comparison with the discrete analysis, but it differs from the bargaining game. The key distinction lies in the nature of preferences: in the BoS game, agents have opposing preferences over the available actions, whereas in the bargaining game, both agents derive utility from sharing the same good. It would be straightforward to modify the formalisation and make the game identical to the bargaining game[7].

To analyse the game, the first step is to identify the best response functions. For agent j, their utility from choosing the same strategy as individual k is given by $U^j_{(xj=xk)}=\alpha^j x^k+\varphi^j x_M$, while the utility from choosing their most preferred action independently (i.e., without coordination) is $U^j_{(T=0)}=\alpha^j x_M$. comparing these utilities yields agent j's best response:

$$x^{*j}=x^k \quad \text{for } x^k \geq (1-\varphi^j/\alpha^j)x_M \qquad \text{and} \quad x^{*j} = x_M \text{ otherwise} \qquad (3)$$

The reaction function of agent k is obtained analogously as follows:

$$x^{*k} = x^j \quad \text{for } x^j \leq (\varphi^k/\alpha^k)x_M \qquad \text{and} \quad x^{*k} = x_M \text{ otherwise} \qquad (4)$$

The potential outcomes of the game are determined by two thresholds, which capture the agents' 'willingness to compromise'. These thresholds, however, are not effective when $\varphi^i \geq \alpha^i$ (where i is a generic agent), that is, when "it is *more* important that they go out together than that each see the preferred entertainment" (Luce and Raiffa 1957, p. 91). This condition corresponds to the original BoS game, in which

[7] To this purpose, the mapping between the bids and the strategies needs to be different for each player. For instance, the Ballet would correspond to a bid b=4 for the woman and to a bid b=0 for the man. This would allow us to assume that the utility function increases with the bids for both agents. Hence b^k=2.5 and b^j=1.5 would imply that both play the same strategy, a strategy which is closer to the most preferred strategy by k.

the equilibria reflect a mutual preference for always choosing the same strategy. In this case, the reaction curves are

$x^{*r} = x^s$ for two generic *r* and *s* agents (5)

When being ‘*together* is not essential–i.e., when $\varphi<\alpha$–the thresholds associated with the willingness to compromise come into play, and a variety of outcomes becomes possible. Figures 7a and 7b display the resulting reaction curves. For individual *j*, the reaction curve is x^{*j}=4 for low values chosen by the other player (x^k). Beyond the threshold, x^{*j} drops and then increases along the 45° line, up to x^*=4. For individual *k*, the reaction curve is x^{*k}=0 for high values of x^j. Below the threshold, $x^{*k} = x^j$; it jumps to the right and then decreases along the 45° line until it returns to x^{*k} =0. (Recall that agent *k* dislikes *x*).

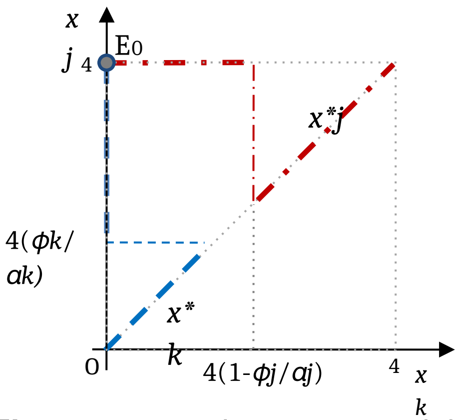

Figure 7a: reaction curves and the equilibrium when the relative importance of the compromise is rather low.

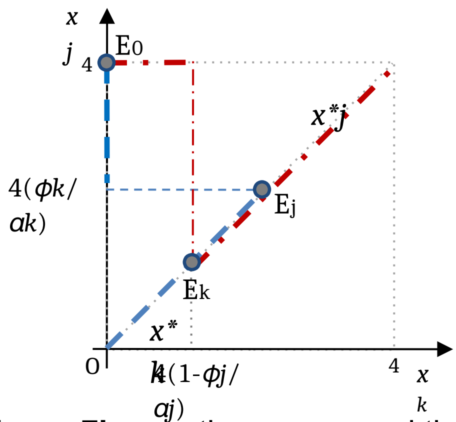

Figure 7b: reaction curves and the equilibria when the relative importance of the compromise is rather high.

