

Stochastics and Statistics

# Selection of a correlated equilibrium in Markov stopping games

David M. Ramsey<sup>a,1</sup>, Krzysztof Szajowski<sup>a,b,\*</sup>

<sup>a</sup> *Institute of Mathematics and Computer Science, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, Wrocław, Poland*

<sup>b</sup> *Department of Statistics, 150 N. University Street, Purdue University, West Lafayette, IN 47907-2068, USA*

Received 22 March 2006; accepted 31 October 2006

Available online 2 January 2007

## Abstract

This paper deals with an extension of the concept of correlated strategies to Markov stopping games. The Nash equilibrium approach to solving nonzero-sum stopping games may give multiple solutions. An arbitrator can suggest to each player the decision to be applied at each stage based on a joint distribution over the players' decisions according to some optimality criterion. This is a form of equilibrium selection. Examples of correlated equilibria in nonzero-sum games related to the best choice problem are given. Several concepts of criteria for selecting a correlated equilibrium are used. © 2006 Elsevier B.V. All rights reserved.

**Keywords:** Stopping time; Markov process; Non-zero sum stopping game; Nash equilibrium; Correlated equilibrium; Secretary problem; Bilateral best choice problem

## 1. Introduction

This paper is devoted to two-person, non-zero sum games based on bilateral stopping of a Markov process. Let  $(X_n, \mathcal{F}_n, \mathbf{P}_x)_{n=0}^N$ ,  $N \in \mathbb{N} \cup \{\infty\}$ , be a homogeneous Markov process defined on a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  with state space  $(\mathbb{E}, \mathcal{B})$ . The following decision problem was considered by Szajowski [51] and Neumann et al. [28]. At each moment  $n$ ,  $n \in \{1, 2, \dots, N\}$ , the decision makers (henceforth called Player 1 and Player 2) are able to observe the Markov chain sequentially. Each player has his own utility function  $G_i : \mathbb{E} \rightarrow \mathfrak{R}$ ,  $i = 1, 2$ , and at each moment  $n$  each decides separately whether to accept or reject the realization  $x_n$  of  $X_n$ . It is assumed that if both players select the same moment  $n$  to accept  $x_n$ , then a lottery decides which player gets the right (priority) of acceptance. According to the lottery at moment  $\tau$ , Player 1 is chosen with probability  $\alpha_\tau$  and Player 2 with probability  $\beta_\tau = 1 - \alpha_\tau$ . The player rejected by the lottery may select any other realization  $x_n$  at a later moment  $n$ ,  $\tau < n \leq N$ . Once accepted a realization cannot be rejected, once rejected it cannot be reconsidered. If a player has not chosen any realization of the Markov process, he gets

\* Corresponding author. Present address: Institute of Mathematics and Computer Science, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, Wrocław, Poland. Tel.: +48 71 3203185; fax: +48 71 328 07 51.

E-mail addresses: [david.ramsey@ul.ie](mailto:david.ramsey@ul.ie) (D.M. Ramsey), [Krzysztof.Szajowski@pwr.wroc.pl](mailto:Krzysztof.Szajowski@pwr.wroc.pl) (K. Szajowski).

URL: <http://neyman.im.pwr.wroc.pl/~szajow> (K. Szajowski).

<sup>1</sup> Present address: Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland.

$G_i^* = \inf_{x \in E} G_i(x)$ . The aim of each player is to choose a realization in such a way as to maximize his expected utility. The problem is formulated as a two-person non-zero sum game. In such a model the players choose their strategies based on their knowledge about the payoff functions and the sets of strategies available to them.

Reviews of alternative approaches have been made by Szajowski and Nowak [32] and Sakaguchi [41]. Examples of such games related to the secretary problem or a version of poker have been solved by Ben Abdelaziz and Krichen [2], Enns and Ferenstein [6], Mazalov [21] and Sakaguchi [37,39,40,38,42]. Different aspects of  $N$  person games, assuming players decide about stopping by majority voting, have been investigated by Kurano, Yasuda, Nakagami [20,56], Szajowski and Yasuda [52] and Ferguson [10].

The concept of Nash equilibrium may lead to multiple solutions and the problem of equilibrium selection appears. The multiplicity of Nash equilibria in strategic form games restricts the predictive power of the concept of Nash equilibrium (see [34]). Many attempts have been made in the literature to reduce the set of (Nash) equilibria either by only considering equilibria which satisfy stricter conditions (refinements), or by making a convincing unique choice out of the set of Nash equilibria (equilibrium selection). The equilibrium selection problem is discussed by Herings and Peeters [18,19]. Some authors start from any refinement and choose an equilibrium using a lottery over all the refined equilibria with the same payoff vector. The method they use to find such lotteries is borrowed from Nash bargaining theory (see [34,35,53]). On the other hand, models of extensions of games with pre-play or intra-play information and communication allow us to define new notions of equilibria. The relevant question is how the outcomes change when communication between players is allowed, or when they are given some kind of pre-play information. This aspect has been discussed by Sorin [49]. Correlated equilibria have been considered, among other things, in non-zero sum matrix games (see [1,7,23,26]), sequential games (see [12,25]), stochastic games (see [47,29,31]) and differential games (see [30,54]). When correlated strategies are adopted, the players may obtain advice from an arbitrator as to which actions should be taken. There is communication between this arbitrator and the players, which is described later.

It is well known (see [15]) that a Nash equilibrium is a complicated concept of solution from a computational point of view, whereas the concept of correlated equilibrium is simple. Correlation means a certain limitation of the freedom of the players in selecting their pure actions, because some process of pre-play communication is needed to realize a correlated strategy. However, any player is free to choose any pure action, regardless of the results of the communication process. The aim of this paper is to adopt this approach to nonzero-sum stopping games.

Section 2.1 gives a review of concepts of solutions to stopping games. Section 2.2 introduces the main concepts of the article: correlated stopping times and the definition of a correlated equilibrium. It is shown that the set of correlated equilibria is non-empty for the model considered. Section 2.3 considers the problem of equilibrium selection and presents some criteria that may be used to select an equilibrium. Section 3 reviews the idea of random priority in stopping games. Section 4 considers the general approach to selecting a correlated equilibrium in a game based on the best choice problem. Section 5 gives examples of equilibrium selection in this game using various selection criteria. It is not the aim of this paper to give a comprehensive guide to such criteria, since many such criteria may be used. However, such a comparison may well be a subject of future research.

## 2. Concepts of a solution in a stopping game

### 2.1. Nash equilibria in stopping games

First of all, basic definitions in the field of optimal stopping of stochastic processes are presented. Next, randomized stopping times are defined and the concept of Nash equilibrium for stopping games is given. Finally, the concept of correlated stopping times is defined.

In the problem of optimal stopping, the basic class of strategies  $\mathcal{T}^N$  are Markov times with respect to  $\sigma$ -fields  $\{\mathcal{F}_n\}_{n=0}^N$ . We permit that  $\mathbf{P}(\tau \leq N) < 1$  for some  $\tau \in \mathcal{T}^N$  when  $N \in \mathbb{N}$  or  $\mathbf{P}(\tau \leq \infty) < 1$  when  $N = \infty$ , which means that it is possible that no state is chosen. Stopping games need a wider class of strategies (see

[55]). In such games randomized stopping times are used. It is assumed that the probability space is rich enough to permit the following constructions.

**Definition 2.1** (see [55]). A random sequence  $p = (p_n)$  such that, for each  $n$ ,

- (i)  $p_n$  is adapted to  $\mathcal{F}_n$ ;
- (ii)  $0 \leq p_n \leq 1$  a.s.

is called a randomized stopping time. The set of all such sequences will be denoted by  $\mathcal{P}^N$ .

**Remark 2.1.** Given a strategy  $p = (p_n) \in \mathcal{P}^N$  and sequence  $A_0, A_1, \dots, A_n$  of i.i.d. r.v. with uniform distribution on  $[0, 1]$  independent of the Markov process  $(X_n, \mathcal{F}_n, \mathbf{P}_x)_{n=0}^N$ ,  $\lambda(p) = \inf\{0 \leq n < N + 1 : A_n \leq p_n\}$  is a stopping time with respect to the  $\sigma$ -fields  $\mathcal{H}_n = \sigma\{\mathcal{F}_n, \{A_0, A_1, \dots, A_n\}\}$ .

If each random variable in  $p$  only takes the values 0 or 1, the strategy is called a pure strategy. In this case the Markov time  $\tau(p)$  is a Markov time with respect to the filtration  $\{\mathcal{F}_n\}_{n=0}^N$  (a pure stopping time). In particular an  $\{\mathcal{F}_n\}$ -Markov time  $\lambda$  corresponds to the strategy  $p = (p_n)$  with  $p_n = \mathbb{1}_{\{\lambda=n\}}$ , where  $\mathbb{1}_A$  is the indicator function for the set  $A$ .

Two players observe a homogeneous Markov process  $(X_n, \mathcal{F}_n, \mathbf{P}_x)_{n=0}^N$  sequentially. Their gain functions  $\mathbf{G}_i : \{0, 1, \dots, N\} \times \mathbb{E} \times \mathbb{E} \rightarrow \mathfrak{R}$  are measurable, bounded, and their strategy sets are the collections of stopping times  $\mathcal{T}_i^N$ ,  $i = 1, 2$ . If  $\tau_1 \in \mathcal{T}_1^N$  and  $\tau_2 \in \mathcal{T}_2^N$  are chosen, the gain of the  $i$ th player is  $G_i(\tau_1, \tau_2) = \mathbf{G}_i(\tau_1 \wedge \tau_2, X_{\tau_1}, X_{\tau_2})$ .

**Definition 2.2**

- (i) A Nash equilibrium in the game  $\mathcal{G}_s = (G_1, G_2, \mathcal{T}_1^N, \mathcal{T}_2^N)$  is a pair  $(\tau_1^*, \tau_2^*) \in \mathcal{T}_1^N \times \mathcal{T}_2^N$  such that for all  $x \in \mathbb{E}$

$$v_1(x) = \mathbf{E}_x G_1(\tau_1^*, \tau_2^*) \geq \mathbf{E}_x G_1(\tau_1, \tau_2^*),$$

$$v_2(x) = \mathbf{E}_x G_2(\tau_1^*, \tau_2^*) \geq \mathbf{E}_x G_2(\tau_1^*, \tau_2),$$

for every  $\tau_1 \in \mathcal{T}_1^N$  and for every  $\tau_2 \in \mathcal{T}_2^N$ .

- (ii) The strategy sets in the game  $\mathcal{G}_s$  can be extended to sets of randomized stopping times  $\mathcal{Q}_1^N \times \mathcal{Q}_2^N$ . A Nash equilibrium  $(q_1^*, q_2^*) \in \mathcal{Q}_1^N \times \mathcal{Q}_2^N$  in the mixed extension  $\mathcal{G}_m = (\tilde{G}_1, \tilde{G}_2, \mathcal{Q}_1^N, \mathcal{Q}_2^N)$  of  $\mathcal{G}_s$ , where  $\tilde{G}_i : \mathcal{Q}_1^N \times \mathcal{Q}_2^N \rightarrow \mathfrak{R}$  is defined as

$$\tilde{G}_i(q_1, q_2) = G_i(\tau(q_1), \tau(q_2)) \tag{1}$$

for  $i = 1, 2$ , such that for all  $x \in \mathbb{E}$

$$\tilde{v}_1(x) = \mathbf{E}_x \tilde{G}_1(q_1^*, q_2^*) \geq \mathbf{E}_x \tilde{G}_1(q_1, q_2^*),$$

$$\tilde{v}_2(x) = \mathbf{E}_x \tilde{G}_2(q_1^*, q_2^*) \geq \mathbf{E}_x \tilde{G}_2(q_1^*, q_2),$$

for every  $q_1 \in \mathcal{Q}_1^N$  and for every  $q_2 \in \mathcal{Q}_2^N$ .

For further consideration, the restricted sets of strategies  $\mathcal{T}_{i,n}^N = \{\tau \in \mathcal{T}_i^N : \tau \geq n\}$ ,  $i = 1, 2$ , are defined. These sets of strategies are used in the recursive construction of a solution when the horizon is finite.

**Definition 2.3.** A pair  $(\tau_{1,n}^*, \tau_{2,n}^*) \in \mathcal{T}_{1,n}^N \times \mathcal{T}_{2,n}^N$  is called an equilibrium point of  $\mathcal{G}_s$  at  $n$ , if for every  $\tau_{1,n} \in \mathcal{T}_{1,n}^N$  and for every  $\tau_{2,n} \in \mathcal{T}_{2,n}^N$

$$v_1(n, X_n) = \mathbf{E}_{X_n} G_1(\tau_{1,n}^*, \tau_{2,n}^*) \geq \mathbf{E}_{X_n} G_1(\tau_{1,n}, \tau_{2,n}^*) \quad \mathbf{P}_x\text{---a.s.},$$

$$v_2(n, X_n) = \mathbf{E}_{X_n} G_2(\tau_{1,n}^*, \tau_{2,n}^*) \geq \mathbf{E}_{X_n} G_2(\tau_{1,n}^*, \tau_{2,n}) \quad \mathbf{P}_x\text{---a.s.}$$

A Nash equilibrium point at  $n = 0$  is a solution of  $\mathcal{G}_s$ . The pair  $(v_1(0, x), v_2(0, x))$  of values is the Nash value corresponding to  $(\tau_{1,0}^*, \tau_{2,0}^*) = (\tau_1^*, \tau_2^*) \in \mathcal{T}_1^N \times \mathcal{T}_2^N$ .

Much research on the solution described by Definition 2.2(i) has been carried out. Let us recall the result by Ohtsubo [33]. Six measurable, bounded functions  $\phi_i, \psi_i$  and  $h_i, i = 1, 2$  on  $\mathbb{E}$  are given. When the  $i$ th player applies stopping time  $\tau_i \in \mathcal{T}_i^N, i = 1, 2$ , his reward function is

$$\mathbf{E}_x[G_i(\tau_1, \tau_2)] = \mathbf{E}_x[\phi(X_{\tau_i})\mathbb{1}_{(\tau_i < \tau_j)} + \psi(X_{\tau_j})\mathbb{1}_{(\tau_j < \tau_i)} + h_i(X_{\tau_i})\mathbb{1}_{(\tau_i = \tau_j < \infty)} + \limsup_{n \rightarrow \infty} h_i(X_{\tau_i = \tau_j = n})],$$

where  $j = 1, 2, j \neq i, x \in \mathbb{E}$ .

**Theorem 2.1** (see [33]). *Let us assume that  $\psi_i \leq h_i \leq \phi_i$  or  $\phi_i \leq h_i \leq \psi_i$  on  $\mathbb{E}$ . There exists a Nash equilibrium  $(\tau_1^*, \tau_2^*)$  of the game  $\mathcal{G}_s$ .*

**Remark 2.2.** Ohtsubo’s proof in [33] gives the construction of a sequence of functions which tends to a Nash value. The Nash value enables the definition of a Nash equilibrium point. In the finite horizon case an equilibrium point and its corresponding Nash value can be calculated by a recursive procedure based on the restricted sets of strategies described immediately above Definition 2.3.

