



HAL
open science

The worst-case payoff in games with stochastic revision opportunities

Yevgeny Tsodikovich

► **To cite this version:**

Yevgeny Tsodikovich. The worst-case payoff in games with stochastic revision opportunities. *Annals of Operations Research*, 2021, 300 (1), pp.205-224. 10.1007/s10479-020-03867-3 . hal-03077847

HAL Id: hal-03077847

<https://amu.hal.science/hal-03077847>

Submitted on 17 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Worst-Case Payoff in Games with Stochastic Revision Opportunities ^{*}

Yevgeny Tsodikovich[†]

November 4, 2020

Abstract

We study infinitely repeated games in which players are limited to subsets of their action space at each stage – a generalization of asynchronous games. This framework is broad enough to model many real-life repeated scenarios with restrictions, such as portfolio management, learning by doing and training. We present conditions under which rigidity in the choice of actions benefits all players in terms of worst-case equilibrium payoff and worst-case payoff. To provide structure, we exemplify our result in a model of a two-player repeated game, where we derive a formula for the worst-case payoff. Moreover, we show that in zero-sum games, lack of knowledge about the timing of the revision can compensate for inability to change the action.

Journal classification number: C73.

Keywords: Asynchronous Games, Rational Minimax, Worst-Case Payoffs, Commitment, Exogenous Timing.

^{*}The author wishes to thank Ehud Lehrer, Eilon Solan, David Lagziel, Rann Smorodinsky, Dotan Persitz, Galit Golan-Ashkenazi and Ilan Nehama and two anonymous referees of the *Annals of Operations Research* journal for their highly valuable comments. The author acknowledges the support of the Israel Science Foundation, Grants #963/15 and #2510/17 and the support of the French National Research Agency Grants ANR-17-EURE-0020.

[†]Aix Marseille Univ, CNRS, AMSE, Marseille, France. e-mail: yevgets@gmail.com

1 Introduction

This paper considers repeated interactions in which players do not necessarily act simultaneously. This is often the case, as different agents have different decision-making schedules that seldom coincide, and their capacity to change their actions often differs. The asynchronous nature of the play can lead to cooperation, similar to the collusion occurring in the price-leadership mechanism [MacLeod, 1985], the commitment involved in the central bank setting the interest rate [Libich and Stehlík, 2010, 2011], and other economic examples presented in the seminal works of Maskin and Tirole [1988] and Lagunoff and Matsui [1997].

In all of these examples, whenever a player cannot revise his action, the action from the previous stage is replayed. However, there are other ways in which actions can be binding; for example, when a revision opportunity is given at every stage, but the new action has to be “close” to the old one (in some metric). To examine such constraints, this paper introduces a generalization of repeated games, namely *Sub-Actions Repeated (SAR) Games*, in which at each stage the players must choose an action from a subset of their mixed action space, stochastically determined by the history of the game. This generalizes the (simultaneous-move) repeated game as well as different models of asynchronous games.

SAR games take place naturally in many real-life situations. One example is the price rigidity caused by publishing prices in an advertisement or a catalog. For some time after the price is published, it is illegal for the vendor to raise it (false advertisement), or at least, costly in terms of reputation. In this situation, the set of available new prices includes only prices lower than the published price.

Another example is portfolio management. In some cases, changing the entire portfolio at one time is cumbersome and prohibitive costly, whereas small deviations from a certain position are feasible and easily implementable. Thus, portfolio management can be modeled as a SAR game where the change in position at each stage is bounded by some feasibility constraint.

A further interesting application is learning. One possible model of “learning by doing” is a model in which repeating a certain action (playing the guitar) makes you proficient but also prevents you from changing career in the future and choosing other actions (like starting a career as an astrophysicist). This can also relate to training models, in which after each stage your skill increases and the subset of available actions that you can take expands.

Finally, SAR games can model interaction between players that affects their rivals' possible actions. For example, Lovo and Tomala [2015] discuss a sail-boat race in which boats can block each other and by doing so prevent the other boats from performing a certain action at the next stage. Such situations can be modeled using a SAR game since the available subset of actions of each player depends on the history and can be affected by the past actions of other players.

We use two solution concepts to evaluate different SAR games (usually to compare them to the simultaneous-move repeated game): the worst-case rational payoff, namely the *rational minimax value*, and the worst-case equilibrium payoff. Our rational minimax value is closely related to the effective minimax value, presented by Wen [1994] and expanded in Wen [2002], Takahashi and Wen [2003] and Yoon [2004]. The effective minimax value captures the idea that players with equivalent utility functions will coordinate their actions rather than punish each other (and themselves in the process). Similarly, the rational effective minimax value captures the idea that players will not punish each other *too much* if it is detrimental to their own payoff.¹ The lower bound we set is the standard minimax value – a player will not punish another player if it causes his own payoff to drop below his standard minimax value. Such behavior can be considered as *irrational*.

We use the rational minimax value because it is a better figure of merit than the standard minimax value, offering a tighter lower bound on equilibrium payoffs. For example, in simultaneous-move repeated games, the rational minimax value is the worst-case equilibrium payoff and any payoff above it can be supported in equilibrium (when the discount factor is close to unity). Moreover, playing asynchronously can have opposite effects on these values, with the standard minimax value decreasing relative to the simultaneous move game and the rational minimax value rising. Thus, our chosen figure of merit better represents the worst-case payoff, in equilibrium and off equilibrium, assuming the other players are rational.

In addition, the rational minimax value is more generally applicable than the effective minimax value, which is only advantageous when there are players with equivalent utility functions. This is rarely the case, however; more frequently, players show partial but not full correlation between utility functions. The rational minimax value can account for this partial correlation, as shown in our paper.

¹An extreme version of this idea is the sub-game perfect equilibrium – punishments are only carried out in a way that ensures the punisher does not lose utility relative to the equilibrium path.

The general nature of SAR games, and their complexity, leads to several technical obstacles when using the rational minimax value and the worst-case equilibrium payoff. First, the Folk Theorem does not hold for SAR games in general, so the worst-case equilibrium payoff is not the rational minimax value (as in standard repeated games). This is because the SAR game is not necessarily periodic, so punishing strategies may disappear and the rational minimax value from a certain time t onward can be different from the rational minimax value from time $t = 0$. For example, if the game is absorbed into some higher payoff, the rational minimax value may no longer be obtained, nor a lower equilibrium payoff.

Another obstacle is that SAR games are non-stationary so the rational minimax value must take into account all the infinite strategies in the game, not only the mixed actions in the one-shot game. This non-standard approach poses computational problems but the nature of the game makes it essential. This led Yoon [2001] to define the lower and upper minimax values, which depend on the order of play when it exists and is periodic. Despite these challenges, we choose to keep the definition as general as possible so as to obtain a main result that is widely applicable. The examples (Section 4) are two periodic SAR games where these problems disappear.

Our contribution is threefold. First, we introduce a new term, the rational minimax value, which provides a tighter lower bound on equilibrium payoff compared to the standard minimax value. Second, we show how the feasible set of *stage games* affects the rational minimax value in the SAR game. This allows us to characterize the stage games for which a simultaneous-move game is inferior to a SAR game, in terms of worst-case equilibrium payoff. In such cases, players will prefer the SAR game over the simultaneous-move repeated game and may try, as far as possible, to limit revision opportunities. Third, our theorem can be applied to produce comparative statics between the SAR game and the simultaneous-move game regarding the worst-case equilibrium payoff, even when the rational minimax value is not computable for all the players.

To exemplify our main result we consider a two-player game with complete information where one plays synchronously (with a revision opportunity at each stage) while the other plays asynchronously and has revision opportunities only at certain randomly chosen stages (“strict-revision process”, as in Tsodikovich and Lehrer [2019]). A deterministic version of this model was explored by Wen [2002] (see also Takahashi and Wen [2003]) and largely inspired this paper. These authors showed how the standard

and the effective minimax values depend on the (constant) number of stages between two consecutive revisions. We generalize this work and determine how standard and rational minimax values change when intra-revision timing is stochastic.

We also consider a slightly different model in which the players do not know the timings of revisions. Similarly to Spiegler [2015], we show that when revision probability is large enough, the minimax value remains the same, despite the lack of agility and the inherent disadvantage of the asynchronous player. Thus, lack of information has a positive effect that compensates for inability to revise actions.

