

Article

# Reactive Strategies: An Inch of Memory, a Mile of Equilibria

Artem Baklanov <sup>1,2</sup> 

<sup>1</sup> Department of Economics, HSE University, 16 Soyuzna Pechatnikov St., 190121 Saint Petersburg, Russian; apbaklanov@hse.ru or baklanov@iiasa.ac.at

<sup>2</sup> International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, A-2361 Laxenburg, Austria

**Abstract:** We explore how an incremental change in complexity of strategies (“an inch of memory”) in repeated interactions influences the sets of Nash equilibrium (NE) strategy and payoff profiles. For this, we introduce the two most basic setups of repeated games, where players are allowed to use only reactive strategies for which a probability of players’ actions depends only on the opponent’s preceding move. The first game is trivial and inherits equilibria of the stage game since players have only unconditional (memory-less) Reactive Strategies (RSs); in the second one, players also have conditional stochastic RSs. This extension of the strategy sets can be understood as a result of evolution or learning that increases the complexity of strategies. For the game with conditional RSs, we characterize all possible NE profiles in stochastic RSs and find all possible symmetric games admitting these equilibria. By setting the unconditional benchmark as the least symmetric equilibrium payoff profile in memory-less RSs, we demonstrate that for most classes of symmetric stage games, infinitely many equilibria in conditional stochastic RSs (“a mile of equilibria”) Pareto dominate the benchmark. Since there is no folk theorem for RSs, Pareto improvement over the benchmark is the best one can gain with an inch of memory.

**Keywords:** Nash equilibrium; 1-memory or memory-one strategies; infinitely repeated games

**MSC:** 91A05; 91A20; 91A10



**Citation:** Baklanov, A. Reactive Strategies: An Inch of Memory, a Mile of Equilibria. *Games* **2021**, *12*, 42. <https://doi.org/10.3390/g12020042>

Academic Editor: Ulrich Berger

Received: 22 March 2021

Accepted: 30 April 2021

Published: 8 May 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the theory of repeated games, restrictions on players’ strategies are not usually imposed (see, for example, [1]). However, an interest in bounded rationality and strategic complexity [2] has led to the study of repeated games where players’ strategies are assumed to be finite automata or to have finite memory (bounded recall). This paper falls into the strategy-restrictions framework: we make an extensive study of the Nash equilibrium in infinitely repeated  $2 \times 2$  games where payoffs are determined by Reactive Strategies (RSs) and players’ payoffs are evaluated according to the limit of means. A player’s reactive strategies are strategies that, for each move of the opponent in the previous round, prescribe stationary probabilities to the player of playing his/her own pure actions.

One may naturally think of two-by-two classification of RSs. The first aspect is related to information. Players may ignore the available information about the action of the opponent in the previous round; hence, we distinguish between unconditional and conditional RSs. The second aspect is related to predictability of actions. Deterministic and semi-deterministic RSs allow deterministic responses (probabilities 1 or 0) to actions in the previous rounds, while for stochastic RSs, all responses have non-degenerate probabilities.

The paper explores how an incremental change of complexity of strategies in repeated interactions influences the sets of Nash equilibrium (NE) strategy and payoff profiles. For this, we introduce, perhaps, the two most basic setups of repeated games. In the first game, which is trivial and inherits equilibria of the stage game, players have only memory-less RSs. In the second game, players in addition have conditional stochastic RSs. This extension of the strategy sets can be understood as a result of evolution or learning,

resulting in increased complexity of strategies. For the game with conditional RSs, we answer the following questions:

- (Q1) What are all possible NE profiles in stochastic RSs?
- (Q2) What are all possible symmetric games admitting NE in stochastic RSs?
- (Q3) Do equilibrium profiles in conditional stochastic RSs Pareto improve over equilibrium profiles in unconditional RSs?

Surprisingly, the answers to (Q1)–(Q3), the most basic results for equilibria in reactive strategies, are still missing in the literature, while similar questions were studied for 1-memory (aka memory-one) strategies with discounted payoffs.

Before we proceed to more technical elements of the paper, let us give a simple example where reactive and 1-memory strategies meet together.

**Example 1.** A textbook example of repeated games is the iterated Prisoner's Dilemma, where the stage game has two strategies for both players, C (cooperate) and D (defect). Any round ends with a possible action profile: CC, CD, DC, or DD. As profile  $i$  of the previous round is known, player 1 chooses C with probability  $p_i$  (hence, D with probability  $1 - p_i$ ). Then a 4-tuple, say  $(p_{CC}, p_{CD}, p_{DC}, p_{DD})$ , would define a 1-memory strategy of player 1. A reactive strategy of player 1 is a 1-memory strategy that only depends on the opponent's action, i.e.,  $p_{CC} = p_{CD}$  and  $p_{DC} = p_{DD}$ ; this reduces 4-tuples to ordered pairs. By omitting the issue of opening moves, one derives 1-memory form for the strategies that were found to be the most popular in experimental study [3]. These strategies are Tit-for-Tat,  $(1, 0, 1, 0)$ ; Always Defect,  $(0, 0, 0, 0)$ ; both are reactive ones, while Grim,  $(1, 0, 0, 0)$ , is not.

### 1.1. Related Literature

Reactive strategies were introduced by Karl Sigmund and Martin Nowak (see [4,5]) as an abstraction of bounded rationality. However, while their research has never been focused directly on the Nash equilibrium, we should note that the authors' comprehensive study of evolutionary stable strategies is indeed relevant to (Q1), as the definition of these strategies requires a symmetric Nash equilibrium to be formed. This is why, for the infinitely repeated Prisoner's Dilemma in [4], a full characterization of symmetric Nash equilibria in the class of reactive strategies was obtained for the limit-of-means payoff. Note that these symmetric equilibria are equilibrium points (see [6]) of adaptive dynamics on the set of all reactive strategies. The concept of adaptive dynamics was introduced in [7] to describe evolution in continuous sets of strategies. In this paper, we significantly extend the results from [4] related to the Nash equilibrium by considering arbitrary  $2 \times 2$  stage games and by characterizing both symmetric and non-symmetric equilibria, with non-symmetric equilibria prevailing over symmetric ones.

For the discounted Prisoner's Dilemma, subgame perfect equilibria in the class of deterministic reactive strategies were studied in [8]. In [9], Appendix 5.5, a brief analysis can be found of equilibrium refinements in the class of deterministic reactive strategies (referred to as "memory zero" strategies) for a repeated Prisoner's Dilemma. In [9], R. Aumann considered eight pure strategies for each player (four stationary reactions multiplied by two actions for the initial round). Note that the concepts of reactive strategies and 1-memory strategies in repeated games with a continuum of pure strategies are related to stationary "reaction functions" [10,11], "immediately reactive strategies" [12], and stationary "single-period-recall strategies" [13]. In contrast to the studies mentioned in this paragraph, we consider limit-of-means payoffs and reactive strategies mixing over (two) pure actions, that is, reactive strategies mapping the previous opponent's moves into distributions over pure actions.

We also study the problem of the existence of a Nash equilibrium. Note that Kryazhimskiy [14] obtained conditions sufficient for the existence of a Nash equilibrium within subsets of the players' 1-memory strategies in repeated bimatrix games with the limit-of-means payoff. The conditions require the sets of players' strategies to have a convexity

property and all strategies to be strictly randomized, so the Kakutani fixed-point theorem becomes applicable.

There is evidently much more literature available on the topic of 1-memory strategies than on the topic of reactive strategies. The recent spike of interest in 1-memory strategies was stimulated by the discovery of Zero-Determinant (ZD) strategies, a special class of 1-memory strategies enabling linear relationships between players' payoffs, irrespective of opponent's strategies (see [15] for the case of the limit-of-means payoff). Remarkably, ZD strategies can fix the opponent's payoff or guarantee some surplus that outperforms the opponent's surplus by the chosen constant factor. Hilbe et al. [16] extended the theory of ZD strategies to the case of the discounted payoff; see [16] for the most recent review of the topic.

Within the literature dealing with 1-memory strategies, one of the most relevant studies for our research was conducted by Dutta and Siconolfi [17]. In that study, the ideas, which were independently proposed in the seminal paper [18] to derive a robust folk theorem for the discounted Prisoner's Dilemma, were extended to generic repeated  $2 \times 2$  games. In other words, Dutta and Siconolfi explored a *Strong Mixed Equilibrium* (SME) in the class of 1-memory strategies. This equilibrium is formed by completely mixed 1-memory strategies and has the salient property that both players are indifferent between their actions after every history. Specifically, in [17], conditions were obtained that are necessary and sufficient for a game to have an SME, and the set of all equilibria payoffs was described. The latter result allowed us to demonstrate that SMEs generally lead to a continuum of equilibrium payoffs but still do not generate a folk theorem. To summarize, Dutta and Siconolfi [17] studied subgame perfect equilibria formed by completely mixed (stochastic) strategies in repeated games with discounted payoffs and 1-memory strategies.

An analysis of subgame perfect equilibria (SPE) in 1-memory strategies focused on the Prisoner's Dilemma was presented in [19]. In particular, SPEs formed by zero-determinant strategies were characterized and the corresponding upper bound for all NE and SPE payoffs was obtained (it is equal to the mutual cooperation payoff).

Moving from 1-memory to finite memory strategies, we should mention finite automata as a common abstraction of bounded rationality; see [2,20,21].

Following [22], we observe different natural levels of complexity, motivating us to distinguish between deterministic and Stochastic Reactive Strategies (SRSs). In this paper, we focus on SRSs as the simplest ones. The higher complexity of some deterministic strategies is due to the fact that the probability of the first action should be a part of the strategy to correctly define expected payoffs, while the payoffs for stochastic strategies do not depend on the first action.

To summarize, the field of finite-memory strategies (and, in particular, reactive strategies) has been studied for the last 30 years. Especially comprehensive results were obtained for 1-memory strategies but even the most basic facts for generic  $2 \times 2$  stage games are still missing for equilibria in RSs. The research not only delivers the missing piece of theory but also demonstrates that even an inch of memory in repeated interactions translates into a mile of new equilibria.

## 1.2. Results and Structure of the Article

First, we provide an expository geometrical representation to draw contrasts between unconditional and conditional SRSs in the repeated games; see Section 2.1. An example of the Prisoner's Dilemma with equal gains from switching is considered in Section 2.2 to illustrate how the representation will be applied to study the Nash equilibrium.

Second, in Section 2.3, we derive a characterization of all Nash equilibria in the class of reactive strategies in terms of solutions to a system of equations and inequalities; see Theorem 1.

Third, in Section 2.4, Theorem 2, we present a characterization of all symmetric games admitting Nash equilibria in the class of reactive strategies; Appendix A contains all related proofs.

Fourth, we elaborate on equilibrium payoff profiles in conditional SRSs; see Section 3. Namely, we demonstrate that:

- If there exists an NE formed by a profile of conditional SRSs, then there are infinitely many NE profiles generated by conditional SRSs that, in general, have distinct payoffs, but we do not have a folk theorem.
- If there exists an NE formed by a profile of unconditional SRSs, then NE profiles in conditional SRSs either Pareto improve over it or provide the same payoff profile.

For symmetric games, the complete analysis of equilibrium payoff profiles becomes feasible in Section 3.2 and Appendix B. In our setting, an NE in unconditional RSs always exists. Among equilibrium payoff profiles in unconditional RSs, we select the symmetric one with the lowest payoffs as the benchmark. We then check if there exists an equilibrium payoff profile formed by conditional stochastic RSs that Pareto dominates this benchmark. Only in the very special games where both players (by choosing the dominating pure strategy) get the highest possible individual payoff are payoff profiles in conditional SRSs dominated by the benchmark unconditional payoff profiles. This means that it is possible to get new equilibria with lower payoffs while the strategies become more advanced; see an example in Section 3.3. In the remaining cases, if equilibrium profiles in SRSs exist and result in distinct payoff profiles, then there exists an NE profile in SRSs dominating the benchmark. In some cases, all NE profiles in conditional SRSs dominate the benchmark.

Fifth, with the extensive use of examples, we highlight some interesting properties of NE profiles in stochastic RSs; see Appendix C.

In this paper, some reactive strategies are excluded from players' strategy sets. In Section 2.5, we discuss how this decision influences the results.

### 1.3. Definitions of Repeated Games

Consider an arbitrary one-shot game  $G$

Game $G$	$L$	$R$
$T$	$A_{T,L}, B_{T,L}$	$A_{T,R}, B_{T,R}$
$B$	$A_{B,L}, B_{B,L}$	$A_{B,R}, B_{B,R}$

The first player chooses rows (possible strategies are  $T$  and  $B$ ); the second player chooses columns (possible strategies are  $L$  and  $R$ );  $T, B, L, R$  stand for *top, bottom, left, right*, respectively. A mixed strategy of player 1 is a probability  $s_1$  of playing action  $T$ ; similarly, a mixed strategy  $s_2$  of player 2 is a probability of playing action  $L$ . Mixed strategies defined by non-degenerate probabilities are called *completely mixed strategies*. Let us introduce payoff functions  $J_1, J_2 : [0, 1]^2 \rightarrow \mathbb{R}$  for the one-shot game  $G$  in mixed strategies by the rules:  $\forall s_1, s_2 \in [0, 1]$

$$J_1(s_1, s_2) \triangleq a_1 s_1 s_2 + b_1 s_1 + c_1 s_2 + A_{B,R}, \quad J_2(s_1, s_2) \triangleq a_2 s_1 s_2 + b_2 s_2 + c_2 s_1 + B_{B,R}, \quad (1)$$

where

$$a_1 \triangleq A_{T,L} - A_{T,R} - A_{B,L} + A_{B,R}, \quad b_1 \triangleq A_{T,R} - A_{B,R}, \quad c_1 \triangleq A_{B,L} - A_{B,R}; \quad (2)$$

$$a_2 \triangleq B_{T,L} - B_{T,R} - B_{B,L} + B_{B,R}, \quad b_2 \triangleq B_{B,L} - B_{B,R}, \quad c_2 \triangleq B_{T,R} - B_{B,R}. \quad (3)$$

#### 1.3.1. Strategies

In the following, we study a game  $G^\infty$  that is a repeated modification of  $G$ . To define  $G^\infty$  in normal form, we first introduce reactive strategies [4,5,23].

**Definition 1.** We define reactive strategies of player 1 (player 2) as arbitrary maps of  $\{L, R\}$  (maps of  $\{T, B\}$ ) into sets of all mixed strategies of player 1 (player 2); players' mixed strategies in the current round are completely defined by the preceding opponent's action. To simplify the notation, we understand any reactive strategy of player 1 as an ordered pair  $u = (u_1, u_2)$  such that the

conditional probability of action T and action B is  $u_1$  and  $(1 - u_1)$ , given that the second player's previous move was L, and the conditional probability of T and B is  $u_2$  and  $(1 - u_2)$ , given that the second player's previous move was R. In a similar way, a reactive strategy of player 2 is a pair  $v = (v_1, v_2)$  such that the conditional probability of action L and action R is  $v_1$  and  $(1 - v_1)$ , given that the first player's previous move was T, and the conditional probability of L and R is  $v_2$  and  $(1 - v_2)$ , given that the first player's previous move was B.

