Repeated Games with Incomplete Information and Short-Run Players

Stephan Waizmann *†

August 2022

I study repeated zero-sum games with incomplete information. In contrast to the canonical setting of Aumann and Maschler (1995), I assume that the uninformed player is a sequence of short–lived players. When monitoring of past actions is perfect, Aumann and Maschler's (1995) "Cav u"-result extends. When monitoring is imperfect, the payoff of the informed player can be strictly higher when facing a sequence of short–lived players than in the canonical setting, depending on parameters. I provide a partial characterization of equilibrium payoffs when monitoring is imperfect.[1]

Keywords: Repeated games with incomplete information, uniform equilibrium, value, short–lived players.
JEL classification: C72, C73.

1 Introduction

The literature on repeated games with incomplete information focuses almost exclusively on the case where both the informed and uninformed player are arbi-

*Department of Economics, Yale University, stephan.waizmann@yale.edu.

†I thank my advisors Johannes Hörner and Larry Samuelson for their invaluable guidance throughout this project. I am grateful for comments and feedback from Jérôme Renault and the audiences at the Micro Theory lunch at Yale and the 33rd Stony Brook International Conference on Game Theory. Part of this work was conducted during my stay at the Toulouse School of Economics whose hospitality I gratefully acknowledge. The usual disclaimer applies.

[1]After completing this paper, I became aware of a paper by Jean-Pierre Beaud and Sylvain Sorin (Beaud and Sorin (2000) "Sequence of opponents and reduced strategies", *International Journal of Game Theory*, 29, pp. 359–64) that studies the same question. All credit belongs to them.

trarily patient. The case where the informed player is arbitrarily impatient but the uninformed player is short-lived has received little attention.[2]

This paper studies this case. It considers a repeated zero-sum game with lack of information on one side played by a patient long-lived player against a sequence of short-lived players. The long-lived player knows the stage game. The short-lived players do not. The focus is on the case where the informed player is arbitrarily patient and the uninformed players live for a single stage only.[3] I refer to this game as the game with short–run players. I compare equilibrium payoffs of this game to the game in which the uninformed player is long–lived as well, which I refer to as the canonical game.

Examples of such situations are common in economics and finance. A middleman with superior information about some asset trades with short–lived investors; see Glosten and Milgrom (1985). A firm that knows its own competence at producing quality sells to myopic consumers; see Mailath and Samuelson (2001).

This paper contains two sets of results. First, under perfect monitoring, Theorem 1 establishes that the payoff to the informed player is the same whether he faces an equally patient long-lived opponent or a sequence of short-lived players. That is, Aumann and Maschler's "Cav u" result applies here as well. The intuition behind this result is standard: by using his information, the informed player reveals it over time.[4] Hence, either he does not use his information past a certain point, or the game becomes one of complete information eventually. Thus, the long-lived player optimally uses his information to "concavify" the payoff that he would obtain in the "non-revealing" game.

Second, this paper shows that this payoff equivalence (between the canonical game and the game considered here) does not uphold when monitoring is imperfect. The informed player receives a higher payoff in equilibrium when facing short-lived opponents than in the canonical game, and this inequality is strict in many games. The crucial difference comes from the uninformed player's experimentation motive: the signals the uninformed player obtains about the informed player's behavior may depend on her own action. Hence, an action that gives a lower payoff stage game payoff may provide information about the long-lived player's actions. There is an informational externality between different stages of the game. While a long-lived player takes this externality into account, a short-lived player does not. Against an opponent who plays a myopic best-response the informed player may be able to use his information without having to reveal it. Under perfect monitoring this effect is absent because the uniformed player's action cannot affect the signal about the informed player's action.[5]

[2]See Section 2.

[3]I discuss the case where the informed player is short–lived in Section 5.4.

[4]I use "he" for the informed and "she" for the uninformed player.

[5]This intuition suggests that the equivalence extends to the case in which the uninformed

As an illustration, consider a signalling structure with the following properties. The signal received by the uninformed player is uninformative for all actions except for one; call that action "explore." When choosing "explore," the uninformed player observes the action of the informed player. However, "explore" is a strictly dominated action in all possible stage games. It is then never optimal for a short-lived player to choose "explore," and the informed player can use his information without revealing it. If the uninformed player is long-lived and patient she may observe the informed player's behavior by occasionally and randomly choosing "explore." In that case, the informed player cannot use his information without revealing it. As a consequence, he may obtain an equilibrium payoff strictly lower than when facing short-lived opponents.

Theorem 2 provides a partial characterization of equilibrium payoffs for the informed player. It is based on the notion of a non-revealing payoff, an extension of Aumann and Maschler's value of the non-revealing game. A non-revealing payoff at a prior probability p is the payoff to a strategy profile of the stage game such that (i) the uninformed player plays a best-response, and (ii) the informed player plays a best response among all strategies that yield the same distribution over signals given the strategy by the uninformed player. The difference to the non-revealing game is in the last qualifier: the informed player plays a best-response among all strategies that yield the same distribution over signals against *all* actions available to the uninformed player.

The rest of the paper is organized as follows. The next section discusses its contribution to the literature. Section 3 introduces the model and the notation. Section 4 states the main results, first for perfect monitoring and then imperfect monitoring. Section 5 provides a discussion of the results. All proofs are contained in the appendix.

2 Related Literature

This paper builds on the work of Aumann and Maschler (1966*, 1967*, 1968*, 1995). They consider repeated zero-sum games with one-sided lack of information in which both players are equally patient. In addition to characterizing the uniform value of such games, they show convergence of the value of the discounted and the finitely repeated games.[6] Their results are the benchmark to which the results of this paper are to be compared.

Lehrer and Yariv (1999) consider the case of two unequally patient players. They

player's actions do not affect the signal. This is indeed the case. Theorem 3 makes this statement precise.

[6] Gensbittel (2015) extends the Aumann-Maschler results to games with infinite action spaces. This paper restricts itself to games with finite action spaces.

characterize equilibrium payoffs as both players discount rates vanish. In contrast, this paper holds the discount rate of the uniformed player fixed at 0. While the payoff bounds in Lehrer and Yariv (1999) apply to the model with short-lived players, those bounds are trivial in this special case. In addition, the focus of this paper is the undiscounted game rather than the limit of the discounted games.

There is a considerable literature on repeated games of incomplete information and non–zero sum payoffs.[7] Hart (1985) charaterizes equilibrium payoffs in two-player general sum games with lack of information on one side and perfect monitoring. Cripps and Thomas (2003) consider discounted repeated games with known own payoffs. They allow for players to have different discount factors. However, both players in their paper are long–lived.