The rest of the paper is organized as follows. Section 2 presents the model, our extension of repeated games and formal definitions of our figures of merit – the rational minimax value and the worst-case equilibrium payoff. In section 3 we discuss the main result and present the conditions required for the rational minimax value to be higher in the asynchronous game than in the simultaneous-move game. The main result is applied to two versions of two-player games with different information structures in Section 4. Conclusions and extensions are presented in Section 5. To improve readability, all proofs are relegated to the Appendix.

2 The model

Let $G = (I, (A_i)_{i=1}^n, (u_i)_{i=1}^n)$ denote a strategic-form game where $I = \{1, \dots, n\}$ is the set of players, A_i is the finite set of pure actions for player i and $u_i : \times_{i=1}^n A_i \rightarrow \mathbb{R}$ is the stage game payoff function for player i . A mixed action α_i of player i is a distribution over A_i , i.e. an element of $\Delta(A_i)$. The expected stage payoff of player i , given the action profile $\alpha = (\alpha_1, \dots, \alpha_n) \in \Delta(A_1) \times \dots \times \Delta(A_n)$ is $u_i(\alpha) = \sum_{a \in \times_{i=1}^n A_i} \alpha(a) u_i(a)$.

We write $-i$ to denote all players except player i and the (*standard*) *minimax value* of the stage game for player i is

$$v_i = \min_{\alpha_{-i} \in \times_{j \neq i} \Delta(A_j)} \max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \alpha_{-i}). \quad (1)$$

The standard minimax value is the lower bound on the payoff that player i can be sure of obtaining when all players try to minimize his payoff in the one-shot game. Finally, let F be the convex hull of the set of feasible payoff vectors, F_{IR} the set of individually rational feasible payoffs, i.e. $F_{IR} = \{(x_1, \dots, x_n) \in F | x_i \geq v_i\}$, and $v_G = (v_1, \dots, v_n)$ the minimax point.

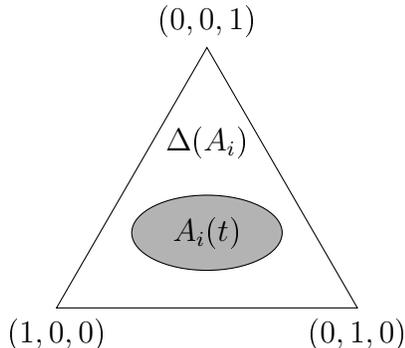


Figure 1: An illustration of $A_i(t)$, the set of available actions to Player i in stage t .

2.1 Sub-Actions Repeated (SAR) Games

To model asynchronous play, we introduce a generalization of the standard repeated game model, namely the sub-actions repeated (SAR) game. In this game, the set of available actions at each stage for each player is a subset of the set of mixed actions. The SAR game unfolds as follows. The stage game is repeated infinitely many times. At stage $t = 0$, each player chooses his respective action from $\Delta(A_i)$. In period $t \geq 1$, a non-empty convex and closed subset $A_i(t) \subseteq \Delta(A_i)$ is randomly chosen² for each player who is restricted to choosing a mixed action from set $A_i(t)$ (see Figure 1).

We require the sets to be convex and closed for technical reasons. Closedness is required because functions over open sets do not necessarily obtain a maximum. If $A_i(t)$ is not closed, all actions may not be optimal (a *best* response) and the discussion should be changed to terms of ϵ -equilibria instead of Nash-equilibria and inf sup instead of min max. Convexity is required to ensure the existence of an equilibrium, as a pre-requirement for the relevant fixed-point theorems. If this condition is not met, there may be no strategy profile in which each action is a best response to the actions of all the others. Another reason for imposing convexity is that we wish to allow mixed actions, and a non-convex sub-actions set implies that some randomizations between mixed actions are precluded.

We do not impose restrictions on the stochastic process that determines the $A_i(t)$ s, nor on what the players observe. This will be further illustrated using specific processes

²A rigorous definition of the probability space is required to well-define the game. The Examples to follow suggest several simple options. Hereafter we assume that after every history this probability space is well-defined.

in the examples below. We assume that each player knows at time t his own $A_i(t)$, the history he observed, and the definition of the stochastic process (“the rules of the game”). Hence, based on the rules, he may also be able to deduce $A_{-i}(\tau)$ for $\tau \geq t$, form beliefs about them or be totally ignorant. A SAR game with a common discount factor δ is denoted by $\Gamma = (G, \{A_i(t)\}_{i \in I}^{t \in \mathbb{N}}, \delta)$ and for simplicity will be referred to as *the repeated game*. This formulation is a generalization of many existing models:

1. Simultaneous-move game (Γ_0): For each player and every t , $A_i(t) = \Delta(A_i)$.
2. One-shot game (Γ_1): For each player, $A_i(t)$ is a singleton containing only the pure action played at stage $t = 0$. Note that the one-shot game is an example of SAR games so the Folk Theorem does not apply to general SAR games.
3. Alternating-moves game [Maskin and Tirole, 1988, Lagunoff and Matsui, 1997]: A two-player game in which $A_1(2t - 1) = \Delta(A_1)$, $A_2(2t) = \Delta(A_2)$ for $t \geq 1$ and at any other stage $A_i(t)$ is a singleton containing the pure action played at time $t - 1$.
4. Asynchronous-moves game with mixed actions [Yoon, 2001, 2004]: At each stage, a random set of players I_t is chosen and only they can revise their action: $A_{i \in I_t}(t) = \Delta(A_i)$. The rest play the *mixed* action they played in the previous stage.
5. Games with a strict-revision process: Some random variable governs the stages at which a player can revise his action. At this stage, $A_i(t) = \Delta(A_i)$. In the rest, $A_i(t)$ is a singleton containing the pure action played at the previous stage. These games are examined in Section 4. Their single-decision maker version was analyzed in Tsodikovich and Lehrer [2019].

In examples 1-3, players know at each time step the $A_i(t)$ s of all the players and observe the *pure* actions played by others. Contrastingly, in example 4, players are able to observe *mixed* actions. This can accommodate situations where mixed actions do not represent randomization over pure action, but rather some distribution of resources (diverse action) as in portfolio management. Two versions of example 5 are presented in Section 4. In both versions, only pure actions are observed, the discrepancy concerns the observability of $A_2(t)$ ($\forall t, A_1(t) = \Delta(A_1)$, and this is common knowledge). In the first version (Section 4.1), both players observe $A_2(t)$ at each stage while in the second version (Section 4.2), Player 1 learns $A_2(t)$ only at time $t + 1$.

Interestingly, with some adjustments, our model can also fit games which are not repeated *per se*. For example, in revision games [Calcagno et al., 2014, Gensbittel et al., 2018]³ the payoff is given to the players only once at a certain deadline according to the action they prepared before it. While awaiting the deadline, the players observe the prepared actions and randomly receive opportunities to revise their actions according to some Poisson process (similar to the pre-opening session in the stock exchange). This can be modeled as a SAR game in which at each time step at most one of the players receives a revision opportunity, while the rest of the players replay their pure action from the previous time stage. As t increases, the probability that no-one receives a revision opportunity increases and converges to 1, to mimic the small probability to revise an action close to the deadline. Unlike in revision games, here all stages are payoff relevant, however for patient enough players ($\delta \rightarrow 1$) the transient stage is negligible and only the “last” actions are important.

2.2 The worst-case payoffs

Denote by Σ_i the set of all possible (behavioral) strategies of player i in the repeated game. Each strategy specifies the action $\alpha_i^t \in A_i(t)$ of player i after every t -length history, which includes the (observed) actions of the players and his $A_i(t)$. Given the strategy profile $\sigma \in \times_{i=1}^n \Sigma_i$, the expected δ -discounted payoff of player i is $u_i(\sigma) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t \mathbb{E}(u_i(\alpha_1^t, \dots, \alpha_n^t))$. For each player i we denote by $\Sigma_{-i} = \times_{j \neq i} \Sigma_j$ the set of all the strategies of the others and by σ_{-i} an element of this set. Hence, the standard minimax value of the repeated game for player i is defined according to

$$v_i(\Gamma) = \min_{\sigma_{-i} \in \Sigma_{-i}} \max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, \sigma_{-i}). \quad (2)$$

The standard minimax is the worst-case payoff that a player can guarantee when other players act as adversaries. This differs from the standard minimax value defined in Eq. (1) since payoffs are maximized and minimized over the entire path of play instead of just in the one-shot game. The value is the same as when the repeated game is a simultaneous-move game ($v_i(\Gamma_0) = v_i(\Gamma_1) = v_i$), but not the same as in the general setting of a SAR game.