Figure 7a depicts a situation in which the importance attached to choosing compatible options is lower than the importance of the individual action. In this case, the only equilibrium is *E0*, in which the individuals choose different strategies.

Figure 7b, by contrast, illustrates a case where the thresholds are such that equilibria involving identical strategies become possible. In addition to the equilibrium *E0*, there also exists a continuum of equilibria spanning from *Ek* to *Ej*.

Let us now return to a discrete example, as shown in Figure 8. Five strategies are available, and the two individuals have different preferences (φ^j=0.6 and φ^k=0.8). The values in parentheses convert labels into the corresponding values used in the continuous representation. Due to the thresholds, not all outcomes in which the two agents go out together are equilibria. Moreover, symmetry is no longer a necessary feature of the outcomes. In this case, the following equilibria arise: (*Fight*; *Ballet*), (*Horse race*, *Horse race*) and (*Movie*; *Movie*) corresponding respectively to E0, Ej, and Ek in Figure 9, which presents the continuous representation of the game.

Note that, since agent *j* attaches relatively less importance to going out together, the couple may end up in equilibrium E*j*, which is closer to agent *j*'s preferred entertainment. Figure 5 shows that the greater the importance attached to being together, the larger the overlapping portion of the reaction curves–hence, the greater the number of equilibria in the discrete case–and *vice-versa*.

Finally, it is worth noting that allowing for decreasing marginal utility of *x*, which is straightforward, does not yield any novel insights. However, introducing satiation alongside decreasing marginal utility slightly alters the picture. In this case, the reaction curves remain parallel to each axis as long as the individual is satiated, and then follow the 45° diagonal up to the threshold level, beyond which each player chooses their preferred (satiating) option independently.

If the satiation point is sufficiently high, the qualitative results remain unchanged, as illustrated in Figure 10a. By contrast, if agents reach satiation quickly, the range of multiple equilibria narrows, as shown by segment E_1E_2 in figure 10b.

k→ *j*↓	*Fight* (4)	*Horse race* (3)	*Movie* (2)	*Comedy* (1)	*Ballet* (0)
Fight (4)	**6.4** 3.2	4 1	4 2	**4** 3	**4 4**
Horse race (3)	3 0	**5.4 4.2**	3 2	3 3	3 4
Movie (2)	2 0	2 1	**4.4 5.2**	2 3	2 4
Comedy (1)	1 0	1 1	1 2	3.4 **6.2**	1 4
Ballet (0)	0 0	0 1	0 2	0 3	2.4 **7.2**

Figure 8. A 5x5 battle of the sexes with non-symmetrical preferences, φ^j=0.6 and φ^k=0.8, $\alpha^j = \alpha^k$ =1, x_M=4.

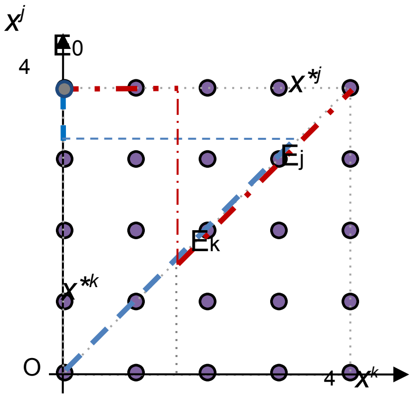

Figure 9. The continuum representation of the 5x5 battle of the sexes of Figure 8.