Fudenberg, Kreps, and Maskin (1990) study nonzero-sum repeated games of complete information and short-lived players. The repetition of zero-sum games with complete information yields the repetition of the statically optimal strategies, irrespective of the patience of the players. However, their results show that it need not be the case that a patient player is better off playing against a sequence of short-run players instead of a single long-run player. More precisely: fix a finite two-player stage game of complete information. Assume perfect monitoring. Let w be a Nash equilibrium payoff to Player 1 in the (finitely repeated or discounted) game in which a sequence short-run players takes the role of Player 2. Then w is a Nash equilibrium payoff to Player 1 in the (finitely repeated or discounted) game in which Player 2 is a single long-lived player. In particular, this holds for the lowest Nash equilibrium payoff in the game with short-run players. Theorem 1 has a similar flavor: there is no difference in equilibrium payoffs in repeated zero-sum games of incomplete information when monitoring is perfect.

The literature often referred to as "reputation"[8] analyzes repeated interaction between a single long-run player and a sequence of short-run players. There are two major differences between this literature and the class of games I consider. First, they usually assume that the long-lived player might be a commitment type, i.e., he has a strategy that is strictly dominant in the repeated game. Here instead, the informed player chooses his strategy freely, i.e., he is not restricted by his type. Second, it is usually assumed that the short-lived players know their own payoffs. Since payoffs are perfectly correlated in the zero-sum case, this second assumption would reduce the games I consider to ones of complete information.

[7] See Forges (1992) for a more detailed discussion of the literature.

[8] Seminal works are Kreps, Milgrom, et al. (1982); Kreps and Wilson (1982), Milgrom and Roberts (1982), Fudenberg and Levine (1989), and Fudenberg and Levine (1992); see Mailath and Samuelson (2015) for a detailed discussion of the literature.

3 Model and Notation

The game is played over infinitely many stages $n = 0, 1, \dots$ There is an infinitely lived player, Player 1 or P1, and a sequence of short–run players, one in each stage.[9] The short-run player in stage n, SR_n, is only active in stage n.

Following the literature, I denote generic action profiles by (i, j) with i (j) the action of Player 1 (SR_n), and let k be the state of nature. Denote by I, J and K the finite sets $\{1, \dots, I\}$, $\{1, \dots, J\}$ and $\{1, \dots, K\}$ as well as their cardinality. Let $(G^k, k \in K)$ be the collection of payoff matrices $G^k \in \mathbb{R}^{I \times J}$. Player 1's payoff in the stage game under action profile (i, j) and the state of nature k is G^k_{ij}, and the payoff to SR_n is $-G^k_{ij}$. The commonly know prior over the state of nature is $p \in \Delta(K)$.

Denote by Y a finite set of public signals. The elements in Y are symbols with no intrinsic meaning. A signal structure is a map $Q : I \times J \to Y$, with the interpretation that $Q(i, j)$ is the signal that obtains when the actions i and j are played. Note that signalling is deterministic. The signal structure can be conveniently represented by a matrix where the (i, j)-th entry $q_{ij} = Q(i, j)$ describes the signal that obtains under the action pair (i, j). In the following, I abuse notation and denote both the signal structure and the signalling matrix (q_{ij}) by Q.

Short–run players do not observe realized stage game payoffs of those short–run players that have played before them. For simplicity, but without loss of generality, I also assume that the long–run player does not observe his stage game payoffs.[10]

Informally, the game is played as follows. Before stage $n = 0$, nature draws a matrix G^k according to the probability distribution $p \in \Delta(K)$. Player 1 is told the draw. The short–run players are not informed about the draw. Player 1 then chooses a row $i \in I$ and SR_n chooses a column $j \in J$. At the end of the stage, the signal q_{ij} is publicly announced. In every stage n, Player 1 and the short-run player SR_n observe the signal from the previous stages $m = 0, \dots, n-1$. In addition, Player 1 recalls his own actions and SR_n observes the actions by all previous short-run players.[11]

Let $h_n = (i_0, j_0, q_{i_0,j_0} \dots, i_{n-1}, j_{n-1}, q_{i_{n-1},j_{n-1}})$ be the complete history at the beginning of stage n, with the convention that $h_0 = \emptyset$. Define H_n to be the set of complete histories at stage n. Similarly, the private history of P1 at stage

[9] I use short-run and short-lived interchangeably. Sometimes I use the abbreviation SR.

[10] This is irrelevant, since he knows k. Note that the assumption that payoffs are not observed is less controversial here than in the canonical set–up, where it is usually made. Indeed, in most applications, short–run players do not see their predecessors' realized payoffs.

[11] While the assumption that the long-run player recalls his action is natural, assuming that the short-run players observe previous short-run players' action may be less plausible, depending on the application. I will discuss this assumption in Section 5.5.

n is $h_n^1 = (i_0, q_{i_0,j_0} \ldots, i_{n-1}, q_{i_{n-1},j_{n-1}})$, and the private history of SR_n is $h_n^2 = (j_0, q_{i_0,j_0} \ldots, j_{n-1}, q_{i_{n-1},j_{n-1}})$. Denote by H_n^1 and H_n^2 the set of private histories of length n of Player 1 and SR_n, respectively.

A behavioral strategy σ for Player 1 is a sequence of maps $(\sigma_n)_{n\in\mathbb{N}}$, where

$$\sigma_n : \{1, \ldots, K\} \times H_n^1 \to \Delta(I).$$

A behavioral strategy τ_n for player SR_n is a map

$$\tau_n : H_n^2 \to \Delta(J).$$

Call the collection of behavioral strategies $\tau = (\tau_n)_{n\in\mathbb{N}}$.

Given a state k and a complete history h_N, the payoff to Player 1 is

$$\gamma_N(k, h_N) = \frac{1}{N+1} \sum_{n=0}^{N} G^k_{i_n j_n},$$

where G^k_{ij} is the (i,j)-th entry of the matrix G^k. The payoff to SR_n given state k and complete history h_n is

$$\gamma_{\mathrm{SR}_n}(k, h_n) = -G^k_{i_n j_n}.$$

Denote the repeated game thus defined by $\Gamma^{\mathrm{SR}}(p)$.

The triple (σ, τ, p) induces a unique probability distribution $\mathbb{P}^{(\sigma,\tau)}$ over the set of plays $\{1, \ldots, K\} \times (I \times J)^{\mathbb{N}}$.[12] Denote by $\mathbb{E}^{(\sigma,\tau)}$ the expectation according to $\mathbb{P}^{(\sigma,\tau)}$.