³See more details in the preceding unpublished papers Kamada and Sugaya [2010], Kamada and Kandori [2011], and Calcagno and Lovo [2010].

| | | Player 2 | |
|----------|-----|----------|---------|
| | | B | S |
| Player 1 | B | (2, 1) | (0, 0) |
| | S | (0, 0) | (1, 2) |
| | P | (0, -1) | (0, -1) |

Table 1: A modified “Battle of the Sexes” where the rational minimax of player 2 is strictly greater than his standard minimax.

In the presence of possible partial equivalence of the utility functions, adversary behavior can harm players and lower their payoff too, sometimes even below *their* minimax value. Thus, we define the *rational minimax value* as the minimum payoff that a player receives when the strategy profile produces an individually rational payoff to all players:

$$v_i^r(\Gamma) = \min_{\sigma \in \Sigma} \{u_i(\sigma_i, \sigma_{-i}) \mid u_j(\sigma_i, \sigma_{-i}) \geq v_j(\Gamma) \text{ for all } j \in I\}. \quad (3)$$

When $v_i^r(\Gamma) > v_i(\Gamma)$, player i cannot guarantee the rational minimax by himself and must rely on the rationality of other players to obtain this value.

Example 1. *Rational minimax in a simultaneous-move game.*

Consider a modified version of the “Battle of the Sexes” game, shown in Table 1. In this game, the row player has an additional action, P , which is dominated by the other actions and does not affect his standard minimax value: $v_1(\Gamma_0) = \frac{2}{3}$. This action can serve as a minimizing strategy against the column player, setting $v_2(\Gamma_0) = -1$. Playing P too often is not individually rational for the row player, as it will lower his payoff below $\frac{2}{3}$. Thus, whenever the row player receives at least $\frac{2}{3}$, the outcome of the game lies inside the gray-shaded area in Figure 2, which results in a payoff of at least $v_2^r(\Gamma_0) = -\frac{1}{3}$ for the column player. The column player cannot guarantee a payoff higher than -1 by himself, but thanks to the rationality of the row player and the partial equivalence of the payoff functions, he receives at least $-\frac{1}{3}$. The payoffs $(\frac{2}{3}, -\frac{1}{3})$ are in this case a tight lower bound on the equilibrium payoffs when the discount factor approaches unity. \triangle

A strategy profile σ is a Nash equilibrium in the game if no player has a profitable deviation, i.e. for each player i and each strategy σ'_i , $u_i(\sigma) \geq u_i(\sigma'_i, \sigma_{-i})$. Let $\mathcal{N}_E(\Gamma)$

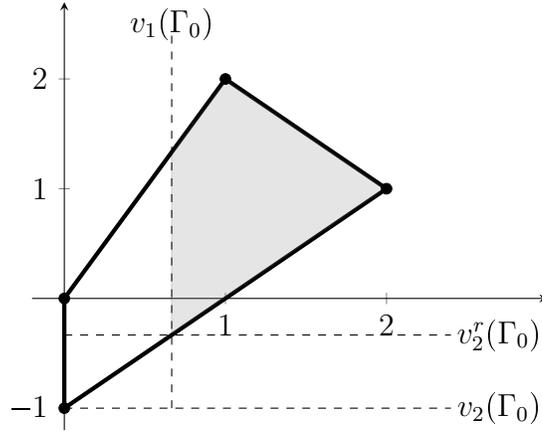


Figure 2: The set of feasible payoffs and the rational minimax of player 2 in the modified “Battle of the Sexes” (Table 1).

be the set of strategy profiles which are Nash Equilibria in the game Γ and for each player i let $v_i^e(\Gamma)$ be the lowest equilibrium payoff, i.e.

$$v_i^e(\Gamma) = \inf_{\sigma \in \mathcal{N}_E(\Gamma)} \{u_i(\sigma)\}.$$

Trivially, the worst-case equilibrium payoff is higher than the rational minimax value, so $v_i^e(\Gamma) \geq v_i^r(\Gamma)$. In the standard simultaneous-move game, they are the same for δ close enough to 1: $v_i^e(\Gamma_0) = v_i^r(\Gamma_0)$.

3 The Main Result

When is a SAR game based on a particular stage game preferred by all the players over the simultaneous-move game, in terms of rational minimax value? We start by considering a general two-player game. The behavior of the rational minimax value in the SAR game is determined by examining the feasible set F , and the location of the standard minimax point v_G relative to it. Generally speaking, this point can lie in one of three areas (A , B , and F), as shown in Figure 3. Note that although both v_1 and v_2 are feasible, they may not be feasible simultaneously (in Example 1, the payoff $(\frac{2}{3}, -1)$ is not feasible), so the minimax point can in fact lie outside F .

When v_G lies within F , the Folk Theorem states that there are equilibria with payoffs arbitrarily close to the standard minimax values. The standard minimax value

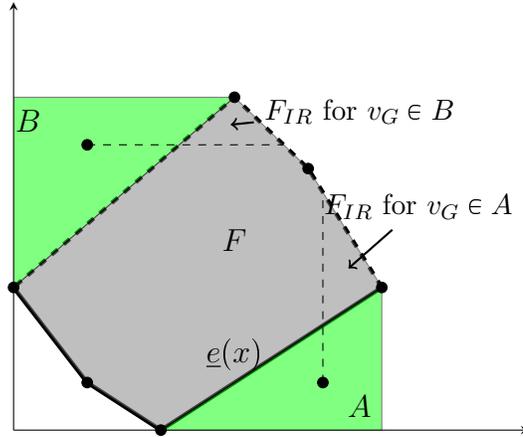


Figure 3: Possible locations of the standard minimax point v_G in the payoff space. The location of v_G determines the feasible individually rational set and the rational minimax value of the players.

is therefore also the rational minimax value, and different SAR games will exhibit different behaviors regarding standard and rational minimax values. Our main result cannot be applied here and we leave this analysis to future studies.

When the minimax point is “below” the feasible set in region A ,⁴ any payoff that is individually rational for player 1 will result in a payoff for player 2 that is strictly larger than his standard minimax. Therefore, the rational minimax of player 1 is equal to his standard minimax, while for player 2 the former is larger: $v_2^r > v_2$. Since the lower envelope of the feasible set, $\underline{e}(x) = \min\{y | (x, y) \in F\}$, is an increasing function of $x \geq v_1$, any SAR game in which the standard minimax value of player 1 is higher than v_1 will result in a higher than v_2^r worst-case rational payoff for player 2 – an improvement for both players. It is this type of equivalence between utility functions that this paper addresses: player 2 will prefer SAR games that limit his ability to lower the payoff of player 1, which increases the minimax value of player 1 and in return, his own rational minimax value.

When generalizing this idea to the n -player game, we impose a similar requirement on the SAR game: that the standard minimax rises for all players whose rational minimax is equal to the standard minimax in the stage game. When there are players that satisfy $v_i^r(\Gamma_0) > v_i(\Gamma_0)$, the minimax point has to be in the region A and the above effect can occur. These players should play asynchronously, increase the minimax

⁴In region B , the players’ roles are reversed.

value of the other players and consequently, increase their own rational minimax value. This result is consistent with previous papers finding that restricting the actions can eliminate unwanted equilibria and serve as a commitment tool, such as Lagunoff and Matsui [1997].

Theorem 1. *Let $\Gamma = (G, \{A_i(t)\}_{i \in I}^{t \in \mathbb{N}}, \delta)$ be a SAR game and let Γ_0 be the corresponding simultaneous-move game. Denote the set of players whose rational minimax in Γ_0 is equal to the standard minimax value in Γ_0 by $I_ =$.*

If $v_i(\Gamma) \geq v_i(\Gamma_0)$ for every $i \in I_ =$ then $v_i^r(\Gamma) \geq v_i^r(\Gamma_0)$ for all players.