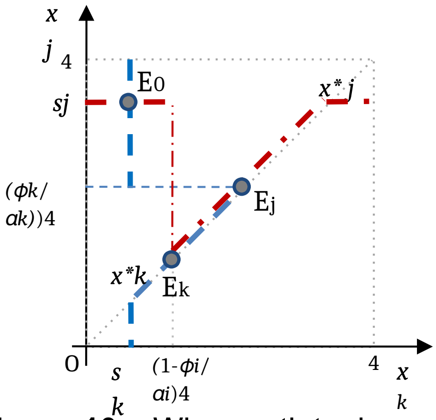

Figure 10a. When satiety does not qualitatively change the picture

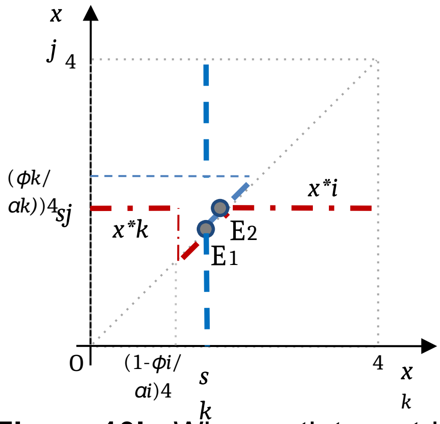

Figure 10b. When satiety restrict the equilibria continuum

4. Beyond the couple

As a final step, we highlight that the modified framework also enables a different type of analysis compared to the original one. The BoS is a game involving agents with conflicting preferences but a stronger desire to coordinate on the same action. The original narrative uses a couple as an illustrative example. However, models of couples in real-life settings should typically allow for pre-play communication and/or repeated interaction. (see, e.g., Lau and Mui, 2008). By contrast, there are many real-world situations where a one-shot BoS framework is more appropriate –for example, political processes such as peace conferences, constitution drafting, or the formation of new governments. The framework proposed here can capture cases in which agents ultimately would prefer to ‘go out separately’. Yet coordination may still emerge either when external factors increase the utility of coordination (U_T) relative to intrinsic preferences, or when new compromise options are introduced.

Moving from conflicting situations–where parties prefer 'to go alone'–to ones in which compromises become available and potentially self-enforcing as equilibrium solutions can be achieved by introducing new options. Such processes require the emergence of compromise solutions that mediate between positions (and preferences) that are initially far apart. At the same time, interest in a joint solution may increase because of a deliberative process that encourages parties to draw more on shared background values than on their differences. Modelling such a deliberation process, particularly as it leads to political compromise, is beyond the scope of this paper and would, in any case, require considering a sequential game and/or pre-play communication. However, as highlighted in Guarnieri (2019) in the discussion of the process that led to agreement on Article 1 of the Italian Constitution, deliberation can explain institutional formation as the outcome of emerging, self-enforcing compromise options[8].

To illustrate the possible shift from a non-coordination to a coordination equilibrium, we can take the formation of the Italian government in spring 2018 as an example. This formation resulted from the agreement known as the *Contratto per il governo del cambiamento* between *Lega* and *Movimento 5 Stelle*, which attracted international attention due to its unexpected nature. The two parties had, and still have, rather different political platforms and electoral constituencies. However, the election outcome produced a highly uncertain situation in which forming a government together was perhaps the only viable alternative to immediately returning to elections–a risky option for both parties.

At the same time, forming a government represented a prime opportunity to demonstrate their ability to break with the past, delivering the change so desired by their constituencies and thereby consolidating their popularity. In terms of the

[8] A theory of compromise-making is also relevant for to the debate on the possibility of reuniting the field of institutional studies (Hindriks & Guala, 2014, 2015; Greif & Kingston, 2011; Aoki, 2001, 2007; Hédoin, 2017; Hodgson, 2006) by epistemologically reconciling the social-ontological rule-based approaches (Searle, 1995, 2005, 2010; Gilbert, 1989) and the game-theoretical equilibrium-based approach (Schotter, 1981; Sugden, 1986; Young, 1998).

proposed framework, the *reward* of 'going out together' increased after the elections and during the subsequent period of uncertainty. Indeed, after lengthy negotiations, the leaders of the two parties signed the *Contratto*, which *outlined* the political agenda they committed to implementing during the legislature. In the context of our discussion, this can be viewed as the emergence of a new intermediate alternative that incorporates elements from both platforms.