In general, there may be no solutions described by Definition 2.2(i) in a stopping game where the given relation between payoff functions  $\phi_i, \psi_i$  and  $h_i, i = 1, 2$  does not hold. However, in the mixed extension  $\mathcal{G}_m$  of the game (see Definition 2.2(ii)), when the gain is defined by (1), the class of stopping games having a solution is wider. The solution of the randomized extension of the game considered by Ohtsubo in [33] has been given by Ferenstein [8].

A general two-person non-zero-sum stopping game has been solved by Shmaya et al. [44], who proved the existence of an  $\epsilon$ -equilibrium for such games. The existence of a Nash equilibrium in a multi-player infinite stage stopping game has been proved by Solan and Vieille [46]. The game considered by Fushimi [13] has been extended to finite horizon stopping games with randomized strategies on a Markov process by Szajowski [51]. The following definition is an extension of the randomized stopping time used in [51] (see also [5,9,6,36,44]).

Let  $\mathcal{Q}_{i,n}^N = \{q = (q_n) \in \mathcal{Q} : q_0 = \dots = q_{n-1} = 0\}$  for  $i = 1, 2$ . If the horizon is finite and  $q_N = 1$  by assumption, then the strategy defines a randomized stopping time which is finite with probability 1.

**Definition 2.4.** A pair  $(q_{1,n}^*, q_{2,n}^*) \in \mathcal{Q}_{1,n}^N \times \mathcal{Q}_{2,n}^N$  is called an equilibrium point of  $\mathcal{G}_m$  at  $n$ , if for every  $q_{1,n} \in \mathcal{Q}_{1,n}^N$  and for every  $q_{2,n} \in \mathcal{Q}_{2,n}^N$

$$\begin{aligned} \tilde{v}_1(n, X_n) &= \mathbf{E}_{X_n} \tilde{G}_1(q_{1,n}^*, q_{2,n}^*) \geq \mathbf{E}_{X_n} \tilde{G}_1(q_{1,n}, q_{2,n}^*) \quad \mathbf{P}_x\text{-a.s.}, \\ \tilde{v}_2(n, X_n) &= \mathbf{E}_{X_n} \tilde{G}_2(q_{1,n}^*, q_{2,n}^*) \geq \mathbf{E}_{X_n} \tilde{G}_2(q_{1,n}^*, q_{2,n}) \quad \mathbf{P}_x\text{-a.s.} \end{aligned}$$

A Nash equilibrium point at  $n = 0$  is a solution of  $\mathcal{G}_m$ . The pair  $(\tilde{v}_1(0, x), \tilde{v}_2(0, x))$  of values is the Nash value corresponding to  $(q_{1,0}^*, q_{2,0}^*) = (q_1^*, q_2^*) \in \mathcal{Q}_1^N \times \mathcal{Q}_2^N$ .

Let the gain functions  $\mathbf{G}_i : \{0, 1, \dots, N\} \times \mathbb{E} \times \mathbb{E} \rightarrow \mathfrak{R}, i = 1, 2$ , be measurable and bounded. Also, the players observe the homogeneous Markov process  $(X_n, \mathcal{F}_n, \mathbf{P}_x)_{n=0}^N$ . The following theorem is a consequence of the results from [51].

**Theorem 2.2.** *There exists a Nash equilibrium  $(q_1^*, q_2^*)$  in the game  $\mathcal{G}_m$  with finite horizon. An equilibrium point and its corresponding Nash value can be calculated by a recursive procedure.*

### 2.2. Correlated equilibria in stopping games

The concept of Nash equilibrium leads to multiple equilibria in stopping games, as in matrix games. These equilibria may give different values of the game to the players. There exists a problem regarding the selection of an equilibrium acceptable to both players. In 1974 Aumann [1] introduced a correlation scheme in randomized strategies for non-zero-sum games extending the concept of Nash equilibrium. Using this approach some process of pre-play communication and correlation mechanisms are needed to realize such a strategy. The idea of correlated equilibria for sequential games has been developed by Forges [12] and Myerson [25]. Various generalisations of correlation devices have been introduced (see [48]). It should be emphasized that in sequen-

tial games pre-play communication can occur either before the start of the game or at each stage before each sequential decision. Communication mechanisms can receive some private message from each player at each stage, and choose for each player a private signal for that stage (*communication device*, see [11,12,25,22]), or send one private signal before the beginning of the game (*correlation device*, see [11]). Private signals may be based on the current state (*weak correlation devices*, see [29]) or on previous signals only (*autonomous correlation devices*, [11]). These approaches are not equivalent (see [11,24,45]). The process of adapting correlated equilibria to stopping games starts with the idea of correlated stopping times. It should be noted that this definition assumes that a weak communication device is used i.e. that players observe a signal at each moment of the game and this signal depends only on the state of the process and not on any of the previous signals. One might consider correlated stopping times which are realized with the aid of a correlation device. In this case, the correlation device is used to choose two stopping times from the set of randomized stopping times. This difference is illustrated in Section 5.4.

**Definition 2.5.** A random sequence  $\hat{q} = \{(q_n^1, q_n^2, q_n^3)\}$  such that, for each  $n$ ,

- (i)  $q_n^i$  is adapted to  $\mathcal{F}_n$  for  $i = 1, 2, 3$ ;
- (ii)  $0 \leq q_n^1 \leq q_n^2 \leq q_n^3 \leq 1$  a.s.

is called a correlated stopping strategy. The set of all such sequences will be denoted by  $\hat{\mathcal{Q}}^N$ .

**Remark 2.3.** Let  $A_1, A_2, \dots, A_N$  be a sequence of i.i.d. r.v. with uniform distribution on  $[0, 1]$  and independent of the Markov process  $(X_n, \mathcal{F}_n, \mathbf{P}_x)_{n=0}^N$ . Denote  $\vec{q}_n = (q_n^1, q_n^2, q_n^3)$ . Correlated stopping times are pairs  $(\lambda^1(\hat{q}), \lambda^2(\hat{q}))$  of Markov times with respect to the  $\sigma$ -fields  $\mathcal{H}_n = \sigma\{\mathcal{F}_n, \{A_0, A_1, \dots, A_n\}\}$  defined by the strategy  $\hat{q} = (\vec{q}_n) \in \hat{\mathcal{Q}}^N$  as follows:

$$\lambda^1(\hat{q}) = \inf\{0 \leq n \leq N : A_n \leq q_n^2\} \tag{2}$$

and

$$\lambda^2(\hat{q}) = \inf\{0 \leq n \leq N : A_n \leq q_n^1 \text{ or } q_n^2 < A_n \leq q_n^3\}. \tag{3}$$

The strategy  $\hat{q}$  will be called the correlation profile and it defines the pair  $(\lambda^1(\hat{q}), \lambda^2(\hat{q}))$ .

In intuitive terms, the vector  $\vec{q}_n = (q_n^1, q_n^2, q_n^3)$  defines the joint distribution of the actions taken by the players at moment  $n$ : with probability  $q_n^1$  both players choose the action “stop”, with probability  $q_n^2 - q_n^1$  Player 1 stops and Player 2 chooses the action “continue”, with probability  $q_n^3 - q_n^2$  Player 1 continues and Player 2 stops and with probability  $1 - q_n^3$  both players continue. A correlated strategy  $\hat{q}$  is assumed to be defined by pre-play communication between the players with the possible aid of an “external judge”. We consider general communication devices. The form of the correlation strategy is known to both players.

If at a given moment the distribution over the pairs of actions available to the players defining a correlated strategy is not concentrated on just one pair of actions, a lottery carried out by the external judge is required to define the pair of actions to be taken. In such a case, this external judge chooses the pair of actions to be taken using an appropriate randomization device and communicates to each player separately the action he should take.

In specific cases, where a player knows the action suggested to the other player given the action suggested to him, such a correlation can be achieved by the players jointly observing the result of an appropriate randomization. This occurs, for example, when the correlation is concentrated on the pairs of actions  $(s, c)$  (Player 1 stops and Player 2 continues) and  $(c, s)$  (Player 1 continues and Player 2 stops).

**Remark 2.4.** It is assumed that the players are using a weak correlation device and continue to receive signals from this device, even after one of them has decided to deviate and play independently. Suppose Player  $j$  deviates from the strategy agreed upon. He may wish to follow the recommendation of the judge in some states (the state is understood as  $\mathcal{F}_n$ ), but decide to ignore the recommendation in others. In states where he decides to deviate from the judge’s recommendation, he uses a mixed strategy over the set of possible actions  $\{c, s\}$ . In this case the actions taken by the players are independent and the probability that the other player,

Player  $-j$ , plays  $c$  is simply the marginal probability that he plays  $c$  according to the correlated strategy agreed upon.

Assume that the players agree on the correlated strategy  $\hat{q}$  and the pair of correlated stopping times associated with this strategy are  $(\lambda^1(\hat{q}), \lambda^2(\hat{q}))$ . In the case Player  $j$  wishes to deviate from this strategy, it follows from the above argument that given his strategy we can define the resulting correlated strategy,  $\hat{q}_j$ , as well as the pair of correlated stopping times associated with this strategy,  $(\lambda^1(\hat{q}_j), \lambda^2(\hat{q}_j))$ . Denote  $\overline{G}_i(\hat{q}) = G_i(\lambda^1(\hat{q}) \wedge \lambda^2(\hat{q}), X_{\lambda^1(\hat{q}) \wedge \lambda^2(\hat{q})})$  and  $\overline{G}_i(\hat{q}_j) = G_i(\lambda^1(\hat{q}_j) \wedge \lambda^2(\hat{q}_j), X_{\lambda^1(\hat{q}_j) \wedge \lambda^2(\hat{q}_j)})$ . The expected payoffs are defined as  $\widehat{G}_i(x, \hat{q}) = \mathbf{E}_x \overline{G}_i(\hat{q})$  and  $\widehat{G}_i(x, (\hat{q}_j)) = \mathbf{E}_x \overline{G}_i(\hat{q}_j)$ , respectively.

If each random vector in  $\hat{q}$  takes values in the set  $\mathcal{S} = \{(0, 0, 0), (0, 0, 1), (0, 1, 1), (1, 1, 1)\}$ , then such a strategy is called a pure strategy. In this case the pair of random variables  $\lambda^1(\hat{q}) = \inf\{n : \vec{q}_n \in \{(0, 1, 1), (1, 1, 1)\}\}$  and  $\lambda^2(\hat{q}) = \inf\{n : \vec{q}_n \in \{(0, 0, 1), (1, 1, 1)\}\}$  are Markov times with respect to the filtration  $\{\mathcal{F}_n\}_{n=0}^N$  (pure stopping times for the players). When restriction is made to the subset  $\mathcal{S}^{eq} = \{(0, 0, 0), (1, 1, 1)\}$ , then the Markov times  $\lambda^1(\hat{q})$  and  $\lambda^2(\hat{q})$  are equal ( $\lambda^1(\hat{q}) = \lambda^2(\hat{q})$ ). In particular an  $\{\mathcal{F}_n\}$ -Markov time  $\lambda$  corresponds to the strategy  $\hat{q} = (\vec{q}_n)$  with  $q_n^i = \mathbb{1}_{\{\lambda=n\}}$ ,  $i = 1, 2, 3, n = 0, 1, \dots, N$ .

**Remark 2.5.** Let  $0 \leq a_n \leq 1$  and  $0 \leq b_n \leq 1$  be such that  $\vec{q}_n = (q_n^1, q_n^2, q_n^3) = (a_n b_n, a_n, a_n + b_n - a_n b_n)$  for  $n = 0, 1, \dots, N$ . Then  $\lambda^1(\hat{q})$  and  $\lambda^2(\hat{q})$  are independent random variables. If  $a_n, b_n \in \{0, 1\}$  then  $\lambda^1(\hat{q})$  and  $\lambda^2(\hat{q})$  are pure stopping times. In particular an  $\{\mathcal{F}_n\}$ -Markov time  $\lambda$  corresponds to the strategy  $\hat{q} = (\vec{q}_n)$  with  $q_n^i = \mathbb{1}_{\{\lambda=n\}}$ ,  $i = 1, 2, 3, n = 0, 1, \dots, N$ .

Each correlated stopping strategy  $\pi = \{\pi_0, \pi_1, \dots, \pi_N\}$  can be presented as

$$\{(\pi_n^{ss}, \pi_n^{sc}, \pi_n^{cs}, \pi_n^{cc})\}_{n=0}^N = \{(q_n^1, q_n^2 - q_n^1, q_n^3 - q_n^2, 1 - q_n^3)\}_{n=0}^N \tag{4}$$

Denote  $\widehat{\mathcal{Q}}_n^N = \{\hat{q} \in \widehat{\mathcal{Q}}^N : q_k^1 = q_k^2 = q_k^3 = 0, k = 0, 1, \dots, n-1, q_n^1 = q_n^2 = q_n^3 = 1\}$ . The policy  $\hat{q} \in \widehat{\mathcal{Q}}_n^N$  will be denoted  $\hat{q}^{(n)}$  and will be called a restricted strategy. The game associated with such a strategy will be called a restricted game.

**Definition 2.6.** A correlated stopping strategy  $\hat{q}^* \in \widehat{\mathcal{Q}}^N$  is called a correlated equilibrium point of  $\mathcal{G}_m$ , if

$$\widehat{G}_i(x, \hat{q}^*) \geq \widehat{G}_i(x, (\hat{q}_i^*)) \tag{5}$$

for every  $x \in \mathbb{E}$ ,  $\hat{q}_i^*$  and  $i = 1, 2$ . This means a player cannot increase his expected reward by switching his decision from the one suggested to him by the arbitrator.