Theorem 1 divides the players into two groups. The first group, $I_ =$, are players whose rational minimax value is equal to the standard minimax value in the simultaneous-move game. When the payoff is individually rational for them, the other players will receive a payoff higher than their standard minimax value – at least the rational minimax value. In any SAR game where the first group receives higher worst-case payoffs, the rational minimax value of the second group is sure to rise as well. In a two-player case, the theorem can be applied to the simultaneous-move game whenever the rational minimax of one of the players is strictly greater than his standard minimax. If the latter occurs, any modification of the repeated game that strictly increases the standard minimax of the other player will also increase the rational minimax of the first player, making this modified game beneficial to both players.

Verifying the conditions of Theorem 1 is not trivial in the general case, as the rational minimax value of all players could be complicated to compute. Nonetheless, the theorem has two prime applications. First, it transforms the question from one about SAR games to a simpler question about the stage game and the geometry of its feasible set. Hence, it identifies which stage games are or are not prone to such an effect, even without focusing on any particular SAR game. Second, when the scope is limited to particular types of SAR game, the Theorem is translated to case-specific conditions that may be more applicable and easier to verify. To demonstrate this, we examine in Section 4 two models of two-player games where only one of the players is asynchronous in a manner previously studied by Tsodikovich and Lehrer [2019]. Applying Theorem 1, we obtain concrete and simple conditions on the inter-revision timings for which the asynchronous game is more favorable than the simultaneous-move game (such as Corollary 1).

Theorem 1 also has direct implications for the worst-case equilibrium payoffs. First,

it should be noted that the rational minimax value is not the worst-case equilibrium payoff, even when players are very patient (δ approaches unity). This can easily be shown from the one-shot game presented in Table 1. Here, $v_1^r(\Gamma_1) = \frac{2}{3}$ and $v_2^r(\Gamma_1) = -\frac{1}{3}$, where the latter can only be achieved when Player 2 plays B and Player 1 chooses B with probability $\frac{1}{3}$ and P with probability $\frac{2}{3}$. Clearly, this is not a Nash Equilibrium. In any equilibrium in this one-shot game, the dominated strategy P is not played and Player 2 therefore receives a positive payoff.

Nevertheless, Theorem 1 can be re-formulated and proven in terms of the worst-case equilibrium payoff for patient enough players. The idea is that for all the players in I_+ , $v_i^e(\Gamma_0) = v_i^r(\Gamma_0) = v_i(\Gamma_0)$ and for all players in general, $v_i^e(\Gamma) \geq v_i^r(\Gamma)$ (with equality for Γ_0). Hence, whenever the theorem conditions hold, v_i^r rises such that $v_i^r(\Gamma) \geq v_i^r(\Gamma_0) = v_i^e(\Gamma_0)$ for all the players, which implies $v_i^e(\Gamma) \geq v_i^e(\Gamma_0)$.

This result strengthens the argument for focusing on the rational minimax value – it provides a tighter lower bound on the equilibrium payoffs and the worst-case equilibrium payoff exhibits the same behavior as the rational minimax value.

Alas, as mentioned above, the Folk Theorem does not hold in general SAR games, so it is possible that $v_i^e(\Gamma) > v_i^r(\Gamma)$. Actually, since we do not impose any regularity or structure conditions on the SAR game, it may not be stationary or repeated in any sense. As in the one-shot game, it is fully possible that after the first stage, some punishing strategies will no longer be available. This would make it impossible to enforce cooperation in order to reach a particular payoff (including the rational minimax value).

Yet, the Folk Theorem does hold in the models presented in Section 4. This is due to the structure of these particular SAR games, which in the main are stationary in expectation. In other words, the expected discounted number of stages until the situation of the first stage is repeated is finite for each player after every history. The proof in this situation is very similar to that presented in Yoon [2001].

4 Two-Player Games with One-Sided Asynchronicity

We study a two-player game in which player 1 can revise her actions at every stage while player 2's revision opportunities depend on some exogenous random variable. Whenever player 2 cannot revise his action, he is forced to repeat the pure action

he chose in the previous stage (“strict-revision process”, see Tsodikovich and Lehrer [2019]). We set two different levels of information for players. In the first model, the players know the schedule of revisions, while in the second model they only know the distribution of the timing of the next revision (clearly, player 2 knows that he has a revision opportunity, but not how long he will subsequently remain inactive; player 1 learns that player 2 had revision opportunities only after choosing her action). This knowledge gap between the models affects the rational minimax values of both players.

Formally, let G be a two-player stage game, $X \in \Delta(\mathbb{N})$ a random variable with finite support,⁵ and x_1, x_2, \dots iid realizations of X . A repeated game with one-sided asynchronous play is a SAR game where player 1 can revise actions at any stage ($A_1(t) = \Delta(A_1)$) and player 2 can revise only at stages $t = x_1, x_1 + x_2, \dots$. At other stages, $A_2(t)$ is a singleton containing the pure action played in stage $t - 1$.

We assume that both players know the distribution of X and consider two models that differ with respect to information about its realizations. In the complete information model, denoted by Γ_X , the realizations of X are known to both players at the outset of the game (equivalently, during each revision they learn the timing of the next revision). In the unknown realizations model, denoted by $\tilde{\Gamma}_X$, player 1 does not know the realizations of X and therefore does not know whether a revision will be possible at each stage. Regardless, based on T , the number of stages since the previous revision, she can compute the conditional probability of a revision at the current stage, $q(T) = \Pr(X = T | X \geq T)$, and base her strategy on it. Similarly, whenever a revision opportunity is granted to Player 2, he does not know the timing of the next revision opportunity so he cannot plan ahead.

Our models are an extension of an example in Wen [2002], who studied the repeated “Battle of the Sexes” (Table 1 without “P”) in which player 1 can revise her actions at every stage while player 2 can revise his only at stages $0, T, 2T, \dots$ for some constant known $T > 1$. In these settings, player 2’s payoff can be lowered to an undiscounted average of $v_2 = \frac{2}{3T}$, while the minimax value of player 1 rises to $v_1 = 1$ regardless of T . Nevertheless, due to the structure of the feasibility set, the equilibrium payoff of player 2 cannot be lower than $v_2^r = \frac{1}{2}$.

⁵This is similar to the Finite Period of Inaction in Expectation condition from Yoon [2001, 2004] and Wen [2002]. This assumption is required for the Folk Theorem to hold in these models. Thus, $v^e = v^r$ for all players.

4.1 One-Sided Asynchronous Games with Complete Information

We approach this problem in two steps. In the first step, we assume that the underlying stage game is a zero-sum game and compute its value. This illustrates how a game with one-sided asynchronicity can unfold and provides a method of computing the standard minimax value for each player in non-zero-sum games. The second step is to compare this calculation to the standard minimax value in the one-shot game and apply the conditions of Theorem 1.

Although this model takes the edge off the randomization by having all the revision stages known in advance, we find that the analysis is still valuable. First, the model is simple enough for the value to be computed directly. Second, it offers a baseline for comparison with more complex models, such as the model where revisions are unknown. Moreover, this is the “worst” model in terms of the restricted player in the zero-sum game, as any unknown information can serve to increase his value.

4.1.1 The Value of the Zero-Sum Repeated Game

Suppose that the underlying stage game is a zero-sum game, where player 1 is the maximizer. Whenever player 2 has no revision opportunity, player 1 will best respond to the anticipated pure action. Thus, at the stages where he has a revision opportunity, player 2 needs to consider not only the stage payoff but also the continuation payoff until his next revision.⁶ The result of this reasoning is an algorithm to compute the optimal strategies and the value of the game, as shown in Proposition 1.

Proposition 1. *The game $\Gamma_X = (G, X, \delta)$ where G is a zero-sum stage game has the value*

$$V_{\Gamma_X} = \frac{(1-\delta)\mathbb{E}(V_X)}{1-\mathbb{E}(\delta^X)}, \quad (4)$$

where V_n is the value of the one-shot zero-sum auxiliary game with the payoff function

$$u_n(a_1, a_2) = u(a_1, a_2) + \frac{\delta - \delta^n}{1 - \delta} \max_{x \in A_1} u(x, a_2), \quad (5)$$

and $\mathbb{E}(V_X) = \sum_n V_n \Pr(X = n)$.