Definition 1 (Uniform equilibrium) *The profile (σ, τ) is a uniform equilibrium with payoff v (for P1) if*

1. *for all n,*
$$\mathbb{E}^{(\sigma,\tau)}[\gamma_{SR_n}] \geq \mathbb{E}^{(\sigma,\tau')}[\gamma_{SR_n t}]$$
for all $\tau' = ((\tau_{m\neq n})_{s\in\mathbb{N}}, \tau'_n)$ holds;

2. *for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that*
$$\mathbb{E}^{(\sigma,\tau)}[\gamma_n] \geq \mathbb{E}^{(\sigma',\tau)}[\gamma_n] - \varepsilon \quad \forall n \geq N,$$
for all σ';

3. $\lim_{N\to\infty} \mathbb{E}^{(\sigma,\tau)}[\gamma_N] = v$.

[12] For notational convenience, I suppress the dependence on the prior p.

The notion of uniform equilibrium adapts the uniform value to the case with short-lived players.[13] Condition 1 states that the short-run players play a best-response. Condition 2 embodies the uniformity property: Player 1's strategy is ε-optimal in every finite but long enough game. Condition 3 requires that the average payoff for P1 converge to v.

This definition focuses on the payoff of the long-run player. This paper does not attempt to characterize the equilibrium sequences of short-run players' payoffs.

Let $\Gamma^{\mathrm{LR}}(p)$ be the game in which the sequence of uninformed players is replaced by a single long–run Player 2. Formally: A behavioral strategy ξ for Player 2 is a collection of maps $(\xi_n)_{n\in\mathbb{N}}$ where

$$\xi_n : H_n^2 \to \Delta(J).$$

Given a state k and a complete history h_N, the payoff to Player 2 is

$$-\gamma_N(k, h_N) = \frac{1}{N+1}\sum_{n=0}^{N} -G^k_{i_n j_n}.$$

Definition 2 (Uniform value (Aumann and Maschler, 1995, p.75)) *The game $\Gamma^{LR}(p)$ admits v_∞ as uniform value if there exists strategies (σ, ξ) such that for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that*

$$\mathbb{E}^{(\sigma,\xi')}[\gamma_n] + \varepsilon \geq v_\infty \geq \mathbb{E}^{(\sigma',\xi)}[\gamma_n] - \varepsilon,$$

for all $n \geq N$ and σ', ξ'.

Let

$$\begin{aligned} G(p) : \Delta(I)^K \times \Delta(J) &\to \mathbb{R} \\ (x, y) &\mapsto \sum_k p^k x^k G^k y. \end{aligned}$$

be the payoff function of the one-shot, incomplete information game. For every $X \subset \Delta(I)^K, Y \subset \Delta(J)$ the triple

$$\langle X, Y, G(p)\rangle$$

defines a zero-sum game. Denote its value by $\mathrm{val}_{X,Y} G(p)$, whenever it exists.[14]

For a function $f : \Delta(K) \to \mathbb{R}$ let $\mathrm{Cav}\, f$ denote the smallest concave function that majorizes f.

[13] See also Hart (1985).

[14] The value need not exist for all X, Y.

4 Results

In this section, I present the main results. Throughout, I compare them to the known results of the canonical game.

4.1 Perfect monitoring

In this section, the focus is on perfect monitoring. Monitoring is perfect if the signal q reveals the action by both players.

Assumption 1 *For all $(i, j), (i', j') \in I \times J$, $q_{ij} = q_{i',j'}$ implies $(i, j) = (i', j')$.*

Under perfect monitoring, the distinction between h_n^1, h_n^2 and h_n is solely one of notation. For the remainder of this subsection, I ignore the notational distinction and write h_n.

Under perfect monitoring, the informed player does not use his information if his play does not depend on the state k, i.e., $\sigma_n(k, h_n) = \sigma_n(k', h_n)$ for every k'. Call such a strategy non-revealing. Restricting himself to such strategies, the game is one of perfect information with payoff matrix $G(p) = \sum_k p^k G^k$. It is clear that Player 1 can guarantee himself a payoff equal to the value of this game.

More precisely, let

$$\mathrm{NR}(p) := \{x | p^k p^{k'} > 0 \implies x^k = x^{k'}\}.$$

A strategy $x \in X$ is non-revealing if and only if $x \in \mathrm{NR}(p)$. Aumann and Maschler (1995) define the non-revealing game to be $\langle \mathrm{NR}(p), \Delta(J), G(p) \rangle$. Denote its value, $\mathrm{val}_{\,\mathrm{NR}(p),\Delta(J)} G(p)$, by $u(p)$.

Aumann and Maschler (1995) show that the value of the canonical game is $\mathrm{Cav}\, u(p)$.

Theorem (Aumann and Maschler, 1995) *Assume monitoring is perfect. The uniform value v_∞ of $\Gamma^{LR}(p)$ is $\mathrm{Cav}\, u(p)$.*

The same result obtains if the informed player faces a sequence of short-lived players instead.

Theorem 1 *Assume monitoring is perfect. The unique uniform equilibrium payoff (to Player 1) of $\Gamma^{SR}(p)$ is $\mathrm{Cav}\, u(p)$.*

Theorem 1 says that the informed player gets the same payoff whether he faces a single long-run player or a sequence of short-run players. It is clear that any uniform equilibrium payoff in $\Gamma^{\mathrm{SR}}(p)$ must be at least v_∞. Equality is less obvious.

The intuition behind Theorem 1 is as follows. By using his information, P1 reveals it. Under perfect monitoring, the uninformed player's action does not influence the signal about P1's action. Hence, the information P1 reveals does not depend on the action of the uninformed player, and consequently, also not on whether the uninformed player is long- or short-lived. In particular, playing a myopic best-response does not influence the information revealed.

However, when P1 uses his information, he reveals some of it. P1 must stop using his information eventually, or he reveals it perfectly. Eventually, either he stops using it, or he fully reveals it. Thus, eventually, the game is akin to a game of complete information whose payoff is just the value of the associated stage game. The proof in the Appendix makes this intuition precise.

The next section presents the results for imperfect monitoring.

4.2 Imperfect monitoring

In this section, Assumption 1 is dropped. Instead, no assumption on the signalling matrix Q is made.

The analysis in this section proceeds as in the preceeding one. First, I discuss non–revealing strategies for imperfect monitoring structures. Second, I present the results for $\Gamma^{\text{LR}}(p)$. Third, I argue that the notion of a non-revealing strategy is not appropriate when the informed player faces a sequence of short-lived opponents. I propose the notion of a non–revealing payoff to address this issue. Lastly, I present the main result of this section, Theorem 2.