For a given  $\hat{q}^{(n)} \in \widehat{\mathcal{Q}}_n^N$  define

$$\begin{aligned} u_i^{ss}(n, x, \hat{q}^{(n+1)}) &= \mathbf{E}_{X_n} \overline{G}_i(\hat{q}^{(n)}) \mathbb{1}_{\{\lambda^1(\hat{q}^{(n)}) = \lambda^2(\hat{q}^{(n)}) = n\}} \mathbb{1}_{\{X_n = x\}}, \\ u_i^{sc}(n, x, \hat{q}^{(n+1)}) &= \mathbf{E}_{X_n} \overline{G}_i(\hat{q}^{(n)}) \mathbb{1}_{\{\lambda^1(\hat{q}^{(n)}) = n, \lambda^2(\hat{q}^{(n)}) > n\}} \mathbb{1}_{\{X_n = x\}}, \\ u_i^{cs}(n, x, \hat{q}^{(n+1)}) &= \mathbf{E}_{X_n} \overline{G}_i(\hat{q}^{(n)}) \mathbb{1}_{\{\lambda^1(\hat{q}^{(n)}) > n, \lambda^2(\hat{q}^{(n)}) = n\}} \mathbb{1}_{\{X_n = x\}}, \\ u_i^{cc}(n, x, \hat{q}^{(n+1)}) &= \mathbf{E}_{X_n} \overline{G}_i(\hat{q}^{(n)}) \mathbb{1}_{\{\lambda^1(\hat{q}^{(n)}) > n, \lambda^2(\hat{q}^{(n)}) > n\}} \mathbb{1}_{\{X_n = x\}}. \end{aligned} \tag{6}$$

For a given correlated profile, at each stage  $n \in \{0, 1, \dots, N\}$ , the players observe the payoffs in the bimatrix game defined by  $(u_1^{ij}, u_2^{ij})_{i,j \in A}$  defined by (6), where  $A = \{s, c\}$ . Based on the concept of a correlated equilibrium for such a bimatrix game, one can define rational behaviour at stage  $n$  and a correlated equilibrium using a weak correlation device.

**Definition 2.7.** A correlated profile  $\hat{q}^* \in \widehat{\mathcal{Q}}^N$  is called a correlated rational strategy of  $\mathcal{G}_m$ , if every restriction  $\hat{q}^{*(n)}$  of  $\hat{q}^*$  to  $\widehat{\mathcal{Q}}_n^N$ ,  $n = 0, 1, \dots, N$ , fulfills

$$\sum_{j \in A} u_1^{sj}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*sj} \geq \sum_{j \in A} u_1^{cj}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*sj} \quad \text{on } \{\omega : \lambda^1(\hat{q}^{*(n)}) = n\}, \tag{7a}$$

$$\sum_{j \in A} u_1^{cj}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*cj} \geq \sum_{j \in A} u_1^{sj}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*cj} \quad \text{on } \{\omega : \lambda^1(\hat{q}^{*(n)}) > n\}, \tag{7b}$$



$$\sum_{j \in A} u_2^{js}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*js} \geq \sum_{j \in A} u_2^{jc}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*jc} \quad \text{on } \{\omega : \lambda^2(\hat{q}^{*(n)}) = n\}, \tag{7c}$$

$$\sum_{j \in A} u_2^{jc}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*jc} \geq \sum_{j \in A} u_2^{js}(n, X_n, \hat{q}^{*(n+1)}) \pi_n^{*js} \quad \text{on } \{\omega : \lambda^2(\hat{q}^{*(n)}) > n\}. \tag{7d}$$

**Theorem 2.3.** A correlated rational strategy  $\hat{q}^*$  for the finite horizon game defined by Definition 2.7 is equivalent to a correlated equilibrium point as defined in Definition 2.6.

**Proof.** The equivalence of these definitions will be proved by recursion. Define

$$\widehat{G}_i^{\mathcal{F}^n}(x, \hat{q}) = \mathbf{E}_x[\overline{G}_i(\hat{q}) | \mathcal{F}_n]. \tag{8}$$

From the properties of conditional expectation  $\widehat{G}_i^{\mathcal{F}^n}(x, \hat{q}) = \mathbf{E}_{X_n} \overline{G}_i(\hat{q}^{(n)}) = \widehat{G}_i(X_n, \hat{q}^{(n)})$ . For  $n = N$ , from the form of a correlated equilibrium of a bimatrix game, condition (5) is equivalent to the conditions in Definition 2.7. Let us assume that the equivalence is established for  $n + 1, n + 2, \dots, N$ . Suppose that at moment  $n$  Player  $j$  wishes to deviate from  $\hat{q}^*$  and thereafter plays according to  $\hat{q}^*$ . In order to consider the stability of  $\hat{q}^*$ , we may assume that he takes the action that was not recommended to him. The resulting correlated strategy is denoted  $\hat{q}_j$ . Clearly,  $\hat{q}_j^{*(n+1)} = \hat{q}^{*(n+1)}$ .

On  $\{\omega : \lambda^1(\hat{q}^{*(n)}) = n\}$  by (8)

$$\widehat{G}_1(X_n, \hat{q}^{*(n)}) = \sum_{j \in A} u_1^{sj}(n, X_n, \hat{q}^{*(n+1)}) \overline{\pi}_n^{*sj}, \tag{9a}$$

$$\widehat{G}_1(X_n, \hat{q}_1^*) = \sum_{j \in A} u_1^{cj}(n, X_n, \hat{q}^{*(n+1)}) \overline{\pi}_n^{*sj}, \tag{9b}$$

where  $\overline{\pi}_n^{*sj} = \frac{\pi_n^{*sj}}{\pi_n^{*ss} + \pi_n^{*sc}}$  and

$$\widehat{G}_2(X_n, \hat{q}^{*(n)}) = \sum_{j \in A} u_2^{js}(n, X_n, \hat{q}^{*(n+1)}) \underline{\pi}_n^{*js}, \tag{10a}$$

$$\widehat{G}_2(X_n, \hat{q}_2^*) = \sum_{j \in A} u_2^{jc}(n, X_n, \hat{q}^{*(n+1)}) \underline{\pi}_n^{*js}, \tag{10b}$$

where  $\underline{\pi}_n^{*js} = \frac{\pi_n^{*js}}{\pi_n^{*ss} + \pi_n^{*cs}}$ , for some correlated strategy  $\hat{q}^*$  (the convention  $\frac{0}{0} = 0$  is adopted). Let us assume that  $\hat{q}^*$  fulfills (5) for  $j = 1$ , then from (9a) and (9b) condition (7a) must be satisfied. Similarly, for  $j = 2$ , from (10a) and (10b) condition (7c) must be satisfied.

On  $\{\omega : \lambda^1(\hat{q}^{*(n)}) > n, \lambda^2(\hat{q}^{*(n)}) = n\}$  it follows from the definition of a correlated equilibrium that conditions (7b) and (7c) are satisfied for Player 1 and 2, respectively. On  $\{\omega : \lambda^1(\hat{q}^{*(n)}) = n, \lambda^2(\hat{q}^{*(n)}) > n\}$  this definition implies (7a) and (7d) for Player 1 and 2, respectively. Finally, on  $\{\omega : \lambda^1(\hat{q}^{*(n)}) > n, \lambda^2(\hat{q}^{*(n)}) > n\}$  this definition implies (7c) and (7d) for Player 1 and 2, respectively. If (7a) and (7d) are fulfilled on  $\mathcal{Q}_n^N$ , then (5) is fulfilled on  $\{\omega : \lambda^1(\hat{q}^{*(n)}) \geq n, \lambda^2(\hat{q}^{*(n)}) \geq n\}$ . It follows that, on  $\mathcal{Q}_n^N$ , the strategy  $\hat{q}^*$  fulfills Definition 2.7, iff it fulfills Definition 2.6. By induction these definitions are equivalent.  $\square$

A correlated strategy constitutes a correlated equilibrium, if at each moment Player  $i$  cannot gain by carrying out a different action to the action suggested to him, when the other player chooses his action from the appropriate conditional distribution given the action suggested to Player  $i$ .

We treat such solutions as solutions to a noncooperative game. Thus, players are assumed not to make any binding agreements or threats regarding reactions to any breaking of an agreement.

**Definition 2.8.** The set of all correlated equilibrium points is denoted  $\mathbb{CE}$ . For  $\hat{q}^* \in \mathbb{CE}$  we define the correlated value of the game to the players as  $\hat{v}_i(x, \hat{q}^*) = \mathbf{E}_x \overline{G}_i(\hat{q}^*)$  for  $i = 1, 2$ .

**Theorem 2.4.** The set of correlated equilibrium points  $\mathbb{CE}$  for the finite horizon stopping games with bounded payoffs is not empty.

**Proof.** This follows from the fact that any Nash equilibrium satisfies the conditions of a correlated equilibrium.  $\square$

This may also be proven directly using recursion by showing that there exists  $\hat{q}^{*(n)}$  which satisfies Conditions (7a) and (7d) for all possible realizations of  $X_n$  for  $n = N, N - 1, \dots, 0$ .

**Remark 2.6.** A proof of the existence of Nash equilibria in two player finite horizon stopping games has been recursively constructed based on the existence of equilibria in two-by-two bimatrix games (see [51]). Such solutions of two-by-two bimatrix games are used in the construction of correlated equilibria in the finite horizon stopping game. The structure of correlated solutions are described in many papers. Forges [11] has shown that the set of payoffs associated with this solution concept is a convex polyhedron. The relation between Nash equilibria and correlated equilibria in bimatrix games has been investigated by Hendrickx et al. [17], Nau et al. [27] and in two-by-two bimatrix games by Calvó-Armengol [4]. However, these results are not implemented in the construction of solution for stopping game considered here.

### 2.3. Selection of a correlated rational equilibria

We consider selection from the following four classes of strategy:

- (i) The class of pure equilibria.
- (ii) The class of mixed equilibria.
- (iii) The class obtained by randomization over mixed equilibria.
- (iv) The class of correlated equilibria.

The first three of these classes are subclasses of the class of correlated equilibria.

When selecting an equilibrium from a given class of equilibria, players communicate at the beginning of the game to decide which strategies should be used. When selecting from the set of pure equilibria no further communication takes place and randomization devices cannot be used. Selecting an equilibrium from the set of mixed equilibria, no further communication takes place, but players can independently use a randomization device to choose their actions at each stage. Selecting from the class obtained by randomization over mixed equilibria, at each stage the players can jointly observe the result of a lottery to decide which mixed equilibrium should be played. Finally, selecting from the class of correlated equilibria, an arbitrator may be used. The arbitrator chooses the actions to be taken according to the appropriate joint distribution and informs each player individually which action he should take.

Hence, in practical terms pure strategies are the easiest class of strategy to play and correlated equilibria are the most complex. However, the mathematical properties of the set of correlated equilibrium often make the problem of selection more tractable.

Theory for selecting an equilibrium in stochastic games has been developed e.g. in [18]. Since the set of Nash equilibria is a subset of the set of correlated equilibria, it is clear that whenever the problem of the selection of a Nash equilibrium exists, the problem of the selection of a correlated equilibrium also exists. However, the notion of correlated equilibrium assumes that communication takes place. Such communication can be used to make the aims of the players precise and to define the criteria used by players to select a correlated equilibrium. We now formulate various criteria for selecting correlated equilibria. These criteria select subsets of  $\mathbb{C}\mathbb{E}$ . The concepts which are used here do not come from the concepts of solution to Nash's problem of cooperative bargaining, but were used by Greenwald and Hall [16] for computer learning of equilibria in Markov games. These four concepts have been chosen by us because they describe natural aims or behaviour of two decision makers in real life. There are many concepts of equilibrium selection and it is beyond the scope of this article to formulate a general method or theory of correlated equilibria selection. It is a contribution to the discussion about selection of solutions in models of competitive decision processes.

**Definition 2.9.** Let us formulate four different selection criteria for correlated equilibria in a stopping game. Denote  $\Pi_n^\bullet = \{\bar{\pi}_n : (\bar{\pi}_n, \hat{q}_\bullet^{*(n+1)}) = \hat{q}_\bullet^{*(n)}\}$ .



(SU) A Stepwise Utilitarian Correlated Equilibrium  $\hat{q}_U^* = \{(\pi_n^{*ss}, \pi_n^{*sc}, \pi_n^{*cs}, \pi_n^{*cc})\}_{n=0}^N$  is an equilibrium  $\hat{q} = \{(\pi_n^{ss}, \pi_n^{sc}, \pi_n^{cs}, \pi_n^{cc})\}_{n=0}^N \in \mathbb{CE}$  such that for every  $n < N + 1$ , the sum of the values of the restricted game to the players is maximized, given this strategy is used at stages  $k, n < k < N + 1$ .

$$\max_{\bar{\pi}_n \in \Pi_n^U} \sum_{\gamma \delta \in B} \sum_{i=1}^2 \pi_n^{\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_U^{*(n+1)}) = \sum_{\gamma \delta \in B} \sum_{i=1}^2 \pi_n^{*\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_U^{*(n+1)}) \quad \text{on } \{\omega : X_n = x\}, \tag{11}$$

where  $B = \{ss, sc, cs, cc\}$ .

(SE) A Stepwise Egalitarian Correlated Equilibrium  $\hat{q}_E^* = \{(\pi_n^{*ss}, \pi_n^{*sc}, \pi_n^{*cs}, \pi_n^{*cc})\}_{n=0}^N$  is an equilibrium  $\hat{q} = \{(\pi_n^{ss}, \pi_n^{sc}, \pi_n^{cs}, \pi_n^{cc})\}_{n=0}^N \in \mathbb{CE}$  such that for every  $n < N + 1$ , the minimum value is maximized, given this strategy is used at stages  $k, n < k < N + 1$ ,

$$\max_{\bar{\pi}_n \in \Pi_n^E} \min_{i \in \{1,2\}} \sum_{\gamma \delta \in B} \pi_n^{\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_E^{*(n+1)}) = \min_{i \in \{1,2\}} \sum_{\gamma \delta \in B} \pi_n^{*\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_E^{*(n+1)}) \quad \text{on } \{\omega : X_n = x\}. \tag{12}$$

(SR) A Stepwise Republican Correlated Equilibrium  $\hat{q}_R^* = \{(\pi_n^{*ss}, \pi_n^{*sc}, \pi_n^{*cs}, \pi_n^{*cc})\}_{n=0}^N$  is an equilibrium  $\hat{q} = \{(\pi_n^{ss}, \pi_n^{sc}, \pi_n^{cs}, \pi_n^{cc})\}_{n=0}^N \in \mathbb{CE}$  such that for every  $n < N + 1$  the maximum value of the restricted game is maximized given this strategy is used at stages  $k, n < k < N + 1$ .

$$\max_{\bar{\pi}_n \in \Pi_n^R} \max_{i \in \{1,2\}} \sum_{\gamma \delta \in B} \pi_n^{\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_R^{*(n+1)}) = \max_{i \in \{1,2\}} \sum_{\gamma \delta \in B} \pi_n^{*\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_R^{*(n+1)}) \quad \text{on } \{\omega : X_n = x\}. \tag{13}$$

(SL- $i$ ) A Stepwise Libertarian  $i$  Correlated Equilibrium  $\hat{q}_L^* = \{(\pi_n^{*ss}, \pi_n^{*sc}, \pi_n^{*cs}, \pi_n^{*cc})\}_{n=0}^N$  is an equilibrium  $\hat{q} = \{(\pi_n^{ss}, \pi_n^{sc}, \pi_n^{cs}, \pi_n^{cc})\}_{n=0}^N \in \mathbb{CE}$  such that for every  $n < N + 1$  the value of the restricted game to Player  $i$  is maximized, given this strategy is used at stages  $k, n < k < N + 1$ .

$$\max_{\bar{\pi}_n \in \Pi_n^L} \sum_{\gamma \delta \in B} \pi_n^{\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_L^{*(n+1)}) = \sum_{\gamma \delta \in B} \pi_n^{*\gamma \delta} u_i^{\gamma \delta}(n, x, \hat{q}_L^{*(n+1)}) \quad \text{on } \{\omega : X_n = x\}. \tag{14}$$

If the horizon is finite, then a correlated equilibrium which fulfills one of the criterion (11)–(14) can be constructed recursively.