Reactive strategies form a subset of 1-memory strategies (see [15–18]). Figure 1 shows the classification of RSs that we follow in this paper.

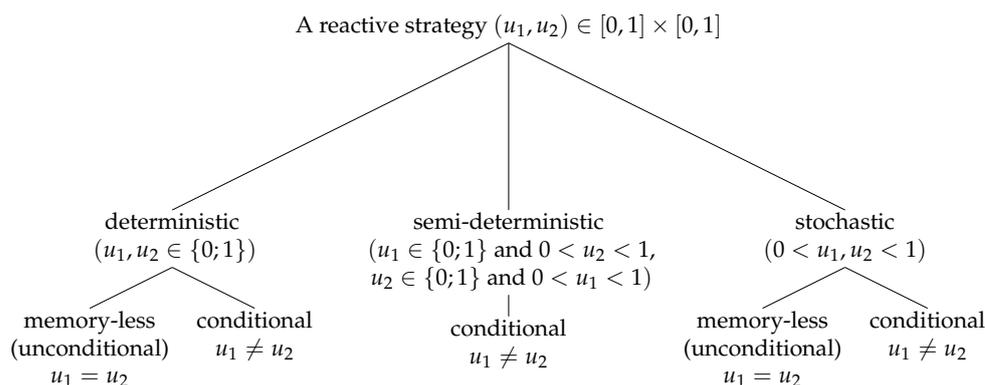


Figure 1. A possible classification of reactive strategies.

To avoid ambiguity, we denote an open interval  $\{x : a < x < b\}$  as  $]a, b[$ . For player 1 (player 2), we introduce the set of *Stochastic Reactive Strategies* (SRSs):  $U \triangleq ]0, 1[^2$  ( $V \triangleq ]0, 1[^2$ ).

We start with the repeated setting where players use only unconditional strategies (both deterministic and stochastic), game  $G_{un}^\infty$ . Game  $G_{un}^\infty$  is basically memory-less infinite repetition of the stage game. Then we slightly increase the complexity of players' behavior; i.e., we allow players to use also SRSs. This changes  $G_{un}^\infty$  into  $G^\infty$ . Thus, the set of all possible strategies of player 1 (of player 2) in game  $G^\infty$  is defined as  $\tilde{U} \triangleq U \cup \{(0, 0), (1, 1)\}$  (as  $\tilde{V} \triangleq U \cup \{(0, 0), (1, 1)\}$ ).

### 1.3.2. Payoffs

Given any initial action profile  $(i_0, j_0)$  corresponding to the first round and a profile of *stochastic* RSs, we obtain an infinite discrete-time Markov *stochastic* process (hence the name for strategies) defined on four possible states:  $(T, L)$ ,  $(T, R)$ ,  $(B, L)$ , and  $(B, R)$  (states 1, 2, 3, and 4, correspondingly). For every actual  $l$ -step trajectory of actions  $(i_1, j_1) \rightarrow (i_2, j_2) \rightarrow \dots \rightarrow (i_l, j_l)$ , the average payoffs are random variables. According to [14], for  $l \rightarrow \infty$ , the limits of expected  $l$ -round average payoffs are well defined for all SRSs and initial profiles. As usual, we call these limits limit-of-means payoffs.

The paper studies equilibrium profiles formed only by SRSs due to two reasons:

1. In contrast to semi-deterministic and deterministic RSs, the payoffs for profiles of SRSs do not depend on the opening move;
2. SRSs capture non-deterministic behavior that is the most natural for the domain of evolutionary game theory, where RSs originated to model real-life processes.

Every profile  $(u, v)$  of SRSs ensures nice properties for the induced Markov process with the transition probability matrix

$$M = \begin{pmatrix} u_1 v_1 & u_1(1 - v_1) & (1 - u_1)v_1 & (1 - u_1)(1 - v_1) \\ u_2 v_1 & u_2(1 - v_1) & (1 - u_2)v_1 & (1 - u_2)(1 - v_1) \\ u_1 v_2 & u_1(1 - v_2) & (1 - u_1)v_2 & (1 - u_1)(1 - v_2) \\ u_2 v_2 & u_2(1 - v_2) & (1 - u_2)v_2 & (1 - u_2)(1 - v_2) \end{pmatrix}.$$

That is, there exists a stationary distribution  $\pi$  (a row vector) such that  $\pi = \pi M$ ; note that  $\pi$  depends on  $u$  and  $v$  but not on the initial profile. In [4,5,23], it was proven that stationary distribution  $\pi : \tilde{U} \times \tilde{V} \rightarrow \mathbb{R}^4$  allows the following representation:

$$\pi(\cdot) = \left( \mathbf{s}_1(\cdot)\mathbf{s}_2(\cdot), \mathbf{s}_1(\cdot)(1 - \mathbf{s}_2(\cdot)), (1 - \mathbf{s}_1(\cdot))\mathbf{s}_2(\cdot), (1 - \mathbf{s}_1(\cdot))(1 - \mathbf{s}_2(\cdot)) \right)$$

where mappings  $\mathbf{s}_1 : \tilde{U} \times \tilde{V} \rightarrow [0, 1]$ ,  $\mathbf{s}_2 : \tilde{U} \times \tilde{V} \rightarrow [0, 1]$  are defined by the rules

$$\mathbf{s}_1(u, v) \triangleq \frac{v_2(u_1 - u_2) + u_2}{1 - (u_1 - u_2)(v_1 - v_2)}, \quad \mathbf{s}_2(u, v) \triangleq \frac{u_2(v_1 - v_2) + v_2}{1 - (u_1 - u_2)(v_1 - v_2)}. \tag{4}$$

Note that the trajectories of actions for profiles of deterministic RSs contain a deterministic cycles (hence the name of the strategies) of action profiles that may depend on the opening moves.

In what follows, we omit arguments of  $\mathbf{s}_1, \mathbf{s}_2$  where it is possible. We also refer to  $(\mathbf{s}_1, \mathbf{s}_2)$  as a *stationary distribution*, since the original stationary distribution  $\pi$  can be easily recovered.

For a profile of SRSs, the limit-of-means payoffs are determined by the stationary distribution that is independent from the initial actions. Hence, for a profile  $(u, v)$  of SRSs, players' payoffs in  $\mathbf{G}^\infty$  are the exact same payoffs as in  $\mathbf{G}$  if player 1 uses mixed strategy  $\mathbf{s}_1(u, v)$  and player 2 uses mixed strategy  $\mathbf{s}_2(u, v)$ . This helps to specify  $J_1^\infty$  and  $J_2^\infty : \forall (u, v) \in \tilde{U} \times \tilde{V}$  that denote the first and second player's payoff function in  $\mathbf{G}^\infty$ , respectively:

$$J_1^\infty(u, v) = J_1(\mathbf{s}_1(u, v), \mathbf{s}_2(u, v)), \quad J_2^\infty(u, v) = J_2(\mathbf{s}_1(u, v), \mathbf{s}_2(u, v)). \tag{5}$$

See the summary for all introduced games in Table 1.

**Table 1.** Summary for two-player games considered in the paper.

Game	Setting	Strategies	Payoffs	Description
$\mathbf{G}$	One-shot	Mixed strategies ([0, 1], [0, 1])	$J_1, J_2$	A one-shot $2 \times 2$ game
$\mathbf{G}_{un}^\infty$	Repeated	Unconditional RSs	$J_1^\infty, J_2^\infty$	the memory-less play of $\mathbf{G}$ that is infinitely repeated; $\mathbf{G}_{un}^\infty$ is 'equivalent' to $\mathbf{G}$ but formalised as repeated interaction.
$\mathbf{G}^\infty$	Repeated	Stochastic and unconditional RSs ( $\tilde{U}, \tilde{V}$ )	$J_1^\infty, J_2^\infty$	The repeated modification of $\mathbf{G}$ where probabilities of actions can be conditioned by the preceding opponent's action. In addition to memory-less strategies from $\mathbf{G}_{un}^\infty$ , players get conditional ones.

### 1.3.3. Equilibria

To summarize, we study a game  $\mathbf{G}^\infty$  with two players; the sets of strategies are  $\tilde{U}$  and  $\tilde{V}$ ; the payoff functions are  $J_1^\infty$  and  $J_2^\infty$ .

**Definition 2.** A profile  $(\hat{u}, \hat{v}) \in U \times V$  is called a Nash equilibrium (NE) in  $\mathbf{G}^\infty$  if the following conditions hold:  $\forall (u, v) \in \tilde{U} \times \tilde{V}$

$$J_1^\infty(u, \hat{v}) \leq J_1^\infty(\hat{u}, \hat{v}) \quad \text{and} \quad J_2^\infty(\hat{u}, v) \leq J_2^\infty(\hat{u}, \hat{v}). \tag{6}$$

In what follows, a pure (completely mixed) Nash equilibrium stands for a Nash equilibrium in a one-shot game  $\mathbf{G}$  formed by pure (completely mixed) strategies. In  $\mathbf{G}^\infty$ , the stationary distribution generated by an NE is called an *Equilibrium Stationary Distribution* (ESD).

## 2. Characterization of Nash Equilibria

### 2.1. Geometric Intuition and Attainable Sets

Our immediate goal is to illustrate for  $G^\infty$  the difference between conditional and unconditional SRSs. First, let us introduce the concept of attainable sets.

**Definition 3.** *The set of all stationary distributions feasible for a fixed opponent’s reactive strategy is called an attainable set for a player. To be precise,  $\forall u \in U, v \in V$*

$$\mathcal{S}_I(v) \triangleq \{(\mathbf{s}_1(\bar{u}, v), \mathbf{s}_2(\bar{u}, v)) : \bar{u} \in U\}, \mathcal{S}_{II}(u) \triangleq \{(\mathbf{s}_1(u, \bar{v}), \mathbf{s}_2(u, \bar{v})) : \bar{v} \in V\}.$$

Here,  $\mathcal{S}_I(v)$  is an attainable set for player 1 if player 2 chooses strategy  $v$ ; see Figure 2b. Similarly,  $\mathcal{S}_{II}(u)$  is an attainable set for player 2 if player 1 chooses  $u$ ; see Figure 2a. Obviously,  $\{(\mathbf{s}_1(u, v), \mathbf{s}_2(u, v))\} = \mathcal{S}_I(v) \cap \mathcal{S}_{II}(u)$ ; see Figure 2c. Using (4), we obtain

$$\mathbf{s}_1(u, v) = u_2 + \mathbf{s}_2(u, v)(u_1 - u_2), \quad \mathbf{s}_2(u, v) = v_2 + \mathbf{s}_1(u, v)(v_1 - v_2). \tag{7}$$

**Lemma 1.** *The following representations hold true  $\forall u \in U \forall v \in V$*

$$\begin{aligned} \mathcal{S}_I(v) &= \{(s_1, s_1(v_1 - v_2) + v_2) : s_1 \in ]0, 1[ \}, \\ \mathcal{S}_{II}(u) &= \{(s_2(u_1 - u_2) + u_2, s_2) : s_2 \in ]0, 1[ \}. \end{aligned} \tag{8}$$

**Proof.** We will check only the first representation, as, using similar arguments, it easy to demonstrate the second one. Fix an arbitrary  $v \in V$ ; let

$$\Omega \triangleq \{(s_1, s_1(v_1 - v_2) + v_2) : s_1 \in ]0, 1[ \}.$$

Take  $\acute{s}_1 \in ]0, 1[$ ; then  $(\acute{s}_1, \acute{s}_1(v_1 - v_2) + v_2) \in \Omega$ . Let  $\acute{u}$  be the reactive strategy of player 1 such that  $\acute{u}_1 = \acute{u}_2 = \acute{s}_1$ . It follows from (7) that  $\mathbf{s}_1(\acute{u}, v) = \acute{s}_1$  and  $\mathbf{s}_2(\acute{u}, v) = \acute{s}_1(v_1 - v_2) + v_2$ . Thus,  $(\mathbf{s}_1(\acute{u}, v), \mathbf{s}_2(\acute{u}, v)) = (\acute{s}_1, \acute{s}_1(v_1 - v_2) + v_2)$ . Then  $(\acute{s}_1, \acute{s}_1(v_1 - v_2) + v_2) \in \mathcal{S}_I(v)$ ; therefore,

$$\Omega \subset \mathcal{S}_I(v). \tag{9}$$

Fix  $\grave{u} \in U$ ; then  $\mathbf{s}_1(\grave{u}, v) \in ]0, 1[$ . It follows from (7) that  $\mathbf{s}_2(\grave{u}, v) = \mathbf{s}_1(\grave{u}, v)(v_1 - v_2) + v_2$ . Then  $(\mathbf{s}_1(\grave{u}, v), \mathbf{s}_2(\grave{u}, v)) \in \Omega$ ; thus,  $\mathcal{S}_I(v) \subset \Omega$ . Combining this with (9), we obtain  $\Omega = \mathcal{S}_I(v)$ .  $\square$

**Remark 1.** *Equation (7) implies that for player 1 to achieve any element of  $\mathcal{S}_I(v) \forall v \in V$ , say  $(s_1^*, s_2^*) \in \mathcal{S}_I(v)$ , she may use the corresponding unconditional RS  $(s_1^*, s_1^*)$ ; for player 2, the similar statement holds true. Figure 3a shows us the corresponding example.*

We now see the key difference between unconditional and conditional SRSs. For a memory-less strategy of the first player (i.e.,  $u_1 = u_2$ ), the attainable set of the second player is a horizontal line. Similarly, for a memory-less strategy of the second player (i.e.,  $v_1 = v_2$ ), the attainable set of player 1 is a vertical line; this is similar for mixed strategies in one-shot games. Hence, memory-less strategies only allow intercepts to be chosen. By contrast, conditional SRSs (i.e.,  $u_1 \neq u_2$  and  $v_1 \neq v_2$ ) also allow slopes of attainable sets to be set. Thus, conditional SRSs have more flexibility in adjusting slopes of attainable sets to gradients of payoff functions at the stationary distributions (the points where players’ attainable sets intersect). This flexibility provides significantly more opportunities for equilibria to emerge (compared to memory-less mixed strategies).