Informally, a strategy x of Player 1 in the game $\langle \Delta(I)^K, \Delta(J), G(p) \rangle$ is non-revealing if the signals observed by Player 2 do not let her draw inferences about the state k regardless of the action she plays. Formally, x is non–revealing if it is an element of

$$\text{NR}(p) := \{x | p^k p^{k'} > 0 \implies x^k Q_{\cdot j} = x^{k'} Q_{\cdot j} \text{ for all } j \in J\}.$$

Here, $zQ_{\cdot j}$ denotes the probability distribution over elements in the j-th column of Q for the probability distribution $z \in \Delta(I)$.

To illustrate the definition, consider the game in Figure 1.

A strategy $x = (x^1, x^2)$ is non–revealing at non-degenerate priors if and only if x^1 and x^2 put the same weight on the first row. Observe that for this signalling matrix, the set $\text{NR}(p)$ is the same as under perfect monitoring. However, this is not true in general.

With this definition, one can extend the value of the non–revealing game from the previous section. Let the value of the non-revealing game be $u(p) = \text{val}_{\,\text{NR}(p),\Delta(J)} G(p)$. The main result for games with two long-lived players is the following.

$$G^1 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} \qquad G^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix} \qquad Q = \begin{pmatrix} a & a & b \\ a & a & c \end{pmatrix}$$

Figure 1: A game in which the uninformed player has a strictly dominated action, and signals are uninformative if she chooses any of her other actions.

Theorem (Aumann and Maschler, 1995) *The uniform value v_∞ of $\Gamma^{LR}(p)$ is* $\operatorname{Cav} u(p)$*.*

The intuition behind this result is similar to the one under perfect monitoring: by using his information, the informed player gives the uninformed player the opportunity to learn it.

However, the theorem does not extend to the case in which the informed player faces short–lived opponents. As an illustration, consider the previous game, as shown in Figure 1. The value of the non–revealing game is $u(p) = p^1(1-p^1) = \operatorname{Cav} u(p)$. This is not an uniform equilibrium payoff in $\Gamma^{\mathrm{SR}}(p)$. Note that playing the third column is a strictly dominated strategy for the second player. Hence, a short-run player never plays the third column on the equilibrium path. The only signal on path is then a. Signals are completely uninformative. Consequently, the informed player can use his information without revealing it against short-lived opponents. The payoff to P1 in any uniform equilibrium of $\Gamma^{\mathrm{SR}}(p)$ is $\min\{p^1, 1-p^1\}$. It is strictly higher than the uniform value of $\Gamma^{\mathrm{LR}}(p)$ for non–degenerate priors.

The example illustrates that the notion of non–revealing play in $\Gamma^{\mathrm{LR}}(p)$ is too strong for the game with short-lived uninformed players. Intuitively, play is non–revealing in the presence of short–lived players when the informed player reveals no information given that the short-lived opponent plays a myopic best–response. The following definition captures this notion.

Definition 3 *A non–revealing payoff at the prior p is a number v^* such that*

1. $v^* = \sum_k p^k x^{*k} G^k y^*$*;*
2. $v^* \leq \sum_k p^k x^{*k} G^k y$ *for all* $y \in \Delta(J)$*;*
3. $v^* \geq \sum_k p^k x^k G^k y^*$ *for all x such that*

$$x^k Q_{\cdot j} = x^{k'} Q_{\cdot j} \text{ for all } j \in \operatorname{supp}(y^*) \text{ whenever } p^k p^{k'} > 0.$$

Say v^* is suppported on $J' \subset J$ if $\operatorname{supp}(y^*) = J'$*.*[15] *Let $v^{*k} = x^{*k} G^k y^*$.*

[15] $\operatorname{supp}(y)$ denotes the support of the mixed strategy $y \in \Delta(J)$.

The definition has the following interpretation. Condition 1 says that v^* is the expected payoff given the strategy profile (x^*, y^*) in the stage game. Condition 2 requires that y^* be a best-response against x^*. Condition 3 requires that x^* be a best-response against y^* among all strategies that yield the same distribution over public signals when the uninformed player uses y^*. This encompasses the idea that a strategy by the informed player must not reveal any information about the state against best-replies by the uninformed player.

Define the set of non–revealing strategies against $J' \subset J$ as

$$\mathrm{NR}(p, J') := \{x | p^k p^{k'} > 0 \implies x^k Q_{\cdot j} = x^{k'} Q_{\cdot j} \text{ for all } j \in J'\}.$$

This definition is the same as the one by Aumann and Maschler (1995) if the uninformed player is restricted to the subset J' of her actions. Clearly,

$$\mathrm{NR}(p, J) = \mathrm{NR}(p)$$

so that the two definitions of non–revealing coincide when there is no restriction on the uninformed player's strategies. Moreover, when monitoring is perfect, $\mathrm{NR}(p, J') = \mathrm{NR}(p)$ for all J'. The non–revealing payoff in games with perfect monitoring equals the value of the non–revealing game.

The following lemma relates the previous two definitions.

Lemma 1 *Let v^* be a non–revealing payoff at p supported on J'. Then*

$$v^* = \mathrm{val}_{\,\mathrm{NR}(p,J'),J'} G(p).$$

Observe that the reverse is not true: $\mathrm{val}_{\,\mathrm{NR}(p,J'),J'} G(p)$ need not be a non–revealing payoff for all subsets J'. The reason is that the best-response by the uninformed player need not lie in J'. Indeed, a non–revealing payoff need not exist for some priors p. Consider the game in Figure 2. In this game, a non–revealing payoff does not exist for $1/4 < p^1 < 3/4$. Intuitively, when the uninformed player chooses the left column, the informed player best-responds by using his information. He does so without revealing it. However, it is then no longer a best-response for the uninformed player to choose the left column. If the informed player's strategy is restricted to be independent of the state of the world, i.e., non–revealing in the sense of Aumann and Maschler (1995), the optimal strategies are such that the uninformed player chooses the left colum with probability 1.

I now state the main result of the section.

Theorem 2 *Assume*

1. *v^* is a non–revealing payoff at p supported on $J' \subset J$,*

2. $v^* = \mathrm{Cav}\,\mathrm{val}_{\,\mathrm{NR}(p,J'),\Delta(J')} G(p)$,

$$G^1 = \begin{pmatrix} 1 & 1/2 \\ -1 & 1/2 \end{pmatrix} \qquad G^2 = \begin{pmatrix} -1 & 1/2 \\ 1 & 1/2 \end{pmatrix} \qquad Q = \begin{pmatrix} a & b \\ a & c \end{pmatrix}$$

Figure 2: A game in which a non–revealing payoff does not exist for $1/4 < p^1 < 3/4$.

3. $\sum_k q^k v^{*k} \geq \operatorname{Cav} u(q)$ *for all* $q \in \Delta(K)$,

then v^* *is a uniform equilibrium payoff in* $\Gamma^{SR}(p)$.