A solution which fulfills one of these stepwise criterion does not have to be optimal according to the corresponding global criterion. This is illustrated in the example below in which the players assess the value of the objects differently and further explained in Section 5.4, which considers egalitarian equilibria. In the latter case, the objective function is non-linear and hence stepwise optimization does not necessarily lead to global optimization, even when players have the same assessment of the value of an object. In the case of the Libertarian  $i$  criterion, it follows from the properties of dynamic programming that a global optimal correlated strategy must be a stepwise optimal strategy. However, there may not be a unique stepwise optimal strategy and in this case a stepwise optimal strategy is not necessarily globally optimal. This is illustrated in Section 5.3, which considers utilitarian equilibria.

**Example 2.7.** Let the driven process  $\zeta_t$  of the game have three states 1, 2, 3 and the game will be played by two players at  $t = 1, 2, 3$ . The process starts at  $t = 1$  with state  $\zeta_1 = 1$ , and takes  $\zeta_2 = 2$  and  $\zeta_3 = 3$ . The utility functions  $G_t(t, i, j)$  are results of adaptation of the individual utility function defined on the state space to evaluation of the observed states of the process and equal priority. The individual evaluation of the states are  $\{3.6; 2.2; 4\}$  by Player 1 and  $\{2.6; 2.4; 4\}$  by Player 2 for states  $\{1, 2, 3\}$ , respectively. The stepwise utilitarian value according to Definition 2.9 is  $(4, 2.6)$ . Global optimization over all correlated strategies leads to the correlated value  $(3.6, 4)$ .

For bounded payoff functions the set of randomized equilibria is not empty (see [9,33]). This implies that the set of correlated equilibria is non-empty, convex and compact. Based on these facts we come to the following conclusion.

**Theorem 2.5.** *The set of correlated equilibrium points satisfying any one of the given criteria above is not empty.*

Using these criteria in the finite horizon case, the appropriate correlated equilibria can be obtained by recursively solving a set of linear programming problems. In the case of libertarian and utilitarian equilibria, the linear objective functions are given by Eqs. (14) and (11), respectively. The constraints are those used in Definition 2.7. Since the feasible set of this linear programming problem is non-empty (a correlated rational strategy always exists), such solutions always exist. In the case of the republican equilibrium, at each stage of the recursion we can solve the following two linear programming problems:

- (1) The maximization of the value of the game to Player 1 subject to the constraints from Definition 2.7, together with the constraint that the value of the game to Player 1 is at least the value of the game to Player 2.
- (2) The maximization of the value of the game to Player 2 subject to the constraints from Definition 2.7, together with the constraint that the value of the game to Player 2 is at least the value of the game to Player 1.

In order to find a correlated strategy satisfying the republican criterion, it suffices to choose an appropriate solution from the solutions to these two problems (maximizing the maximum value). It should be noted that the union of the feasible sets of these two linear programming problems is the set of correlated rational strategies. Hence, at least one of these feasible sets must be non-empty. In the case where only one feasible set is non-empty, the solution of the corresponding linear programming problem maximizes the maximum value.

An analogical procedure using two linear programming problems can be used to find a correlated equilibrium satisfying the egalitarian criterion. When the objective function is defined by the maximization of the value of the game to Player  $i$ , then the additional constraint in the linear programming problem is that the value of the game to Player  $i$  is not greater than the value of the game to the other Player.

It should be noted that although the existence of correlated equilibria of the types given above is guaranteed, it is possible that such equilibria are not unique.

### 3. Random priority in Markov stopping games

The problem of bilateral stopping of a Markov process can be used to model the decision process in which two decision makers observe a Markov sequence sequentially to choose a state. If at some moment both would like to accept the same state, then a random device selects one of them (see [28,51]). Let us recall the mathematical formulation of the problem.

Let  $(X_n, \mathcal{F}_n, \mathbf{P}_x)_{n=0}^N$  be a homogeneous Markov process defined on a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  with a state space  $(\mathbb{E}, \mathcal{B})$ . At each moment  $n = 0, 1, \dots, N$  the decision makers (henceforth called Player 1 and Player 2) are able to observe the Markov chain sequentially. Each player has his own utility function  $G_i : \mathbb{E} \rightarrow \mathfrak{R}$ ,  $i = 1, 2$ , and at each moment  $n$  each decides separately whether to accept or reject the realization  $x_n$  of  $X_n$ . We assume the  $G_i$  are measurable and bounded. If it happens that both players have selected the same moment  $n$  to accept  $x_n$ , then a lottery decides which player gets the right (priority) of acceptance. According to the lottery at moment  $\tau$ , Player 1 is chosen with probability  $\alpha_\tau$  and Player 2 with probability  $\beta_\tau = 1 - \alpha_\tau$ . The player rejected by the lottery may select any other realization  $x_n$  at any later moment  $n$ ,  $\tau < n \leq N$ . Once accepted a realization cannot be rejected, once rejected it cannot be reconsidered. If a player does not choose any realization of the Markov process, he gets  $G_i^* = \inf_{x \in \mathbb{E}} G_i(x)$ . The aim of each player is to choose a realization which maximizes his expected utility. As in Section 2.2, correlated equilibria are adapted to this problem.

$\mathcal{S}_k^N = \{\tau \in \mathcal{S}^N : \tau \geq k\}$ , where  $\mathcal{S}^N$  is the set of stopping times with respect to  $\{\mathcal{F}_n\}_{n=0}^N$ . One can define the set of strategies  $\tilde{\mathbf{\Lambda}}^N = \{(\lambda^1(\hat{q}), \{\sigma_n^1\}) : \hat{q} \in \mathcal{Q}^N, \{\sigma_n^1\} \in \mathcal{S}_{n+1}^N \text{ for every } n\}$  and  $\tilde{\mathbf{M}}^N = \{(\lambda^2(\hat{q}), \{\sigma_n^2\}) : \hat{q} \in \mathcal{Q}^N, \{\sigma_n^2\} \in \mathcal{S}_{n+1}^N \text{ for every } n\}$  for Player 1 and Player 2, respectively.

Let  $\xi_0, \xi_1, \dots, \xi_n$  be i.i.d.r.v. uniformly distributed on  $[0, 1]$  and independent of  $\bigvee_{n=0}^N \mathcal{H}_n$  and the lottery be given by  $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_N)$ . Denote  $\tilde{\mathcal{H}}_n = \sigma\{\mathcal{H}_n, \xi_0, \xi_1, \dots, \xi_n\}$ .

**Definition 3.1.** For every pair  $(s, t) \in \tilde{\mathbf{\Lambda}}^N \times \tilde{\mathbf{M}}^N$  define

$$\tau_1(s, t) = \lambda^1(\hat{q}) \mathbb{1}_{\{\lambda^1(\hat{q}) < \lambda^2(\hat{q})\}} + (\lambda^1(\hat{q}) \mathbb{1}_{\{\xi_{\lambda^1(\hat{q})} \leq \alpha_{\lambda^1(\hat{q})}\}} + \sigma_{\lambda^2(\hat{q})}^1 \mathbb{1}_{\{\xi_{\lambda^1(\hat{q})} > \alpha_{\lambda^1(\hat{q})}\}}) \mathbb{1}_{\{\lambda^1(\hat{q}) = \lambda^2(\hat{q})\}} + \sigma_{\lambda^2(\hat{q})}^1 \mathbb{1}_{\{\lambda^1(\hat{q}) > \lambda^2(\hat{q})\}}$$

and

$$\tau_2(\mathbf{s}, \mathbf{t}) = \lambda^2(\hat{q}) \mathbb{1}_{\{\lambda^1(\hat{q}) > \lambda^2(\hat{q})\}} + (\lambda^2(\hat{q}) \mathbb{1}_{\{\xi_{i,2}(\hat{q}) > \alpha_{i,2}(\hat{q})\}} + \sigma_{\lambda^1(\hat{q})}^2 \mathbb{1}_{\{\xi_{i,2}(\hat{q}) \leq \alpha_{i,2}(\hat{q})\}}) \mathbb{1}_{\{\lambda^1(\hat{q}) = \lambda^2(\hat{q})\}} + \sigma_{\lambda^1(\hat{q})}^2 \mathbb{1}_{\{\lambda^1(\hat{q}) < \lambda^2(\hat{q})\}}.$$

The Markov times  $\tau_1(\mathbf{s}, \mathbf{t})$  and  $\tau_2(\mathbf{s}, \mathbf{t})$  are the selection times of Player 1 and Player 2 when they use strategies  $\mathbf{s} \in \tilde{\mathbf{A}}$  and  $\mathbf{t} \in \tilde{\mathbf{M}}$ , respectively, and the lottery is  $\bar{\alpha}$ .

For each  $(\mathbf{s}, \mathbf{t}) \in \tilde{\mathbf{A}}^N \times \tilde{\mathbf{M}}^N$  and given  $\bar{\alpha}$  the payoff function for the  $i$ th player is defined as  $f_i(\mathbf{s}, \mathbf{t}) = G_i(X_{\tau_i(\mathbf{s}, \mathbf{t})})$ . Let  $\tilde{R}_i(x, \mathbf{s}, \mathbf{t}) = \mathbf{E}_x f_i(\mathbf{s}, \mathbf{t}) = \mathbf{E}_x G_i(X_{\tau_i(\mathbf{s}, \mathbf{t})})$  be the expected gain of Player  $i$ , when the players use  $(\mathbf{s}, \mathbf{t})$ .

Let  $h_i(n, X_n) = \text{ess sup}_{\tau \in \mathcal{S}_n^N} \mathbf{E}_{X_n} G_i(X_\tau)$  and  $\sigma^{*i}$  be a stopping time such that  $h_i(0, x) = \mathbf{E}_x G_i(X_{\sigma^{*i}})$  for every  $x \in \mathbb{E}$ ,  $i = 1, 2$ . Let  $\Gamma_n^i = \{x \in \mathbb{E} : h_i(n, x) = G_i(x)\}$ . We have  $\sigma^{*i} = \inf\{n : X_n \in \Gamma_n^i\}$  (cf. [43]). Denote  $\sigma_k^{*i} = \inf\{n > k : X_n \in \Gamma_n^i\}$ . The restricted sets of strategies  $\hat{\mathbf{A}}^N = \{(\lambda^1(\hat{q}), \{\sigma_n^{*1}\}) : \hat{q} \in \mathcal{Q}^N, \{\sigma_n^{*1}\} \in \mathcal{S}_{n+1}^N \text{ for every } n\}$  and  $\hat{\mathbf{M}}^N = \{(\lambda^2(\hat{q}), \{\sigma_n^{*2}\}) : \hat{q} \in \mathcal{Q}^N, \{\sigma_n^{*2}\} \in \mathcal{S}_{n+1}^N \text{ for every } n\}$  for Player 1 and Player 2, respectively, in the random priority game will be considered, taking into account the definition of optimal behaviour of the players when one player has chosen a state. For a given  $\hat{q} \in \mathcal{Q}^N$  the pair of strategies  $(\hat{\mathbf{s}}, \hat{\mathbf{t}}) \in \hat{\mathbf{A}}^N \times \hat{\mathbf{M}}^N$  is defined. Let

$$\begin{aligned} \varphi_1(\hat{q}) &= \varphi_1(\hat{\mathbf{s}}, \hat{\mathbf{t}}) = G_1(X_{\lambda^1(\hat{q})}) \mathbb{1}_{\{\lambda^1(\hat{q}) < \lambda^2(\hat{q})\}} + \tilde{h}_1(\lambda^2(\hat{q}), X_{\lambda^2(\hat{q})}) \mathbb{1}_{\{\lambda^1(\hat{q}) > \lambda^2(\hat{q})\}} \\ &\quad + \left[ G_1(X_{\lambda^1(\hat{q})}) \alpha_{\lambda^1(\hat{q})} + \tilde{h}_1(\lambda^1(\hat{q}), X_{\lambda^1(\hat{q})}) (1 - \alpha_{\lambda^1(\hat{q})}) \right] \mathbb{1}_{\{\lambda^1(\hat{q}) = \lambda^2(\hat{q})\}}, \end{aligned} \tag{15a}$$

$$\begin{aligned} \varphi_2(\hat{q}) &= \varphi_2(\hat{\mathbf{s}}, \hat{\mathbf{t}}) = G_2(X_{\lambda^2(\hat{q})}) \mathbb{1}_{\{\lambda^2(\hat{q}) < \lambda^1(\hat{q})\}} + \tilde{h}_2(\lambda^1(\hat{q}), X_{\lambda^1(\hat{q})}) \mathbb{1}_{\{\lambda^2(\hat{q}) > \lambda^1(\hat{q})\}} \\ &\quad + \left[ G_2(X_{\lambda^1(\hat{q})}) (1 - \alpha_{\lambda^1(\hat{q})}) + \tilde{h}_2(\lambda^1(\hat{q}), X_{\lambda^1(\hat{q})}) \alpha_{\lambda^1(\hat{q})} \right] \mathbb{1}_{\{\lambda^1(\hat{q}) = \lambda^2(\hat{q})\}} \end{aligned} \tag{15b}$$

for each  $\hat{q} \in \mathcal{Q}^N$ , where  $\tilde{h}_i(n, X_n) = \text{ess sup}_{\tau \in \mathcal{S}_{n+1}^N} \mathbf{E}_{X_n} G_i(X_\tau) = \mathbf{E}_{X_n} h_i(n+1, X_{n+1})$ . Denote  $R_i(x, \hat{q}) = \mathbf{E}_x \varphi_i(\hat{q})$  for every  $x \in \mathbb{E}$ ,  $i = 1, 2$ . For a given  $\hat{q} \in \hat{\mathcal{Q}}_n^N$ ,  $n = 0, 1, \dots, N$ , define

$$\begin{aligned} u_i^{ss}(n, x, \hat{q}) &= \mathbf{E}_{X_n} \varphi_i(\hat{q}) \mathbb{1}_{\{\lambda^1(\hat{q}) = \lambda^2(\hat{q}) = n, X_n = x\}}, \\ u_i^{sc}(n, x, \hat{q}) &= \mathbf{E}_{X_n} \varphi_i(\hat{q}) \mathbb{1}_{\{\lambda^1(\hat{q}) = n, \lambda^2(\hat{q}) > n, X_n = x\}}, \\ u_i^{cs}(n, x, \hat{q}) &= \mathbf{E}_{X_n} \varphi_i(\hat{q}) \mathbb{1}_{\{\lambda^1(\hat{q}) > n, \lambda^2(\hat{q}) = n, X_n = x\}}, \\ u_i^{cc}(n, x, \hat{q}) &= \mathbf{E}_{X_n} \varphi_i(\hat{q}) \mathbb{1}_{\{\lambda^1(\hat{q}) > n, \lambda^2(\hat{q}) > n, X_n = x\}}. \end{aligned} \tag{16}$$

**Definition 3.2.** A pair of strategies  $(\hat{\mathbf{s}}^*, \hat{\mathbf{t}}^*) \in \hat{\mathbf{A}}^N \times \hat{\mathbf{M}}^N$  is a correlated equilibrium in the random priority stopping game if  $\hat{q}^* \in \mathcal{Q}^N$  fulfills the conditions given in Definition 2.7 with  $u_i$  defined by (16).