**Proposition 1.** *For a one-shot game  $G$ , any NE profile  $(u, v)$  in mixed strategies becomes the corresponding NE profile  $((u, u), (v, v))$  in both  $G^\infty$  and  $G_{un}^\infty$ .*

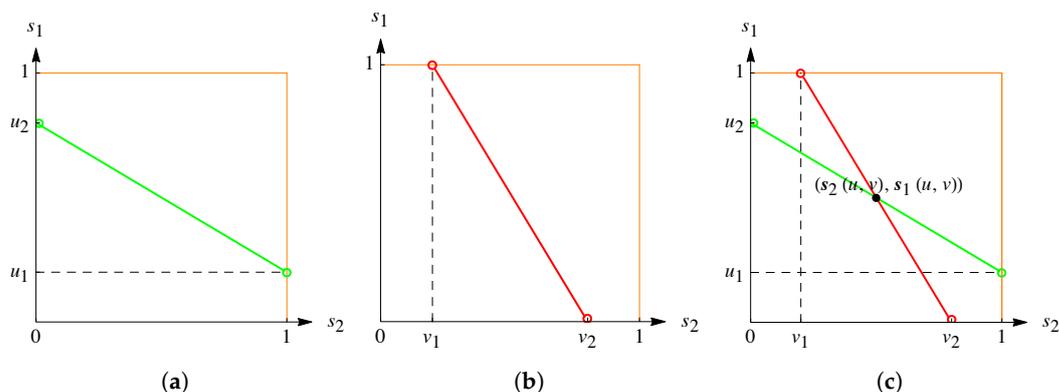
**Proof.** Fix the memory-less RS  $(u, u)$  of player 1 corresponding to an equilibrium  $(u, v)$  in a one-shot game  $G$ . We will show that  $(v, v)$  is a best response to  $(u, u)$  in  $G^\infty$ . By Remark 1, for player 2, deviations in unconditional strategies are “operationally equivalent” to deviations in conditional strategies – that is, if player 2 can deviate, then he can deviate with a strategy of the form  $(\tilde{v}, \tilde{v})$ . If player 2 picks an unconditional RS  $(\tilde{v}, \tilde{v}), \tilde{v} \in [0, 1]$ , we arrive at the stationary distribution  $(u, \tilde{v})$ . Combining the last fact with (5), we obtain that

$$J_2^\infty((u, u), (\tilde{v}, \tilde{v})) = J_2(u, \tilde{v}) \quad \forall \tilde{v} \in [0, 1]. \tag{10}$$

From the definition of NE in mixed strategies in  $G$ , it follows that  $J_2(u, \cdot)$ , as a function defined on  $[0, 1]$ , attains its maximum at  $v$ . By (10) and the above argument on the operational equivalence,

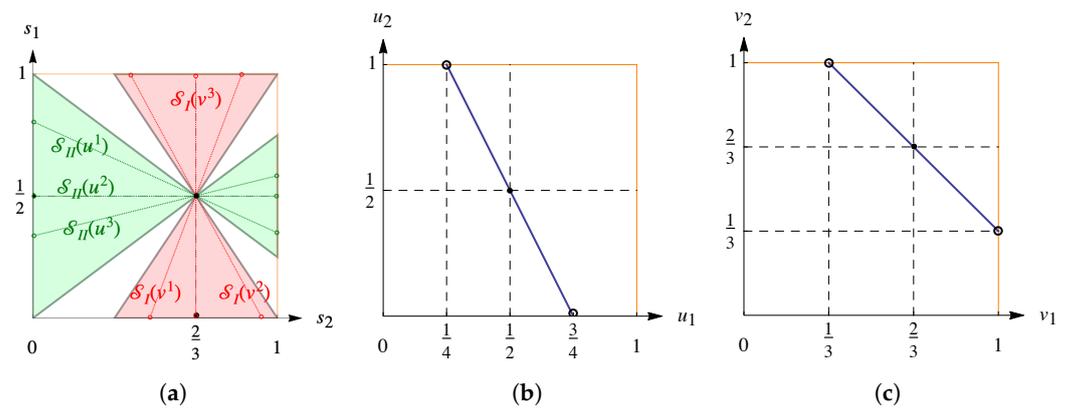
$$\max_{\tilde{v} \in [0, 1] \times [0, 1]} J_2^\infty((u, u), \tilde{v}) = \max_{\tilde{v} \in [0, 1]} J_2^\infty((u, u), (\tilde{v}, \tilde{v})) = \max_{\tilde{v} \in [0, 1]} J_2(u, \tilde{v}) = J_2^\infty((u, u), (v, v))$$

Thus, unconditional RS  $(v, v)$ , which emerges from the stage game NE, “remains” the best reply to  $(u, u)$  in  $G^\infty$  (and hence in  $G_{un}^\infty$ ). Symmetrically,  $(u, u)$  is a best response to  $(v, v)$ , and hence the profile is an NE in  $G^\infty$  (hence, in  $G_{un}^\infty$ ).  $\square$



**Figure 2.** Let  $u = v = (0.2, 0.8)$ . Then, plot (a) corresponds to attainable set  $S_{II}(u)$ ; plot (b) corresponds to  $S_I(v)$ . Plot (c) depicts the corresponding stationary distribution  $(s_1(u, v), s_2(u, v)) = (1/2, 1/2)$  as the intersection of the attainable sets.

**Remark 2.** Note that players are free to move within their attainable sets from one point (stationary distribution) to another in a variety of ways. Moreover, for a stationary distribution within an attainable set of one player, there are infinitely many strategies leading to this stationary distribution for the given opponent’s strategy. In turn, the choice among these strategies leads to distinct attainable sets for the opponent. For example, Figure 3 shows the set of all possible strategy profiles leading to the stationary distribution  $(s_1, s_2) = (1/2, 2/3)$ . Figure 3a depicts all possible slopes of players’ attainable sets. Note that the lines circumscribing the shaded regions have simple analytic representations. This allows us to infer all possible players’ strategies generating this stationary distribution. Namely, any pair of  $u$  from Figure 3c and  $v$  from Figure 3c leads to the same stationary distribution  $(1/2, 2/3)$ . All  $(u, v)$  such that the stationary distribution equals  $(s_1, s_2) = (1/2, 2/3)$  induce the attainable sets inside the shaded regions.



**Figure 3.** For stationary distribution  $(s_1, s_2) = (\frac{1}{2}, \frac{2}{3})$ , plot (a) shows the range of all possible attainable sets generating it. Namely, the green region (the red region) is a ‘combination’ of all possible attainable sets of player 2 (player 1), which are lines similar to Figure 2a (to Figure 2b). Notice that all profiles  $(u^i, v^j)$  result in the same stationary distribution but different attainable sets  $S_I(v^j), S_{II}(u^i)$ ; here  $u^2 = (\frac{1}{2}, \frac{1}{2}), v^3 = (\frac{2}{3}, \frac{2}{3})$  are unconditional RSs. Plots (b,c) depict all SRSs of players that joined in profiles generate this stationary distribution

2.2. Prisoner’s Dilemma with Equal Gains from Switching

We consider  $G^{es}$ , a Prisoner’s Dilemma with equal gains from switching (see [4]),

Game $G^{es}$	I	II
I	3,3	1,4
II	4,1	2,2

Equal gains from switching implies that  $a_1 = a_2 = 0$ ; clearly, we have  $b_1 = b_2 = -1, c_1 = c_2 = 2$ . For the repeated version of  $G^{es}$ , we have (see (1) and (5))

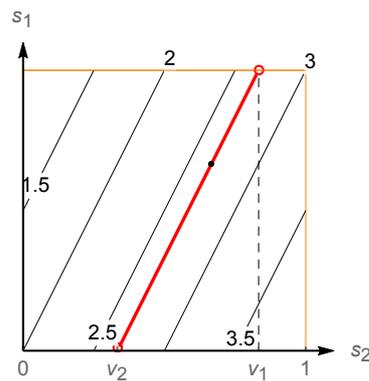
$$J_1^\infty(u, v) \triangleq -s_1(u, v) + 2s_2(u, v) + 2, \quad J_2^\infty(u, v) \triangleq -s_2(u, v) + 2s_1(u, v) + 2.$$

Lemma A2 tells us that every pair of well-defined SRSs such that  $u^* = (u_1^*, u_2^*) = (u_2^* - \frac{b_1}{c_1}, u_2^*)$  and  $v^* = (v_1^*, v_2^*) = (v_2^* - \frac{b_1}{c_1}, v_2^*)$  forms a Nash equilibrium. Figure 4 illustrates that for player 1, curves-of-equal-payoff coincide with attainable sets if player 2 uses the above equilibrium strategies, meaning that equilibrium strategies fix the opponent’s payoff:

$$J_1^\infty(u, v^*) = 2v_2^* + 2 \forall u \in U, \quad J_2^\infty(u^*, v) = 2u_2^* + 2 \forall v \in V.$$

The reactive strategy  $(1, 1 + \frac{b_1}{c_1})$  is known as *generous tit-for-tat* [24]; it leads to the most cooperative symmetric equilibrium.

Note that an equilibrium can only be sustained if both players choose strategies  $u$  and  $v$  such that  $u_1 > u_2$  and  $v_1 > v_2$ . Value  $v_1 - v_2$  directly encourages player 1 to cooperate. Specifically, if  $v_1 - v_2 > -\frac{b}{c} = 1/2$ , then player 1 has an incentive to climb the attainable set by choosing more cooperative strategies increasing  $s_1$  and  $s_2$ . If  $v_1 - v_2 < 1/2$ , then player 1 has an incentive to step down the attainable set; this decreases  $s_1$  and  $s_2$ . Clearly, memory-less SRSs can not support Nash equilibria. For example, the attainable sets of player 1 become vertical lines, giving her clear incentives to defect since her payoff increases moving from top to bottom of the attainable sets. Note that, generally, curves-of-equal-payoff are non-linear; see the example in Appendix C.1.



**Figure 4.** Thin lines correspond to curves-of-equal-payoff for player 1. As before, the red line corresponds to the attainable set  $\mathcal{S}_I(v)$  of player 1 for  $v = (\frac{5}{6}, \frac{1}{3})$ . Here,  $\mathcal{S}_I(v)$  coincides with the curve of payoffs equal to  $8/3$ . Strategies  $u = v = (\frac{5}{6}, \frac{1}{3})$  form a Nash equilibrium with the ESD  $(\frac{2}{3}, \frac{2}{3})$ .

### 2.3. Characterization of Nash Equilibria in $\mathbf{G}^\infty$

Since every NE is formed by strategies that are best responses to each other, we proceed further with a standard analysis using derivatives. Let  $f_x$  stand for the derivative of a function  $f(\cdot)$  w.r.t.  $x$ .

**Lemma 2** (The necessary condition). *If  $\hat{u} = (\hat{u}_1, \hat{u}_2) \in U, \hat{v} = (\hat{v}_1, \hat{v}_2) \in V$  form a Nash equilibrium and  $(\hat{s}_1, \hat{s}_2)$  is the corresponding ESD, then the following holds true:*

$$\begin{aligned} a_1(2(\hat{v}_1 - \hat{v}_2)\hat{s}_1 + \hat{v}_2) + b_1 + c_1(\hat{v}_1 - \hat{v}_2) &= 0 \quad , \quad 2a_1(\hat{v}_1 - \hat{v}_2) \leq 0, \\ a_2(2(\hat{u}_1 - \hat{u}_2)\hat{s}_2 + \hat{u}_2) + b_2 + c_2(\hat{u}_1 - \hat{u}_2) &= 0 \quad , \quad 2a_2(\hat{u}_1 - \hat{u}_2) \leq 0. \end{aligned}$$

**Proof.** Assume that  $(\hat{u}, \hat{v})$  is a Nash equilibrium with the corresponding ESD  $(\hat{s}_1, \hat{s}_2)$ . Combining (5), (6), and (8), we have

$$\begin{aligned} \max_{u \in U} J_1^\infty(u, \hat{v}) &= \max_{u \in U} J_1(s_1(u, \hat{v}), s_2(u, \hat{v})) = \max_{(s_1, s_2) \in \mathcal{S}_I(\hat{v})} J_1(s_1, s_2) = \\ &= \max_{s_1 \in ]0, 1[} J_1(s_1, s_1(\hat{v}_1 - \hat{v}_2) + \hat{v}_2) = J_1(\hat{s}_1, \hat{s}_2), \end{aligned} \tag{11}$$

$$\begin{aligned} \max_{v \in V} J_2^\infty(\hat{u}, v) &= \max_{v \in V} J_2(s_1(\hat{u}, v), s_2(\hat{u}, v)) = \max_{(s_1, s_2) \in \mathcal{S}_{II}(\hat{u})} J_2(s_1, s_2) = \\ &= \max_{s_2 \in ]0, 1[} J_2(s_2(\hat{u}_1 - \hat{u}_2) + \hat{u}_2, s_2) = J_2(\hat{s}_1, \hat{s}_2). \end{aligned} \tag{12}$$

It follows from (1)–(3) that

$$J_1(s_1, s_1(\hat{v}_1 - \hat{v}_2) + \hat{v}_2) = a_1(\hat{v}_1 - \hat{v}_2)s_1^2 + s_1(b_1 + c_1(\hat{v}_1 - \hat{v}_2) + a_1\hat{v}_2) + c_1\hat{v}_2 + A_{B,R},$$

$$J_2(s_2(\hat{u}_1 - \hat{u}_2) + \hat{u}_2, s_2) = a_2(\hat{u}_1 - \hat{u}_2)s_2^2 + s_2(b_2 + c_2(\hat{u}_1 - \hat{u}_2) + a_2\hat{u}_2) + c_2\hat{u}_2 + B_{B,R}.$$

Define the functions  $J_1^*[\hat{v}], J_2^*[\hat{u}] : [0, 1] \rightarrow \mathbb{R}$  by the rules:  $\forall s_1, s_2 \in [0, 1]$

$$J_1^*[\hat{v}](s_1) \triangleq a_1(\hat{v}_1 - \hat{v}_2)s_1^2 + s_1(b_1 + c_1(\hat{v}_1 - \hat{v}_2) + a_1\hat{v}_2) + c_1\hat{v}_2 + A_{B,R}, \tag{13}$$

$$J_2^*[\hat{u}](s_2) \triangleq a_2(\hat{u}_1 - \hat{u}_2)s_2^2 + s_2(b_2 + c_2(\hat{u}_1 - \hat{u}_2) + a_2\hat{u}_2) + c_2\hat{u}_2 + B_{B,R}.$$

We now obtain the first and second derivatives of these functions:

$$J_1^*[\hat{v}]_{s_1} = a_1(2(\hat{v}_1 - \hat{v}_2)s_1 + \hat{v}_2) + b_1 + c_1(\hat{v}_1 - \hat{v}_2), \tag{14}$$

$$J_1^*[\hat{v}]_{s_1, s_1} = 2a_1(\hat{v}_1 - \hat{v}_2), \tag{15}$$

$$J_2^*[\hat{u}]_{s_2} = a_2(2(\hat{u}_1 - \hat{u}_2)s_2 + \hat{u}_2) + b_2 + c_2(\hat{u}_1 - \hat{u}_2), \tag{16}$$

$$J_2^*[\hat{u}]_{s_2, s_2} = 2a_2(\hat{u}_1 - \hat{u}_2). \tag{17}$$

Using (11)–(17) and the second derivative test, we complete the proof.  $\square$

**Theorem 1** (Characterization of all Nash equilibria). *The set of all Nash equilibria coincides with the set of all  $\hat{u}, \hat{v} \in \mathbb{R}^2$  solving the following system:*

$$\begin{cases} a_1(2(\hat{v}_1 - \hat{v}_2)\hat{s}_1 + \hat{v}_2) + b_1 + c_1(\hat{v}_1 - \hat{v}_2) = 0 & \text{(i)} \\ a_2(2(\hat{u}_1 - \hat{u}_2)\hat{s}_2 + \hat{u}_2) + b_2 + c_2(\hat{u}_1 - \hat{u}_2) = 0 & \text{(ii)} \\ a_1(\hat{v}_1 - \hat{v}_2) \leq 0, \quad a_2(\hat{u}_1 - \hat{u}_2) \leq 0 & \text{(iii)} \\ \hat{s}_1 = \mathbf{s}_1(\hat{u}, \hat{v}), \quad \hat{s}_2 = \mathbf{s}_2(\hat{u}, \hat{v}), & \text{(iv)} \\ 0 < \hat{u}_1, \hat{u}_2, \hat{v}_1, \hat{v}_2 < 1 & \text{(v)} \end{cases}. \tag{18}$$

**Proof.** Let  $\mathbf{N}$  denote the set of all Nash equilibria in  $\mathbf{G}^\infty$ ;  $\mathbf{S}$  stands for the set of all  $(\hat{u}, \hat{v})$  solving (18). Let us prove that  $\mathbf{S} = \mathbf{N}$ .