Corollary 1 *If in addition to Assumptions 1–3 of Theorem 2,* $v^* > \operatorname{Cav} u(p)$ *holds, then there is a uniform equilibrium of* $\Gamma^{SR}(p)$ *in which the informed player receives a strictly higher payoff than in* $\Gamma^{LR}(p)$.

The conditions of Theorem 2 have the following interpretation. Condition 1 says that a non–revealing payoff exists. Condition 2 says that, in addition to not revealing any information when playing a strategy profile that yields a payoff of v^*, the informed player does not want to use his information in a way that is detectable by the uninformed player. Here, the informed player could use his information to concavify his payoffs in the restricted game in which the uninformed player chooses actions only in J', the support of v^*. Note that the informed player can guarantee himself a payoff of $\operatorname{Cav} \operatorname{val}_{\mathrm{NR}(p,J'),\Delta(J')} G(p)$ when the uninformed players restrict themselves to actions in J'.

Condition 3 states that v^* is an individually rational payoff of the informed player. It needs to be satisfied by any uniform equilibrium payoff.

The main bite in the assumptions of Theorem 2 lies in the existence of the non–revealing payoff and it being equal to the concavification of the value of the restricted game.

While Theorem 2 is an existence theorem, its interest comes from the corollary: it gives an easy way to check for uniform equilibrium payoffs in $\Gamma^{\mathrm{SR}}(p)$ that are higher than the value of $\Gamma^{\mathrm{LR}}(p)$. In such games, the informed player can be strictly better off facing a sequence of short–lived players instead of an equally patient long–lived player. The game depicted in Figure 1 is such a game: the unique uniform equilibrium payoff of $\Gamma^{\mathrm{SR}}(p)$ is strictly higher than the value of the canonical game.

Even if the assumptions of the Corollary are satisfied, one cannot conclude that the informed player is strictly better off playing against short-lived opponents. This is because $\Gamma^{\mathrm{SR}}(p)$ can have multiple uniform equilibria with different payoffs. The game in Figure 3 is an example of such a game.

The value of the non–revealing game in Figure 3 is concave in p^1 with $u(1/2) = 2/7$. Thus, the value of $\Gamma^{\mathrm{LR}}(1/2)$ is $2/7$. However, there are two non–revealing payoffs: $v^* = 2/7$ with full support and $v^* = 1/2$ with support on the left and

$$G^1 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 2 \\ \frac{1}{3} & \frac{1}{3} & -2 \end{pmatrix} \qquad G^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 2 \\ \frac{1}{3} & \frac{1}{3} & -2 \end{pmatrix} \qquad Q = \begin{pmatrix} a & a & b \\ a & a & c \\ a & a & d \end{pmatrix}$$

Figure 3: A game with multiple non–revealing payoffs for priors $p^1 \approx 1/2$.

middle column. Both non–revealing payoffs satisfy the assumptions of Theorem 2 and hence are uniform equilibrium payoffs of $\Gamma^{\mathrm{SR}}(1/2)$.

Theorem 2 gives sufficient condition for the existence of a uniform equilibrium in $\Gamma^{\mathrm{SR}}(p)$. However, the characterization is incomplete. Theorem 2 is mute on necessary conditions for uniform equilibrium payoffs. A complete characterization of uniform equilibrium payoffs would involve both necessary and sufficient conditions. I hope to obtain such a complete characterization in future work.

A special, yet important class of monitoring structures, are those in which the public signal is independent of the uninformed player's action. In such games, the informed player receives the same payoff in the game with short–run players as in the canonical game.

Assumption 2 *The monitoring structure satisfies, for all $i, i' \in I, j, j' \in J$,*

$$q_{ij} = q_{i'j} \implies q_{ij'} = q_{i'j'}.$$

Theorem 3 *Under Assumption 2, the payoff to Player 1 is* $\operatorname{Cav} u(p)$ *in any uniform equilibrium of* $\Gamma^{SR}(p)$.

The condition on the monitoring structure means that when the uninformed player cannot distinguish between two actions for some of her actions, she cannot do so for any of her other actions as well. The information she learns from a signal is independent of her own action. The uninformed players do not have an incentive to forego payoffs in the stage game in order to better monitor the actions of the informed player: the experimentation motive is absent.

With such a monitoring structure, the informed player's payoff in the game with short–run players is equal to the value of the canonical game. The intuition behind Theorem 3 is the same as for Theorem 1. The only difference is that P1 can use some of his information without revealing it, no matter the behavior of the uninformed players.

Theorem 3 emphasizes the role of experimentation and informational externalities between short–run players in different stages on the distinction of the canonical game and the game with short–run players. The payoff equivalence between the canonical game and the game with short–run players obtains when the experimentation motive is absent.

5 Discussion

This section discusses the results in the preceding section.

5.1 Convergence of finite and discounted games

For a strategy profile (σ,τ) let $u_\lambda(\sigma,\tau) = \mathbb{E}^{(\sigma,\tau)}\left[\sum_{n=0}^{\infty}\lambda(1-\lambda)^n G^k_{i_n j_n}\right]$. Together with the repeated game form in Section 3, u_λ defines a repeated game, the discounted game with short–lived players. Call this game $\Gamma^{\mathrm{SR}}_\lambda(p)$. Similarly, let $\Gamma^{\mathrm{SR}}_n(p)$ be the n–times repeated game with a long–lived informed player and short–lived uninformed players.

Let $\overline{v}_\lambda(p)$ ($\underline{v}_\lambda(p)$) be the supremum (infimum) of the Nash equilibrium payoffs to Player 1 in $\Gamma^{\mathrm{SR}}_\lambda(p)$.

Theorem 4 *Under Assumption 1, $\overline{v}_\lambda(p)$ and $\underline{v}_\lambda(p)$ converge as $\lambda \to 0$. Their common limit is* $\operatorname{Cav} u(p)$.

Moreover, the convergence holds for the finitely repeated game as well: Let $\overline{v}_n(p)$ ($\underline{v}_n(p)$) be the supremum (infimum) of the Nash equilibrium payoffs to Player 1 in $\Gamma^{\mathrm{SR}}_n(p)$.

Theorem 5 *Under Asumption 1, $\overline{v}_n(p)$ ($\underline{v}_n(p)$) converges to* $\operatorname{Cav} u(p)$ *as $n \to \infty$.*

These results resemble the familiar convergence results in games with two long–lived players. The extension of the preceeding theorems to imperfect monitoring is ongoing work.