The auxiliary game represents the situation a player faces when a revision opportunity is given: player 2 chooses an action for his entire period of inactivity and player

⁶Except for trivial stage games where player 2 can achieve the value in pure actions.

| | | Player 2 | |
|----------|---|----------|---|
| | | B | S |
| Player 1 | B | 2 | 0 |
| | S | 0 | 1 |
| | P | 0 | 0 |

| | | Player 2 | |
|----------|---|----------------------------|-------------------------|
| | | B | S |
| Player 1 | B | $2(1 + \delta + \delta^2)$ | $\delta + \delta^2$ |
| | S | $2(\delta + \delta^2)$ | $1 + \delta + \delta^2$ |
| | P | $2(\delta + \delta^2)$ | $\delta + \delta^2$ |

Table 2: The payoffs of the row player in the auxiliary one-shot zero-sum games derived from the game presented in Table 1 for $n = 1$ (left) and $n = 3$ (right).

1 only needs to choose an action for this stage. Subsequently, the pure action is fixed and she can best respond to it. The value in Eq. (4) is larger than the value of the one-shot game (which is v_1) unless player 2 has a pure optimal strategy.

For large realizations of X and sufficiently patient players, the first term in Eq. (5) is insignificant relative to the second term, which means that the play is in effect sequential: player 2 chooses first and player 1 responds. Therefore, when X is a “large” random variable (meaning that $\mathbb{E}(\delta^X)$ is very small) the value of the game reaches its limit – the minimax value of the stage game in *pure* actions. In addition, for a fixed X we show in Proposition 3 in the Appendix that the value is continuous in δ when approaching patience, and its limit is the value of the undiscounted game.

Example 1 (Continued). *The minimax values in “Battle of the Sexes” when the realizations of X are common knowledge.*

Consider the game presented in Table 1 and assume player 2 revises his actions according to the random variable

$$X = \begin{cases} 1 & \text{w.p. } \frac{1}{2}, \\ 3 & \text{w.p. } \frac{1}{2}. \end{cases} \quad (6)$$

Again, the row player can play “P” in every stage, setting $v_2(X) = -1$. To compute the standard minimax value of the row player, we consider only her payoffs and, according to Proposition 1, compute the value of two one-shot auxiliary games presented in Table 2. For $n = 1$ this is the regular one-shot game with $V_1 = \frac{2}{3}$. For $n = 3$ and for a large enough discount factor ($\delta > \frac{\sqrt{5}-1}{2}$) the optimal strategies are $[S, S]$ and the value is $V_3 = 1 + \delta + \delta^2$. Plugging these numbers into Eq. (4) leads to the standard minimax

value of the row player in the repeated zero-sum game:

$$v_1(X) = (1 - \delta) \frac{\frac{5}{3} + \delta + \delta^2}{2 - \delta - \delta^3}. \quad (7)$$

For patient enough players ($\delta \rightarrow 1$), the standard minimax value is $v_1(X) = \frac{11}{12}$, which is also the result in the undiscounted case (see Proposition 3 in the Appendix): when $X = 1$ the payoff is $\frac{2}{3}$ during one stage and when $X = 3$ the payoff per stage is 1 and is given during 3 stages. The total payoff in these 4 stages is $3 + \frac{2}{3}$ and the expected payoff per stage is $\frac{3 + \frac{2}{3}}{4} = \frac{11}{12}$.

△

4.1.2 The Rational Minimax Value of the Non-Zero-Sum Game

Eq. (4) is the value of the zero-sum game which can be used to determine the standard minimax value. We showed that unless player 2 has a pure minimax strategy against player 1, her minimax value will increase. The last feature required to apply Theorem 1 is that in the one-shot game, the rational minimax value of player 2 must be larger than his standard minimax value. This happens only if the minimax point in the stage game lies in region “A”, shown in Figure 3. These conditions combined define the class of stage games where one-sided asynchronous play is better for both players than the simultaneous-move repeated game in terms of rational minimax values.

Corollary 1. *If G is a two-player stage game such that*

1. *Player 2 has no pure minimaxing strategy against player 1; and*
2. *$v_2 < \underline{e}(v_1) = \min\{y \mid (v_1, y) \in F\}$ (i.e., $(v_1, v_2) \in A$),*

then for every random variable $X \in \Delta(\mathbb{N})$ and every discount factor $\delta \in (0, 1)$, the rational minimax value for both players in Γ_X is larger than the rational minimax value in the simultaneous-move game.

The proof is omitted as it is a direct result of Theorem 1, Proposition 1 and the discussion above. Note that the strong inequality in the second condition imposes a strong inequality in the result, despite the fact that in Theorem 1 the inequality is weak. Instead, we show how this corollary affects the “Battle of the Sexes” example above.

Example 1 (Continued). *The rational minimax values in “Battle of the Sexes”.*

Both conditions of Corollary 1 are satisfied for the game shown in Table 1. Thus, for any X , the rational minimax value of both players in Γ_X is larger than the rational minimax value in Γ_0 . For example, when considering X presented in Eq. (6) we showed that for patient enough players $v_1^r(\Gamma_X) = \frac{11}{12}$, which yields $v_2^r(\Gamma_X) = -\frac{1}{12} > v_2^r(\Gamma_0)$. \triangle

Note that this result does not hold in the symmetric version of the “Battle of the Sexes”, where action P is unavailable since the minimax point lies within F . This statement holds for all symmetric games – if the game is symmetric, the standard minimax value for all players is equal and v_G must lie within F . Thus, the rational minimax is also the standard minimax and the conditions of Corollary 1 (and Theorem 1 in the general case) never hold.

Nevertheless, this model enables us to directly study the effect of X on the rational minimax payoff. Suppose that the game is the symmetric “Battle of the Sexes”, X is the random variable that is either 1 w.p. $1 - p$ or 2 w.p. p , and for simplicity we consider the undiscounted case ($\delta \rightarrow 1$). In this case, the standard minimax values are $v_1 = \frac{2+4p}{3+3p}$ and $v_2 = \frac{1}{2+p}$. For $0 \leq p \leq 0.5$, the minimax point is an interior point of F and the standard minimax value is both the rational minimax value and the tight lower bound on the equilibrium payoffs. As p increases further, the minimax point leaves F and moves diagonally right and down (see Figure 4). In this region, the standard minimax value of player 1 is also his rational minimax value, but this is not true for player 2. His rational minimax value is $v_2^r = \underline{v}(v_1) = \frac{v_1}{2}$, the corresponding value of v_1 on the lower envelope of F , and it is now increasing with p . Eventually, when $p = 1$, $v_1 = 1$, $v_2 = \frac{1}{3}$ and $v_2^r = \frac{1}{2}$ as was calculated by Wen [2002] (see Figure 5).

4.2 One-Sided Asynchronous Games with Partial Information

We now consider the second model, where the realizations of X are unknown and only the distribution of X is common knowledge. We assume that revision opportunities are observable after the actions are chosen and, as before, start by studying the value when the stage game is a zero-sum game. Unlike Section 4.1.1, there is no simple formula for the new value of the game. Moreover, it is not clear whether the value changes, even if the optimal strategy of player 2 in the zero-sum stage game is mixed.

Consider some stage $t \neq 0$ and suppose that the pure action a was played at past

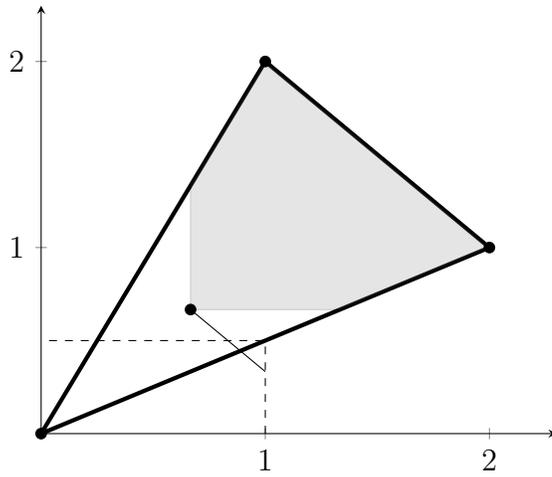


Figure 4: The movement of the minimax point for the symmetric “Battle of the Sexes” game, as the probability of revision every turn decreases for the column player. For a certain probability ($p = 0.5$), the minimax point leaves F , and the rational effective minimax starts to rise for both players along the lower envelope of F .