It should also be noted that the two parties were not in a symmetrical position. *Lega* belonged to the right wing and ran its campaign alongside three allies: *Forza Italia*, *Fratelli d'Italia,* and *Noi con l'Italia-UDC*. *Movimento 5 stelle*, by contrast, originated as a movement opposed to traditional parties and therefore could run a campaign announcing political allies–least of all with the *Partito Democratico,* with which it had always maintained very conflictual relations. Hence, it is plausible that the reward from the agreement (U_T) was higher for *Movimento 5 Stelle*, and that the agreement resulted closer to *Lega*'s priorities, as in Figure 8 *(Horse race, Horse race)* or Figure 9 (*equilibrium E1*). This may partly explain why opinion polls reported that, about six months after the elections, *Lega* had doubled its popularity and surpassed *Movimento 5 Stelle*, which dropped well below the nearly 33% it had achieved at the elections. After one year in government, *Lega*'s rising political consensus made the prospect of 'going alone' to new elections attractive for its leadership, who initiated a government crisis in August 2019[9].

To summarise, the proposed framework can prove useful for illustrating compromise-making which can be conceptualized as one-shot games. In the case of the *Lega-Movimento 5 Stelle* government formation, the election outcome strongly shifted the importance the parties attributed to the agreement and induced them to solve the coordination problem by envisioning new compromising options that were not considered prior to the interaction itself. Two parties with very different political platforms reached an agreement that was difficult to anticipate before the elections, and this agreement endured as long as they assigned high value to 'going out together'.

[9] This was a strategic mistake because elections have been avoided by the formation of a new government based on an agreement between *5 Stelle* and *Partito Democratico* in September 2019. The agreement was possible because of a certain affinity of their political programmes, and the advantage of "going together", that is, avoiding the momentum of Lega.

5. Concluding remarks

The Battle of the Sexes is a classic example in game theory, illustrating the challenges of (non-cooperative) coordination between agents who have different preferences over outcomes but a strong mutual desire to coordinate on the same action. The purpose of this note was to highlight a minor inconsistency between the narrative and the associated payoffs. Restoring this consistency does not alter the outcome of the BoS game, but it opens the door to a more general framework that helps students interpret a variety of real-life situations.

We also pointed out that the issue of (in)consistency is closely linked to the realism of the model. First, the original payoffs rely on an unrealistic assumption: if the partners go out separately, they are indifferent between attending their preferred or disliked event. This assumption appears especially questionable today, as partners tend to be more autonomous and freer than when the game was originally proposed. Second, it is unrealistic to assume that the domain of choice is limited to only two options. Third, modelling the BoS as a one-shot game makes it ill-suited for representing situations–such as those involving couples–that typically involve repeated interactions.

Because of these features, the message conveyed by the BoS game is that coordination is difficult when agents have differing preferences but a strong desire or interest in acting jointly. The game has multiple equilibria, with no clear indication of how players should coordinate. Moreover, the pure strategy equilibria are unfair. It may even happen, as an out-of-equilibrium outcome, that each partner ends up attending the event they prefer the least. Embedding the BoS game within a more general framework helps convey a different message–namely, that various outcomes are possible, including compromise, and it clarifies the conditions under which they may emerge.

***

Funding: ECOESIONE "Progetto finanziato dal Ministero dell'Ambiente e della Tutela del Territorio e del Mare nell'ambito della Strategia Nazionale per lo Sviluppo Sostenibile" and institutional activity at the Università di Pisa.

Conflicts of interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aoki, M. (2007). 'Endogenizing institutions and institutional changes', *Journal of Institutional Economics*, 3(01), pp. 1-31.

Aoki, M. (2011). 'Institutions as cognitive media between strategic interactions and individual beliefs', *Journal of Economic Behavior & Organization*, 79(1), pp. 20-34.

Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. Journal of mathematical Economics, 1(1), pp. 67-96.

Binmore, K. (2007a). *Game theory: a very short introduction* (Vol. 173). Oxford University Press.