The correctness of the definition is a consequence of Theorem 2.3.

#### 4. Selection of equilibria in a game based on the best choice problem

We now consider examples based on a game theoretic version of the best choice problem. There are many versions of the problem considered by many authors, e.g. Ben Abdelaziz and Krichen [2], Fushimi [13], Mazalov [21], Sakaguchi [37,39,40], Szajowski [51] (see [32,41] for reviews of alternative approaches). Two players observe a sequence of  $N$  objects presented in random order. The  $n$ th object is observed at moment  $n$ , ( $1 \leq n \leq N$ ). The objective of each player is to obtain the most valuable object. They can only observe the relative rank of an object with respect to the objects already seen. If the  $n$ th object is the  $k$ th best seen so far, then its relative rank  $R_n$  is equal to  $k$ . At each moment  $n$  both players have two possible actions, reject the object and continue inspecting (denoted  $c$ ) and accept the object (denoted  $s$ ). Each player can obtain at most one object and on obtaining an object ceases searching. If just one player wishes to accept an object, then he obtains that object. If both players wish to obtain an object, then Player 1 obtains it with probability  $\alpha$ ,  $\alpha \in (0, 1)$ , otherwise Player 2 obtains it. The player not obtaining the object is free to carry on inspecting

the sequence. Without loss of generality we can assume  $\alpha \geq 0.5$ . Players should only accept an object, if its relative rank is 1 (it is the best object seen so far). Such objects will be known as candidates.

The payoff matrix for the game played when both players are still searching and a candidate appears at moment  $n$  is denoted by  $M_n$ . A vector of values of such a game to the players is denoted by  $(v_1(n, \hat{q}), v_2(n, \hat{q}))$ . Let  $W_n$  denote the optimal expected reward of a lone searcher after rejecting  $n$  objects. Considering the arrival time of the first candidate after time  $n$ , denoted by  $k$ ,  $u_i^{cc}(n, \hat{q}) = \sum_{k=n+1}^N \frac{nv_i(k, \hat{q})}{k(k-1)}$ ,  $i \in \{1, 2\}$ . It should be noted that the functions  $u_i^{cc}$  and  $v_i$  are independent of  $x_n$  and this is reflected in the simplified notation.  $M_n$  is given by

$$M_n = \begin{pmatrix} u_1^{ss}(n, \hat{q}), u_2^{ss}(n, \hat{q}) & u_1^{sc}(n, \hat{q}), u_2^{sc}(n, \hat{q}) \\ u_1^{cs}(n, \hat{q}), u_2^{cs}(n, \hat{q}) & u_1^{cc}(n, \hat{q}), u_2^{cc}(n, \hat{q}) \end{pmatrix} = \begin{pmatrix} \frac{n\alpha}{N} + (1 - \alpha)W_n, \frac{n(1-\alpha)}{N} + \alpha W_n & \frac{n}{N}, W_n \\ W_n, \frac{n}{N} & u_1^{cc}(n, \hat{q}), u_2^{cc}(n, \hat{q}) \end{pmatrix}. \tag{17}$$

It can be seen that in order to solve the 2-player game we must first solve the problem faced by a lone searcher. This is the standard secretary (best choice) problem. For finite  $N$  the optimal reward function  $W_n$  can be calculated by recursion. We have  $W_N = 0$  and

$$W_{n-1} = \frac{(n-1)W_n}{n} + \frac{1}{n} \max \left\{ \frac{n}{N}, W_n \right\}.$$

It can be shown that a lone searcher should accept a candidate iff  $n \geq n^*$ , where  $n^*$  is the smallest  $n$  satisfying  $\frac{n}{N} \geq W_n$ . The values of the game played at a decision point,  $v_i(n, \hat{q})$ , and the expected rewards from future search,  $u_i^{cc}(n, \hat{q})$ , are also calculated recursively. The criteria for selecting an equilibrium are based on the expected rewards from future search  $u_i^{cc}(n, \hat{q})$ . Conditioning on whether the  $n$ th object is a candidate or not, it follows that

$$u_i^{cc}(n-1, \hat{q}) = \frac{1}{n} v_i(n, \hat{q}) + \frac{n-1}{n} u_i^{cc}(n, \hat{q}).$$

For ease of defining an equilibrium, it is assumed that if a player is indifferent to accepting and rejecting an object, then he accepts that object. The pair of actions  $(s, s)$  constitutes a Nash equilibrium in the game given by  $M_n$  when  $u_1^{ss}(n, \hat{q}) \geq u_1^{cs}(n, \hat{q})$  and  $u_2^{ss}(n, \hat{q}) \geq u_2^{sc}(n, \hat{q})$ . It follows that  $(s, s)$  is a Nash equilibrium when  $n/N \geq W_n$ . The strategy pair  $(s, c)$  is a Nash equilibrium when  $u_1^{sc}(n, \hat{q}) \geq u_1^{cc}(n, \hat{q})$  and  $u_2^{sc}(n, \hat{q}) > u_2^{ss}(n, \hat{q})$ . It follows that  $u_1^{sc}(n, \hat{q}) \leq n/N < W_n$ . Similarly,  $(c, s)$  is a Nash equilibrium when  $u_2^{sc}(n, \hat{q}) \leq n/N < W_n$ . The strategy pair  $(c, c)$  is a Nash equilibrium when  $n/N < u_i^{cc}(n, \hat{q})$  for  $i \in \{1, 2\}$ .

It can be shown by recursion that the (intuitively clear) inequality  $u_i^{cc}(n, \hat{q}) \leq W_n$  holds for  $n \leq N$ . Hence, considering the conditions given above, there are 2 pure Nash equilibria when  $u_1^{cc}(n, \hat{q}) \leq n/N < W_n$  and  $u_2^{cc}(n, \hat{q}) \leq n/N$ . In all other cases there is exactly one Nash equilibrium in pure strategies. In this case this is the unique correlated equilibrium. Hence, there is no problem of equilibrium selection.

The conditions for a correlated equilibrium in the game given by  $M_n$  can be expressed as follows:

$$\begin{aligned} \pi_n^{ss} [u_1^{ss}(n, \hat{q}) - u_1^{cs}(n, \hat{q})] + \pi_n^{sc} [u_1^{sc}(n, \hat{q}) - u_1^{cc}(n, \hat{q})] &\geq 0, \\ \pi_n^{cs} [u_1^{cs}(n, \hat{q}) - u_1^{ss}(n, \hat{q})] + \pi_n^{cc} [u_1^{cc}(n, \hat{q}) - u_1^{sc}(n, \hat{q})] &\geq 0, \\ \pi_n^{ss} [u_2^{ss}(n, \hat{q}) - u_2^{sc}(n, \hat{q})] + \pi_n^{cs} [u_2^{cs}(n, \hat{q}) - u_2^{cc}(n, \hat{q})] &\geq 0, \\ \pi_n^{sc} [u_2^{sc}(n, \hat{q}) - u_2^{ss}(n, \hat{q})] + \pi_n^{cc} [u_2^{cc}(n, \hat{q}) - u_1^{cs}(n, \hat{q})] &\geq 0. \end{aligned} \tag{18}$$

When both  $(c, s)$  and  $(s, c)$  constitute Nash equilibria it follows that  $\pi_n^{sc}$  and  $\pi_n^{cs}$  are always associated with non-negative coefficients in the inequalities given above. Hence, in this case, any distribution satisfying  $\pi_n^{sc} + \pi_n^{cs} = 1$  defines a correlated equilibrium (this is a sufficient, but not necessary, condition).

For finite  $N$  the definition of a correlated equilibrium based on one of the criteria given leads to a set of linear programming problems to be solved by recursion. The constraints are given by the inequalities (18), together with the conditions required for  $\pi_n^{ss}, \pi_n^{sc}, \pi_n^{cs}$  and  $\pi_n^{cc}$  to define a probability distribution referred to as  $\pi_n = (\pi_n^{ss}, \pi_n^{sc}, \pi_n^{cs}, \pi_n^{cc})$ . The objective function is defined by the criterion used. In the remainder of this article, distributions which are concentrated at one point will often be described by the pair of actions taken. For example the pair of actions  $(c, s)$  taken at moment  $n$  corresponds to  $\pi_n = (0, 0, 1, 0)$ .

## 5. Examples of equilibrium selection

In this section we consider equilibrium selection from the following sets

- (i) The class of pure equilibria.
- (ii) The class of mixed equilibria.
- (iii) The class obtained by randomization over mixed equilibria.
- (iv) The class of correlated equilibria.

The four selection criteria introduced in Section 2.3 are adopted.

### 5.1. Libertarian equilibria

#### 5.1.1. Libertarian 1 equilibria

At a Stepwise Libertarian 1 Equilibrium,  $u_1^{cc}(n-1, \hat{q})$  is maximized at each step in the recursion procedure. It can be seen from Eq. (4) that this is equivalent to the maximization of  $v_1(n, \hat{q})$ .

We first consider selection from the class of pure strategies. From the conditions for a Nash equilibria, it follows that for  $n \geq n^*$ , the only equilibrium is given by the pair  $(s, s)$ . This strategy pair constitutes the equilibrium chosen for  $n \geq n^*$ , independently of the criterion used.

Independently of  $n$  and  $\alpha$ , it follows from Eq. (17) and the equilibrium conditions that whenever  $(c, s)$  and  $(s, c)$  are both Nash equilibria, Player 1 obtains his maximum payoff when the strategy pair  $(c, s)$  is chosen. Since  $u_1^{cc}(n, \hat{q}_{L1}^*) > u_2^{cc}(n, \hat{q}_{L1}^*)$ , the general form of the Stepwise Libertarian 1 Equilibrium is:  $(c, c)$  for  $n < n_1$ ,  $(c, s)$  for  $n_1 \leq n < n^*$  and  $(s, s)$  for  $n \geq n^*$ .

Since  $(c, s)$  is a Nash equilibrium, it is also a correlated equilibrium. Whenever the problem of equilibrium selection exists, arguing as above, Player's 1 payoff is maximized by selecting this strategy pair. Hence, the Stepwise Libertarian 1 Equilibrium from the class of pure strategies is also the Stepwise Libertarian 1 Equilibrium from the class of correlated equilibria (and hence from any of the classes considered). Since this is the unique Stepwise Libertarian 1 Equilibrium, it follows by the principles of dynamic programming, that it is also the Global Libertarian 1 Equilibrium.

#### 5.1.2. Libertarian 2 equilibria

The form of the Stepwise Libertarian 2 Equilibrium depends on the value of  $\alpha$ . Arguing as above, whenever two pure equilibria exist the stepwise choice criterion chooses the strategy pair  $(s, c)$ , which maximizes the expected payoff of Player 2. Define  $A_N$  to be the set of  $\alpha$  for which the expected payoff of Player 1 is greater than the expected payoff of Player 2 at any equilibrium when the horizon is  $N$  (see [51]). If  $\alpha \in A_N$ , the Stepwise Libertarian 2 Equilibrium is of the following form:  $(c, c)$  for  $n < n_1$ ,  $(c, s)$  for  $n_1 \leq n < n_2$ ,  $(s, c)$  for  $n_2 \leq n < n^*$  and  $(s, s)$  for  $n \geq n^*$ . Otherwise, the Stepwise Libertarian 2 Equilibrium is of the form:  $(c, c)$  for  $n < n_1$ ,  $(s, c)$  for  $n_1 \leq n < n^*$  and  $(s, s)$  for  $n \geq n^*$ . Since the Stepwise Libertarian 2 Equilibrium is unique, it is also the Global Libertarian 2 Equilibrium. Henceforth, such equilibria will simply be referred to as Libertarian Equilibria.

It should be noted that at moments when there are two Nash equilibria, at least one of the players accepts a candidate. At such points the sum of the expected payoffs of the players on the appearance of a candidate is  $n/N + W_n$ , regardless of which player obtains such a candidate. It follows that the stepwise maximization of one player's expected payoff is equivalent to the stepwise minimization of the other player's expected payoff subject to the equilibrium constraints.

#### 5.1.3. Asymptotic solutions

Let  $t$  be the proportion of objects already observed ( $t$  is referred to as the time). It follows that the asymptotic forms of these equilibria when  $N \rightarrow \infty$  are of the form (see [51]):

- (1) Libertarian 1:  $(c, c)$  for  $t < t_1$ ,  $(c, s)$  for  $t_1 \leq t < e^{-1}$  and  $(s, s)$  for  $t \geq e^{-1}$ .
- (2) Libertarian 2 for  $\alpha \leq \alpha_0 \approx 0.5299$ :  $(c, c)$  for  $t < t_2$ ,  $(s, c)$  for  $t_2 \leq t < e^{-1}$  and  $(s, s)$  for  $t \geq e^{-1}$ .
- (3) Libertarian 2 for  $\alpha > \alpha_0$ :  $(c, c)$  for  $t < t_3$ ,  $(c, s)$  for  $t_3 \leq t < t_2$ ,  $(s, c)$  for  $t_2 \leq t < e^{-1}$  and  $(s, s)$  for  $t \geq e^{-1}$ ,

where  $t_1 = e^{-(2+\alpha)/2}$ ,  $t_2 = e^{-(3-\alpha)/2}$ ,  $t_3 = \exp[e^{(1-\alpha)/2} - 5/2]$ . The values of the games to the players at these equilibria are as follows:

- (1) Libertarian 1:  $u_1^{cc}(0, \hat{q}_L^*) = e^{-1} - \frac{t_1(1-\alpha)}{2}$ ,  $u_2^{cc}(0, \hat{q}_L^*) = t_1$
- (2) Libertarian 2 for  $\alpha \leq \alpha_0$ :  $u_1^{cc}(0, \hat{q}_L^*) = t_2$ ,  $u_2^{cc}(0, \hat{q}_L^*) = e^{-1} - \frac{\alpha t_2}{2}$
- (3) Libertarian 2 for  $\alpha > \alpha_0$ :  $u_1^{cc}(0, \hat{q}_L^*) = e^{-1} - [e^{(1-\alpha)/2} - 1]t_3$ ,  $u_2^{cc}(0, \hat{q}_L^*) = t_3$ .