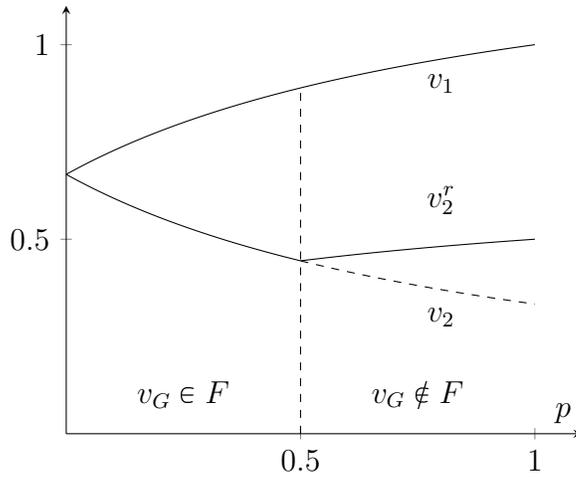


Figure 5: The standard and rational minimax value in the symmetric “Battle of the Sexes”. When the minimax point leaves F , the rational minimax value of Player 2 rises (solid) due to the nature of the game, while the standard minimax value continues to fall (dashed).

$T - 1$ stages by player 2. Action a will be played at this stage either if player 2 has no revision opportunity (with probability $1 - q(T)$) or has one and decides to replay it (with some probability). Thus, this stage is equivalent to a stage game where player 2 can choose any action in $\Delta(A_2)$ that satisfies $\Pr(a) \geq 1 - q(T)$. If there exists an optimal strategy satisfying this condition, player 2 can secure the value of the one-shot game for stage t . Otherwise, he must choose a mixed action from a set that does not contain any optimal action, which means getting a lower payoff for this stage and, as a result, for the entire repeated game.

The existence of a maximin strategy that chooses a with probability higher than $1 - q(T)$ for any T is not enough to secure the value at other stages. The maximin strategy may have a non-zero probability of choosing an action that is never played in an optimal strategy with probability higher than $1 - q(T)$ for some T . In this case, there is some positive probability that this action will be chosen and that, at one of the subsequent stages, player 2 will be forced to play non-optimally. To guarantee the value at each stage, it is essential that every action chosen under the optimal strategy can be played under some optimal strategy with probability higher than $1 - q(T)$ for all possible T s. Therefore, the maximin strategies should choose this action with probability higher than

$$p_X = \sup_{T \leq \max(\text{supp}(X))} (1 - q(T)). \quad (8)$$

The next definition formally defines the set of pure actions that can be used in this manner.

Definition 1. *A non-empty set of actions, $B_i \subseteq A_i$, is said to be a p -min optimal set (for player i) in the one-shot zero-sum game G if for every $a \in B_i$ there exists an optimal strategy $\alpha \in \Delta(B_i)$ that satisfies $\alpha(a) \geq p$.*

It follows from the above that the existence of a p_X -min optimal set for player 2 is sufficient to achieve the value of the one-shot game in the repeated game. The next proposition formalizes this discussion and proves that this is also a necessary condition. Otherwise, player 2 will receive less in the repeated game than in the value of the stage game.

Proposition 2. *Let G be a zero-sum two-player stage game, $X \in \Delta(\mathbb{N})$ and $\tilde{\Gamma}_X$ the corresponding repeated game where the realizations of X are unknown. The value of the repeated game, $V_{\tilde{\Gamma}_X}$ is equal to the value of the one-shot stage game, $v(\Gamma_1)$, iff player 2 has a p_X -min optimal set, where p_X is defined in Eq. (8). Otherwise, $V_{\tilde{\Gamma}_X} > V_{\Gamma_1}$.*

Finally, consider a repeated game $\tilde{\Gamma}_X$ where G is a non-zero-sum stage game. According to Proposition 2, the standard minimax value of player 1 can remain the same as in the one-shot game whenever player 2 has a p_X -min optimal set against player 1, despite the latter's advantage. In this case, Theorem 1 cannot be applied and the rational minimax value of player 2 could also remain the same. The lack of information about the realizations of X benefits player 2 when the underlying game is a zero-sum game but prevents him from increasing the worst-case rational payoff otherwise.

Example 2. *“Battle of the Sexes” with unknown realizations of X .*

We return to Example 1, this time assuming that the realizations of X are unknown. For the standard minimax value of the row player to rise, it is necessary and sufficient to show that in the zero-sum game shown in Table 2, the column player has no p_X -min action set. This is true, for example, for the X defined in Eq. (6) since $q(2) = 0$. At every stage where the column player has no revision opportunity, the row player knows that $X = 3$ and the revision will take place only at the next stage. She will best respond to the anticipated pure action and the value of the game increases.

On the other hand, for $X \sim \text{Geom}(\frac{2}{3})$ and for every $T \in \mathbb{N}$, the probability of maintaining the previous action is at least $1 - q(T) = \frac{1}{3}$, thus $p_X = \frac{1}{3}$. The optimal strategy in the stage game is $[\frac{1}{3}(B), \frac{2}{3}(S)]$ thus $\{B, S\}$ is $\frac{1}{3}$ -min action set and player 2 can obtain the value by using the following strategy: at every revision opportunity, change actions with probability 0.5. Neither the standard minimax value of the row player nor the rational minimax value of both players changes for this X . \triangle

The last example shows that the important property of X in this model is the conditional probabilities denoted by $q(T)$ and not the actual probabilities or the support. In the example, Player 2 prefers a revision process with infinite support that reveals no information over a random variable with finite and small support whose conditional probabilities reveal information about the realizations.

5 Concluding Remarks

This paper addresses a dual objective: to introduce a generalization of repeated games and define the rational minimax value. Our new model for repeated games, SAR games, covers more general frameworks that encompass asynchronicity and other obstacles to revising actions. Here, we examined its impact on the worst-case rational payoff. We leave it to future research to explore other potential uses.

Our new rational minimax value was used to achieve better bounding of the equilibrium payoffs. This value deserves greater attention as it holds interesting information on payoffs, on strategies in simultaneous-move games (and SAR games in general) and on whether cooperation between players is possible. This too is left for the future.

References

- Riccardo Calcagno and Stefano Lovo. Preopening and equilibrium selection. 2010.
- Riccardo Calcagno, Yuichiro Kamada, Stefano Lovo, and Takuo Sugaya. Asynchronicity and coordination in common and opposing interest games. *Theoretical Economics*, 9(2):409–434, 2014.
- Fabien Gensbittel, Stefano Lovo, Jérôme Renault, and Tristan Tomala. Zero-sum revision games. *Games and Economic Behavior*, 108:504–522, 2018.
- Yuichiro Kamada and Michihiro Kandori. Asynchronous revision games. 2011.
- Yuichiro Kamada and Takuo Sugaya. Asynchronous revision games with deadline: Unique equilibrium in coordination games. *Unpublished paper*, 2010.
- Roger Lagunoff and Akihiko Matsui. Asynchronous choice in repeated coordination games. *Econometrica: Journal of the Econometric Society*, pages 1467–1477, 1997.
- Jan Libich and Petr Stehlík. Incorporating rigidity and commitment in the timing structure of macroeconomic games. *Economic Modelling*, 27(3):767–781, 2010.
- Jan Libich and Petr Stehlík. Endogenous monetary commitment. *Economics Letters*, 112(1):103–106, 2011.
- Stefano Lovo and Tristan Tomala. Markov perfect equilibria in stochastic revision games. *HEC Paris Research Paper No. ECO/SCD-2015-1093*, 2015.
- W Bentley MacLeod. A theory of conscious parallelism. *European Economic Review*, 27(1):25–44, 1985.
- Eric Maskin and Jean Tirole. A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles. *Econometrica: Journal of the Econometric Society*, pages 571–599, 1988.