Take  $(\hat{u}, \hat{v}) \in \mathbf{S}$ . As (v) in (18) holds,  $(\hat{u}, \hat{v})$  is a pair of reactive strategies;  $\mathbf{S} \subset U \times V$ . Obviously, (iv) ensures that  $(\hat{s}_1, \hat{s}_2)$  is the corresponding stationary distribution;  $(\hat{s}_1, \hat{s}_2) = (\mathbf{s}_1(\hat{u}, \hat{v}), \mathbf{s}_2(\hat{u}, \hat{v}))$ . As (iii) in (18) holds true, we have three cases:

1.  $a_1(\hat{v}_1 - \hat{v}_2) < 0$  and  $a_2(\hat{u}_1 - \hat{u}_2) < 0$ ,
2.  $a_1(\hat{v}_1 - \hat{v}_2) = 0$  and  $a_2(\hat{u}_1 - \hat{u}_2) < 0$  (or, symmetrically,  $a_1(\hat{v}_1 - \hat{v}_2) < 0$  and  $a_2(\hat{u}_1 - \hat{u}_2) = 0$ ),
3.  $a_1(\hat{v}_1 - \hat{v}_2) = 0$  and  $a_2(\hat{u}_1 - \hat{u}_2) = 0$ .

All cases are composed from similar elements: equalities and strong inequalities. We consider only two of these elements. The analysis for the others is similar. Suppose  $a_1(\hat{v}_1 - \hat{v}_2) < 0$ . Then  $J_1^*[\hat{v}](\cdot)$  is a concave quadratic function (see (13)). Conditions (i) and (iii) ensure that  $\hat{s}_1$  is a local maximum of  $J_1^*[\hat{v}](\cdot)$ ; see (14), (15). As local maxima of concave functions are also global maxima, we have

$$\begin{aligned} \max_{\bar{s}_1 \in ]0,1[} J_1(\bar{s}_1, \bar{s}_1(\hat{v}_1 - \hat{v}_2) + \hat{v}_2) &= \max_{\bar{s}_1 \in ]0,1[} J_1^*[\hat{v}](\bar{s}_1) = J_1^*[\hat{v}](\hat{s}_1) = \\ J_1(\hat{s}_1, \hat{s}_1(\hat{v}_1 - \hat{v}_2) + \hat{v}_2) &= J_1(\hat{s}_1, \hat{s}_2) = \max_{\bar{u} \in U} J_1^\infty(\mathbf{s}_1(\bar{u}, \hat{v}), \mathbf{s}_2(\bar{u}, \hat{v})). \end{aligned}$$

Assume  $a_1(\hat{v}_1 - \hat{v}_2) = 0$ . Thus,  $a_1 = 0$  and/or  $\hat{v}_1 = \hat{v}_2$ . If  $\hat{v}_1 = \hat{v}_2$ , then player 2 uses the repeated version of her mixed Nash equilibrium strategy. If  $a_1 = 0$ , then it follows from (13) and (i) in (18) that  $\forall s_1 \in [0, 1] \quad J_1^*[\hat{v}](s_1) = c_1\hat{v}_2 + A_{B,R}$ . We have already observed this effect in the example of Section 2.2. Thus, in all cases, player 1 has a payoff independent of her strategy.

Generalizing the last two paragraphs, from (11) and (12), we obtain that

$$\begin{aligned} \max_{\bar{u} \in U} J_1^\infty(\mathbf{s}_1(\bar{u}, \hat{v}), \mathbf{s}_2(\bar{u}, \hat{v})) &= J_1(\hat{s}_1, \hat{s}_2) = J_1^\infty(\mathbf{s}_1(\hat{u}, \hat{v}), \mathbf{s}_2(\hat{u}, \hat{v})), \\ \max_{\bar{v} \in V} J_2^\infty(\mathbf{s}_1(\hat{u}, \bar{v}), \mathbf{s}_2(\hat{u}, \bar{v})) &= J_2(\hat{s}_1, \hat{s}_2) = J_2^\infty(\mathbf{s}_1(\hat{u}, \hat{v}), \mathbf{s}_2(\hat{u}, \hat{v})). \end{aligned}$$

Consequently,  $\hat{u}$  and  $\hat{v}$  are the best responses to each other, meaning that

$$\mathbf{S} \subset \mathbf{N}. \tag{19}$$

Take  $(u, v) \in \mathbf{N}$ . Trivially, (v) holds true. Taking into account (4) and Lemma 2, we see that (i)–(iv) are fulfilled. Thus,  $\mathbf{N} \subset \mathbf{S}$ ; from (19), we have  $\mathbf{S} = \mathbf{N}$ .  $\square$

#### 2.4. Existence of Nash Equilibria in Symmetric Games

In this section, we obtain conditions for the existence of NE when players are identical (one-shot games are symmetric). We fix a symmetric game  $\mathbf{G}$

Game G	I	II
I	$A_1, A_1$	$A_2, A_3$
II	$A_3, A_2$	$A_4, A_4$

Recall that for G we have the following versions of (2) and (3):

$$a_1 = A_1 - A_2 - A_3 + A_4, b_1 = A_2 - A_4, c_1 = A_3 - A_4;$$

$$a_2 = A_1 - A_2 - A_3 + A_4, b_2 = A_2 - A_4, c_2 = A_3 - A_4.$$

This allows us to introduce the following convenient definitions:

$$a \triangleq a_1 = a_2, b \triangleq b_1 = b_2, c \triangleq c_1 = c_2. \tag{20}$$

If  $a \neq 0$ , then

$$\mathbf{b} \triangleq -\frac{b}{a}, \mathbf{c} \triangleq -\frac{c}{a}. \tag{21}$$

To reach the goal of the section, we must derive conditions on  $a, b$ , and  $c$  that ensure the existence of a solution of (18). Taking into account inequality (iii) in (18), in Appendix A, we study three possible cases:  $a = 0, a < 0, a > 0$ . The next theorem combines Lemmas A1, A7 and A10 (see Appendices A.1, A.2 and A.3, respectively).

**Theorem 2.** *For repeated symmetric games, there exists a reactive Nash equilibrium iff one of the following conditions holds true:*

1.  $a = b = c = 0$ ;
2.  $a = 0$  &  $-1 < -\frac{b}{c} < 1$ ;
3.  $a \neq 0$  &  $0 < \mathbf{b} < 1$ ;
4.  $a < 0$  &  $\mathbf{c} < 2 - \mathbf{b}$  &  $\mathbf{b} \geq 1$ ;
5.  $a < 0$  &  $\mathbf{c} > -\mathbf{b}$  &  $\mathbf{b} \leq 0$ ;
6.  $a > 0$  &  $\mathbf{c} < \mathbf{b}$  &  $\mathbf{b} \leq 0$ ;
7.  $a > 0$  &  $\mathbf{c} > \mathbf{b}$  &  $\mathbf{b} \geq 1$ .

In this paper, we focus on symmetric stage games due to tractability of the corresponding results. Appendix C.2 shows an example of a non-symmetric stage game.

### 2.5. If All RSs Are Available

In this paper, players' strategy sets do not include conditional deterministic and semi-deterministic RSs, and we focus only on equilibria generated by SRSs. Thus, in  $G^\infty$ , the set of all possible profiles was  $\tilde{U} \times \tilde{V}$  instead of  $[0, 1]^2 \times [0, 1]^2$ , the set of all profiles in RSs. Let us discuss how this decision influences the NE profiles in SRSs.

Recall that for  $G^\infty$ , the premises of the Kakutani fixed-point theorem do not hold, as the strategy set is not compact, although payoff functions are continuous. Thus, the existence of an NE in an arbitrary game  $G^\infty$  is not guaranteed. By simply allowing all RSs, the strategy set becomes compact, but the payoff functions become dependent on initial action profiles, making payoff functions no longer continuous and properly defined. The last fact is due to payoff-relevance of the initial round for some strategy profiles. Thus, one may decide to complement strategies with one extra variable, an initial action, leading to more complicated analysis. There are two approaches that can be used to avoid the complication. The first approach, the one used for discounted payoffs, is to assume that nature draws initial actions. In this case, one may consider only equilibria that hold independently of initial moves (see [17,18]); actual equilibrium payoffs may depend on the initial actions. The second approach, which is possible for the-limit-of-means setting and is used in this article, is to restrict the strategy set such that players' payoffs are independent of initial actions. Both approaches ensure that any equilibrium holds for all opening moves.

Starting with the restricted strategy sets, one may ask, what if we extend the set of strategies to include all RSs? How would it influence the set of all equilibria? One straightforward aspect is that new equilibria may form. The second (less obvious) aspect is that we increase players’ abilities to deviate; this may erase some existing equilibria.

Regarding the first aspect, the precise answer requires a thorough analysis that can not be included in this paper since, as we mentioned above, the results will depend on initial actions. More importantly, the value of such results is not obvious.

Regarding the second aspect, the characterization of all Nash equilibria generated by SRSs remains valid independently of initial moves, even if one considers the setting where all RSs are available for players. Let us show this. Consider player 1. If player 2 picks an SRS  $\hat{v}$ , then the corresponding stationary distribution is well defined by (4) for all RSs of player 1 and does not depend on initial actions. For player 1, in terms of payoff-relevant deviations, the set of RSs is equivalent to  $\tilde{U}$ .

### 3. Equilibrium Payoffs in Conditional SRSs

In this section, we compare equilibrium payoffs in conditional SRSs versus unconditional RSs in the following simple way. First, we fix a stage game  $G$  such that Theorem 2 gives us an NE in SRSs. The repeated version  $G_{inf}^\infty$  inherits all equilibrium payoff profiles from  $G$  that are generated by the corresponding unconditional RSs. Among these payoff profiles, we select the symmetric one with the lowest payoffs as the benchmark. We then compare the best symmetric NE payoff profile in  $G^\infty$  formed by SRSs against this benchmark. Since we have existence results only for symmetric games, the complete analysis is only feasible for this special case. Nevertheless, the existence of an NE in memory-less SRSs implies the existence of an NE in conditional SRSs that allows us to compare the corresponding payoffs even for non-symmetric games. The next subsection presents this partial result; afterwards, we complete analysis for symmetric games.

#### 3.1. Payoffs for NE Profiles of Unconditional and Conditional SRSs

From (4) and Theorem 1, it follows that any  $(u, v) \in \mathbb{R}^2$  admitting solutions of the system

$$\begin{cases} y = v_1 - v_2, & a_1(s_1y + s_2) + b_1 + c_1y = 0, & (i) \\ x = u_1 - u_2, & a_2(s_1 + s_2x) + b_2 + c_2x = 0, & (ii) \\ s_1 = \frac{v_2x + u_2}{1 - xy}, & s_2 = \frac{u_2y + v_2}{1 - xy}, & 0 \geq a_2x, 0 \geq a_1y, \\ 0 < u_1, u_2, v_1, v_2 < 1, & -1 < x, y < 1 \end{cases} \quad (22)$$

forms a NE. Note that (22) is a system of six equations with eight variables. We consider  $(x, y)$  or  $(s_1, s_2)$  as free parameters to express the remaining variables;  $x$  and  $y$  are introduced to simplify the notation. We omit the rather simple case  $a_1a_2 = 0$  and start solving the system under the assumption that

$$a_1a_2 \neq 0.$$

Suppose  $-1 < x, y < 1$ ; let us express  $u, v, s_1, s_2$  from (22) in terms of  $x$  and  $y$

$$\begin{cases} u_2 = \frac{2a_2x(b_1 + c_1y) - a_1(b_2 + c_2x)(1 + xy)}{a_1a_2(1 - xy)}, & s_1 = \frac{a_1(b_2 + c_2x) - a_2x(b_1 + c_1y)}{a_1a_2(-1 + xy)}, \\ v_2 = \frac{2a_1y(b_2 + c_2x) - a_2(b_1 + c_1y)(1 + xy)}{a_1a_2(1 - xy)}, & s_2 = \frac{-a_1(b_2 + c_2x)y + a_2(b_1 + c_1y)}{a_1a_2(-1 + xy)}, \\ u_1 = u_2 + x, & v_1 = v_2 + y, & 0 < u_1, u_2, v_1, v_2 < 1, \\ 0 \geq a_2x, & 0 \geq a_1y. \end{cases} \quad (23)$$

Assume that a completely mixed NE exists in a one-shot game  $\mathbf{G}$ . Then  $(-\frac{b_2}{a_2}, -\frac{b_1}{a_1})$  is the corresponding NE profile since  $a_1 a_2 \neq 0$ . Furthermore,  $u = (-\frac{b_2}{a_2}, -\frac{b_2}{a_2})$  and  $v = (-\frac{b_1}{a_1}, -\frac{b_1}{a_1})$  form an NE in  $\mathbf{G}^\infty$  by Proposition 1. Indeed (see (5)),  $\forall \tilde{u} \in \tilde{U} \forall \tilde{v} \in \tilde{V}$

$$J_1^\infty(\tilde{u}, v) = a_1 s_1(\tilde{u}, v) \left(-\frac{b_1}{a_1}\right) + b_1 s_1(\tilde{u}, v) - \frac{c_1 b_1}{a_1} + A_{B,R} = A_{B,R} - \frac{c_1 b_1}{a_1} = J_1^\infty(u, v), \quad (24)$$

$$J_2^\infty(u, \tilde{v}) = a_2 \left(-\frac{b_2}{a_2}\right) s_2(u, \tilde{v}) + b_2 s_2(u, \tilde{v}) - \frac{c_2 b_2}{a_2} + B_{B,R} = B_{B,R} - \frac{c_2 b_2}{a_2} = J_2^\infty(u, v). \quad (25)$$

The next proposition shows that for both players, equilibrium payoffs in conditional SRSs are at least equilibrium payoffs in memory-less SRSs.