### 5.2. Republican equilibria

A Stepwise Republican Equilibrium can be obtained by choosing the appropriate Libertarian Equilibrium. When  $\alpha = 0.5$ , the Libertarian 1 and Libertarian 2 Equilibria are both Stepwise Republican Equilibria. When  $\alpha > 0.5$ , the Libertarian 1 Equilibrium defines the unique Stepwise Republican Equilibrium. At each moment the future expected reward of the player whose payoff is being maximized is at least as great as the future expected reward of the other player. It follows that these stepwise equilibria are the only Global Republican Equilibria.

### 5.3. Utilitarian equilibria

The sum  $\sum_{i=1}^2 u_i^{cc}(n-1, \hat{q}_U^*)$  can be maximized by maximizing the sum  $\sum_{i=1}^2 v_i(n, \hat{q}_U^*)$  recursively.

#### 5.3.1. Stepwise utilitarian equilibria in the class of pure equilibria

Suppose  $(c, s)$  and  $(s, c)$  are both Nash equilibria at moment  $n$ . From the form of the payoff matrix, it follows that the sum of the expected payoffs is independent of the Nash equilibrium chosen. Therefore, in the class of pure equilibria any pure equilibrium is a Stepwise Utilitarian Equilibrium. However, the value of such a game (and hence the global form of the equilibrium) depends on the equilibrium chosen. Hence, a Stepwise Utilitarian Equilibrium need not be a Global Utilitarian Equilibrium.

#### 5.3.2. Stepwise utilitarian equilibria in the classes of mixed equilibria and randomization over mixed equilibria

Now consider the class of mixed strategies. Suppose both  $(c, s)$  and  $(s, c)$  are Nash equilibria at moment  $n$ . At any essentially mixed equilibrium there is a positive probability that the strategy pair  $(c, c)$  is played. It follows that at such an equilibrium the sum of the expected payoffs is less than the sum of the expected payoffs at a pure equilibrium. Hence, the set of Stepwise Utilitarian Equilibria is simply the set of pure equilibria.

Arguing in a similar manner, the set of Stepwise Utilitarian Equilibria in the class of randomizations over the set of mixed equilibria are randomizations satisfying  $\pi_n^{sc} + \pi_n^{cs} = 1$  (randomizations concentrated on the pure equilibria).

#### 5.3.3. Stepwise utilitarian equilibria in the class of correlated equilibria

Whenever both  $(c, s)$  and  $(s, c)$  are Nash equilibria, the linear programming problem to be solved is the maximization of

$$\sum_{i=1}^2 v_i(n, \hat{q}_U^*) = \left(W_n + \frac{n}{N}\right) (\pi_n^{ss} + \pi_n^{sc} + \pi_n^{cs}) + [u_1^{cc}(n, \hat{q}_U^*) + u_2^{cc}(n, \hat{q}_U^*)] \pi_n^{cc}$$

subject to the constraints given by the system of inequalities given by (18). The maximum (obtained by setting  $\pi_n^{ss} + \pi_n^{sc} + \pi_n^{cs} = 1$ ) is attainable, since any correlation with  $\pi_n^{sc} + \pi_n^{cs} = 1$  satisfies the constraints. It follows that  $\pi_n^{cc} = 0$ . Also, from the equilibrium conditions

$$\pi_n^{ss} \leq \min \left\{ \frac{\pi_n^{sc} [u_1^{sc}(n, \hat{q}) - u_1^{cc}(n, \hat{q})]}{u_1^{cs}(n, \hat{q}) - u_1^{ss}(n, \hat{q})}, \frac{\pi_n^{cs} [u_2^{cs}(n, \hat{q}) - u_2^{cc}(n, \hat{q})]}{u_2^{sc}(n, \hat{q}) - u_2^{ss}(n, \hat{q})} \right\}.$$

It should be noted that at each moment there are an infinite number of correlations satisfying the optimality criterion.



### 5.3.4. Global utilitarian equilibria

Now we derive the form of a correlated equilibrium which is always a Global Utilitarian Equilibrium. Consider two equilibria where at least one player always accepts a candidate for  $n > n_0$ . Furthermore, suppose that when  $n = n_0$  at the first equilibrium (denoted  $\hat{q}^1$ ) at least one player accepts a candidate and at the other equilibrium (denoted  $\hat{q}^2$ ) there is a positive probability that neither candidate accepts a candidate. From the arguments presented above  $u_1^{cc}(n_0, \hat{q}^1) + u_2^{cc}(n_0, \hat{q}^1) = u_1^{cc}(n_0, \hat{q}^2) + u_2^{cc}(n_0, \hat{q}^2)$ . Conditional on the event that both players reject a candidate at moment  $n_0$  at the second equilibrium, the sum of expected payoffs is given by  $u_1^{cc}(n_0 - 1, \hat{q}^2) + u_2^{cc}(n_0 - 1, \hat{q}^2) = u_1^{cc}(n_0, \hat{q}^2) + u_2^{cc}(n_0, \hat{q}^2)$ . At the first equilibrium

$$u_1^{cc}(n_0 - 1, \hat{q}^1) + u_2^{cc}(n_0 - 1, \hat{q}^1) = \frac{n_0 - 1}{n} [u_1^{cc}(n_0, \hat{q}^1) + u_2^{cc}(n_0, \hat{q}^1)] + \frac{1}{n_0} \left[ \frac{n_0}{N} + W_{n_0} \right].$$

At this equilibrium at least one of the two pairs of actions  $(c, s)$  and  $(s, c)$  must constitute a Nash equilibrium. Suppose  $(c, s)$  constitutes a Nash equilibrium. From the stability conditions it follows that  $n_0/N \geq u_2^{cc}(n_0, \hat{q}^1)$ . It can also be shown by induction that  $W_{n_0} > u_1^{cc}(n_0, \hat{q}^1)$ . Hence, it follows that the sum of the expected future payoffs at moment  $n_0 - 1$  is greater at the first equilibrium than at the second equilibrium. The proof that this inequality is also satisfied when  $(s, c)$  constitutes a Nash equilibrium is analogous.

It follows that an equilibrium which maximizes the period of time over which at least one player accepts a candidate is a Global Utilitarian Equilibrium. Considering the form of the payoff matrices and the stability conditions for at least one of  $(c, s)$  and  $(s, c)$  to be a Nash equilibrium, it can be seen that this can be achieved by minimizing the expected reward of the player with the least expected reward subject to the equilibrium conditions. This is equivalent to maximising the expected reward of the player with the greatest expected reward (i.e. the Libertarian 1 Equilibrium). However, it can be shown that, for example, the Libertarian 2 Equilibrium for  $N = 20$ ,  $\alpha = 0.6$  is also a Global Utilitarian Equilibrium. However, the asymptotic form of a Global Utilitarian Equilibrium for  $N \rightarrow \infty$  is unique (ignoring deviations on a set of zero measure). This is due to the fact that deviations will increase the payoff of the player obtaining the smaller payoff and hence decrease the length of the interval on which at least one player accepts a candidate.

It follows that the Libertarian 1 Equilibrium is a Global Utilitarian Equilibrium for all the classes of strategies considered.

### 5.4. Egalitarian equilibria

Szajowski [51] presents examples of equilibria which equalize the values of the game to the players. Such equilibria can be treated as a correlated equilibrium according to the definition of Solan and Vohra [48]. These strategies derived are not Stepwise Egalitarian Equilibria, since the appropriate maximization is not carried out at each moment  $n$ .

#### 5.4.1. Stepwise egalitarian equilibria in the class of pure equilibria

As in the case of Libertarian Equilibria, the form of such an equilibrium depends on  $\alpha$ . Suppose  $\alpha \in A_N$ . It follows that the Libertarian 2 Equilibrium is the unique Stepwise Egalitarian Equilibrium and hence the unique Global Egalitarian Equilibrium.

Now suppose  $\alpha \notin A_N$  and both  $(c, s)$  and  $(s, c)$  are pure Nash equilibrium for  $n_1 < n < n^*$ . In this case, stepwise application of the choice criterion selects  $(c, s)$  when  $u_2^{cc}(n, \hat{q}) > u_1^{cc}(n, \hat{q})$  and  $(s, c)$  when  $u_2^{cc}(n, \hat{q}) < u_1^{cc}(n, \hat{q})$ . When  $u_2^{cc}(n, \hat{q}) = u_1^{cc}(n, \hat{q})$ , either equilibrium can be chosen. We now show that the players alternate between  $(c, s)$  and  $(s, c)$  for  $n_1 < n \leq n_2$  (the initial strategy pair to be used on this set depends on  $n$  and  $\alpha$  and can be determined by the recursive procedure, it is assumed that the final strategy pair used on this set is  $(c, s)$ ) and  $(s, c)$  is used when  $n_2 < n < n^*$ . To prove this statement, it should be noted that from the selection criterion  $n_2$  is defined to be the first moment at which Player 2's expected payoff from future search exceeds Player 1's expected payoff from future search at the Libertarian 2 Equilibrium. Hence,  $(s, c)$  satisfies the selection criterion for  $n_2 < n < n^*$ . We now show that the strategy pair used should alternate on the set  $n_1 < n \leq n_2$ . First note that by assumption  $u_1^{cc}(n_2, \hat{q}) < u_2^{cc}(n_2, \hat{q})$  and  $u_1^{cc}(n_2 + 1, \hat{q}) \geq u_2^{cc}(n_2 + 1, \hat{q})$ . Assume that  $u_1^{cc}(n, \hat{q}) < u_2^{cc}(n, \hat{q})$  and  $u_1^{cc}(n + 1, \hat{q}) \geq u_2^{cc}(n + 1, \hat{q})$ . It follows that  $(c, s)$  and  $(s, c)$  should be chosen at moments  $n$  and  $n + 1$ , respectively. Thus,

$$u_1^{cc}(n, \hat{q}) - u_2^{cc}(n, \hat{q}) = \frac{1}{n+1} \left( \frac{n+1}{N} - W_{n+1} + n[u_1^{cc}(n+1, \hat{q}) - u_2^{cc}(n+1, \hat{q})] \right) > \frac{1}{N} - \frac{W_{n+1}}{n+1}.$$

Hence,

$$\begin{aligned} u_1^{cc}(n-1, \hat{q}) - u_2^{cc}(n-1, \hat{q}) &= \frac{1}{n} \left( W_n - \frac{n}{N} + (n-1)[u_1^{cc}(n, \hat{q}) - u_2^{cc}(n, \hat{q})] \right) \\ &> \frac{1}{n} \left( W_n - \frac{n}{N} + (n-1) \left[ \frac{1}{N} - \frac{W_{n+1}}{n+1} \right] \right) = \frac{2W_{n+1}}{n(n+1)} - \frac{1}{nN} > 0. \end{aligned}$$

This follows from the equilibrium conditions  $W_n = W_{n+1} > \frac{n+1}{N}$ .

Similarly, it can be shown that if  $u_1^{cc}(n, \hat{q}) > u_2^{cc}(n, \hat{q})$  and  $u_1^{cc}(n+1, \hat{q}) \leq u_2^{cc}(n+1, \hat{q})$ , then  $u_1^{cc}(n-1, \hat{q}) < u_2^{cc}(n-1, \hat{q})$ . It follows by reverse induction that the equilibrium is of the required form. It should be noted that since, in general, values of the game to the players cannot be equalized, at moment  $n_1$  either  $(c, s)$  or  $(s, c)$  could be a unique Nash equilibrium.

#### 5.4.2. Global egalitarian equilibria in the class of pure equilibria

In the case where  $\alpha \notin A_N$ , a Stepwise Egalitarian Equilibrium is not necessarily a Global Egalitarian Equilibrium. For example for  $N = 35$  and  $\alpha = 0.5$ , the following strategy is a Stepwise Egalitarian Equilibrium (by symmetry the other can be obtained by reversing the role of the players):  $(s, s)$  for  $n \geq 14$ ,  $(c, s)$  for  $n = 11$  and  $n = 13$ ,  $(s, c)$  for  $n = 12$  and  $(c, c)$  for  $n \leq 10$ . The values of the game to the players are  $u_1^{cc}(0, \hat{q}) = 0.2994$ ,  $u_2^{cc}(0, \hat{q}) = 0.2959$ .

In general, to derive the Global Egalitarian Equilibrium in pure equilibria, one should consider a set of order  $2^k$  pure equilibria, where  $k$  is the maximum number of moments at which two pure equilibria exist. In particular cases the number of equilibria to be considered can be reduced by simple arguments.

Considering the example given above, taking into account the symmetry of the game and the fact that the Global Utilitarian Equilibrium cannot be a Global Egalitarian Equilibrium, we must consider the three equilibria at which between  $n = 11$  and  $n = 13$ ,  $(c, s)$  is played twice and  $(s, c)$  once. Numerical calculation of these equilibria indicates that one of the two Global Egalitarian Equilibria is as follows (the other is obtained by reversing the role of the players):  $(s, s)$  for  $n \geq 14$ ,  $(c, s)$  for  $n = 12, 13$ ,  $(s, c)$  for  $n = 11$ ,  $(c, c)$  for  $n \leq 10$ . The values of the game to the players are  $u_1^{cc}(0, \hat{q}) = 0.2962$ ,  $u_2^{cc}(0, \hat{q}) = 0.2990$ .

Due to the discrete nature of the search space, it seems quite possible that any algorithm for finding the Global Egalitarian Equilibrium will be of exponential complexity with respect to  $N$  (since any stepwise equilibrium can be calculated by recursion with a bounded number of operations at each step, the calculation of such an equilibrium will be of linear complexity).

#### 5.4.3. Stepwise egalitarian equilibria in the class of mixed equilibria

Arguing as before, if  $\alpha \in A_N$ , then the Libertarian 2 Equilibrium is the unique Stepwise (and Global) Egalitarian Equilibrium.