5.2 Equilibrium notion

Instead of considering uniform equilibria, one could have defined payoffs on the normal form and then studied Nash equilibria of that game. For instance, fix a Banach limit L. Given strategies (σ,τ), define the payoff to Player 1 as

$$u(\sigma,\tau) = L[\{\mathbb{E}^{(\sigma,\tau)}[\gamma_N(k,h_N)]\}_N].$$

The payoff to SR_n is

$$u_n(\sigma,\tau) = \mathbb{E}^{(\sigma,\tau)}[\gamma_{\mathrm{SR}_n}(k,h_n)].$$

Call the game thus defined $\Gamma^{\mathrm{SR}}_L(p)$.[16] Then the main results obtain mutatis mutandis.

[16]Every Banach limit defines a different game. However, none of the statements depends on which Banach limit is chosen.

Theorem 6 *Assume*

1. v^* *is a non–revealing payoff at* p *supported on* $J' \subset J$,
2. $v^* = \operatorname{Cav}\operatorname{val}_{\mathrm{NR}(p,J'),\Delta(J')} G(p)$,
3. $\sum_k q^k v^{*k} \geq \operatorname{Cav} u(q)$ *for all* $q \in \Delta(K)$,

then v^* *is a payoff in* $\Gamma_L^{SR}(p)$.

5.3 Monitoring structures

Two implicit assumption on the monitoring structures were made. First, signals are publicly observed. Second, the signal structure is independent of the state k. Neither assumption is crucial for the results. Modifying the definition of non–revealing payoffs in a straightforward manner, Theorem 2 obtains under private and state–dependent signal structures.

5.4 Informed short–run players

Assume in this section that the short–run players observe the state. The long–lived player does not observe the state.

The next proposition partially characterizes equilibrium payoffs in this setting.

Proposition 1 *Maintain Assumption 1. Assume that there is no* $j \in J$ *such that* j *is strictly dominant in* G^k *and* $G^{k'}$, $k \neq k'$. *Then there exists an equilibrium in which the payoff to Player 1 is*

$$\sum_k p^k \operatorname{val} G^k.$$

The intuition behind Proposition 1 is as follows. Player 1 can learn the state from the informed short–run players' actions. During an initial phase, he plays strategies such that the short–run player's best–response depends on the realized state. The short–run players have no incentive to withhold information about the state by playing a myopic but uninformative best–response. Hence, Player 1 knows the state after finitely many stages. The game is then one of complete information.

The conclusion of Proposition 1 fails if either hypothesis do not hold. Consider the game depicted in Figure 4. Choosing the first column is a dominant action in both states. Hence, it is the only action chosen by the short–run players on the equilibrium path. Player 1 then obtains a payoff of $\max\{1+p^1, 2-p^1\}$. The value of both stage games is 2. Thus, the long–lived player receives a strictly lower payoff when he is not informed of the state.

$$G^1 = \begin{pmatrix} 2 & 3 \\ 1 & 3 \end{pmatrix} \qquad G^2 = \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix}$$

Figure 4: A game in which the uninformed long–lived player receives a payoff less than $\sum_k p^k \text{val}\, G^k$.

$$G^1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad G^2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad Q = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

Figure 5: An example of a game that shows that the assumption that short–run players observe the actions of preceeding short–run players is crucial.

A slight modification of the game depicted in Figure 4 shows that Proposition 1 is not a complete characterization of equilibrium payoffs. There exist games in which the hypotheses of Proposition 1 are satisfied, but the game has an equilibrium in which Player 1 receives a payoff strictly less than $\sum_k p^k \text{val}\, G^k$.

5.5 Non–observability of actions

Suppose that the short-run players observe the public signals from previous stages. However, they do not observe the actions by previous short–run players. The game in Figure 5 illustrates that this assumption is crucial.

The value of the non–revealing game is $u(p^1) = p^1(1-p^1)$. This is also the payoff in $\Gamma^{\text{SR}}(p)$, according to Theorem 3, in which short–run players observe actions by previous short–run players. If they do not, there is a uniform equilibrium with payoff equal to $1/2$ when the prior is $p^1 = 1/2$.

Appendix

Proof of Theorem 1. Assumption 1 is a special case of Assumption 2. Theorem 1 thus follows from Theorem 3. ■

Proof of Lemma 1. A straightforward rewriting of the definitions. ■

Proof of Theorem 2. The proof is by construction. Denote by $\Gamma^{\text{LR}}(p, J')$ the restriction of $\Gamma^{\text{LR}}(p)$ in which the uninformed player is restricted to actions in $J' \subset J$. Let τ^* be an optimal strategy of the uninformed player in $\Gamma^{\text{LR}}(p, J')$.

Define strategies as follows. Player 1 plays according to x^* after every history.

At stage n, SR_n plays as follows. If h_n^2 consists only of elements $j \in J', q \in \{q_{ij} \mid j \in J', i \in \mathrm{supp}(x^*)\}$, SR_n plays according to τ^* with probability $(1 - 1/n^2)$ and mixes equiprobably over J' with probability $1/n^2$. If h_n^2 contains $j \notin J'$ or $q \notin \{q_{ij} \mid j \in J', i \in \mathrm{supp}(x^*)\}$, SR_n plays according to a strategy for the uninformed player in $\Gamma^{\mathrm{LR}}(p)$ that defends the uniform value.

The payoff to Player 1 from this strategy profile is v^*. Since

$$v^* = \mathrm{Cav}\,\mathrm{val}_{\,\mathrm{NR}(p,J'),\Delta(J')} G(p),$$

this is the highest payoff Player 1 can achieve against the strategies of the short–run players. Moreover, it is easy to see that Condition 2 of Definition 1 is satisfied.

SR_n plays a best–response: on path, only actions in J' are played. However, $J' = \mathrm{supp}(y^*)$, and y^* is a best–response against x^* in the stage game. Hence, any $j \in J'$ is optimal. ■

Proof of Theorem 1.

The proof adapts a sequence of results that are well established.[17]

For the proof, I assume that the signal q_{ij} contains the action of the uninformed players, i.e., $\forall i \in I, q_{ij} = q_{ij'} \implies j = j'$. Define a equivalence relation on $I \times I$ as follows. $i \sim i'$ if $q_{ij} = q_{i'j}$ for some j. Under Assumption 2 it is without loss of generality to assume that $q_{ij} = ([q_i], j)$, where $[q_i]$ is the equivalence class containing i. Under these assumptions, the following holds. Let (σ', τ) be a uniform equilibrium of $\Gamma^{\mathrm{SR}}(p)$ with payoff v. Then there exists a uniform equilibrium (σ, τ) with payoff v such that σ is measurable with respect $\{K \times H_n\}$, i.e. σ depends only on the public history and the state k.