**Proposition 2.** *Suppose that  $a_1 a_2 \neq 0$  and there exists an NE profile  $(\hat{u}, \hat{v})$  of memory-less SRSs in  $\mathbf{G}_{un}^\infty$  (hence in  $\mathbf{G}^\infty$ ); then, for any NE profile  $(u, v)$  of conditional SRSs in  $\mathbf{G}^\infty$ , the following holds true*

$$J_1^\infty(\hat{u}, \hat{v}) - J_1^\infty(u, v) = \frac{y(a_2(c_1 + b_1 x) - a_1(b_2 + c_2 x))^2}{a_1 a_2^2 (-1 + xy)^2} \leq 0,$$

$$J_2^\infty(\hat{u}, \hat{v}) - J_2^\infty(u, v) = \frac{x(a_1(c_2 + b_2 y) - a_2(b_1 + c_1 y))^2}{a_2 a_1^2 (-1 + xy)^2} \leq 0$$

where  $x = u_1 - u_2, y = v_1 - v_2$ .

Let us outline the proof. From (1) and (5), we have  $J_i^\infty$  as a function of stationary distributions. Since we deal with ESDs, we use formulas for  $s_1$  and  $s_2$  from (23). Hence,  $J_i^\infty, i = 1, 2$ , become functions of only  $x$  and  $y$  that are defined by the NE profile  $(u, v)$ ; these equilibrium profiles exist by our assumption. Using (24) and (25), it remains only to perform some algebra and to recall that  $a_2 x \leq 0, a_1 y \leq 0$  (see the last line in (23)).

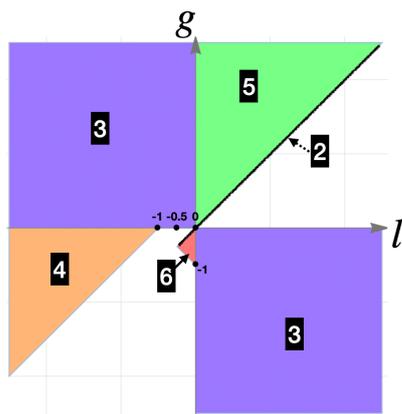
### 3.2. Symmetric Games

For symmetric games, results in Section 2.4 allow us to gain more understanding on how equilibrium payoffs in conditional SRSs compare to the benchmark payoffs in memory-less RSs. Note that a generic one-shot symmetric game with distinct payoffs on the leading diagonal is strategically equivalent to

Game $G^{10}$	C	D
C	1, 1	-l, 1 + g
D	1 + g, -l	0, 0

Generic one-shot symmetric games with identical payoffs on the leading diagonal are considered in Appendix B. Using these generic games, we study equilibrium payoffs for games satisfying the conditions of Theorem 2. For  $G^{10}$ , Figure 5 shows all pairs of  $(l, g)$  such that there exists an equilibrium according to Theorem 2. For example, the first condition results in trivial stage games that  $G^{10}$  cannot be. The second condition (region 2) restricted to Quadrant I, includes prisoner’s dilemmas with equal gains from switching. Quadrant I contains all prisoner’s dilemmas and the entire region 5 that embodies the ones where cycles  $(C, D) \rightarrow (D, C) \rightarrow (C, D) \rightarrow \dots$  are more profitable than cycles  $(C, C) \rightarrow (D, D) \rightarrow (C, C) \rightarrow \dots$ .

Region 3 completely covers Quadrant II and IV where all stage games with two pure equilibria are located; here, we have completely mixed NE for stage games and, hence, an equilibrium profile in memory-less SRSs for the repeated versions. If  $-l \neq 1 + g$ , by Proposition 2, there exists an NE profile of conditional SRSs that Pareto dominates the benchmark of NE payoffs for memory-less SRSs; else all ESDs (hence, NE payoff profiles) for conditional and unconditional SRSs coincide.



**Figure 5.** The numbered regions are derived from the conditions of Theorem 2 that ensures existence of NE for the repeated modification of generic game  $G^{10}$  for possible values of  $l$  and  $g$ .

Quadrant III contains all deadlock games, which have one dominant strategy with the equilibrium payoff being Pareto optimal. This property makes deadlock games of the least interest for theory since self-interest and mutual benefit are fully compatible. The repeated versions of games from region 4 have NE payoff profiles Pareto dominating this unconditional benchmark. Thus, in all cases considered up to this point, an equilibrium profile in conditional SRSs provides greater or equal payoffs than the benchmark equilibrium profile in memory-less RSs. The exceptions to this rule are region 2, restricted to Quadrant III, and region 6 that contain very special deadlock games: both players (by choosing dominating pure strategy) get the highest possible individual payoff. Because of this unique property, the repeated versions of these one-shot games are the only cases when all NE profile in conditional SRSs are Pareto dominated by any equilibrium profile in memory-less RSs. We provide an example of such game in Section 3.3. Finally, all one-shot games  $G^{10}$  in region 6, after we swap strategies for both players, comply with condition 7 in Theorem 2; i.e., conditions 6 and 7 cover strategically equivalent deadlock games.

Appendix B provides the similar analysis for generic symmetric games with identical payoffs on the leading diagonal, where we conclude that in the cases when payoffs of the stage game are not diverse enough, and payoffs of NE profiles in SRSs coincide with the memory-less benchmark. Otherwise, there exists an NE profile in SRSs that Pareto dominates the benchmark.

Let us summarize both cases of the generic games. Only in the very special deadlock games where both players (by choosing dominating pure strategy) get the highest possible individual payoff, payoff profiles in conditional SRSs are dominated by the benchmark unconditional payoff profiles; see the next subsection for an example. In the remaining cases, if distinct ESDs do exist, then there exists an NE profile in SRSs dominating the benchmark.

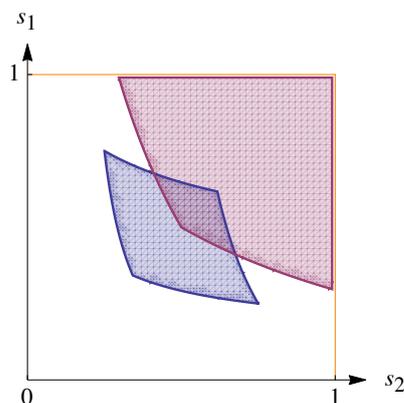
### 3.3. A Game with Pareto-Efficient Equilibrium and Dominant Strategies

The following example highlights one counterintuitive property of NE in SRSs by showing that it is possible to establish equilibria with lower payoffs by introducing a more complex class of strategies. Consider the following symmetric game  $G^{NB}$ .

Game $G^{NB}$	I	II
I	19, 19	2, 7
II	7, 2	0, 0

Once we scale all payoffs into the range  $[0, 1]$ , the stage game falls into region 6 of Figure 5. Here  $a = 10$ ,  $b = -0.2$ , and  $c = -0.7$ . According to Theorem 2, a NE in SRSs exists; we depict the set of all ESDs in Figure 6. A mixed NE in the stage game does not exist since  $b < 0$ . The profile of the memory-less RSs  $((1, 1), (1, 1))$  forms an NE and corresponds to the Pareto-efficient pure NE in the stage game. This equilibrium is excluded from our characterization as we are focusing on equilibria in SRSs. Remarkably, there

is a noticeable gap between this equilibrium payoff, serving as memory-less benchmark, and payoffs corresponding to ESDs in the repeated game. For any  $s \in ]\frac{1+\sqrt{161}}{40}, \frac{1+\sqrt{561}}{40}[$ , there exists a Nash equilibrium such that  $(s, s)$  is the ESD. For example, equilibrium strategies  $u = (\frac{4}{55}, \frac{34}{55})$  and  $v = (\frac{4}{55}, \frac{34}{55})$  generate the ESD  $(\frac{2}{5}, \frac{2}{5})$  with payoffs  $\frac{26}{5}$  (compare with 19). Thus, the example shows that it is possible to establish equilibria with lower payoffs by introducing a more complex class of strategies.



**Figure 6.** The blue set is the set of all ESDs for  $G^{NB}$ ; the red set is the set of all stationary distributions such that both players get payoffs greater than 7 (the second-highest payoff in the one-shot game).

#### 4. Conclusions

We show analytically that even an incremental change of complexity of strategies (an inch of memory) in repeated interactions has a huge influence on the sets of equilibrium payoff profiles for some classes of stage games that are circumscribed by Theorem 2. The first thing to note is that except for some trivial cases, payoff-relevant indeterminacy holds true [17]; i.e., there now exists a continuum of equilibria in “new” conditional strategies with distinct payoff profiles. Clearly, there is still no folk theorem (see the example in Appendix C.1); this still holds [17] for the next incremental improvement, 1-memory strategies. It is also now evident that one should not expect some consistency while incrementally increasing the memory of strategies. For example, in RSs, there exist equilibria with higher payoffs than the upper bound obtained independently of the discount factor in [17] (see the example in Appendix C.1) and the reverse dominance condition [17], which is important for the 1-memory case, does not influence the existence of NE in RSs.

In this paper, we focused on symmetric stage games due to tractability of the corresponding results. Appendix C.2 shows the example of a non-symmetric stage game where the set of all ESDs looks like a special combination of the corresponding sets for two symmetric stage games derived from the payoff matrix. This may be a way to study a broad class of non-symmetric stage games. An alternative approach may be to study specific classes of non-symmetric stage games, e.g., strategically zero-sum games [25].

In all examples considered so far, the region of all ESDs was a connected set. Appendix C.3 demonstrates the example with disconnected regions of equilibria. A branch of literature assumes that we may have some local dynamics on the set of all Nash equilibria, e.g., in the spirit of adaptive dynamics [4] or local cooperation patterns [26]. This example calls for new models to tackle the issue of disconnectedness that may be an obstacle on the way to higher equilibrium payoffs.

As for future work, we note that the idea of a reactive strategy was recently generalized to “reactive learning strategy” [27] allowing a player to gradually change the probability of actions depending on the past actions of the opponent. It is not known how the change from reactive strategies to reactive learning strategies would influence the equilibrium outcomes.

**Funding:** The study was supported by a grant from the Russian Science Foundation (project No. 20-71-00034). Note: all results were obtained at HSE University according to the tasks of the project (grant proposal) except for numerical simulations of sets of ESDs with Wolfram Mathematica, Appendices C.2 and C.3 that were not specified as tasks or deliverables in the project and were done during my visits in IIASA.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** The author would like to thank all referees and my colleagues, especially Daniel T. Jessie and Egor Ivanovski, for their valuable comments. The author would like to express his deepest gratitude to academician Arkady Kryazhimskiy and Elena Rovenskaya for encouraging this research.

**Conflicts of Interest:** The author declares no conflict of interest.

**Abbreviations**

The following abbreviations are used in this manuscript:

ESD	Equilibrium stationary distribution
NE	Nash equilibrium
RS (SRS)	Reactive strategy (stochastic reactive strategy)
SME	Strong mixed equilibrium
SPE	Subgame perfect equilibrium
ZD	Zero-determinant (strategies)

**Appendix A. Existence of Equilibria in Symmetric Games**

Let us derive conditions for  $a, b, c$  (see (20)) that ensure the existence of solutions in (18). Inequalities (iii) in (18) lead to three possible cases:  $a = 0, a < 0, a > 0$ .

*Appendix A.1. Case  $a = 0$*

Assuming  $a = 0$ , we obtain the specific version of (18):

$$\begin{cases} b + c(v_1 - v_2) = 0, \\ b + c(u_1 - u_2) = 0, \\ 0 < u_1, u_2, v_1, v_2 < 1. \end{cases} \tag{A1}$$

Recall that solutions of this system generate the set of all Nash equilibria.

If  $c = 0$ , then  $b = 0$ . If  $a = 0, c = 0$ , and  $b = 0$ , then  $A_1 = A_2 = A_3 = A_4$ . Trivially, any pair of reactive strategies is a Nash equilibrium. If  $c \neq 0$ , then we simplify (A1)

$$\begin{cases} (u_1 - u_2) = (v_1 - v_2) = -\frac{b}{c} \\ 0 < u_1, u_2, v_1, v_2 < 1 \end{cases}.$$

The recent statements lead us to the following lemma.

**Lemma A1.** *Suppose that in a symmetric game  $a = 0$ . Then, there exists a reactive Nash equilibrium iff  $-1 < -\frac{b}{c} < 1$  or  $c = 0, b = 0$ .*

Note that in this case, a mixed Nash equilibrium exists only if  $b = 0$ .

**Lemma A2.** *Suppose that in a symmetric game  $a = 0$ . If  $c = 0$  and  $b = 0$ , then any pair of reactive strategies is a Nash equilibrium. If  $-1 < \frac{b}{c} < 1$ , then the set of all Nash equilibria coincides with the set of all pairs  $(u, v)$  such that*

$$u = (\tilde{u} - \frac{b}{c}, \tilde{u}), \quad v = (\tilde{v} - \frac{b}{c}, \tilde{v})$$

where  $\tilde{u}, \tilde{v} \in [\frac{b}{c}, 1 + \frac{b}{c}] \cap ]0, 1[$ .

If  $a = 0$  and  $\frac{b}{c} = 0$ , then we have the following game:

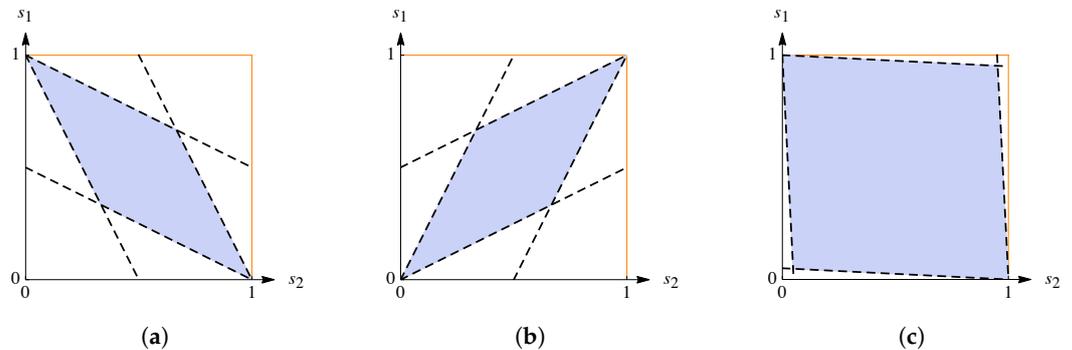
	I	II
I	$A_1, A_1$	$A_4, A_1$
II	$A_1, A_4$	$A_4, A_4$

Here  $A_1 \neq A_4$ . Trivially, in this case any stationary distribution  $(s_1, s_2) \in ]0, 1[^2$  is an ESD; the Nash equilibrium is formed by the strategies  $u = (s_1, s_1)$  and  $v = (s_2, s_2)$ , which are repeated mixed strategies. These equilibria are not strict.