- Ran Spiegler. Agility in repeated games: An example. *Economics Letters*, 131:47–49, 2015.
- Satoru Takahashi and Quan Wen. On asynchronously repeated games. *Economics Letters*, 79(2):239–245, 2003.
- Yevgeny Tsodikovich and Ehud Lehrer. Stochastic revision opportunities in markov decision problems. *Annals of Operations Research*, 279(1-2):251–270, 2019.
- Quan Wen. The “folk theorem” for repeated games with complete information. *Econometrica: Journal of the Econometric Society*, pages 949–954, 1994.
- Quan Wen. Repeated games with asynchronous moves. *Department of Economics Vanderbilt University Working Papers*, 204, 2002.
- Kiho Yoon. A folk theorem for asynchronously repeated games. *Econometrica*, 69(1): 191–200, 2001. ISSN 1468-0262.
- Kiho Yoon. The effective minimax value of asynchronously repeated games. *International Journal of Game Theory*, 32(4):431–442, 2004.

A Proofs of Theorems

Theorem 1. *Let $\Gamma = (G, \{A_i(t)\}_{i \in I}^{t \in \mathbb{N}}, \delta)$ be a SAR game and let Γ_0 be the corresponding simultaneous-move game. Denote the set of players whose rational minimax in Γ_0 is equal to the standard minimax value in Γ_0 by I_* .*

If $v_i(\Gamma) \geq v_i(\Gamma_0)$ for every $i \in I_$ then $v_i^r(\Gamma) \geq v_i^r(\Gamma_0)$ for all players.*

Proof of Theorem 1. The theorem trivially holds for all the players in I_* . Assume $I_* \neq I$ and assume by contradiction that there exists a player j for whom $v_j^r(\Gamma_0) > v_j^r(\Gamma)$. Therefore, there exists a feasible and individually rational payoff \underline{x} in the game Γ so that $v_j^r(\Gamma_0) > x_j \geq v_j^r(\Gamma)$. In addition, by the closedness of the feasible and individually rational set of the game Γ_0 , there exists a payoff $\underline{y} \in F_{IR}$ so that $y_j = v_j^r(\Gamma_0)$.

Define $\underline{z} = \epsilon \underline{y} + (1 - \epsilon) \underline{x}$. Since F is convex, $\underline{z} \in F$ for all $\epsilon \in [0, 1]$. For players in I_* , the payoff \underline{z} is individually rational in Γ_0 since $z_i = \epsilon y_i + (1 - \epsilon) x_i \geq \epsilon v_i(\Gamma_0) + (1 - \epsilon) v_i(\Gamma) \geq v_i(\Gamma_0)$. It is possible to choose ϵ very close to 1 so that the payoff will be rational also for the rest of the players, since $z_i = \epsilon y_i + (1 - \epsilon) x_i \geq \epsilon v_i^r(\Gamma_0) + (1 - \epsilon) x_i$ and $v_i^r(\Gamma_0) > v_i(\Gamma_0)$. However, $z_j = \epsilon v_j^r(\Gamma_0) + (1 - \epsilon) x_j < v_j^r(\Gamma_0)$ which is a contradiction since \underline{z} is an individually rational payoff in Γ_0 with a lower payoff for player j than its rational minimax value.

Therefore, in the game Γ the rational minimax of all the players in $I \setminus I_*$ is greater than in Γ_0 and the proof is complete. \blacksquare

Proposition 1. *The game $\Gamma_X = (G, X, \delta)$ where G is a zero-sum stage game has the value*

$$V_{\Gamma_X} = \frac{(1-\delta)\mathbb{E}(V_X)}{1-\mathbb{E}(\delta^X)}, \quad (4)$$

where V_n is the value of the one-shot zero-sum auxiliary game with the payoff function

$$u_n(a_1, a_2) = u(a_1, a_2) + \frac{\delta - \delta^n}{1 - \delta} \max_{x \in A_1} u(x, a_2), \quad (5)$$

and $\mathbb{E}(V_X) = \sum_n V_n \Pr(X = n)$.

Proof of Proposition 1. First, note that the SAR game has a value according to standard arguments of contraction, so the rest of the proof deals with finding it. Second, note that if the realization of X at $t = 0$ is n , then starting from $t = n$ the situation is the same situation as at $t = 0$, so the continuation payoff is $\delta^n V_{\Gamma_X}$, where

V_{Γ_X} is the value of the game. At the first n stages the actions and payoffs can be found by writing explicitly the optimization problem of the payoff:

$$(1 - \delta) \min_{\alpha_2 \in \Delta(A_2)} \left(\max_{\alpha_1^0 \in \Delta(A_1)} u(\alpha_1^0, \alpha_2) + \sum_{a_2 \in A_2} \alpha_2(a_2) \left(\sum_{t=1}^{n-1} \max_{a_1^t \in A_1} \delta^t u(a_1^t, a_2) \right) \right), \quad (9)$$

where a_2 is the pure action that was chosen by the mixed action α_2 with probability $\alpha_2(a_2)$ at $t = 0$ and a_1^t is the pure action of player 1 in stage t . For $1 \leq t \leq n - 1$ the maximizing action would be a pure best response to a_2 , since a_2 is already known in advance in those stages. We denote by $b : A_2 \rightarrow A_1$ to be some best response function (it does not matter which best response to a_2 is chosen, if there are several, as all yield the same payoff and we are interested only in the payoff and it is the same for all best responses and all $1 \leq t \leq n - 1$).

We can rewrite the payoff of the mixed actions as the expected value of the pure actions and use the fact that every $\alpha_1 \in \Delta(A_1)$ is a probability distribution over A_1 to turn the last formula into

$$(1 - \delta) \min_{\alpha_2 \in \Delta(A_2)} \max_{\alpha_1 \in \Delta(A_1)} \sum_{a_2 \in A_2} \sum_{a_1 \in A_1} \alpha_2(a_2) \alpha_1(a_1) \left(u(a_1, a_2) + \frac{\delta - \delta^n}{1 - \delta} u(b(a_2), a_2) \right). \quad (10)$$

For every $n \in \mathbb{N}$, consider the auxiliary zero-sum stage game with the same players and actions with the modified payoff function

$$u_n(a_1, a_2) = u(a_1, a_2) + \frac{\delta - \delta^n}{1 - \delta} u(b(a_2), a_2) = u(a_1, a_2) + \frac{\delta - \delta^n}{1 - \delta} \max_{x \in A_1} u(x, a_2). \quad (11)$$

This auxiliary game has a value, V_n , which is the minimax value in Eq. (10). Therefore, the payoff of the first n stages is $(1 - \delta)V_n$. The expected δ -discounted payoff from $t = 0$ until the next revision opportunity is therefore $\mathbb{E}(V_X) = \sum_n V_n \cdot \Pr(X = n)$ and the continuation payoff from this revision onward is $\mathbb{E}(\delta^X)V_{\Gamma_X}$.

To conclude, the value of the repeated zero-sum game must satisfy

$$V_{\Gamma_X} = (1 - \delta)\mathbb{E}(V_X) + \mathbb{E}(\delta^X)V_{\Gamma_X}, \quad (12)$$

which leads to Eq. (4) and the proof is complete. \blacksquare

Proposition 2. *Let G be a zero-sum two-player stage game, $X \in \Delta(\mathbb{N})$ and $\tilde{\Gamma}_X$ the corresponding repeated game where the realizations of X are unknown. The value of the repeated game, $V_{\tilde{\Gamma}_X}$ is equal to the value of the one-shot stage game, $v(\Gamma_1)$, iff player 2 has a p_X -min optimal set, where p_X is defined in Eq. (8). Otherwise, $V_{\tilde{\Gamma}_X} > V_{\Gamma_1}$.*

Proof of Proposition 2. Suppose that player 2 has a p_X -min action set B and consider the following strategy for him:

1. At $t = 0$ choose an action according to some maximin strategy with support over B .
2. Whenever a revision opportunity arrives after T consecutive stages playing the pure action a , play each pure action with the probability:

$$\alpha'_2(a') = \mathbf{1}_{a' \neq a} \frac{\alpha_2^*(a')}{q(T)} + \mathbf{1}_{a'=a} \frac{\alpha_2^*(a') - (1-q(T))}{q(T)}, \quad (13)$$

where α_2^* is some minimax strategy with support over B .⁷

Assume a was played T stages in a row, and player 2 follows this strategy. The probability that the pure action a' will be played next is

$$\begin{aligned} \Pr(a') &= \Pr(a'|\text{no revision})\Pr(\text{no revision}) + \Pr(a'|\text{revision})\Pr(\text{revision}) \\ &= \mathbf{1}_{a'=a}(1 - q(T)) + \alpha'_2(a')q(T) = \alpha_2^*(a'). \end{aligned} \quad (14)$$

In this stage, player 1 plays against the mixed action α_2^* which yields the expected payoff of at least V_{Γ_1} . This is true for every stage, so the payoff for player 2 in the repeated game is at least V_{Γ_1} as well. Player 1, however, can guarantee to receive at least V_{Γ_1} by playing the minimaxing action of the stage game in every stage, which sets the value of the game to be exactly $V_{\tilde{\Gamma}_X} = V_{\Gamma_1}$.