Binmore, K. (2007b). *Playing for real: a text on game theory*. Oxford University Press.

Buchanan, J. M. (1977). *Freedom in constitutional contract : perspectives of a political economist* (1st. ed). Texas A & M University Press.

Davis, M. D. (2012). *Game theory: A nontechnical introduction*. Courier Corporation.

Farrell, J., & Rabin, M. (1996). Cheap talk. *Journal of Economic perspectives*, *10*(3), pp. 103-118.

Fudenberg, D., & Tirole, J. (1991). Game theory, MIT press Cambridge*, MA*.

Gilbert, M. (1989). *On social facts*, Princeton (NJ): Princeton University Press.

Gintis, H. (2009). *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. Second Edition. Princeton University Press.

Greif, A. and Kingston, C. (2011). Institutions: Rules or equilibria? In Schoefield N. and Caballero G. (eds.), *Political economy of institutions, democracy and voting*, Basel (CH): Springer, pp. 13-43

Guarnieri, P. (2019). Interactive intentionality and norm formation. *Journal of Institutional Economics*, *15*(4), pp. 579-593.

Grüne-Yanoff, T. and Schweinzer, P. (2008). "The role of stories in applying game theory", *Journal of Economic Methodology*, vol. 15, 2, pp. 131-146.

Harsanyi, J. C., & Selten, R. (1972). A generalized Nash solution for two-person bargaining games with incomplete information. *Management Science,* 18(5-part-2), pp. 80-106.

Harsanyi, J. C., & Selten, R. (1988). *A general theory of equilibrium selection in games*. MIT Press.

Hédoin, C. (2017). 'Institutions, Rule-Following and Game Theory', *Economics & Philosophy*, *33*(1), pp. 43-72.

Helbing, D., M.Schönhof, H.-U.Stark, and J. A.Hołyst. 2005. How individuals learn to take turns: emergence of alternating cooperation in a congestion game and the prisoner's dilemma. Adv. Compl. Syst. 8, pp. 87–116.

Hindriks, F. and Guala, F. (2014). 'Institutions, rules, and equilibria: a unified theory', *Journal of Institutional Economics*, 11(3), pp. 459-480.

Hindriks, F. and Guala, F. (2015). 'A unified social ontology', *The Philosophical Quarterly*, 65(259), pp. 177-201.

Hodgson, G. M. (2006). 'What are institutions?', *Journal of Economic Issues*, 40(1), pp. 1-23.

Hwang, I. (2018). A theory of bargaining deadlock. Games and Economic Behavior, 109, pp. 501-522.

Lang, S. H. P., & Mui, V. L. (2008). 'Using turn taking to mitigate coordination and conflict problems in the repeated battle of the sexes game'. *Theory and Decision*, 65(2), pp. 153-183.

Luce, R. D., & Raiffa, H. (1957). Games and Decisions: Introduction and Critical Survey.

Morgan, M. S. (2001). "Models, stories and the economic world", *Journal of Economic Methodology*, vol. 8, 3, pp. 361.

Nash Jr, J. F. (1950). The bargaining problem, *Econometrica*, pp. 155-162.

Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. MIT press.

Rubinstein, A. (2006). *Lecture notes in microeconomic theory.* Princeton, NJ: Princeton University.

Schelling, T. C. (1960). *The strategy of conflict*. Harvard university press.

Schotter, A. (1981). *The economic theory of social institutions*, Cambridge (UK): Cambridge University Press.

Searle, J. R. (1995). *The construction of social reality*, New York: Simon and Schuster.

Searle, J. R. (2005). 'What is an institution', *Journal of institutional economics*, 1(1), pp. 1-22.

Searle, J. R. (2010). *Making the social world: The structure of human civilization*. Oxford (UK): Oxford University Press.

Sugden, R. (1986). *The economics of rights, co-operation and welfare*, Oxford: Basil Blackwell, 2004 2nd edition.

Tadelis, S. (2013). *Game theory: an introduction*. Princeton University Press.

Young, H. P. (1998). *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton (NJ): Princeton University Press.