The set of all Nash equilibria from Lemma A2 generates the parallelogram of ESDs; see Figure A1. The area of the shaded regions tends to 1 as  $\frac{b}{c} \rightarrow 0$ ; this is the idea behind the following propositions.

**Proposition A1.** Every pair of reactive strategies  $(u, v)$  such that  $u_1 - u_2 = v_1 - v_2$  forms Nash equilibria in symmetric games with  $a = 0$  and  $\frac{b}{c} = u_2 - u_1$ .

**Proposition A2.** For every stationary distribution  $(s_1, s_2) \in ]0, 1[^2$ , there exists  $\gamma \in ]0, 1[$  such that  $(s_1, s_2)$  is an ESD in all symmetric games with  $a = 0$  and  $|\frac{b}{c}| < \gamma$ .



**Figure A1.** The case  $a = 0$ . The shaded regions in (a–c) depict the set of all ESDs for the games with  $\frac{b}{c} = \frac{1}{2}, -\frac{1}{2}, \frac{1}{20}$ , respectively.

Though the case  $a = 0$  is degenerate, it is of interest in connection with ZD strategies [15]. These strategies are 1-memory ones. Since reactive strategies can also be treated as 1-memory strategies, we investigate the relationship between equalizer ZD and reactive strategies.

**Lemma A3.** For symmetric games with  $c \neq 0$ , a first player’s reactive strategy  $u$  is an equalizer ZD strategy if  $a = 0$  and  $u_2 - u_1 = \frac{b}{c}$ .

The proof is based on [15], (8). Here, for the 1-memory strategy  $(p_1, p_2, p_3, p_4)$ , one has to substitute by the reactive one  $(u_1, u_2, u_1, u_2)$ , and for the payoff profile  $(R, S, T, P)$  one has to substitute  $(A_1, A_2, A_3, A_4)$ . This gives the system

$$\begin{cases} (u_2 - u_1)c = a + b \\ (u_2 - u_1)(a + c) = b \end{cases}$$

Subtracting the first equation from the second one, we obtain  $(u_2 - u_1)a = -a$ . The final part of the proof is trivial.

**Proposition A3.** In symmetric games with  $a = 0$ , all reactive Nash equilibria are generated by equalizer ZD strategies.

Appendix A.2. Case  $a < 0$

We now proceed with the case  $a < 0$ . As before, we derive conditions (on  $a, b, c$ ) that ensure the existence of solutions of (18). Before we proceed, let us express (22) for the

symmetric games in terms of  $s_1, s_2; 0 < s_1, s_2 < 1$ . If  $c + as_2 \neq 0, c + as_1 \neq 0$ , and  $(s_1, s_2)$  is an ESD, then the corresponding NE profile  $(u, v)$  is defined by the following system:

$$\begin{cases} u_2 = \frac{cs_1 + bs_2 + 2as_1s_2}{c + as_2}, x = -\frac{b + as_1}{c + as_2}, \\ v_2 = \frac{bs_1 + cs_2 + 2as_1s_2}{c + as_1}, y = -\frac{b + as_2}{c + as_1}, \\ u_1 = u_2 + x, v_1 = v_2 + y, \\ 0 \geq ax, 0 \geq ay, 0 < u_1, u_2, v_1, v_2 < 1. \end{cases} \tag{A2}$$

From Theorem 1 we derive that for any equilibrium pair  $(u, v)$ , the following holds true:

$$(u_1 - u_2 \geq 0) \& (v_1 - v_2 \geq 0).$$

Since  $s_2 = s_1(v_1 - v_2) + v_2$  and  $s_1 = s_2(u_1 - u_2) + u_2$ , we rewrite (i) and (ii) in (18)

$$\begin{cases} (as_1 + c)(v_1 - v_2) + as_2 + b = 0, & \text{(i)} \\ (as_2 + c)(u_1 - u_2) + as_1 + b = 0, & \text{(ii)} \\ u_1 - u_2 \geq 0, v_1 - v_2 \geq 0, & \text{(iii)} \\ s_1 = \mathbf{s}_1(u, v), s_2 = \mathbf{s}_2(u, v), & \text{(iv)} \\ 0 < u_1, u_2, v_1, v_2 < 1. & \text{(v)} \end{cases} \tag{A3}$$

First, we pay attention to a degenerate case. Suppose that there exists a solution to (A3) such that  $as_1 + c = 0$ . This is only possible if  $as_2 + b = 0$ . Assuming this, we get

$$\begin{cases} (c - b)(u_1 - u_2 - 1) = 0, \\ u_1 - u_2 \geq 0, v_1 - v_2 \geq 0, \\ -\frac{c}{a} = \mathbf{s}_1(u, v), -\frac{b}{a} = \mathbf{s}_2(u, v), \\ 0 < u_1, u_2, v_1, v_2 < 1. \end{cases} \tag{A4}$$

Since  $u_1 - u_2 \neq 1 \forall u \in \mathcal{S}_I$ , a solution of (A4) may exist only if  $c - b = 0$ . Assuming this, we get

$$-\frac{c}{a} = -\frac{b}{a} = \mathbf{s}_1(u, v) = \mathbf{s}_2(u, v). \tag{A5}$$

Then there exists a solution iff  $0 < -\frac{b}{a} < 1$ ; this corresponds to the existence of the mixed Nash equilibrium. Clearly, the strategies  $u = (-\frac{b}{a}, -\frac{b}{a})$  and  $v = (-\frac{b}{a}, -\frac{b}{a})$  form a Nash equilibrium. We stress that this is not a unique solution to (A5) under the assumptions. However, all Nash equilibria lead to the same ESD and payoffs. Let us now combine our findings in the following:

**Lemma A4.** *If in a symmetric game  $a \neq 0$  and there is an ESD  $(\hat{s}_1, \hat{s}_2)$  such that*

$$(a\hat{s}_1 + c = 0) \vee (a\hat{s}_2 + c = 0),$$

*then  $b = c, \hat{s}_2 = \hat{s}_1, 0 < -\frac{b}{a} < 1$ , and the strategies  $u = (-\frac{b}{a}, -\frac{b}{a})$  and  $v = (-\frac{b}{a}, -\frac{b}{a})$  form a Nash equilibrium. Moreover,  $(\hat{s}_1, \hat{s}_2)$  is the only possible ESD for the game.*

For the remainder of this section, we suppose that  $b \neq c$ . We define the useful functions  $\mathbf{x}, \mathbf{y} : ]0, 1[ \times ]0, 1[ \rightarrow ]-1, 1[$  by the rules motivated by (A2)

$$\mathbf{x}(s_1, s_2) \triangleq -\frac{b + as_1}{c + as_2}, \mathbf{y}(s_1, s_2) \triangleq -\frac{b + as_2}{c + as_1}. \tag{A6}$$

Lemma A4 implies that if  $a \neq 0, b \neq c$ , then  $\mathbf{x}, \mathbf{y}$  are well defined for any ESD. Obviously,

$$\mathbf{x}(s, s) = \mathbf{y}(s, s) \forall s \neq -\frac{c}{a}. \tag{A7}$$

**Lemma A5.** *In a symmetric game with  $a < 0, b \neq c$ , for  $s \in ]0, 1[$ , there exists a Nash equilibrium with the corresponding ESD  $(s, s)$  iff  $0 \leq \mathbf{x}(s, s) < 1$ .*

**Proof.** Assume that strategies  $u, v$  lead to the ESD  $(s, s)$ . Combining (A6) with (i) and (ii) in (A3), we get that  $u_1 - u_2 = v_1 - v_2 = -\frac{b+as}{c+as} = \mathbf{x}(s, s) = \mathbf{y}(s, s)$ . Since  $u, v$  are well defined strategies, it follows from (iii) in (A3) that  $0 \leq \mathbf{x}(s, s) = \mathbf{y}(s, s) < 1$ .

Suppose that  $0 \leq \mathbf{x}(s, s) < 1$  for a stationary distribution  $(s, s)$ ; recall that  $\mathbf{x}(s, s) = \mathbf{y}(s, s)$ . To show that  $(s, s)$  is an ESD, one has to ensure validity (i)–(v) in (A3) by demonstrating well-defined strategies  $u^*$  and  $v^*$  generating  $(s, s)$  such that  $u_1^* - u_2^* = v_1^* - v_2^* = \mathbf{x}(s, s)$ . Let us recall the geometric representation of attainable sets; the difference between the first and second component of reactive strategies completely determines the slopes of the opponent’s attainable set (see Figures 2 and 3). We see that it is always possible to draw an attainable set of player 1 (straight line) through the point  $(s, s)$  with the slope defined by  $\mathbf{y}(s, s)$  if  $0 \leq \mathbf{y}(s, s) < 1$ . As a result we obtain the well-defined reactive strategy  $(v_1^*, v_2^*)$ . In the same way, we derive well-defined  $(u_1^*, u_2^*)$  such that  $u_1^* - u_2^* = \mathbf{x}(s, s)$ . By construction,  $u^*$  and  $v^*$  generate  $(s, s)$ . □

**Lemma A6.** *Suppose that in a symmetric game with  $a < 0, b \neq c$ , a Nash equilibrium exists with the ESD  $(s_1, s_2)$ . Then,  $(s_2, s_1)$  and  $(\frac{s_1+s_2}{2}, \frac{s_1+s_2}{2})$  are ESDs too.*

**Proof.** There are two cases:  $s_1 = s_2$  and  $s_1 \neq s_2$ . The proof is trivial in the first case. In the second case, by using the symmetry of the game, it is easy to prove the existence of ESD  $(s_2, s_1)$ . As  $(s_1, s_2)$  corresponds to some Nash equilibrium and  $a < 0$ , for  $x = \mathbf{x}(s_1, s_2)$  and  $y = \mathbf{y}(s_1, s_2)$ , we have  $0 \leq x < 1, 0 \leq y < 1$  (see (A2)). Let us now consider  $x^*$  such that  $x^* \triangleq \mathbf{x}(\frac{s_1+s_2}{2}, \frac{s_1+s_2}{2})$ . It is easy to show that in this case,

$$x^* = \frac{-x - y + 2xy}{-2 + x + y}.$$

Straightforwardly,  $x^* > 0$ . Moreover,  $x^* - 1 = \frac{2(1-x)(1-y)}{-2+x+y} < 0$ . Thus,  $0 \leq x^* < 1$ . Combining this with Lemma A5, we obtain that  $(\frac{s_1+s_2}{2}, \frac{s_1+s_2}{2})$  is an ESD. □

**Corollary A1.** *In symmetric games with  $a < 0, b \neq c$ , there exists a Nash equilibrium iff there exists a pair of identical reactive strategies forming a Nash equilibrium. The last is equivalent to the existence of an ESD  $(s, s)$  for some  $0 < s < 1$ .*

The next theorem provides a necessary and sufficient condition for the existence of a Nash equilibrium.

**Theorem A1.** *In symmetric games with  $a < 0, b \neq c$ , there exists a Nash equilibrium iff there exists a stationary distribution  $(s, s)$  such that*

$$0 \leq \mathbf{x}(s, s) < 1. \tag{A8}$$

The proof is a combination of Lemma A5, Lemma A6, and the last corollary. Further, we derive conditions such that there exists a solution in (A8). By definition, put

$$\mathbf{b} \triangleq -\frac{b}{a}, \mathbf{c} \triangleq -\frac{c}{a}.$$

We stress that if  $a \neq 0$ , then  $\mathbf{b}, \mathbf{c}$ , and sign of  $a$  completely characterize Nash equilibria in symmetric games. The solution of (A8) is the union of two regions:

$$s > c \ \& \ s \leq b \ \& \ s > \frac{b+c}{2}, \tag{A9}$$

$$s < c \ \& \ s \geq b \ \& \ s < \frac{b+c}{2}. \tag{A10}$$

To have some  $s \in ]0, 1[$  satisfying (A9) or (A10), we must require, respectively,

$$1 > c \ \& \ b > 0 \ \& \ \frac{b+c}{2} < 1 \ \& \ b > c \ \& \ b > \frac{b+c}{2}, \tag{A11}$$

$$0 < c \ \& \ b < 1 \ \& \ \frac{b+c}{2} > 0 \ \& \ b < c \ \& \ b < \frac{b+c}{2}. \tag{A12}$$

Elaborating (A11) and (A12) and adding Lemma A4, we arrive at the next

**Lemma A7.** *For symmetric games with  $a < 0$ , there exists a Nash equilibrium iff*

$$(c < 2 - b \ \& \ b \geq 1) \vee (c > -b \ \& \ b \leq 0) \vee (0 < b < 1). \tag{A13}$$

Remarkably, some symmetric games allow  $(s, s)$  to be an ESD for every  $s \in ]0, 1[$ .

**Proposition A4.** *In a symmetric game,  $(s, s)$  is an ESD for all  $s \in ]0, 1[$  if  $a < 0$  and*

$$(c \leq 0 \ \& \ b \geq 1 \ \& \ c + b \leq 0) \vee (c \geq 1 \ \& \ b \leq 0 \ \& \ c + b \geq 2).$$

To prove this, we find all  $b, c$  such that (A9) or (A10) hold for all  $s \in ]0, 1[$ .

*Appendix A.3. Case  $a > 0$*

An analysis in this case is more involved, but quite similar. It follows from Theorem 1 that for every equilibrium  $(u, v)$  the following holds:  $(u_1 - u_2 \leq 0)$  and  $(v_1 - v_2 \leq 0)$ . For the remainder of this section, we suppose that  $b \neq c$ , as Lemma A4 was obtained independently of the sign of  $a$ . We start with an analog of Lemma A5; see also (A7).

**Lemma A8.** *In symmetric games with  $a > 0, b \neq c$ , for  $s \in ]0, 1/2]$  there exists an ESD  $(s, s)$  iff*

$$-\frac{s}{1-s} < x(s, s) \leq 0. \tag{A14}$$

*In symmetric games with  $a > 0, b \neq c$ , for  $s \in ]1/2, 1[$  there exists an ESD  $(s, s)$  iff*

$$-\frac{1-s}{s} < x(s, s) \leq 0. \tag{A15}$$

We omit the proof based on the inference of analytic formulas for tangents (slopes) of border lines similar to the ones in Figure 3a. These formulas provide the lower bounds for  $x(s, s)$  in (A14) and (A15). The next step is to derive an analog of Lemma A6.