Fix a uniform equilibrium (σ, τ) such that σ is measurable with respect to the public history. For any public history h_n that has positive probability under $\mathbb{P}^{(\sigma,\tau)}$ let $\rho_n(h_n) \in \Delta(K)$ be the posterior probability. For histories that occur with zero probability, define $\rho_n(h_n) \in \Delta(K)$ arbitrarily. Note that the process $\rho = \{\rho_n\}$ is a martingale. Moreover, $\rho_{n+1}(h_n, [q_i], j) = \rho_{n+1}(h_n, [q_i], j')$ for all $j, j' \in J$ almost surely.

The local payoff to Player 1 at the history h_n is given by

$$U_n(\sigma, \tau)(h_n) = \sum_k \rho_n^k(h_n) \sigma_n(k, h_n) G^k \tau_n(h_n).$$

Let $y(\rho)$ be an optimal strategy of the column player in $\langle \mathrm{NR}(p), \Delta(J), G(\rho) \rangle$. Denote $\overline{\sigma}_n(k, h_n) = \sum_k \rho_n^k(h_n) \sigma_n(k', h_n)$. Observe that $\overline{\sigma}_n(k', h_n)$ is indepedent of k. Let $\overline{\sigma}_n(h_n)$ be the strategy that plays $\overline{\sigma}_n(k', h_n)$ irrespective of the state k.

[17] See, for example, Sorin (2002), Propositions 3.2, 3.3, 3.5, Lemmata 3.4, 3.5, 3.13.

Fix a history h_n on path. Let $||G|| = \max_{i,j,k} G^k_{ij}$. Then

$$
\begin{aligned}
& U_n(\rho_n(h_n)) = \\
& \sum_k \rho^k_n(h_n)\sigma_n(k,h_n)G^k\tau_n(h_n) \\
& \le \sum_k \rho^k_n(h_n)\sigma_n(k,h_n)G^k y(\rho_n(h_n)) \\
& = \sum_k \rho^k_n(h_n)(\sigma_n(k,h_n) - \overline{\sigma}_n(h_n) + \overline{\sigma}_n(h_n))G^k\tau'(\rho_n(h_n)) \\
& \le u(\rho_n(h_n)) + \sum_k \rho^k_n(h_n)(\sigma_n(k,h_n) - \overline{\sigma}_n(h_n))G^k y(\rho_n(h_n)) \\
& \le u(\rho_n(h_n)) + \sum_k \rho^k_n(h_n) \sum_{[q_i]} \left| \sum_{i\in[q_i]} \sigma_n(k,h_n)(i) - \overline{\sigma}_n(h_n)(i) \right| ||G||.
\end{aligned}
$$

Here, $\sigma_n(h_n)(i)$ is the i-th entry of the vector $\sigma_n(h_n)$, i.e., the probablity that $\sigma_n(h_n)$ assigns to the row i. The first inequality follows because SR_n plays a best-response. The second inequality holds because $\overline{\sigma}_n(h_n) \in \text{NR}(p)$.

The posterior probabilities evolve according to

$$
\rho^k_{n+1}(h_n,[q_i]) = \frac{\sum_{i\in[q_i]} \rho^k_n(h_n)\sigma_n(k,h_n)(i)}{\sum_k \sum_{i\in[q_i]} \rho^k_n(h_n)\sigma_n(k,h_n)(i)} = \frac{\sum_{i\in[q_i]} \rho^k_n(h_n)\sigma_n(k,h_n)(i)}{\sum_{i\in[q_i]} \overline{\sigma}_n(h_n)(i)},
$$

where the last equality follows from the definition of $\overline{\sigma}_n(h_n)$.

It is:

$$
\begin{aligned}
& \sum_k \rho^k_n(h_n) \sum_{[q_i]} \left| \sum_{i\in[q_i]} \sigma_n(k,h_n)(i) - \overline{\sigma}_n(h_n)(i) \right| \\
& = \sum_{[q_i]} \sum_k \left| \sum_{i\in[q_i]} \rho^k_n(h_n)(\sigma_n(k,h_n)(i) - \overline{\sigma}_n(h_n)(i)) \right| \\
& = \sum_{[q_i]} \left(\sum_{i\in[q_i]} \overline{\sigma}_n(h_n)(i) \right) \sum_k |\rho^k_{n+1}(h_n,[q_i]) - \rho^k_n(h_n)| \\
& = \sum_{[q_i]} \left(\sum_{i\in[q_i]} \overline{\sigma}_n(h_n)(i) \right) ||\rho_{n+1}(h_n,[q_i]) - \rho_n(h_n)||_1.
\end{aligned}
$$

Hence, the local payoff to Player 1 at h_n is bounded by

$$U_n(\rho_n(h_n)) \leq u(\rho_n(h_n)) + ||G|| \sum_{[q_i]} \left(\sum_{i \in [q_i]} \overline{\sigma}_n(h_n)(i) \right) ||\rho_{n+1}(h_n, [q_i]) - \rho_n(h_n)||_1.$$

The payoff to Player 1 can then be upper bounded as follows.

$$\begin{aligned}
&\sum_{n=0}^{N-1} \mathbb{E}^{(\sigma,\tau)}[U_n(\rho_n(h_n))] \\
&\leq \sum_{n=0}^{N-1} \mathbb{E}^{(\sigma,\tau)} \left[u(\rho_n(h_n)) + ||G|| \sum_{[q_i]} \left(\sum_{i \in [q_i]} \overline{\sigma}_n(h_n)(i) \right) ||\rho_{n+1}(h_n, [q_i]) - \rho_n(h_n)||_1 \right] \\
&\leq \sum_{n=0}^{N-1} \mathbb{E}^{(\sigma,\tau)} [\mathrm{Cav}\, u(\rho_n(h_n))] + ||G|| \sum_{n=0}^{N-1} \mathbb{E}^{(\sigma,\tau)} [||\rho_{n+1} - \rho_n||_1] \\
&\leq \sum_{n=0}^{N-1} \mathrm{Cav}\, u \left(\mathbb{E}^{(\sigma,\tau)}(\rho_n(h_n)) \right) + ||G|| \sum_{n=0}^{N-1} \sum_k \mathbb{E}^{(\sigma,\tau)} \left[|\rho^k_{n+1} - \rho^k_n| \right] \\
&\leq N \mathrm{Cav}\, u(p) + ||G|| \sum_k \sqrt{p^k(1-p^k)} \sqrt{N}.
\end{aligned}$$

The first inequality uses the bound from the previous display. The second inequality uses that $\mathrm{Cav}\, u$ majorizes u. The third inequality is a consequence of Jensen's inequality. The last inequality uses the fact that $\{\rho_n\}$ is a martingale and Lemma 2. The claim follows. ■