When both  $(c, s)$  and  $(s, c)$  are Nash equilibria at moment  $n$ , there also exists a mixed equilibrium at which Player 1 plays  $s$  with probability  $p_n$  and Player 2 plays  $s$  with probability  $r_n$  ( $0 < p_n, r_n < 1$ ,  $p_n$  and  $r_n$  depend on the exact form of the equilibrium used, but for notational ease this dependence is not reflected in the notation). From the Bishop–Cannings Theorem [3], it can be shown that when a candidate appears at moment  $n$  the expected payoff of Player 1 is given by

$$r_n W_n + (1 - r_n) u_1^{cc}(n, q) = r_n \left[ \frac{\alpha n}{N} + (1 - \alpha) W_n \right] + \frac{(1 - r_n) n}{N} > \frac{n}{N}.$$

It follows that

$$r_n = \frac{n - N u_1^{cc}(n, q)}{\alpha n + N(1 - \alpha) W_n - N u_1^{cc}(n, q)}.$$

It can be shown in a similar way that the expected payoff of Player 2 at such a decision point is strictly greater than  $\frac{n}{N}$  (the minimum payoff at either of the pure Nash equilibria) and

$$p_n = \frac{n - Nu_2^{cc}(n, q)}{(1 - \alpha)n + N\alpha W_n - Nu_2^{cc}(n, q)}.$$

Hence, at the unique Stepwise Egalitarian Equilibrium, when  $\alpha = 0.5$  the mixed equilibrium is chosen whenever more than one Nash equilibrium exists. For the problem with  $N = 35$  and  $\alpha = 0.5$  it can be shown that  $(s, s)$  is used for  $n \geq 14$ ,  $p_{13} = 0.9706$ ,  $p_{12} = 0.7668$  and  $p_{11} = 0.4101$ ,  $r = p_i$ ,  $i = 11, 12, 13$ ,  $(c, c)$  is used for  $n \leq 10$  and  $u_i^{cc}(0, q) = 0.2956$ ,  $i = 1, 2$ . It follows that extending the class of strategies to mixed strategies actually lowers the value of the game to the players at the Stepwise Egalitarian Equilibrium!

Now consider  $\alpha \notin A_N$ , but  $\alpha \neq 0.5$  and suppose there are two pure Nash equilibria. Due to the discrete nature of the class of equilibria, it is difficult to make general statements about the form of the Stepwise Egalitarian Equilibrium without extensive calculations. However, we present an intuitive idea of the form of such an equilibrium.

Whenever

$$\frac{W_n}{n} + \frac{n-1}{n}u_2^{cc}(n, q) < \frac{1}{N} + \frac{n-1}{n}u_1^{cc}(n, q),$$

the stepwise mechanism chooses  $(s, c)$  [the expression on the left (right) hand side is the maximum (minimum) future expected reward of Player 2 (Player 1) before observing the  $n$ th object]. It follows that calculating recursively, the stepwise mechanism will initially choose  $(s, c)$  until the future expected payoffs of the two players are very similar. Thereafter, a mixed strategy will be chosen. Suppose equilibrium selection is required for  $n_1 \leq n < n^*$ . Then for  $n_1 \leq n \leq n_2$  a mixed strategy will be used and for  $n_2 < n < n^*$   $(s, c)$  will be used. It should be noted that at moment  $n_1 - 1$  it is possible that either  $(c, s)$  or  $(s, c)$  is the unique Nash equilibrium, since such a procedure cannot ensure equalization of the value of the game to the players.

#### 5.4.4. Global egalitarian equilibria in the class of mixed equilibria

In general the number of equilibria to select from is of order  $3^k$ , where  $k$  is the maximum number of moments at which equilibrium selection is required. Consider the problem with  $N = 35$  and  $\alpha = 0.5$ . In this class of equilibria when  $\alpha = 0.5$  the maximum number of moments at which equilibrium selection is required is the number of such moments at the Stepwise Egalitarian Equilibrium (at such an equilibrium the maximum expected payoff is minimized subject to the equilibrium conditions). The minimum number is achieved at the Republican Equilibrium. Since at both equilibria selection is required three times, there are  $3^3 = 27$  equilibria to select from. Using the symmetry of the game, this number is reduced to 14. Furthermore, it can be seen that an equilibrium at which only  $(s, c)$  and mixed equilibria are used gives Player 1 a lower expected payoff than at the Stepwise Egalitarian Equilibrium. It follows that it is sufficient to consider the following seven cases:

- (a) the three cases for which  $(c, s)$  is chosen twice and  $(s, c)$  once.
- (b) the three cases for which  $(c, s)$ ,  $(s, c)$  and a mixed equilibrium are all chosen once.
- (c) a mixed equilibrium is chosen for  $n = 11, 12, 13$ .

The four cases mentioned in (a) and (c) have already been considered. Consideration of the three remaining equilibria leads to the conclusion that the two Global Egalitarian Equilibria in the class of pure equilibria are also Global Egalitarian Equilibria in the class of mixed equilibria.

It seems likely that a Global Egalitarian Equilibrium in the class of mixed strategies will consist of a sequence of pure equilibria chosen in such a way as to minimize the difference between the payoffs of the two players, since using a mixed equilibrium risks the choice of  $(c, c)$ , which is the worst outcome for both players. It is possible that a mixed equilibrium is used at a small number of points, since switching between pure Nash equilibria does not ensure equalization of the players' payoffs.

#### 5.4.5. Stepwise egalitarian equilibria in the class of correlated equilibria

In order to derive a Stepwise Egalitarian Equilibrium, we have to recursively solve the following linear programming problem of maximizing the expected payoff of Player 2, namely

$$v_2(n, \hat{q}_E^*) = \pi_n^{ss}u_2^{ss}(n, \hat{q}_E^*) + \pi_n^{sc}u_2^{sc}(n, \hat{q}_E^*) + \pi_n^{cs}u_2^{cs}(n, \hat{q}_E^*) + \pi_n^{cc}u_2^{cc}(n, \hat{q}_E^*)$$

subject to the equilibrium conditions (18), together with the conditions stating that  $\pi_n$  is a probability distribution and that if both players are still searching at moment  $n$ , then the future expected reward of Player 2 is not greater than the future expected reward of Player 1. This condition can be expressed as

$$\frac{1}{n}[v_2(n, \hat{q}_E^*) + (n-1)u_2^{cc}(n, \hat{q}_E^*)] \leq \frac{1}{n}[v_1(n, \hat{q}_E^*) + (n-1)u_1^{cc}(n, \hat{q}_E^*)]. \quad (19)$$

It should be noted that, in general, we should also consider the analogous linear programming problem of maximizing the future expected reward of Player 1 subject to his future reward being not greater than the future expected reward of Player 2 and choose the appropriate solution from these two solutions. However, from the assumption that  $\alpha \geq 0.5$ , the definition of such an equilibrium and the convexity and closedness of the set of correlated equilibria, it follows that the future expected reward of Player 1 will never be smaller than the future expected reward of Player 2.

Suppose a non-trivial correlated equilibrium exists at moment  $n$ . The payoff of Player 2 is maximized by setting  $\pi_n^{sc} = 1$ . If at such an equilibrium Player 2's future expected payoff is less than Player 1's future expected payoff, then this is the unique Egalitarian correlation at moment  $n$ . Suppose the future expected payoffs at moment  $n$  are equal. A Stepwise Egalitarian Equilibrium can be obtained by simultaneously maximizing the sum of the payoffs and equalizing them. This may be achieved by setting  $\pi_j^{cs} = \pi_j^{sc} = 0.5$  at moments  $j$  ( $j < n$ ) where equilibrium selection is necessary. Note that this is not the only solution.

Finally, there may exist a moment  $j$  at which the payoffs become equalized. This may be achieved by setting  $\pi_j^{cs} = 1 - \pi_j^{sc}$ ,  $\pi_j^{ss} = \pi_j^{cc} = 0$  and calculating the value of  $\pi_j^{sc}$ , for which equality is obtained in Condition (19) (again, this will not be the unique solution).

If  $\alpha \in A_N$ , it follows that the Libertarian 2 Equilibrium is the unique Stepwise Egalitarian Equilibrium.

It should be noted that in cases where a Stepwise Egalitarian Equilibrium is not uniquely defined, such an equilibrium must equalize the expected future payoffs of the players subject to the condition that at least one of the players accepts a candidate. It follows that such a recursive definition of the value functions is well defined.

Considering the class obtained by randomization over mixed equilibria, the Stepwise Egalitarian Equilibrium is given by choosing one of the two pure Nash equilibria according to the procedure described above.

#### 5.4.6. Global egalitarian equilibria in the class of correlated equilibria

For  $\alpha \in A_N$  the Libertarian 2 Equilibrium is the unique Global Egalitarian Equilibrium. Whenever  $\alpha \notin A_N$ , one can derive the Global Egalitarian Equilibrium by considering the properties of Libertarian Equilibria. The Libertarian 1 Equilibrium is the Global Utilitarian Equilibrium (maximizes the sum of the values) and the Libertarian 2 Equilibrium maximizes the sum of the values subject to the condition that Player 2 obtains at least the same expected payoff as Player 1 (this strategy maximises the length of the period over which at least one player accepts a candidate given that Player 2 obtains at least the same payoff as Player 1). It follows that using a correlation device to randomize over the Libertarian Equilibria in such a way that the expected payoff of the players is equal constitutes the Global Egalitarian Equilibrium.

**Remark 5.1.** It should be noted that such a correlated strategy is a randomization over the set of pure Nash equilibria, that is to say that it can be obtained by using a correlating device. It is not of the form given by Definition 2.5. This definition assumes that the players choose their action at each moment based on a signal which is dependent only on the present state of the process and not on any of the previous signals. Hence, when using such a weak correlating device, the players have no recall of the signals sent at previous moments and so cannot base their actions on a signal sent before the game commences.

For example consider the game with  $N = 100$  and  $\alpha = 0.52$ . Calculating recursively, it can be shown that the following strategy is a Stepwise Egalitarian Equilibrium:  $(s, s)$  for  $n \geq 38$ ,  $(s, c)$  for  $32 \leq n < 38$ ,  $(0, 0.9437, 0.0563, 0)$  for  $n = 31$ ,  $(0, 0.5, 0.5, 0)$  for  $n = 30$  and  $(c, c)$  for  $n \leq 29$ . The value of the game to both players is 0.2930.

To calculate the Global Egalitarian Equilibrium, we first derive the Libertarian Equilibria. The Libertarian 1 Equilibrium is as follows:  $(s, s)$  for  $n \geq 38$ ,  $(c, s)$  for  $29 \leq n < 38$  and  $(c, c)$  for  $n \leq 28$ . The values of the game are given by  $u_1^{cc}(0, \hat{q}_{L1}) = 0.3027$ ,  $u_2^{cc}(0, \hat{q}_{L1}) = 0.2860$ . The Libertarian 2 Equilibrium is as follows:  $(s, s)$  for

$n \leq 38$ ,  $(s, c)$  for  $30 \leq n < 38$  and  $(c, c)$  for  $n \leq 29$ . The values of the game are given by  $u_1^{cc}(0, \hat{q}_{L2}) = 0.2918$ ,  $u_2^{cc}(0, \hat{q}_{L2}) = 0.2943$ .

The Global Egalitarian Equilibrium is obtained by randomizing over these equilibria in such a way as to equalize the values of the game to the players. Suppose that Libertarian 1 is played with probability  $p$ , otherwise Libertarian 2 is played. It follows that

$$p = \frac{u_2^{cc}(0, \hat{q}_{L2}^*) - u_1^{cc}(0, \hat{q}_{L2}^*)}{u_2^{cc}(0, \hat{q}_{L2}^*) - u_1^{cc}(0, \hat{q}_{L2}^*) + u_1^{cc}(0, \hat{q}_{L1}^*) - u_1^{cc}(0, \hat{q}_{L1}^*)}. \tag{20}$$

It follows that for this game  $p = 0.1340$  and  $u_i^{cc}(0, \hat{q}_E^*) = 0.2932$ ,  $i = 1, 2$ .

It can be seen that the selection criterion for a Global Egalitarian Equilibrium can be satisfied using a randomization over global Nash equilibria. Hence, this is the Global Egalitarian Equilibrium in the class obtained by randomization over mixed equilibria.

#### 5.4.7. Asymptotic results for the class of correlated equilibria

Szajowski [51] showed that the payoffs of the players can be equalized at equilibrium, iff  $\alpha \leq \alpha_0 \approx 0.5299$ . It follows that when  $\alpha > \alpha_0$  the Libertarian 2 Equilibrium presented earlier is the unique Stepwise (and Global) Egalitarian Equilibrium. For  $\alpha \in (0.5, \alpha_0)$ , a strategy of the following form can be used as a Stepwise Egalitarian Equilibrium:

$$\begin{aligned} \pi(t) &= (1, 0, 0, 0), & t \geq e^{-1}, \\ \pi(t) &= (0, 1, 0, 0), & t_2 \leq t < e^{-1}, \\ \pi(t) &= \left(0, \frac{1}{2}, \frac{1}{2}, 0\right), & t_1 \leq t < t_2, \\ \pi(t) &= (0, 0, 0, 1), & t < t_1. \end{aligned}$$

As far as we know, no equilibrium of such a form has yet been derived, hence we give a sketch of the derivation.

Let  $W(t)$  be the expected reward of a lone searcher after rejecting a proportion  $t$  of the objects. Gilbert and Mosteller [14] (see also [50]) showed that for  $t \geq e^{-1}$ ,  $W(t) = -t \ln t$  and for  $t < e^{-1}$ ,  $W(t) = e^{-1}$ . Define the future expected reward of Player  $i$  given that both players are still inspecting at time  $t$  to be  $V_i(t)$ . The payoff matrix for such a game on the appearance of a candidate is of the form

$$M(t) = \begin{pmatrix} \alpha t + (1 - \alpha)W(t), t(1 - \alpha) + \alpha W(t) & t, W(t) \\ W(t), t & V_1(t), V_2(t) \end{pmatrix}. \tag{21}$$

From the definition of  $u_i^{cc}(n, \hat{q})$  it follows that

$$V_i(t) = \int_t^1 \frac{tv_i(r) dr}{r^2}, \tag{22}$$

where  $(v_1(t), v_2(t))$  is the appropriate vector of the values of the game with payoff matrix  $M(t)$  to Players 1 and 2, respectively.

Considering the conditions for a Nash equilibrium, the strategy pair  $(s, s)$  constitutes the unique Nash equilibrium when  $t \geq W(t)$ . Thus,  $(s, s)$  is the Nash equilibrium when  $t \geq e^{-1}$ . On this time interval  $v_1(t) = \alpha t + (1 - \alpha)W(t)$  and  $v_2(t) = (1 - \alpha)t + \alpha W(t)$ . By dividing Eq. (22) by  $t$  and then differentiating with respect to  $t$ , we obtain two differential equations for  $V_1(t)$  and  $V_2(t)$ . Together with the boundary conditions  $V_1(1) = V_2(1) = 0$ , these differential equations lead to

$$\begin{aligned} V_1(t) &= \frac{-t \ln t}{2} [2\alpha - (1 - \alpha) \ln t], \\ V_2(t) &= \frac{-t \ln t}{2} [2(1 - \alpha) - \alpha \ln t]. \end{aligned}$$

This part of the solution is common to all the types of correlated equilibrium under investigation. For  $t_2 < t < e^{-1}$ , the pair of actions carried out is  $(s, c)$ . It follows that  $v_1(t) = t$  and  $v_2(t) = e^{-1}$ . Solving Eq. (22) and using the boundary conditions at  $t = e^{-1}$ , we obtain

$$V_1(t) = -\frac{(1-\alpha)t}{2} - t \ln t,$$

$$V_2(t) = e^{-1} - \frac{\alpha t}{2}.$$

At time  $t_2$  the future expected payoffs of the two players are equal. Hence,  $t_2$  satisfies  $V_1(t_2) = V_2(t_2)$ , which can be solved numerically. For  $t_1 < t < t_2$ ,  $v_1(t) = v_2(t) = 0.5(t + e^{-1})$ . Solving Eq. (22), together with the boundary conditions at  $t = t_2$ , we obtain

$$V_1(t) = V_2(t) = -\frac{t}{4} + \frac{e^{-1} - t \ln t}{2}.$$

At time  $t_1$ ,  $V_1(t_1) = t_1$ . This equation can be solved numerically. For  $t < t_1$ ,  $V_1(t) = V_2(t) = t_1 \approx 0.2908$ .