**Lemma A9.** *If in symmetric games with  $a > 0, b \neq c$ , a Nash equilibrium exists with the ESD  $(s_1, s_2)$ , then  $(s_2, s_1)$  and  $(\frac{s_1+s_2}{2}, \frac{s_1+s_2}{2})$  are ESDs too.*

**Proof.** There are two cases:  $s_1 = s_2$  and  $s_1 \neq s_2$ . The proof is trivial in the first case. In the second case, by using symmetry, it is easy to prove the existence of a Nash equilibrium with the ESD  $(s_2, s_1)$ . We consider three cases:  $s_1 + s_2 > 1, s_1 + s_2 = 1$ , and  $s_1 + s_2 < 1$ . Analyses of the cases are similar; therefore, we proceed only with case  $s_1 + s_2 > 1$ . Without loss of generality, we assume that  $s_1 < s_2$ . Since  $(s_1, s_2)$  is a Nash equilibrium, there exists the pair

$(x, y)$  such that  $x = \mathbf{x}(s_1, s_2)$  and  $y = \mathbf{y}(s_1, s_2)$ ;  $x \leq 0, y \leq 0$ . We infer analytic formulas for the tangents of border lines similar to those in Figure 3a to derive the following restriction:

$$-\frac{1-s_1}{s_2} < x, \quad -\frac{1-s_2}{s_1} < y.$$

We stress that any pair of reactive strategies  $(u, v)$  leading to  $(s_1, s_2)$  complies with

$$-\frac{1-s_1}{s_2} < u_1 - u_2, \quad -\frac{1-s_2}{s_1} < v_1 - v_2.$$

Let us now consider value  $x^* = \mathbf{x}(\frac{s_1+s_2}{2}, \frac{s_1+s_2}{2})$ ; we have  $x^* = \frac{-x-y+2xy}{-2+x+y}$ . Straightforwardly,  $x^* < 0$ . Moreover,  $x^* + 1 = \frac{2(xy-1)}{-2+x+y} > 0$ . Thus,  $-1 < x^* < 0$ . In contrast to Lemma A6, the last statement is insufficient to apply Lemma A8. We also need to prove a stronger result (see (A15)):

$$-\frac{1-\frac{s_1+s_2}{2}}{\frac{s_1+s_2}{2}} = \frac{s_1+s_2-2}{s_1+s_2} < x^*. \tag{A16}$$

First, we obtain derivatives of  $x^*$  w.r.t.  $x, y$ :

$$x_x^* = 2\left(\frac{y-1}{x+y-2}\right)^2, \quad x_y^* = 2\left(\frac{x-1}{x+y-2}\right)^2.$$

For all  $(x, y) \in ]-1, 0]^2$ , functions  $x^*(\cdot, y)$  and  $x^*(x, \cdot)$  are monotonically nondecreasing on  $]-1, 0]$ . Thus,  $x^*$  attains minimum on  $[-\frac{1-s_1}{s_2}, 0] \times [-\frac{1-s_2}{s_1}, 0]$  at  $(-\frac{1-s_1}{s_2}, -\frac{1-s_2}{s_1})$ ;

$$x^*\left(-\frac{1-s_1}{s_2}, -\frac{1-s_2}{s_1}\right) = \frac{2+(s_1-s_2)-(s_1-s_2)^2}{(s_1-s_2)^2-(s_1+s_2)}.$$

Recall that we are proving (A16). Since  $s_1 + s_2 > 1$  and  $(s_1 - s_2)^2 < 1$ , we obtain that

$$\begin{aligned} x^*(x, y) - \frac{s_1+s_2-2}{s_1+s_2} &> \frac{2+(s_1-s_2)-(s_1-s_2)^2}{(s_1-s_2)^2-(s_1+s_2)} - \frac{s_1+s_2-2}{s_1+s_2} = \\ &= \frac{2(s_1-s_2)^2(1-s_1-s_2)}{(s_1+s_2)((s_1-s_2)^2-(s_1+s_2))} > 0 \quad \forall (x, y) \in ]-\frac{1-s_1}{s_2}, 0] \times ]-\frac{1-s_2}{s_1}, 0]. \end{aligned}$$

It follows from Lemma A8 that  $(\frac{s_1+s_2}{2}, \frac{s_1+s_2}{2})$  is an ESD.  $\square$

**Corollary A2.** *In symmetric games with  $a > 0, b \neq c$ , there exists a Nash equilibrium iff there exists a pair of identical reactive strategies forming a Nash equilibrium. The latter is identical to the existence of an ESD  $(s, s)$  for some  $0 < s < 1$ .*

The next theorem provides a necessary and sufficient condition for the existence of a Nash equilibrium.

**Theorem A2.** *In symmetric games with  $a > 0, b \neq c$ , there exists a Nash equilibrium iff there exists an ESD  $(s, s), 0 < s < 1$ , such that*

$$\max\left\{-\frac{1-s}{s}; -\frac{s}{1-s}\right\} < \mathbf{x}(s, s) \leq 0. \tag{A17}$$

The proof is a combination of Lemma A8, Lemma A9, and the last corollary. Let us derive conditions ensuring the existence of a solution in (A17). As before, we use the notation  $\mathbf{b} = -\frac{b}{a}, \mathbf{c} = -\frac{c}{a}$ . We decompose (A17) into two systems of the inequalities:

$$\left(\frac{s-1}{s} < \frac{\mathbf{b}-s}{s-\mathbf{c}} < 0\right) \& \left(\frac{1}{2} \leq s < 1\right), \tag{A18}$$

$$\left(-\frac{s}{s-1} < \frac{\mathbf{b}-s}{s-\mathbf{c}} \leq 0\right) \& \left(0 < s \leq \frac{1}{2}\right). \tag{A19}$$

We begin with (A18). Arguing as in the case  $a < 0$ , we obtain that the conditions

$$\left(\mathbf{c} < \mathbf{b} \& \mathbf{b} \leq \frac{1}{2} \& \mathbf{c} < \frac{1}{2}\right) \vee \left(\mathbf{c} < \mathbf{b} \& \frac{1}{2} < \mathbf{b} < 1\right) \vee \left(\mathbf{c} > \mathbf{b} \& \mathbf{b} \geq \frac{1}{2}\right) \tag{A20}$$

admit solutions of (A18). We proceed to (A19). First, (A19) can be transformed in

$$\frac{\hat{s}-1}{\hat{s}} < \frac{\hat{\mathbf{b}}-\hat{s}}{\hat{s}-\hat{\mathbf{c}}} \leq 0 \& \frac{1}{2} \leq \hat{s} < 1 \tag{A21}$$

(an instance of (A18)) by applying the following rules:

$$\hat{s} \triangleq 1-s, \hat{\mathbf{b}} \triangleq 1-\mathbf{b}, \hat{\mathbf{c}} \triangleq 1-\mathbf{c}. \tag{A22}$$

Substituting  $\hat{s}, \hat{\mathbf{b}}$ , and  $\hat{\mathbf{c}}$  in (A21), we get

$$\left(\frac{1-s-1}{1-s} < \frac{1-\mathbf{b}-1+s}{1-s-1+\mathbf{c}} \leq 0\right) \& \left(\frac{1}{2} \leq 1-s < 1\right).$$

Simplifying, we have  $\left(\frac{-s}{1-s} < \frac{s-\mathbf{b}}{\mathbf{c}-s} \leq 0\right) \& \left(-\frac{1}{2} \leq -s < 0\right)$ ; this chain of inequalities is identical to (A19). Thus, bijective rules (A22) provide an easy way for one-to-one transformation of problem (A19) into problem (A18), which was solved in (A20). Using (A20), we solve the problem for (A21):

$$\left(\hat{\mathbf{c}} < \hat{\mathbf{b}} \& \hat{\mathbf{b}} \leq \frac{1}{2} \& \hat{\mathbf{c}} < \frac{1}{2}\right) \vee \left(\hat{\mathbf{c}} < \hat{\mathbf{b}} \& \frac{1}{2} < \hat{\mathbf{b}} < 1\right) \vee \left(\hat{\mathbf{c}} > \hat{\mathbf{b}} \& \hat{\mathbf{b}} \geq \frac{1}{2}\right).$$

Using rules (A22), we obtain the solution of problem (A19):

$$\left(\mathbf{c} > \mathbf{b} \& \mathbf{b} \geq \frac{1}{2} \& \mathbf{c} > \frac{1}{2}\right) \vee \left(\mathbf{c} > \mathbf{b} \& \frac{1}{2} > \mathbf{b} > 0\right) \vee \left(\mathbf{c} < \mathbf{b} \& \mathbf{b} \leq \frac{1}{2}\right). \tag{A23}$$

It is time to combine (A20), (A23), and Lemma A4.

**Lemma A10.** *If  $a > 0$ , then there exists a reactive Nash equilibrium iff*

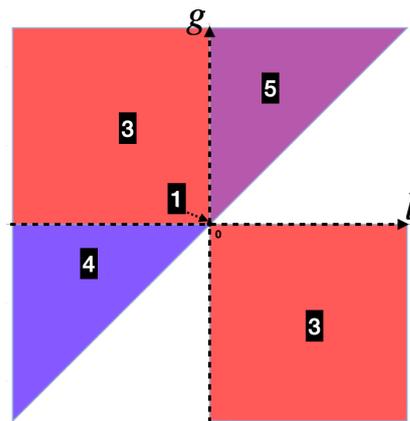
$$\left(\mathbf{c} < \mathbf{b} \& \mathbf{b} \leq 0\right) \vee \left(\mathbf{c} > \mathbf{b} \& \mathbf{b} \geq 1\right) \vee \left(0 < \mathbf{b} < 1\right).$$

### Appendix B. Theorem 2 for Symmetric Games with Equal Payoffs on the Leading Diagonal

A generic one-shot symmetric game with identical payoffs on the leading diagonal is strategically equivalent to

Game $G^{00}$	C	D
C	0,0	-l,g
D	g,-l	0,0

Figure A2 shows all pairs of  $(l, g)$  such that there exists an equilibrium in the repeated version of  $G^{00}$  in SRSs by Theorem 2.



**Figure A2.** The numbered regions are derived from the corresponding conditions of Theorem 2. The regions specify all  $l$  and  $g$  such that there exists an NE profile in SRSs for the repeated modification of  $G^{00}$ .

Table A1 contains the corresponding case analysis. In general, there exists an NE profile in conditional SRSs that Pareto dominates a benchmark NE profile in memory-less reactive strategies. Thus, an inch of memory for players (upgrading memory-less RSs to conditional SRSs) brings new equilibria where they both better off. Only a very special type of this one-shot games is the exception to this rule, i.e., games with  $-l = g$ .

**Table A1.** Summary for Theorem 2 for symmetric games with equal payoffs on the leading diagonal.

Condition of the Theorem	One-Shot Game Description	The Benchmark in Memory-Less RSs	NE Payoffs in Conditional SRSs
1	A trivial stage game with identical payoffs	Any profile of memory-less RSs forms a NE with payoffs $(0,0)$	Any profile of conditional RSs forms an NE with payoffs $(0,0)$
2, 6, and 7	Can not hold for $G^{00}$		
3	Coordination and anti-coordination stage games with two pure equilibria	There is an NE profile in memory-less SRSs with payoffs $(\frac{l_g}{1-g}, \frac{l_g}{1-g})$	If $-l \neq g$ , then there exists an NE profile of conditional SRSs Pareto dominating the memory-less benchmark. If $-l = g$ , then there is an unique ESD $(\frac{1}{2}, \frac{1}{2})$ ; all payoff profiles of equilibria in SRSs coincide.
4 and 5	Stage games having one dominant pure strategy; any symmetric profile of mixed strategies Pareto improves the NE payoffs	Payoff profile $(0,0)$ of dominant strategies	Any NE profile of conditional SRSs Pareto dominates the memory-less benchmark.

**Appendix C. Additional Examples**

In this section, we provide examples illustrating properties of reactive Nash equilibria. In what follows, we use definitions given in (2), (3), and (20). Recall that for symmetric games,  $\mathbf{b} \triangleq -b/a, \mathbf{c} \triangleq -c/a$  if  $a \neq 0$ .

Appendix C.1 presents two examples of the Prisoner’s Dilemma to demonstrate non-symmetric equilibria. In the first example we observe equilibrium strategies leading to payoffs that are higher than the mutual cooperation payoff. In the second example, one

can see that there exist equilibria that admit any level of mutual cooperation. We then consider an example of non-symmetric games admitting equilibria that may be incorrectly recognized as symmetric ones; see Appendix C.2. Finally, in Appendix C.3, we highlight the example that makes it possible to observe disconnected regions of equilibria. Assume that we have some dynamics on the set of all Nash equilibria in the spirit of adaptive dynamics [4] or local cooperation patterns [26]. This disconnectedness becomes an obstacle on the way to higher equilibrium payoffs.

Appendix C.1. Non-Symmetric Equilibria in Prisoner’s Dilemma and Folk Theorem

Reactive strategies were introduced to study the evolution of cooperation in the repeated Prisoner’s Dilemma. Recall that in [4,5] symmetric equilibria in the repeated Prisoner’s Dilemma were characterized. The following examples illustrate an abundance of non-symmetric equilibria in repeated Prisoner’s Dilemmas. Consider two games:

Game $G^{P1}$	C	D	Game $G^{P2}$	C	D
C	5,5	-1,15	C	5,5	-2,8
D	15,-1	0,0	D	8,-2	0,0

A mixed Nash equilibrium does not exist in both games.

*Game  $G^{P1}$ .* Here  $a = -9, \mathbf{b} = -1/9, \mathbf{c} = 5/3$ . This Prisoner’s Dilemma is different from the standard versions since  $\frac{A_2+A_3}{2} > A_1$ . Hence, contrary to the standard setting, infinite cycles like  $CD, DC, CD, \dots$  are more profitable for both players than the infinite mutual cooperation. As a result, in some equilibria, both payoffs are strictly better off than the mutual cooperation payoff; see Figure A3a. For example, equilibrium strategies  $u = (\frac{13}{15}, \frac{1}{5})$  and  $v = (\frac{13}{15}, \frac{1}{5})$  lead to the ESD  $(\frac{3}{5}, \frac{3}{5})$  with the payoffs  $\frac{129}{25}$ , which are greater than the mutual cooperation payoff 5.

*Game  $G^{P2}$ .* Here,  $a = -1, \mathbf{b} = -2, \mathbf{c} = 8$ . We depict the set of all ESDs in Figure A3b. It follows from Proposition A4 that for any  $s \in ]0, 1[$  there exists a Nash equilibrium such that  $(s, s)$  is the ESD. Therefore, there are equilibria admitting any level of mutual cooperation. For example, equilibrium strategies  $u = v = (\frac{3498499}{3500500}, \frac{1998999}{3500500})$  lead to the ESD  $(\frac{999}{1000}, \frac{999}{1000})$ .