The existence of a p_X -min action set for player 2 is also a necessary condition. Otherwise, there is a positive probability to play in one of the stages an action that is played in maxmin strategies in lower than p_X probability. When such action is played, for large enough realization of X , there will be a stage in which $1 - q(T)$ is greater than the probability to play this action in equilibrium, forcing player 2 to choose this action with probability higher than the optimal probability. In this stage, the best response of player 1 to the subset of mixed actions in which the former action is played with probability of at least $1 - q(T)$ would yield a higher payoff than V_{Γ_1} , resulting in $V_{\tilde{\Gamma}_1} > V_{\Gamma_1}$. ■

B The Continuity of the Value Function For $\delta = 1$

In the following proposition we show that for the zero-sum game introduced in Section 4.1.1, the value function is continuous as $\delta \uparrow 1$. Thus, for patient enough players, it

⁷Thus $\alpha_2^*(a) \geq 1 - q(T)$ and α_2^* is indeed a probability.

is possible to approximately calculate the value in a simpler manner than the one described in Proposition 1, by considering the undiscounted value instead of a discounted value. An example of such computation can be found in the paragraph that follows Eq. (7).

Proposition 3. *Consider the game $\Gamma_X = (G, X, \delta)$, with G being a zero-sum game and assume the players evaluate their undiscounted average payoff instead of the undiscounted stream of payoffs. If both players know the realizations of X then the game has a value and the value is $\frac{\mathbb{E}(XV_X^0)}{\mathbb{E}(X)}$, where for every n , V_n^0 is the value of the one-shot game defined by*

$$u_n^0(a_1, a_2) = \frac{1}{n}u(a_1, a_2) + \frac{n-1}{n} \max_{x \in A_1} u(x, a_2), \quad (15)$$

and $V_X^0 = \sum V_n^0 \Pr(X = n)$.

Proof The proof of is done in a similar manner to proposition 1, by dividing the game to “blocks” of inactivity and analysing the value of each such block.

Consider a stage of the game in which player 2 can revise the action and assume that $X = n$. Player 2 strives to minimize the payoff of the n next rounds so he should choose a mixed action that minimizes the maximum of his expected payoff (per round):

$$\frac{1}{n} \min_{\sigma_2 \in \Delta(A_2)} \left(\max_{\sigma_1^0 \in \Delta(A_1)} u(\sigma_1^0, \sigma_2) + \sum_{k=1}^{n-1} \sum_{a_2 \in A_2} \max_{\sigma_1^k \in \Delta(A_1)} \sigma_2(a_2) u(s_1^k, a_2) \right), \quad (16)$$

where a_2 is the pure action that was chosen by the mixed action σ_2 with probability $\sigma_2(a_2)$. As in the proof of Proposition 1, for $1 \leq k \leq n - 1$ the maximizing action for player 1 would be some pure action denoted $b(a_2)$, since a_2 is already predetermined in those stages. We can rewrite the payoff of the mixed actions as the expected value of the pure actions and use the fact that every $\sigma_1 \in \Delta(A_1)$ is a probability function over A_1 to turn the last formula into

$$\begin{aligned} & \min_{\sigma_2 \in \Delta(A_2)} \max_{\sigma_1 \in \Delta(A_1)} \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \sigma_1(a_1) \sigma_2(a_2) \left(\frac{1}{n} u(a_1, a_2) + \frac{n-1}{n} u(b(a_2), a_2) \right) \\ & = \min_{\sigma_2 \in \Delta(A_2)} \max_{\sigma_1 \in \Delta(A_1)} \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \sigma_1(a_1) \sigma_2(a_2) u_n(a_1, a_2) = V_n^0. \end{aligned} \quad (17)$$

The above arguments can be applied for the maximin strategy for player 1, which would yield the same result as in Equation (17). Thus, the expected payoff per round if $X = n$ is V_n^0 for the next n rounds. Since the above are the optimal mixed actions for both players in each “block” of n stages and since the play in each “block” is

independent of the other stages, the strategy of playing the appropriate optimal action in each block according to the realization of X is optimal in the repeated game. To compute the value of the game we need to compute the average expected payoff when playing those strategies.

Let $m \in \mathbb{N}$ and consider the $m + 1$ stages $(t_i)_{0 \leq i \leq m}$ in which player 2 had revision opportunities. Each t_i for $i \geq 1$ is a random variable with the property that the random variables $t_i - t_{i-1} := X_i$ are independent and have distribution of X . Thus, the expectation of the number of stage after m revisions, t_m , is $\mathbb{E}(X_1 + \dots + X_m) = m\mathbb{E}(X)$.

To compute the value of the game we can look at the average payoff

$$x_M(\sigma_1, \sigma_2) = \frac{1}{M} \sum_{t=0}^M \mathbb{E}(u(\sigma_1^t, \sigma_2^t)) \quad (18)$$

and rewrite it as

$$t_m \bar{x}_{t_m} = \sum_{t=0}^{t_m} x_t. \quad (19)$$

By taking expectations on both sides, the right hand side would be the total expected payoff upto stage t_m which is $\mathbb{E}(X_1 V_{X_1} + \dots + X_m V_{X_m}) = m\mathbb{E}(X V_X^0)$ and the left hand side is $\mathbb{E}(t_m \bar{x}_{t_m}) = \mathbb{E}(t_m) \mathbb{E}(\bar{x}_{t_m}) + Cov(t_m, \bar{x}_{t_m}) = m\mathbb{E}(X) \mathbb{E}(\bar{x}_{t_m}) + Cov(t_m, \bar{x}_{t_m})$. The value of the game is the expected average payoff when taking the limit $m \rightarrow \infty$:

$$\mathbb{E}(\bar{x}_{t_m}) = \frac{\mathbb{E}(X V_X^0)}{\mathbb{E}(X)} + \frac{Cov(t_m, \bar{x}_{t_m})}{m\mathbb{E}(X)}. \quad (20)$$

We can bound the covariance by the product of the appropriate standard deviations. First, note that $\sqrt{\mathbb{V}(t_m)} = \sqrt{m\mathbb{V}(X)}$. Second, the standard deviation of the average is bounded since the average itself is bounded by the largest and lowest possible payoffs. Thus, the covariance is $O(\sqrt{m})$ and when taking the limit $m \rightarrow \infty$ the covariance term drops and the right hand side turns to be $\frac{\mathbb{E}(X V_X^0)}{\mathbb{E}(X)}$. The limit of the left hand side of the equation is a limit point of the average expected payoff.

The above calculation was done only for time periods in which player 2 had a revision, and it is sufficient to do so since in any other stage of the game, the expected deviation from the last stage in which player 2 had a revision is bounded and goes to zero as $m \rightarrow \infty$. Thus, the series of the average expected payoffs has only one limit point which is, by definition, guarantee-able by both players and thus it is the value of the game: $\frac{\mathbb{E}(X V_X^0)}{\mathbb{E}(X)}$.

Note that indeed $V_{\Gamma_X} = \frac{(1-\delta)\mathbb{E}(V_X)}{1-\mathbb{E}(\delta^X)} \rightarrow \frac{\mathbb{E}(X V_X^0)}{\mathbb{E}(X)}$ as $\delta \uparrow 1$. This calculation is straightforward using l'Hôpital's rule when minding the slightly different definition of the auxiliary game in the discounted and undiscounted case. ■