Calculating in a similar way, for  $\alpha = \alpha_0$  the unique Stepwise Egalitarian Equilibrium strategy is as follows:  $(s, s)$  for  $t \geq e^{-1}$ ,  $(s, c)$  for  $t_1 \leq t < e^{-1}$  and  $(c, c)$  for  $t < t_1$ .

Finally for  $\alpha = 0.5$  the following is a Stepwise Egalitarian Equilibrium:  $\pi(t) = (1, 0, 0, 0)$  for  $t \geq e^{-1}$ ,  $\pi(t) = (0, 0.5, 0.5, 0)$  for  $t_1 \leq t < e^{-1}$  and  $\pi(t) = (0, 0, 0, 1)$  for  $t < t_1$ .

It should be noted that  $t_1$ , the value of the game to both players, is independent of  $\alpha$  for  $\alpha \in [0.5, \alpha_0]$ .

Now we consider the Global Egalitarian Equilibrium for  $\alpha < \alpha_0$ . Such equilibria can be obtained by randomly choosing one of the two Libertarian Equilibria in such a way as to equalize the values of the game to the players. Considering the asymptotic forms of these equilibria, it follows from Eq. (20) that the probability  $p$  of selecting the Libertarian 1 Equilibrium is given by

$$p = \frac{2e^{-1} - t_2(2 + \alpha)}{4e^{-1} - t_2(2 + \alpha) - t_1(3 - \alpha)},$$

where

$$t_1 = \exp\left(-\frac{2 + \alpha}{2}\right); \quad t_2 = \exp\left(-\frac{3 - \alpha}{2}\right).$$

When  $\alpha = 0.5$ ,  $p = 0.5$ . The value of the game at the Global Egalitarian Equilibrium is approximately 0.2914 and the relative gain compared to the stepwise equilibrium is approximately 0.2%. In order to achieve such an equilibrium, it suffices to observe a signal at the beginning of the game. Forges [11] refers to such a method as a correlation device.

### 5.5. Conclusions and final remarks

It can be seen that for each of the criteria considered the class obtained by randomization over mixed equilibria is rich enough to contain stepwise and global solutions for the two-person best choice problems considered. This is due to the form of the payoff matrices, no advantage can be obtained by choosing  $(s, s)$  with a positive probability. However, in other games this may not be the case (see [1]). Also, even if an optimal solution is from a “simpler” class of strategies, selecting from the set of correlated strategies is always mathematically tractable, as it involves solving a set of linear programming problems.

Stepwise derivation of equilibria is simpler than finding a global optimum according to a given criterion, but may well lead to equilibria which are not globally optimal. However, the final example shows that for certain problems global optima can be derived by combining stepwise equilibria in an appropriate way.

Concerning selection criteria, there are other concepts that could be applied. For example, solutions to the bargaining problem. In this paper our intention has been to present the concepts of correlated equilibria and some new ideas of equilibrium selection in stopping games. We hope that our approach will induce researchers to compare such approaches. The important feature of these criteria is that they should be a consequence of



some pre-play discussion between players concerning their aims. Other concepts and comparison of methods will be subject of further research by the authors.

## References

- [1] R.J. Aumann, Subjectivity and correlation in randomized strategies, *Journal of Mathematical Economics* 1 (1974) 67–96.
- [2] F. Ben Abdelaziz, S. Krichen, An interactive method for the optimal selection problem with two decision makers, *European Journal of Operational Research* 162 (3) (2005) 602–609.
- [3] D.T. Bishop, C. Cannings, A generalized war of attrition, *Journal of Theoretical Biology* 70 (1978) 85–124.
- [4] A. Calvó-Armengol, The set of correlated equilibria of  $2 \times 2$  games, Tech. report, Department of Economics, Edifici B, 08193 Bellaterra (Barcelona), Spain, 2004, <http://selene.uab.es/acalvo/correlated.pdf>.
- [5] V. Domansky, Dynkin's games with randomized optimal stopping rules, in: K. Szajowski, A.S. Nowak (Eds.), *Advances in Dynamic Games: Applications to Economics, Finance, Optimization, and Stochastic Control* (Boston), *Annals of the International Society of Dynamic Games*, vol. 7, Birkhuser, 2005, pp. 247–262.
- [6] E.G. Enns, E.Z. Ferenstein, On a multi-person time-sequential game with priorities, *Sequential Analysis* 6 (1987) 239–256.
- [7] Fe S. Evangelista, T.E.S. Raghavan, A note on correlated equilibrium, *International Journal of Game Theory* 25 (1) (1996) 35–41.
- [8] E.Z. Ferenstein, Two-person non-zero-sum games with priorities, in: F. Thomas Bruss, Thomas S. Ferguson, Stephen M. Samuels (Eds.), *Strategies for Sequential Search and Selection in Real Time*, *Proceedings of the AMS-IMS-SIAM Joint Summer Research Conferences held June 21–27, 1990*, University of Massachusetts at Amherst, *Contemporary Mathematics*, vol. 125. Oxford University Press, 1992, pp. 119–133.
- [9] E.Z. Ferenstein, On randomized stopping game, in: K. Szajowski, A.S. Nowak (Eds.), *Advances in Dynamic Games: Applications to Economics, Finance, Optimization, and Stochastic Control* (Boston), *Annals of the International Society of Dynamic Games*, vol. 7, Birkhäuser, 2005, pp. 223–233.
- [10] T.S. Ferguson, Selection by committee, in: K. Szajowski, A.S. Nowak (Eds.), *Advances in Dynamic Games: Applications to Economics, Finance, Optimization, and Stochastic Control* (Boston), *Annals of the International Society of Dynamic Games*, vol. 7, Birkhäuser, 2005, pp. 203–209.
- [11] F. Forges, An approach to communication equilibria, *Econometrica* 54 (1986) 1375–1385.
- [12] F. Forges, Communication equilibria in repeated games with incomplete information, *Mathematics of Operations Research* 13 (2) (1988) 191–231.
- [13] M. Fushimi, The secretary problem in a competitive situation, *Journal of the Operations Research Society of Japan* 24 (1981) 350–358.
- [14] J.P. Gilbert, F. Mosteller, Recognizing the maximum of a sequence, *Journal of American Statistical Association* 61 (313) (1966) 35–73.
- [15] I. Gilboa, E. Zemel, Nash and correlated equilibria: Some complexity considerations, *Games and Economic Behavior* 1 (1) (1989) 80–93.
- [16] A. Greenwald, K. Hall, Correlated Q-learning, in: Tom Fawcett, Nina Mishra (Eds.), *Proceedings of Twentieth International Conference on Machine Learning (ICML-2003)*, August 21–24, The AAAI Press, Washington, DC, 2003, pp. 242–249.
- [17] R. Hendrickx, R. Peeters, J. Potters, A relation between Nash equilibria and correlated equilibria, *International Game Theory Review* 4 (4) (2002) 405–413.
- [18] P. Jean-Jacques Herings, Ronald J.A.P. Peeters, Equilibrium selection in stochastic games, *International Game Theory Review* 5 (4) (2003) 307–326.
- [19] P. Jean-Jacques Herings, Ronald J.A.P. Peeters, Stationary equilibria in stochastic games: Structure, selection, and computation, *Journal of Economic Theory* 118 (1) (2004) 32–60.
- [20] M. Kurano, M. Yasuda, J. Nakagami, Multi-variate stopping problem with a majority rule, *Journal of the Operations Research Society of Japan* 23 (1980) 205–223.
- [21] V.V. Mazalov, A game related to optimal stopping of two sequences of independent random variables having different distributions, *Mathematica Japonica* 43 (1996) 121–128.
- [22] J.-F. Mertens, Correlated- and communication equilibria, in: J.-F. Mertens et al. (Eds.), *Game-theoretic methods in general equilibrium analysis*. *Proceedings of the NATO Advanced Study Institute, Long Island, NY (USA), July 1–12, 1991*, NATO ASI Series. Series D. Behavioural and Social Sciences, vol. 77, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994, pp. 243–248.
- [23] H. Moulin, *Game Theory for the Social Sciences*, New York University Press, New York, 1986.
- [24] R.B. Myerson, Acceptable and predominant correlated equilibria, *International Journal of Game Theory* 15 (1986) 133–154.
- [25] R.B. Myerson, Multistage games with communication, *Econometrica* 54 (1986) 323–358.
- [26] R.B. Myerson, Communication, correlated equilibria and incentive compatibility, in: R.J. Aumann et al. (Eds.), *Handbook of Game Theory with Economic Applications*, *Handbooks in Economics*, vol. 2, Elsevier, Amsterdam, 1994, pp. 827–847.
- [27] R. Nau, S.G. Canovas, P. Hansen, On the geometry of Nash equilibria and correlated equilibria, *International Journal of Game Theory* (2004) 443–453.
- [28] P. Neumann, D. Ramsey, K. Szajowski, Randomized stopping times in Dynkin games, *ZAMM* 82 (11–12) (2002) 811–819.

- [29] A.S. Nowak, Existence of correlated weak equilibria in discounted stochastic games with general state space, *Stochastic Games and Related Topics. In Honor of Prof. L.S. Shapley, Proc. Workshop, Chicago/IL (USA) 1987* Theory Decis. Libr., Ser. C 7, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991, pp. 135–143.
- [30] A.S. Nowak, Correlated relaxed equilibria in nonzero-sum linear differential games, *Journal of Mathematical Analysis and Applications* 163 (1992) 104–112.
- [31] A.S. Nowak, T.E.S. Raghavan, Existence of stationary correlated equilibria with symmetric information for discounted stochastic games, *Mathematics of Operations Research* 17 (3) (1992) 519–526.
- [32] A.S. Nowak, K. Szajowski, Nonzero-sum stochastic games, in: T. Parthasarathy, M. Bardi, T.E.S. Raghavan (Eds.), *Stochastic and Differential Games. Theory and Numerical Methods* (Boston), Annals of the International Society of Dynamic Games, Birkhäuser, 1998, pp. 297–342.
- [33] Y. Ohtsubo, A nonzero-sum extension of Dynkin's stopping problem, *Mathematics of Operations Research* 12 (1987) 277–296.
- [34] H. Peters, K. Vrieze, Nash refinements of equilibria, *Journal of Optimization Theory and Applications* 83 (2) (1994) 355–373.
- [35] T. Radzik, On a new solution concept for bargaining problems, *Applied Mathematics* 25 (3) (1998) 285–294.
- [36] D. Rosenberg, E. Solan, N. Vieille, Stopping games with randomized strategies, *Probability Theory and Related Fields* 119 (2001) 433–451.
- [37] M. Sakaguchi, Non-zero-sum games related to the secretary problem, *Journal of the Operations Research Society of Japan* 23 (1980) 287–293.
- [38] M. Sakaguchi, A simplified two-person multi-stage poker with optimal stopping, *Mathematica Japonica* 28 (1983) 287–303.
- [39] M. Sakaguchi, Bilateral sequential games related to the no-information secretary problem, *Mathematica Japonica* 29 (1984) 961–974.
- [40] M. Sakaguchi, Some two-person bilateral games in the generalized secretary problem, *Mathematica Japonica* 33 (1988) 637–654.
- [41] M. Sakaguchi, Optimal stopping games—a review, *Mathematica Japonica* 42 (1995) 343–351.
- [42] M. Sakaguchi, Optimal stopping games where players have weighted privilege, *Game Theory and Application* 4 (2001) 116–131.
- [43] A.N. Shiryaev, *Optimal Stopping Rules*, Springer-Verlag, New York, Heidelberg, Berlin, 1978.
- [44] E. Shmaya, E. Solan, Two-player nonzero-sum stopping games in discrete time, *Annals of Probability* 32 (3B) (2004) 2733–2764.
- [45] E. Solan, Characterization of correlated equilibria in stochastic games, *International Journal of Game Theory* 30 (2) (2001) 259–277.
- [46] E. Solan, N. Vieille, Quitting games, *Mathematics of Operations Research* 26 (2) (2001) 265–285.
- [47] E. Solan, N. Vieille, Correlated equilibrium in stochastic games, *Games and Economic Behavior* 38 (2002) 362–399.
- [48] E. Solan, R.V. Vohra, Correlated equilibrium in quitting games, *Mathematics of Operations Research* 26 (3) (2001) 601–610.
- [49] S. Sorin, Communication, correlation and cooperation, in: S. Hart et al. (Eds.), *Cooperation: Game-theoretic approaches, Proceedings of the NATO ASI, SUNY, Stony Brook, NY, USA, July 18–29, 1994, NATO ASI Ser., Ser. F, Comput. Syst. Sci., vol. 155*, Springer, Berlin, 1997, pp. 199–218.
- [50] W.E. Stein, D.A. Seale, A. Rapoport, Analysis of heuristic solutions to the best choice problem, *European Journal of Operational Research* 151 (1) (2003) 140–152.
- [51] K. Szajowski, Markov stopping games with random priority, *Zeitschrift für Operations Research* 39 (1) (1994) 69–84.
- [52] K. Szajowski, M. Yasuda, Voting procedure on stopping games of Markov chain, in: Anthony H. Christer Shunji Osaki, C. Thomas Lyn (Eds.), *UK–Japanese Research Workshop on Stochastic Modelling in Innovative Manufacturing, July 21–22, 1995, Moller Centre, Churchill College, University of Cambridge, UK, Lecture Notes in Economics and Mathematical Systems, vol. 445*, Springer, 1996, pp. 68–80.
- [53] W. Thomson, Cooperative models of bargaining, in: R.J. Aumann et al. (Eds.), *Handbook of Game Theory with Economic Applications, Handbooks in Economics, vol. 2*, Elsevier, Amsterdam, 1994, pp. 1237–1284.
- [54] B. Tolwiński, A. Haurie, G. Leitmann, Cooperative equilibria in differential games, *Journal of Mathematical Analysis and Applications* 119 (1986) 182–202.
- [55] M. Yasuda, On a randomized strategy in Neveu's stopping problem, *Stochastic Processes and their Applications* 21 (1985) 159–166.
- [56] M. Yasuda, J. Nakagami, M. Kurano, Multi-variate stopping problem with a monotone rule, *Journal of the Operations Research Society of Japan* 25 (1982) 334–350.