Lemma 2 *For a martingale* $\mathbf{q} = \{q_n\}$ *with values in* $[0, 1]$ *and* q_1 *almost surely constant it holds that*

$$\mathbb{E} \left[\sum_{n=1}^{m} |q_{n+1} - q_n| \right] \leq \sqrt{q_1(1-q_1)} \sqrt{m}.$$

Proof. See Sorin (2002, p. 32). ■

Proof of Theorem 4. The proof of Theorem 4 follows the logic of the proof of Theorem 3. It uses Lemma 3. The details are omitted. ■

Lemma 3 *For a martingale* $\mathbf{q} = \{q_n\}$ *with values in* $[0, 1]$ *and* q_1 *almost surely constant it holds that*

$$\mathbb{E} \left[\sum_{n=1}^{\infty} \lambda(1-\lambda)^{n-1} |q_{n+1} - q_n| \right] \leq \sqrt{q_1(1-q_1)} \sqrt{\frac{\lambda}{2-\lambda}}.$$

Proof. See Sorin (2002, p. 32). ■

Proof of Theorem 5. A direct consequence of the proof of Theorem 3. ■

Proof of Theorem 6. The proof is a minor variation of the proof of Theorem 2, and therefore omitted. ■

Proof of Proposition 1. I begin by stating, without proof, the following

Claim *Assume that there is no $j \in J$ such that j is striclty dominant in G^k and $G^{k'}$ for some $k \neq k'$. Then, for any $k \neq k'$, there exists $x_{kk'} \in \Delta(I)$ such that there exist $j^k_{kk'} \neq j^{k'}_{kk'} \in J$ such that*

$$j^k_{kk'} \in \arg\min_{y \in \Delta(J)} x_{kk'} G^k y$$

and

$$j^{k'}_{kk'} \in \arg\min_{y \in \Delta(J)} x_{kk'} G^{k'} y.$$

Enumerate the pairs $(k, k'), k \neq k'$ such that number assigned to (k, k') equals the one corresponding to (k', k). Denote by $e(k, k')$ the number assigned to the pair (k, k'). Define a strategy for Player 1 as follows. In the first $1/2K(K-1)$ stages, play the strategy $x_{kk'}$ when the stage is $e(k, k')$. For all stages $n \geq 1/2K(K-1)$: if there is a k such that $j^k_{kk'}$ was played in stage $e(k, k')$ for all $k' \neq k$, play an optimal strategy in G^k; otherwise play an optimal strategy in G^1.

The strategy by SR_n is as follows. For $n = e(k, k')$, play $J^k_{kk'}$ is the state is k, $j^{k'}_{kk'}$ if the state is k', and an arbitrary element of $\arg\min_{y \in \Delta(J)} x_{kk'} G^{k''} y$ if the state is $k'' \neq k, k'$. For $n \geq 1/2K(K-1)$, SR_n plays an optimal strategy in G^k when the state is k.

By definition of $j^k_{kk'}$, SR_n plays a best–response. It is straightforward to verify that the strategy by Player 1 is a best–response as well. ■

References

Aumann, Robert J. and Michael B. Maschler (1966*). "Game theoretic aspects of gradual disarmament". In: *Report of the U.S. Arms Control and Disarmament Agency ST-80*. Washington, D.C. Chap. V, pp. V1–V55.

– (1967*). "Repeated games with incomplete information: a survey of recent results". In: *Report of the U.S. Arms Control and Disarmament Agency ST-116*. Washington, D.C. Chap. III, pp. 287–403.

– (1968*). "Repeated games of incomplete information: the zero-sum extensive case". In: *Report of the U.S. Arms Control and Disarmament Agency ST-143*. Washington, D.C. Chap. III, pp. 37–116.

Aumann, Robert J. and Michael B. Maschler with Richard E. Stearns (1995). *Repeated Games with Incomplete Information*. Cambridge, MA, London, England: The MIT Press.

Cripps, Martin W. and Jonathan P. Thomas (2003). “Some Asymptotic Results in Discounted Repeated Games of One–Sided Incomplete Information”. In: *Mathematics of Operations Research* 28.3, pp. 433–462.

Forges, Françoise (1992). “Repeated Games of Incomplete Information: Non–Zero–Sum”. In: *Handbook of Game Theory with Economic Applications*. Ed. by Robert J. Aumann and Sergiu Hart. Vol. 1. Amsterdam: North–Holland: Elsevier Science Publishers B.V. Chap. 6, pp. 155–177.

Fudenberg, Drew, David M. Kreps, and Eric S. Maskin (1990). “Repeated Games with Long-run and Short-run Players”. In: *Review of Economic Studies* 57, pp. 555–573.

Fudenberg, Drew and David Levine (1989). “Reputation and equilibrium selection in games with a patient player”. In: *Econometrica* 57, pp. 759–778.

– (1992). “Maintaining a reputation when strategies are imperfectly observed”. In: *The Review of Economic Studies* 59, pp. 561–579.

Gensbittel, Fabien (2015). “Extensions of the Cav(u) Theorem for Repeated Games with Incomplete Information on One Side”. In: *Mathematics of Operations Research* 40.1, pp. 80–104.

Glosten, Lawrence and Paul Milgrom (1985). “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders”. In: *Journal of Financial Economics* 14, pp. 71–100.

Hart, Sergiu (1985). “Nonzero-Sum Two-Person Repeated Games with Incomplete Information”. In: *Mathematics of Operations Research* 10.1, pp. 117–153.

Kreps, David M., Paul Milgrom, et al. (1982). “Rational cooperation in the finitely repeated prisoners’ dilemma”. In: *Journal of Economic Theory* 27, pp. 245–252.

Kreps, David M. and Robert Wilson (1982). “Reputation and Imperfect Information”. In: *Journal of Economic Theory* 27, pp. 253–279.

Lehrer, Ehud and Leeat Yariv (1999). “Repeated Games with Incomplete Information on one Side: The Case of Different Discount Factors”. In: *Mathematics of Operations Research* 24.1, pp. 204–218.

Mailath, George J. and Larry Samuelson (2001). “Who wants a good reputation?” In: *The Review of Economic Studies* 68.2, pp. 415–441.

– (2015). “Reputations in Repeated Games”. In: *Handbook of Game Theory with Economic Applications*. Ed. by H. Peyton Young and Shmuel Zamir. Vol. 4. Amsterdam: North–Holland: Elsevier Science Publishers B.V. Chap. 4, pp. 165–238.

Milgrom, Paul and John Roberts (1982). “Predation, Reputation and Entry Deterrence”. In: *Journal of Economic Theory* 27, pp. 280–312.

Sorin, Sylvain (2002). *A First Course on Zero-Sum Repeated Games.* Paris, Berlin, Heidelberg: Springer.

[17]An asterix behind the year denotes work that has not been published.