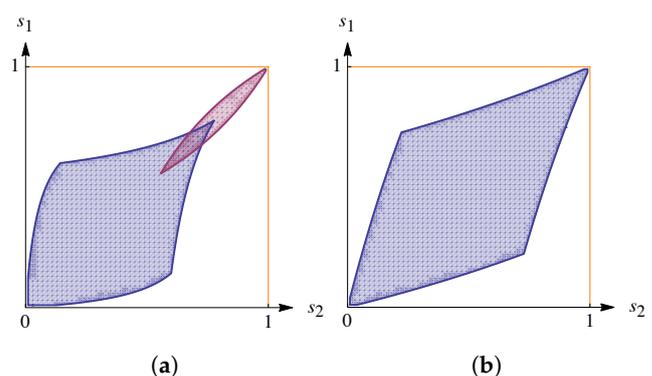
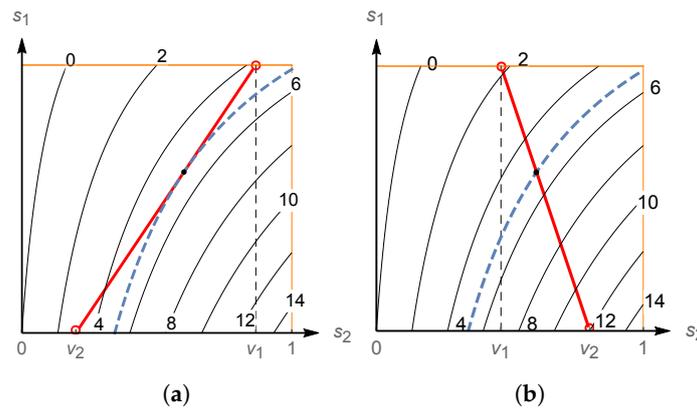


Figure A3. The blue sets in (a,b) depict the sets of all ESDs for  $G^{P1}$  and  $G^{P2}$ , respectively. The red region corresponds to all stationary distributions such that both players get more than the mutual cooperation payoff (this is possible only in  $G^{P1}$ ).

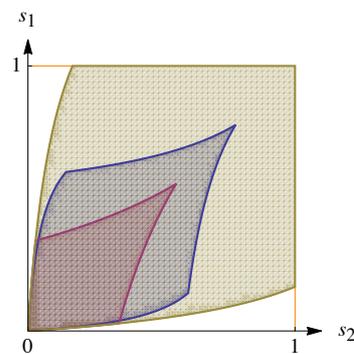
Figure A4 shows that to ensure an equilibrium, player 2 must choose  $v_1 > v_2$  and not the other way round. For an equilibrium with a higher rate of cooperation, player 2 must increase  $v_1$  even further. Generally, the greater the difference  $(v_1 - v_2)$ , the less attractive the defection for player 1. In (a), the strategy  $v = (\frac{13}{15}, \frac{3}{5})$  sets the attainable set of player 1 (the straight line) to be a tangent to the curve of payoffs equal to  $\frac{129}{25}$  at the ESD  $(\frac{3}{5}, \frac{3}{5})$ . In (b), the strategy  $v = (\frac{7}{15}, \frac{4}{5})$  incentivizes player 1 to descend in the attainable set (the straight line) to the unconditional defection.



**Figure A4.** For  $G^{P1}$ , plots (a,b) depict the first player’s attainable sets (the straight lines) for  $v = (\frac{13}{15}, \frac{3}{5})$  and  $v = (\frac{7}{15}, \frac{4}{5})$ , respectively. The thin lines correspond to the payoff levels equal to 0, 2, ..., 14.

The range of all possible payoffs for SMEs was obtained in [17] proving that there is no folk theorem (see [17], Corollary 2, Theorem 4). For example, for *game*  $G^{P1}$ , all SME payoffs are in  $]0, 5]$  for any discount factor. In [28], Section 4, the repeated Prisoner’s Dilemma was considered as an example of a game with non-rich action spaces. It was rigorously shown that there are no efficient payoff vectors corresponding to subgame perfect equilibria in deterministic 1-memory strategies, even if the discount factor is close to one. However, by introducing an additional pure action, it becomes possible to obtain an efficient outcome, which is higher than the upper bound for SME payoffs.

Figure A5 shows that there are Nash equilibria in  $G^{P1}$  with payoffs higher than would be ensured by the mutual cooperation, although not all individually rational payoffs are supported by reactive equilibria. First, this means that there are ESDs with payoffs for both players that are higher than the upper bound shown in [28] (for equilibria in deterministic 1-memory strategies) and in [17] (for SMEs). Second, similarly to [17], in our case there is no folk theorem.



**Figure A5.** The blue set corresponds to all ESD. The brown set corresponds to distributions with individually rational payoffs (greater or equal to 0); the red set corresponds to distributions such that both payoffs are in range  $[0, 5]$ .

*Appendix C.2. Symmetric ESD in Non-Symmetric Games*

Consider the non-symmetric game  $G^{ns}$

Game $G^{ns}$	L	R
T	3, -6	-2, -3
B	6, -4	0, 0

admitting a continuum of ESDs of the form  $(s, s)$ . We say that ESDs of this form are *symmetric ESDs* (though this may be not a good choice of words). Symmetric ESDs have an intriguing property. Imagine that in a repeated game, players are in an equilibrium, and we observe only action frequencies (but not conditional frequencies). If the frequencies

are symmetric, then it would be natural to conclude that we are observing a symmetric game. This following example, however, shows that the last assumption may not be true.

It is easy to see that  $a_1 = -1, b_1/a_1 = 2, c_1/a_1 = -6, a_2 = 1, b_2/a_2 = -3, c_1/a_1 = -4$ . Note that  $a_1a_2 < 0$ ; this is impossible for symmetric games. Figure A6c depicts the set of all ESDs; we see an abundance of symmetric ESDs. One can also see a clear connection between the non-symmetric game  $G^{ns}$  and two symmetric games (see Figure A6a,b) constructed directly from players' payoff matrices.

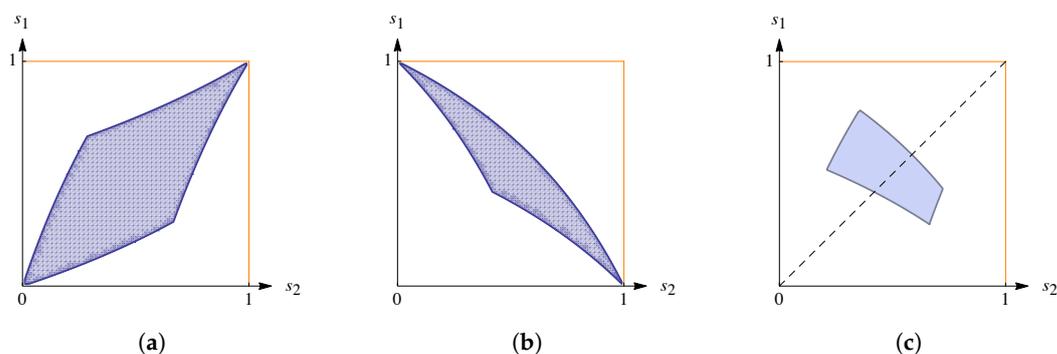


Figure A6. Plots (a–c) depict the set of all ESDs for the symmetric game with  $a < 0, \mathbf{b} = -2, \mathbf{c} = 6$ , for the symmetric game with  $a > 0, \mathbf{b} = 3, \mathbf{c} = 4$ , and non-symmetric  $G^{ns}$ , respectively.

### Appendix C.3. Games with Disconnected Regions of ESDs

Let us demonstrate that in the case  $a > 0$ , it is possible to observe disconnected regions of ESDs. Consider the stage game

Game $G^{D1}$	C	D
C	-12.5, -12.5	-9.5, -13
D	-13, -9.5	0, 0

Note that there exists a mixed NE in the stage game. For  $G^{D1}$ , we have  $a = 10, \mathbf{b} = 0.95, \mathbf{c} = 1.3$ . We depict the set of all ESDs in Figure A7a and all possible symmetric equilibria in Figure A7b.

Assume that we have some dynamics on the set of all Nash equilibria in the spirit of the adaptive dynamics [4] or local cooperation patterns [26]. This disconnectedness then becomes an obstacle on the way to higher equilibrium payoffs.

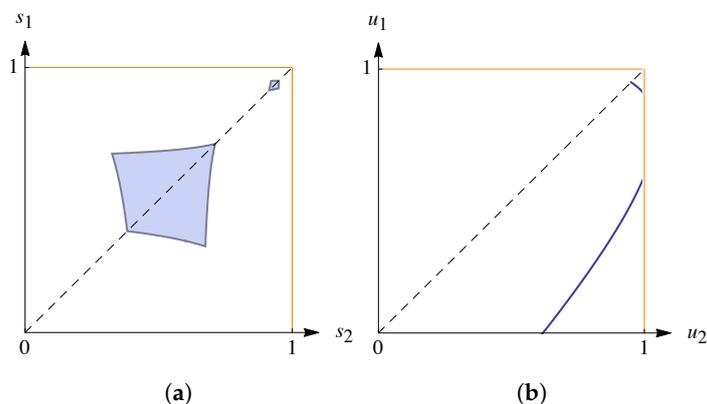


Figure A7. Plots (a,b) depict all ESDs and all symmetric equilibria for  $G^{D1}$ , respectively.

### References

1. Mailath, G.J.; Samuelson, L. *Repeated Games and Reputations: Long-Run Relationships*; Oxford University Press: Oxford, UK, 2006.
2. Kalai, E. Bounded Rationality and Strategic Complexity in Repeated Games. In *Game Theory and Applications*; Academic Press, Inc.: San Diego, CA, USA, 1990; Chapter V, pp. 131–157.

3. Dal Bó, P.; Fréchette, G.R. Strategy Choice in the Infinitely Repeated Prisoner's Dilemma. *Am. Econ. Rev.* **2019**, *109*, 3929–3952. [[CrossRef](#)]
4. Nowak, M.; Sigmund, K. The evolution of stochastic strategies in the Prisoner's Dilemma. *Acta Appl. Math.* **1990**, *20*, 247–265. [[CrossRef](#)]
5. Nowak, M. Stochastic strategies in the Prisoner's Dilemma. *Theor. Popul. Biol.* **1990**, *38*, 93–112. [[CrossRef](#)]
6. Nowak, M.; Sigmund, K. Evolutionary Dynamics of Biological Games. *Science* **2004**, *303*, 793–799. [[CrossRef](#)] [[PubMed](#)]
7. Hofbauer, J.; Sigmund, K. Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.* **1990**, *3*, 75–79. [[CrossRef](#)]
8. Kalai, E.; Samet, D.; Stanford, W. A note on reactive equilibria in the discounted prisoner's dilemma and associated games. *Int. J. Game Theory* **1988**, *17*, 177–186. [[CrossRef](#)]
9. Aumann, R.J. Repeated Games. In *Issues in Contemporary Microeconomics and Welfare*; Feiwel, G.R., Ed.; Palgrave Macmillan: London, UK, 1985; Chapter 5, pp. 209–242. [[CrossRef](#)]
10. Friedman, J.W. Reaction Functions and the Theory of Duopoly. *Rev. Econ. Stud.* **1968**, *35*, 257–272. [[CrossRef](#)]
11. Stanford, W.G. Subgame perfect reaction function equilibria in discounted duopoly supergames are trivial. *J. Econ. Theory* **1986**, *39*, 226–232. [[CrossRef](#)]
12. Kamihigashi, T.; Furusawa, T. Global dynamics in repeated games with additively separable payoffs. *Rev. Econ. Dyn.* **2010**, *13*, 899–918. [[CrossRef](#)]
13. Friedman, J.W.; Samuelson, L. Continuous Reaction Functions in Duopolies. *Games Econ. Behav.* **1994**, *6*, 55–82. [[CrossRef](#)]
14. Kryazhimskiy, A. Equilibrium stochastic behaviors in repeated games. In *Abstracts of International Conference Stochastic Optimization and Optimal Stopping*; Steklov Mathematical Institute: Moscow, Russia, 2012; pp. 38–41.
15. Press, W.H.; Dyson, F.J. Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10409–10413. [[CrossRef](#)] [[PubMed](#)]
16. Hilbe, C.; Traulsen, A.; Sigmund, K. Partners or rivals? Strategies for the iterated prisoner's dilemma. *Games Econ. Behav.* **2015**, *92*, 41–52. [[CrossRef](#)]
17. Dutta, P.K.; Siconolfi, P. Mixed strategy equilibria in repeated games with one-period memory. *Int. J. Econ. Theory* **2010**, *6*, 167–187. [[CrossRef](#)]
18. Ely, J.C.; Välimäki, J. A Robust Folk Theorem for the Prisoner's Dilemma. *J. Econ. Theory* **2002**, *102*, 84–105. [[CrossRef](#)]
19. D'Arcangelo, C. Too Much of Anything Is Bad for You, Even Information: How Information Can Be Detrimental to Cooperation and Coordination. Ph.D. Thesis, University of Trento, Trento, Italy, 2018.
20. Abreu, D.; Rubinstein, A. The Structure of Nash equilibrium in Repeated Games with Finite Automata. *Econom. J. Econom. Soc.* **1988**, *56*, 1259–1281. [[CrossRef](#)]
21. Marks, R. Repeated games and finite automata. In *Recent Developments in Game Theory*; Edward Elgar Publishing: Cheltenham, UK, 1992; pp. 43–64.
22. Baek, S.K.; Jeong, H.C.; Hilbe, C.; Nowak, M.A. Comparing reactive and memory-one strategies of direct reciprocity. *Sci. Rep.* **2016**, *6*, 1–13. [[CrossRef](#)] [[PubMed](#)]
23. Nowak, M. An evolutionarily stable strategy may be inaccessible. *J. Theor. Biol.* **1990**, *142*, 237–241. [[CrossRef](#)]
24. Nowak, M.; Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **1992**, *355*, 250–253. [[CrossRef](#)]
25. Moulin, H.; Vial, J.P. Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. *Int. J. Game Theory* **1978**, *7*, 201–221. [[CrossRef](#)]
26. Kleimenov, A.; Kryazhimskiy, A. *Normal Behavior, Altruism and Aggression in Cooperative Game Dynamics*; IIASA: Laxenburg, Austria, 1998.
27. McAvoy, A.; Nowak, M.A. Reactive learning strategies for iterated games. *Proc. R. Soc. A* **2019**, *475*, 20180819. [[CrossRef](#)]
28. Barlo, M.; Carmona, G.; Sabourian, H. Repeated games with one-memory. *J. Econ. Theory* **2009**, *144*, 312–336. [[CrossRef](#)]